G. L. Mullen
H. Stichtenoth
H. Tapia-Recillas
*Editors*

# Finite Fields with Applications to Coding Theory, Cryptography and Related Areas

Springer

# Finite Fields with Applications to Coding Theory, Cryptography and Related Areas

Springer
*Berlin*
*Heidelberg*
*New York*
*Hong Kong*
*London*
*Milan*
*Paris*
*Tokyo*

Gary L. Mullen
Henning Stichtenoth
Horacio Tapia-Recillas
*Editors*

# Finite Fields with Applications to Coding Theory, Cryptography and Related Areas

Proceedings of the Sixth International Conference on Finite Fields and Applications, held at Oaxaca, México, May 21–25, 2001

Springer

*Editors*

Gary L. Mullen
The Pennsylvania State University
Department of Mathematics
16802 University Park, PA, USA
e-mail: mullen@math.psu.edu

Henning Stichtenoth
Universität Essen
FB6 Mathematik und Informatik
45117 Essen, Deutschland
e-mail: stichtenoth@uni-essen.de

Horacio Tapia-Recillas
Universidad Autónoma Metropolitana, México
Departamento de Matemáticas
Iztapalapa 09340
Av. San Rafael Atlixco 186
09340 México, D.F., México
e-mail: htr@xanum.uam.mx

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Finite fields with applications to coding theory, cryptography and related areas: proceedings of the Sixth International Conference on Finite Fields and Applications, held at Oaxaca, México, May 21–25, 2001 / Gary L. Mullen... ed. – Berlin; Heidelberg; New York; Hong Kong; London; Milan; Paris; Singapore; Tokyo: Springer, 2002
ISBN-13:978-3-642-63976-0 e-ISBN-13:978-3-642-59435-9
DOI: 10.1007/978-3-642-59435-9

Mathematics Subject Classification (2000):
primary: 11Txx; secondary: 05-XX, 51Exx, 94Axx, 94Bxx

ISBN-13:978-3-642-63976-0



Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science+Business Media GmbH

http://www.springer.de



Softcover reprint of the hardcover 1st edition 2002



production: PRO EDIT GmbH, Heidelberg
Typeset by the authors using a Springer TEX macro package
Cover design: *design & production* GmbH, Heidelberg

Printed on acid-free paper SPIN 10881880 46/3142hs – 5 4 3 2 1 0

# Preface

This volume represents the refereed proceedings of the "**Sixth International Conference on Finite Fields and Applications** ($Fq6$)" held in the city of Oaxaca, México, between 22–26 May 2001. The conference was hosted by the Departmento do Matemáticas of the Universidad Autónoma Metropolitana-Iztapalapa, México. This event continued a series of biennial international conferences on Finite Fields and Applications, following earlier meetings at the University of Nevada at Las Vegas (USA) in August 1991 and August 1993, the University of Glasgow (Scotland) in July 1995, the University of Waterloo (Canada) in August 1997, and at the University of Augsburg (Germany) in August 1999. The Organizing Committee of $Fq6$ consisted of Dieter Jungnickel (University of Augsburg, Germany), Neal Koblitz (University of Washington, USA), Alfred Menezes (University of Waterloo, Canada), Gary Mullen (The Pennsylvania State University, USA), Harald Niederreiter (National University of Singapore, Singapore), Vera Pless (University of Illinois, USA), Carlos Rentería (IPN, México), Henning Stichtenoth (Essen University, Germany), and Horacio Tapia-Recillas, Chair (Universidad Autónoma Metropolitan-Iztapalapa, México).

The program of the conference consisted of four full days and one half day of sessions, with 7 invited plenary talks, close to 60 contributed talks, basic courses in finite fields, cryptography and coding theory and a series of lectures at local educational institutions.

Finite fields have an inherently fascinating structure and they are important tools in discrete mathematics. Their applications range from combinatorial design theory, finite geometries, and algebraic geometry to coding theory, cryptology, and scientific computing. A particularly fruitful aspect is the interplay between theory and applications which has led to many new perspectives in research on finite fields. This interplay has been a dominant theme in earlier $F_q$ conferences and was very much in evidence at $Fq6$. Applied or applications-oriented topics accounted for a significant part of the program.

These proceedings reflect the wide variety of topics represented at the conference. Most invited talks and a good proportion of the contributed talks are on permanent record here. All contributed talks were screened before the conference and all full papers were carefully refereed. We would like to take this opportunity to thank the members of the Organizing Committee and all referees who helped in these tasks. These colleagues contributed enormously to the quality of the conference presentations and to guaranteeing high standards for these proceedings.

We greatly appreciate the generous financial support received for the conference. A fair portion of the funds were provided by a grant from the Consejo

Nacional de Ciencia y Tecnología (CONACYT), México and from various offices of the host institution. We also thank Universidad Benito Juárez de Oaxaca, Instituto Tecnológico de Oaxaca, Dirección General de Servicios do Cómputo Académico-UNAM, Instituto Politécnico Nacional, Sociedad Matemática Mexicana, Certicom Corp., Institute of Combinatorics and Applications, and Red de Criptología (CONACYT-UAM) for diverse kinds of support.

We are grateful to various offices of the state of Oaxaca who helped with additional funds and organizational issues. Thanks are also due to the Governor of the state of Oaxaca, who gave a reception for the participants in the splendid setting of the Centro Cultural Santo Domingo in the city of Oaxaca. Last but not least, the highly efficient and friendly manner in which the conference took place would not have been possible without the enthusiasm and hard work by the assistants, secretaries and students who saw to many details involved in such a major event; we are grateful to all of them.

Regarding the present proceedings, we thank Dr. Martin Peters of Springer-Verlag who gave us the opportunity to edit this volume with a top publisher and in an attractive form. Working with him and all the staff at Springer-Verlag is always a pleasure.

Finally, we are pleased to confirm that the *Fq* series will continue with *Fq*7 in Toulouse, France in May 2003. We expect another lively and stimulating meeting there, which should, like the previous conferences, serve as an important meeting place for theoretical as well as applied aspects of finite fields. We hope to see you there!

May 2002

Horacio Tapia-Recillas
Gary Mullen
Henning Stichtenoth

# Contents

# Commutative Semifields of Rank 2 Over Their Middle Nucleus

Simeon Ball[1]* and Michel Lavrauw[2]

[1] Queen Mary, University of London, London, E1 4NS, United Kingdom
[2] Eindhoven University of Technology, Eindhoven, 5600MB, The Netherlands

**Abstract.** This article is about finite commutative semifields that are of rank 2 over their middle nucleus, the largest subset of elements that is a finite field. These semifields have a direct correspondence to certain flocks of the quadratic cone in $PG(3,q)$ and to certain ovoids of the parabolic space $Q(4,q)$. We shall consider these links, the known examples and non-existence results.

## 1 Semifields

A *finite semifield* $\mathcal{S}$ is a finite algebraic system that possesses two binary operations, addition and multiplication, which satisfy the following axioms.

(S1) Addition is a group with identity 0.
(S2) $a(b+c) = ab + ac$ and $(a+b)c = ac + bc$ for all $a$, $b$, $c \in \mathcal{S}$.
(S3) There exists an element $1 \neq 0$ such that $1a = a = a1$ for all $a \in \mathcal{S}$.
(S4) If $ab = 0$ then either $a = 0$ or $b = 0$.

Throughout this article the term semifield will refer to a finite semifield. The additive group of a semifield must be commutative. By (S2),

$$(ac + ad) + (bc + bd) = (a + b)(c + d) = (ac + bc) + (ad + bd).$$

Hence, $ad + bc = bc + ad$ and any elements that can be written as products commute under addition. By (S4) and finiteness, any element of $\mathcal{S}$ can be written as a product and so it follows that the additive group is abelian. Moreover it is not difficult to show that the group is elementary abelian. Let $a \neq 0$, and let $p$ be the additive order of $a$. If $p$ is not prime then we can write $p = rs$ for $r$ and $s$ integers not equal to 1, and by observing that $0 = (pa)a = (rsa)a = (ra)(sa)$ we get a contradiction from (S4). The fact that every nonzero element has prime order suffices to show that the group is elementary abelian, and that all nonzero elements have the same prime order $p$. This number $p$ is the *characteristic* of the semifield. An elementary abelian group can be viewed as a vector space over a finite field. In particular $\mathcal{S}$ has $p^n$ elements where $n$ is the dimension of $\mathcal{S}$ over the field $GF(p)$. There are

* The author acknowledges the support of an EPSRC (UK) Advanced Research Fellowship AF/990 480.

many examples of semifields known and some standard constructions can be found in Knuth [19]. If the order is $p$, the semifield must be $GF(p)$. If the order is $p^2$, the semifield is $GF(p^2)$. This is not difficult to see. Let $\{1, x\}$ be a basis for the semifield. Multiplication is determined by $x^2 = ax + b$ and the polynomial $x^2 - ax - b$ has no roots in $GF(p)$ else we would have $x^2 - ax - b = (x - r)(x - s) = 0$ contradicting (S4). Thus $x^2 - ax - b$ is irreducible and the multiplication is $GF(p^2)$. This short argument comes again from [19] where it is also determined that the only semifield of order 8 is $GF(8)$. And completing the question of existence Albert [1] and Knuth [19] construct semifields that are not finite fields for every other order $q = p^h$, that is $h \geq 3$ if $p$ is odd and $h \geq 4$ if $p = 2$.

The major motivation to study semifields in the 1960's was their use in the construction of projective planes, see Hughes and Piper [16] or Hall [15]. Every semifield determines a projective plane and the projective plane is Desarguesian if and only if the semifield is a field. The incidence structure constructed from a semifield $\mathcal{S}$ with

$$\begin{array}{lll} \text{Points: } (0,0,1) & \text{Lines: } [0,0,1] & \\ (0,1,a) & [0,1,a] & a \in \mathcal{S} \\ (1,a,b) & [1,a,b] & a,\, b \in \mathcal{S} \end{array}$$

such that the point $(x_1, x_2, x_3)$ is incident with the line $[y_1, y_2, y_3]$ if and only if

$$y_1 x_3 = x_2 y_2 + x_1 y_3$$

is a projective plane $\pi(\mathcal{S})$ of order $|\mathcal{S}|$. It is a simple matter to check that any two points of $\pi(\mathcal{S})$ are incident with a unique line and dually that any two lines of $\pi(\mathcal{S})$ are incident with a unique point and hence that $\pi(\mathcal{S})$ is a projective plane. However it is harder to determine when two semifields $\mathcal{S}$ and $\mathcal{S}'$ determine the same projective plane, i.e. $\pi(\mathcal{S}) \cong \pi(\mathcal{S}')$. In [19] Knuth defines an *isotopism* from $\mathcal{S}$ to $\mathcal{S}'$ and shows that an isotopism is equivalent to a set of three 1-1 maps $(F, G, H)$ linear over $GF(p)$ from $\mathcal{S}$ to $\mathcal{S}'$, such that

$$(ab)H = (aF)(bG)$$

for all $a$, $b$, $c \in \mathcal{S}$. Two semifields $\mathcal{S}$ and $\mathcal{S}'$ are *isotopic* if there is an isotopism from $\mathcal{S}$ to $\mathcal{S}'$. We have the following theorem due to Albert, a proof of which can be found in [19].

**Theorem 1.** *Two semifields coordinatize the same projective plane if and only if they are isotopic.*

In his original work on semifields Dickson [12] considered constructing commutative semifields, that is semifields that satisfy

(S5) $ab = ba$ for all $a$ and $b$ in $\mathcal{S}$.

We define the *middle nucleus* of a commutative semifield to be

$$\mathcal{N} := \{x \mid (ax)b = a(xb),\ \forall a, b \in \mathcal{S}\}.$$

It is clear that $\mathcal{N}$ contains the field $GF(p)$ where $p$ is the characteristic and that $\mathcal{N}$ is itself a finite field. Moreover, $\mathcal{S}$ can be viewed as a vector space over its middle nucleus. Dickson [13] gave a construction of a commutative semifield of rank 2 over its middle nucleus. It is as follows. Let $\mathcal{S} := \{(x, y) \mid x, y \in GF(q)\}$ and let $\sigma$ be an automorphism of $GF(q)$ where $q$ is odd. Addition is defined component-wise and multiplication by

$$(x, y)(u, v) = (xv + yu, yv + mx^{\sigma}u^{\sigma})$$

where $m$ is a non-square in $GF(q)$. The only axiom that requires much thought is (S4) and we shall check this in a more general setting shortly. In this article we shall only be concerned with commutative semifields that are of rank 2 over their middle nucleus which have a correspondence with certain useful geometric objects.

Cohen and Ganley [10] made significant progress in the investigation of commutative semifields of rank 2 over their middle nucleus. They put Dickson's construction in the following more general setting. Let $\mathcal{S}$ be a commutative semifield of order $q^2$ with middle nucleus $GF(q)$. Then there is an $\alpha \in \mathcal{S} \setminus GF(q)$ such that $\{1, \alpha\}$ is a basis for $\mathcal{S}$. Addition in $\mathcal{S}$ is component-wise and multiplication is defined as

$$\begin{aligned}(x, y)(u, v) = (x\alpha + y)(u\alpha + v) &= xu\alpha^2 + (xv + yu)\alpha + yv \qquad (1)\\ &= (xv + yu + g(xu), yv + f(ux))\end{aligned}$$

where $x\alpha^2 = g(x)\alpha + f(x)$, $f$ and $g$ are functions from $GF(q) \to GF(q)$. The distributive laws are satisfied if and only if both $f$ and $g$ are linear maps, in other words, $f(x + y) = f(x) + f(y)$ and $g(x + y) = g(x) + g(y)$ for all $x$, $y$ in $GF(q)$. Thus we must check (S4). Suppose that

$$(x\alpha + y)(u\alpha + v) = 0$$

and that $x$, $y$, $u$ and $v$ are non-zero. It follows that

$$g(xu) + xv + yu = 0$$

and

$$f(xu) + yv = 0$$

and eliminating $y$ that

$$xv^2 + vg(xu) - uf(xu) = 0.$$

Writing $xu = z$ and $v/u = w$

$$zw^2 + g(z)w - f(z) = 0.$$

If one or more of $x$, $y$, $u$ or $v$ is zero it follows immediately that at least one of $(x, y)$ or $(u, v)$ is $(0, 0)$. Hence we have proved the following theorem which comes from [10].

**Theorem 2.** *Let $\mathcal{S}$ be a commutative semifield of rank 2 over its middle nucleus $GF(q)$. Then there exist linear functions $f$ and $g$ such that multiplication in $\mathcal{S}$ is defined as in (1) and $zw^2 + g(z)w - f(z) = 0$ has no solutions for all $w$, $z \in GF(q)$ and $z \neq 0$.*

If $q$ is odd then this quadratic in $w$ will have no solutions in $GF(q)$ if and only if

$$g(z)^2 + 4zf(z)$$

is a non-square for all $z \in GF(q)^*$. Cohen and Ganley [10] prove the following theorem for $q$ even.

**Theorem 3.** *For $q$ even the only commutative semifield of rank 2 over its middle nucleus $GF(q)$ is the finite field $GF(q^2)$.*

In light of this theorem we restrict ourselves to the case $q$ is odd.

Let us consider again the example of Dickson. We have $g = 0$ and $f(z) = mz^\sigma$ where $m$ is a non-square. We had only to check that (S4) is satisfied and this is clear since $g(z)^2 + 4zf(z) = 4mz^{\sigma+1}$ is a non-square for all $z \in GF(q)^*$.

## 2 Flocks of the Quadratic Cone

Let $q$ be an odd prime power and let $\mathcal{K}$ be a quadratic cone of $PG(3, q)$ with vertex $v$ and base a conic $\mathcal{C}$. The quadratic cones of $PG(3, q)$ are equivalent under the action of $PGL(4, q)$ so we can assume that $v$ is the point $\langle 0, 0, 0, 1\rangle$ and the conic $\mathcal{C}$ in the plane $\pi$ with equation $X_3 = 0$, is the set of zeros of $X_0X_1 = X_2^2$.

A *flock* $\mathcal{F}$ of $\mathcal{K}$ is a partition of $\mathcal{K} \setminus \{v\}$ into $q$ conics. We call the planes that contain conics of the flock the *planes of the flock*. A flock $\mathcal{F}$ is equivalent to a flock $\mathcal{F}'$ if there is an element in the stabiliser group of the quadratic cone that maps the planes of the flock $\mathcal{F}$ to the planes of the flock $\mathcal{F}'$. If all the planes of the flock share a line then the flock is called *linear*.

Let

$$a_0X_0 + a_1X_1 + a_2X_2 + a_3X_3 = 0$$

be a plane of the flock. Since $\langle 0, 0, 0, 1\rangle$ is disjoint from any plane of the flock $a_3 \neq 0$ and hence we may assume that $a_3 = 1$. The point $\langle 1, 0, 0, -a_0\rangle$ is incident with the quadratic cone and this plane and hence the coefficients of $X_0$ in the planes of the flock are distinct. Hence we can parameterise by the elements of $GF(q)$ so that the planes of the flock are

$$\pi_t \ : \ tX_0 - f(t)X_1 + g(t)X_2 + X_3 = 0$$

where $t \in GF(q)$ and $f$ and $g$ are functions from $GF(q) \to GF(q)$.

The points that are incident with the line that is the intersection of two planes of the flock $\pi_t$ and $\pi_s$ are incident with the plane

$$(t-s)X_0 - (f(t)-f(s))X_1 + (g(t)-g(s))X_2 = 0.$$

The points that are incident with the cone $\mathcal{K}$ satisfy the equation $X_0X_1 = X_2^2$. If the equation

$$(t-s)X_2^2 - (f(t)-f(s))X_1^2 + (g(t)-g(s))X_1X_2 = 0$$

has a solution then we can find a line on the cone, by choosing the $X_0$ coordinate appropriately, that would be contained in the plane above, and hence a point on the cone and incident with both the planes $\pi_t$ and $\pi_s$. The flock property implies that no such point exists and hence that this equation has no solutions. There is no solution with $X_1 = 0$ as this would imply that $X_2 = 0$ and that $t = s$. Hence we can put $w = X_2/X_1$ and we have the forward implication of the following theorem which is due to Thas [26].

**Theorem 4.** *Let $\mathcal{F}$ be a flock of the quadratic cone with vertex $\langle 0,0,0,1\rangle$ and base $X_0X_1 = X_2^2$. Then there exists functions $f$ and $g$ from $GF(q) \to GF(q)$ such that the planes of the flock are*

$$tX_0 - f(t)X_1 + g(t)X_2 + X_3 = 0$$

*where $t \in GF(q)$ and $\mathcal{F}$ is a flock if and only if*

$$(t-s)w^2 + (g(t)-g(s))w - (f(t)-f(s)) = 0$$

*has no solution for all $s$ and $t \in GF(q)$, $s \neq t$.*

If $f$ and $g$ are additive then the condition of the theorem says that $\mathcal{F}$ is a flock if and only if

$$zw^2 + g(z)w - f(z) = 0$$

has no solutions for $w \in GF(q)$ and $z \in GF(q)^*$. A flock with this property is called a *semifield flock* as such a flock is in one-to-one correspondence with a commutative semifield of rank 2 over its middle nucleus. This is clear from Theorem 2. The commutative semifield $\mathcal{S} = \{(x,y) \mid x,y \in GF(q)\}$ where addition is defined component-wise and multiplication is defined by

$$(x,y)(u,v) = (xv + yu + g(xu), yv + f(xu))$$

is the semifield associated to the flock $\mathcal{F}$.

The known examples of semifield flocks up to equivalence are listed in Table 1. In all relevant cases $m$ is taken to be a non-square in $GF(q)$ and $\sigma$ is a nontrivial automorphism of $GF(q)$. Some of the links between the commutative semifields, certain ovoids of $Q(4,q)$, semifield flocks of the quadratic cone and semifield translation planes were not known until recently and hence in

**Table 1.** The known examples of semifield flocks up to equivalence

| name | $g(x)$ | $f(x)$ | $q=p^h$ |
|---|---|---|---|
| linear | 0 | $mx$ | all |
| Dickson [12] Kantor [18] Knuth [19] | 0 | $mx^\sigma$ | |
| Cohen-Ganley [10] Thas-Payne [28] | $x^3$ | $m^{-1}x+mx^9$ | $3^h$ |
| Penttila-Williams [24], Bader-Lunardon-Pinneri [4] | $x^3$ | $x^{27}$ | $3^5$ |

most cases more than one person or persons is accredited with the discovery of the functions $f$ and $g$. In fact in the second case Dickson [12] discovered the semifield, Kantor [18] the ovoid and Knuth [19] the semifield plane. In the third example Cohen and Ganley [10] discovered the semifield while Thas and Payne found the ovoid [28]. And in the fourth example Penttila and Williams discovered the ovoid [24] and details concerning the corresponding flock were investigated by Bader, Lunardon and Pinneri [4]. We shall discuss these equivalent objects in the following sections and explain the links between them and how this can be of use. Firstly however we shall check that the last two examples in Table 1 do indeed satisfy the condition of Theorem 2 and Theorem 4. In the Cohen-Ganley Thas-Payne example

$$g(x)^2+4xf(x)=g(x)^2+xf(x)=x^6+m^{-1}x^2+mx^{10}=m(x^5-m^{-1}x)^2$$

which is a non-square for all $x\in GF(3^h)^*$.

The Penttila-Williams example is somewhat more difficult to prove. The following comes from [2]. We have that

$$g(x)^2+4xf(x)=g(x)^2+xf(x)=x^6+x^{28}=x^6(1+x^{22})$$

and since $3^5-1=242=2.11^2$ we need to show that $1+\epsilon$ is a non-square for all $\epsilon$ such that $\epsilon^{11}=1$. Now $(q-1)/2=121=1+3+3^2+3^3+3^4$ and in $GF(3^5)$

$$(1+\epsilon)^{121}=(1+\epsilon)(1+\epsilon^3)(1+\epsilon^9)(1+\epsilon^{27})(1+\epsilon^{81}).$$

The set $\{1,3,9,27.81\}$ are the squares modulo 11 and each non-zero integer modulo 11 can be written exactly 3 times as the sum of elements of $\{1,3,9,27,81\}$ modulo 11. Hence in $GF(3^5)$

$$(1+\epsilon)^{121}=2=-1$$

and $1+\epsilon$ is a non-square for all $\epsilon$ such that $\epsilon^{11}=1$.

The following theorem comes from [14].

**Theorem 5.** *The projective planes obtained from the flocks $\mathcal{F}$ and $\mathcal{F}'$ are isomorphic if and only if the flocks $\mathcal{F}$ and $\mathcal{F}'$ are equivalent.*

The projective planes in Theorem 5 are constructed, via the Bruck Bose André method, from the spread

$$\{\langle(y, x, 1, 0), (f(x), y + g(x), 0, 1)\rangle \mid x, y \in GF(q)\} \cup \{\langle(1, 0, 0, 0), (0, 1, 0, 0)\rangle\}.$$

This plane is a semifield plane. Following [11, (5.1.2)] the spread comes from the spread set

$$\mathcal{D} = \left\{ \begin{pmatrix} y + g(x) & x \\ f(x) & y \end{pmatrix} \mid x, y \in GF(q) \right\}$$

which has the property that the determinant of $M - N$ is non-zero for all distinct $M$, $N \in \mathcal{D}$. The plane is coordinatised by the semifield whose multiplication is defined by

$$\begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} v + g(u) & u \\ f(u) & v \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} xv + yu + xg(u) \\ yv + xf(u) \end{pmatrix}.$$

We can check that this multiplication defines a semifield. It is only condition (S4) that requires some work. If

$$xg(u) + xv + yu = 0$$

and

$$xf(u) + yv = 0$$

then

$$xv^2 + xg(u)v - xuf(u) = 0.$$

If $x = 0$ one can check that then one of either $(x, y)$ or $(u, v)$ is equal to $(0, 0)$. If $x \neq 0$ then, since $g(u)^2 + 4uf(u)$ is a non-square for all $u \in GF(q)^*$, $u = 0$ and it follows that $(u, v) = (0, 0)$. Hence this is a semifield. Note that this means we can construct a not necessarily commutative semifield from the functions $f$ and $g$. Now semifields that we get from the above multiplication will be isotopic if there corresponding flocks are equivalent by Theorem 1 and Theorem 5. However we have not proved that the commutative semifields that we get from the functions $f$ and $g$ are isotopic if and only if their associated flocks are equivalent.

The following theorem which we prove in Section 6 shows that there is an isotopism between two commutative semifields if their associated flocks are equivalent.

**Theorem 6.** *$\mathcal{F}$ and $\hat{\mathcal{F}}$ are equivalent semifield flocks if and only if there exists a linear one-to-one map $F$ from $\mathcal{S}$ to $\hat{\mathcal{S}}$ and a $GF(p)$-linear map $H$ from $\mathcal{S}$ to $\hat{\mathcal{S}}$ such that*

$$(ab)H = (aF).(bF)$$

*for all $a$, $b \in \mathcal{S}$ where . is multiplication in $\hat{\mathcal{S}}$ and $\mathcal{F}$ and $\hat{\mathcal{F}}$ are the semifield flocks associated to the commutative semifields $\mathcal{S}$ and $\hat{\mathcal{S}}$ of rank 2 over their middle nucleus $GF(q)$, $q = p^n$, respectively.*

Let us consider for the moment a flock that is linear, i.e. with the property that all the planes of the flock contain a common line. The points that are dual to the planes of the flock

$$\{\langle t, -f(t), g(t), 1\rangle \mid t \in GF(q)\}$$

are collinear and so $(f(t)-f(s))/(t-s)$ and $(g(t)-g(s))/(t-s)$ are constant for all $s \neq t$. Hence $f$ and $g$ have polynomial degree 1. The following theorem is from Thas [26].

**Theorem 7.** *A flock whose planes are all incident with a common point is either linear (in which case the planes of the flock share a common line) or equivalent to a semifield flock of Dickson, Kantor, Knuth type.*

*Remark 1.* It follows from this theorem that the semifield flocks we obtain directly from the Cohen-Ganley so-called sporadic example of a semifield and the semifields from [25] are equivalent to a semifield flock of Dickson, Kantor, Knuth type. In [25, Theorem 1] $g(t) = t^{\sqrt{q}}$ and $f(t) = ct$ and it is a simple matter to check that the planes of the flock are all incident with the point $\langle c, 1, 0, 0\rangle$ and in [25, Theorem 2] $g(t) = at + bt^{\sqrt{q}}$ and $f(t) = t$ and the planes of the flock are all incident with the point $\langle 1, 1, 0, 0\rangle$. As mentioned in [14] the sporadic example of Cohen and Ganley over $GF(5^2)$ with $g(t) = t^5$ and $f(t) = 2\sqrt{2}t^5 + t$ the planes of the flock are all incident with the point $\langle 1, 1, 2\sqrt{2}, 0\rangle$. By Theorem 6 their associated commutative semifields are isotopic to a Dickson, Kantor, Knuth semifield. All known examples of commutative semifields rank 2 over their middle nucleus are isotopic to one of the commutative semifields rank 2 over their middle nucleus constructed from the pairs of functions in Table 1.

In the following argument we are going to use the so-called *linear representation* of $PG(2, q)$ so let us recall what we mean by this (for more details see [22]). Let $GF(q_0)$ be a subfield of $GF(q)$, $q = q_0^n$. Let $\mathcal{V}$ be the vector space of rank 3 over $GF(q)$. The projective plane $PG(2, q)$ is the incidence geometry whose points are the subspaces of rank 1 of $\mathcal{V}$ and whose lines are the subspaces of rank 2 of $\mathcal{V}$. However $\mathcal{V}$ is a vector space of rank $3n$ over $GF(q_0)$ and the points of $PG(2, q)$ are subspaces of rank $n$ which are mutually disjoint and cover $\mathcal{V} \setminus \mathbf{0}$, i.e. they form a spread $\Lambda$. The spread $\Lambda$ induces a spread in the subspace generated by any two elements of $\Lambda$ (since this subspace is a line of $PG(2, q)$). We call a spread with this property *normal*.

Let us consider a semifield flock $\mathcal{F}$. The points

$$\{\langle t, -f(t), g(t), 1\rangle \mid t \in GF(q)\}$$

that are dual to the planes of the flock project on to the plane $X_3 = 0$ the set of points

$$\mathcal{W} := \{\langle t, -f(t), g(t), 0\rangle \mid t \in GF(q)\}.$$

Since the functions $f$ and $g$ are additive, they are linear over some subfield $GF(q_0)$ of $GF(q)$. The maximum subfield with this property is often called the kernel of the flock. This kernel is equal to the left nucleus of the semifield (and hence equal to the right nucleus since the semifield is commutative). If we look at the linear representation of the plane $X_3 = 0$ the set $\mathcal{W}$ is a subspace of rank $n$ over $GF(q_0)$ .

The vertex of the quadratic cone is the point $\langle 0, 0, 0, 1\rangle$ and this point is dual to the plane $X_3 = 0$ and the $q + 1$ lines on the quadratic cone are dual to a set of $q + 1$ lines in the plane $X_3 = 0$ that are tangents to some conic $\mathcal{C}'$. The definition of a flock implies that the points in $\mathcal{W}$ are not incident with a tangent to this conic $\mathcal{C}'$, i.e. the set $\mathcal{W}$ is contained in the internal points of the conic $\mathcal{C}'$. If the flock is linear then the set $\mathcal{W}$ is a point of the plane $X_3 = 0$. Theorem 7 implies that the flock is of Dickson, Kantor, Knuth type if and only if the set $\mathcal{W}$ is contained in a line of the plane $X_3 = 0$. In all other cases the set $\mathcal{W}$ in the linear representation contains a subplane $PG(2, q_0)$ that is contained in the internal points of a conic in $PG(2, q)$. However this cannot always occur. The following is from [5].

**Theorem 8.** *If there is a subplane of order $q_0$ contained in the internal points of a conic in $PG(2, q)$ where $q = q_0^n$ then $q_0 < 4n^2 - 8n + 2$.*

The above argument leads immediately to the following corollaries.

**Corollary 1.** *A semifield flock of the quadratic cone of $PG(3, q)$ whose defining functions $f$ and $g$ are linear over the subfield $GF(q_0)$ where $q = q_0^n$ and $q_0 \geq 4n^2-8n+2$ is either a linear flock or a Dickson, Kantor, Knuth semifield flock.*

**Corollary 2.** *A commutative semifield of rank 2 over its middle nucleus $GF(q)$ that has defining functions $f$ and $g$ which are linear over the subfield $GF(q_0)$ where $q = q_0^n$ and $q_0 \geq 4n^2 - 8n + 2$ is either the finite field $GF(q^2)$ or isotopic to a Dickson, Kantor, Knuth semifield.*

*Remark 2.* We may expect something much stronger than this bound to hold. Indeed we can see that the theorem hypothesis requires that their is a subplane in the internal points of the conic. However in fact the set $\mathcal{W}$ is contained in the internal points of a conic and in the linear representation of $PG(2, q)$ it is a subspace of rank $n$ over $GF(q_0)$.

The bound in the theorem for $n = 3$ gives $q_0 < 14$ and by computer Bloemen, Thas and van Maldeghem [7] have checked that there are no other semifield flocks other than the linear flock and the Dickson, Kantor, Knuth flocks. Note also that the only other known examples have $q_0 = 3$.

The following nice result of Bader and Lunardon [3] shows that in some sense the Pentilla-Williams example is sporadic, and any other examples yet to discovered.

**Theorem 9.** *If there is a polynomial $h(t)$ over $GF(q)$ such that for a fixed non-square $m$ in $GF(q)$ the equality*

$$g(t) + 4tf(t) = mh(t)$$

*is a polynomial identity then $f$ and $g$ are one of the first three examples in Table 1.*

## 3 The Generalized Quadrangle $T(\mathcal{E})$

A *generalized quadrangle* is a set of points and a set of lines with an incidence relation that satisfies the following axioms.

(Q1) Every two points are incident with at most one line.
(Q2) For all anti-flags $(p, L)$ (the point $p$ is not incident with the line $L$) there is exactly one point incident with $L$ and collinear with $p$.
(Q3) There is no point collinear with all others.

Let $G$ be a generalized quadrangle in which there is a line incident with at least three points and a point incident with at least three lines. It is not difficult to prove that the number of points incident with a line, and the number of lines incident with a point, are constants. We say $G$ is a generalized quadrangle of order $s$, $t$ if every line is incident with $s + 1$ points and every point is incident with $t + 1$ lines.

An *egg* $\mathcal{E}_{m,n}$ of $PG(2n + m - 1, q)$ is a set of $q^m + 1$ $(n - 1)$-subspaces with the properties that any three elements of $\mathcal{E}_{m,n}$ span a $(3n - 1)$-space and every element of $\mathcal{E}_{m,n}$ is contained in a $(n + m - 1)$-subspace called a *tangent space* that is skew from all other elements of $\mathcal{E}_{m,n}$. We write $\mathcal{E}$ for $\mathcal{E}_{m,n}$ when no confusion is possible.

The following construction of the generalized quadrangle $T(\mathcal{E}_{m,n})$ from an egg is based on a construction due to Tits and comes from Payne and Thas [23]. Let $\mathcal{E}_{m,n}$ be an egg of $\pi = PG(2n+m-1, q)$ and embed the space $\pi$ in $PG(2n + m, q)$. Points are defined as

(i) the points of $PG(2n + m, q) \setminus \pi$,
(ii) the $(n + m)$-spaces of $PG(2n + m, q)$ that contain a tangent space of $\mathcal{E}_{m,n}$ but are not contained in $\pi$,
(iii) a symbol $(\infty)$.

Lines are defined as

(a) the $n$-spaces of $PG(2n+m, q)$ which contain an element of $\mathcal{E}_{m,n}$ but are not contained in $\pi$,
(b) the elements of $\mathcal{E}_{m,n}$.

Incidence is as follows. A point of type (i) is incident with a line of type (a) if they are incident in $PG(2n+m,q)$. A point of type (ii) is incident with the lines of type (a) which it contains and the unique line of type (b) which it contains. The point of type (iii) is incident with all lines of type (b).

$T(\mathcal{E}_{m,n})$ is a generalized quadrangle of order $(q^n, q^m)$, [23, Theorem 8.7.1] or [20, Theorem 3.3.1]. Let $\mathcal{C}$ be a non-singular conic in $PG(2, q_0^n)$. In the linear representation described in the previous section the $q_0^n + 1$ points of $\mathcal{C}$ become $q_0^n + 1$ $(n-1)$-subspaces of $PG(3n-1, q_0)$ which form an egg $\mathcal{E}_{\mathcal{C}}$ whose tangent spaces correspond to the set of tangent lines of $\mathcal{C}$. The generalized quadrangle $T(\mathcal{E}_{\mathcal{C}})$ is the Tits generalized quadrangle $T_2(\mathcal{C})$ of order $(q_0^n, q_0^n)$.

An *ovoid* $\mathcal{O}$ of a generalised quadrangle is a set of points with the property that every line is incident with exactly one point of $\mathcal{O}$. An ovoid of a generalised quadrangle of order $(s,t)$ contains $st+1$ points.

Let us consider an ovoid $\mathcal{O}$ of $T_2(\mathcal{C})$ that contains the point $(\infty)$. The set $\mathcal{O} \setminus \{(\infty)\}$ is a set of $q^{2n}$ points of type (a) with the property that the line of $PG(3n, q_0)$ spanned by any two of them meets $\pi$ in a point not contained in an element of the egg $\mathcal{E}_{\mathcal{C}}$.

Let us consider again the set $\mathcal{W}$ from the previous section which is contained in the internal points of a conic $\mathcal{C}'$. In the linear representation $\mathcal{W}$ is a $(n-1)$-subspace of a $(3n-1)$-space $\pi'$ disjoint from all elements and all tangent spaces of the egg $\mathcal{E}_{\mathcal{C}'}$. In the dual space the space $\mathcal{W}^*$ dual to $\mathcal{W}$ is a $(2n-1)$-subspace of a $(3n-1)$-space $\pi$ disjoint from the $(n-1)$-subspaces dual to the tangent spaces. In the dual setting we have an egg $\mathcal{E}_{\mathcal{C}}$ where $\mathcal{C}$ is the dual of the conic $\mathcal{C}'$. Embed $\pi$ in a $(3n)$-space and let $P$ be any point of $PG(3n,q) \setminus \pi$. The $(2n)$-subspace $\langle \mathcal{W}^*, P \rangle$ has the property that any two of its points span a line that meets $\pi$ in point not in the egg $\mathcal{E}_{\mathcal{C}}$. Hence

$$(\langle \mathcal{W}^*, P \rangle \setminus \pi) \cup \{(\infty)\}$$

is an ovoid of the generalised quadrangle $T_2(\mathcal{C})$.

The above argument was first explained by Thas [27].

## 4 Ovoids of $Q(4,q)$

In this section we shall see that $T_2(\mathcal{C})$ is isomorphic to the classical generalised quadrangle $Q(4,q)$ and hence that commutative semifields of rank 2 over their middle nucleus imply certain ovoids of $Q(4,q)$.

A quadratic form $Q(\mathbf{x})$ on a vector space $\mathcal{V}$ over a field $F$ satisfies the axioms

$$Q(\lambda \mathbf{x}) = \lambda^2 Q(\mathbf{x}) \text{ for all } \mathbf{x} \in \mathcal{V}$$

$$Q(\mathbf{x}+\mathbf{y}) = Q(\mathbf{x}) + Q(\mathbf{y}) + b(\mathbf{x},\mathbf{y})$$

where $b(\mathbf{x},\mathbf{y})$ is a bilinear form. A *totally singular* subspace $\mathcal{S}$ is a subspace with the property that $Q(\mathbf{x}) = 0$, $Q(\mathbf{y}) = 0$ and $b(\mathbf{x},\mathbf{y}) = 0$ for all $\mathbf{x}$, $\mathbf{y} \in \mathcal{S}$.

We restrict ourselves to the case where the field $F = GF(q)$ and the maximum rank of a totally singular subspace is 2. The classification of quadratic forms over a finite field says that there are three such inequivalent non-singular quadratic forms (for more details on the equivalence and singularity of quadratic forms see [8]). Let $\mathcal{G}$ denote the geometry whose points are the totally singular subspaces of rank 1 and whose lines are the totally singular subspaces of rank 2 for one of these quadratic forms. Let $\langle \mathbf{x} \rangle$ and $\mathcal{S}$ be totally singular subspaces of rank 1 and 2 respectively such that $\mathbf{x} \notin \mathcal{S}$, i.e. a non-incident point and line of $\mathcal{G}$. The rank of $\mathcal{S} \cap \mathbf{x}^\perp$ where $\mathbf{x}^\perp := \{\mathbf{z} \in \mathcal{V} \mid b(\mathbf{x}, \mathbf{z}) = 0\}$ is 1 since $\mathbf{x}^\perp$ is a hyperplane not containing $\mathcal{S}$. In terms of the geometry this implies that for a non-incident point $P$ and line $l$ of $\mathcal{G}$ there is a unique point $P'$ incident with $l$ and collinear with $P$. Hence from the three quadratic forms we obtain three generalised quadrangles which are called the classical orthogonal generalised quadrangles. These are listed in Table 2 in which $g$ is an irreducible homogeneous quadratic form.

**Table 2.** The classical orthogonal generalised quadrangles

| name | label | $n$ | Canonical form |
|---|---|---|---|
| Hyperbolic | $Q^+(3,q)$ | 4 | $Q(\mathbf{x}) = x_0x_1 + x_2x_3$ |
| Parabolic | $Q(4,q)$ | 5 | $Q(\mathbf{x}) = x_0x_1 + x_3x_4 - x_2^2$ |
| Elliptic | $Q^-(5,q)$ | 6 | $Q(\mathbf{x}) = x_0x_1 + x_2x_3 + g(x_4,x_5)$ |

An ovoid $\mathcal{O}$ of a classical orthogonal generalised quadrangle of order $(s,t)$ is a set of $st+1$ totally singular subspaces of rank 1 with the property that for all distinct $\langle \mathbf{x} \rangle, \langle \mathbf{y} \rangle \in \mathcal{O}$ the bilinear form $b(\mathbf{x}, \mathbf{y}) \neq 0$.

Let the generalised quadrangle $Q(4,q)$ of order $(q,q)$ be defined by the quadratic form

$$Q(\mathbf{x}) = x_0x_1 + x_3x_4 - x_2^2.$$

An ovoid $\mathcal{O}$ of $Q(4,q)$ has $q^2+1$ points. We may assume that $\langle 0,0,0,0,1 \rangle \in \mathcal{O}$. The associated bilinear form to $Q$ is

$$b(\mathbf{x}, \mathbf{y}) = x_0y_1 + y_0x_1 + x_3y_4 + x_4y_3 - 2x_2y_2.$$

For any $\langle x \rangle \in \mathcal{O}$

$$0 \neq b(\mathbf{x}, \langle 0,0,0,0,1 \rangle) = x_3$$

and hence we can assume that $x_3 = 1$. Moreover if $\mathbf{x} = (x_0, x_1, x_2, 1, x_4)$ and $\mathbf{y} = (x_0, y_1, x_2, 1, y_4)$ where $\langle x \rangle$ and $\langle y \rangle \in \mathcal{O}$ then

$$b(\mathbf{x}, \mathbf{y}) = x_0y_1 + x_0x_1 + y_4 + x_4 - 2x_2^2 = Q(\mathbf{x}) + Q(\mathbf{y}) = 0$$

and so the first and third coordinate pair are distinct pairs for distinct points of the ovoid. Hence there is a polynomial $F(x,y)$ such that the ovoid

$$\mathcal{O} = \{\langle x, F(x,y), y, 1, y^2 - xF(x,y) \rangle \mid x, y \in GF(q)\} \cup \{\langle 0,0,0,0,1 \rangle\}.$$

**Table 3.** The known examples of ovoids of $Q(4,q)$

| name | $F(x,y)$ | $q$ | restrictions |
|---|---|---|---|
| elliptic quadrics | $mx$ | all | |
| Kantor [18] | $mx^\alpha$ | odd | $\alpha \in Aut(GF(q))$ |
| Thas-Payne [28] | $m^{-1}x + (mx)^{1/9} + y^{1/3}$ | $3^h$ | |
| Penttila-Williams [24] | $x^9 + y^{81}$ | $3^5$ | |
| Ree-Tits slice [18] | $x^{2\alpha+3} + y^\alpha$ | $3^{2h+1}$ | $\alpha = \sqrt{3q}$ |
| Tits [29] | $x^{\alpha+1} + y^\alpha$ | $2^{2h+1}$ | $\alpha = \sqrt{2q}$ |

In the article of Penttila and Williams [24] the stabiliser group of each of the known ovoids is calculated. Note that in four examples of Table 3 $F(x,y) = f(x) + g(y)$ where $f$ and $g$ are linear over some subfield of $GF(q)$. In the previous section we constructed an ovoid of $T_2(\mathcal{C})$ from a semifield flock. However the generalised quadrangle $T_2(\mathcal{C})$ is isomorphic to $Q(4,q)$. Let $\phi : Q(4,q) \to T_2(\mathcal{C})$, where $\mathcal{C}$ is the conic $X_0X_1 = X_2^2$, be the map

$$\begin{aligned}
\langle 0,0,0,0,1\rangle &\mapsto (\infty)\\
\langle a,b,c,1,c^2-ab\rangle &\mapsto \langle a,b,c,1\rangle\\
\langle a^2,1,a,0,b\rangle &\mapsto \langle (a^2,1,a,0),(-b,0,0,1)\rangle\\
\langle 1,0,0,0,a\rangle &\mapsto \langle (1,0,0,0),(0,-a,0,1)\rangle.
\end{aligned}$$

This is indeed an isomorphism since collinearity is preserved. The points $\langle \mathbf{x}\rangle = \langle a,b,c,1,c^2-ab\rangle$ and $\langle \mathbf{x}'\rangle = \langle a',b',c',1,c'^2-a'b'\rangle$ are collinear in $Q(4,q)$ if and only if $b(\mathbf{x},\mathbf{x}') = ab' + ba' - ab - a'b' + c^2 - 2cc' + c'^2 = (c-c')^2 - (a-a')(b-b') = 0$ if and only if the point $\langle a-a', b-b', c-c'\rangle$ lies on the conic $X_0X_1 = X_2^2$. One can check that the other incidences are preserved.

Hence from the ovoid of $T_2(\mathcal{C})$ that was constructed in the previous section we get an ovoid of $Q(4,q)$. In the next section we shall use explicit coordinates to calculate $F(x,y)$ from the functions $f$ and $g$ that determine the semifield flock. The following theorem is from Lunardon [22].

**Theorem 10.** *If $\mathcal{F}$ and $\mathcal{F}'$ are semifield flocks of the quadratic cone then the ovoids that come from the flocks are equivalent if and only if the flocks $\mathcal{F}$ and $\mathcal{F}'$ are equivalent.*

## 5 Correspondence Between the Ovoid and the Flock Using Coordinates

As in [20] we follow the argument of Thas [27] using coordinates. Let us see how this works. It may help to refer back to end of Section 3.

The lines on the quadratic cone with vertex $\langle 0,0,0,1\rangle$ and base defined by the equation $X_0X_1 = X_2^2$ dualise with respect to the standard inner product

to lines in the plane $X_3 = 0$ with equation

$$X_0 + a^2 X_1 + aX_2 = 0.$$

These lines are tangents to the conic whose points are the zeros of the quadratic form $\mathcal{Q}' = 4X_0X_1 - X_2^2$. The associated bilinear form is

$$b'(\mathbf{x}, \mathbf{y}) = 4x_0y_1 + 4y_0x_1 - 2x_2y_2.$$

We wish to view the vector space of rank 3 over $GF(q)$ as a vector space of rank $3n$ over $GF(q_0)$ and the bilinear form $b'$ over this vector space is

$$\hat{b}(\mathbf{x}, \mathbf{y}) = Tr_{q\to q_0}(4x_0y_1 + 4y_0x_1 - 2x_2y_2).$$

In Section 3 the set $\mathcal{W}$ is contained in the hyperplane $X_3 = 0$ and is the set of points $\{\langle t, -f(t), g(t)\rangle \mid t \in GF(q)\}$. The functions $f$ and $g$ are linear over some subfield $GF(q_0)$ and so we can write

$$f(t) = \sum_{i=0}^{n-1} c_i t^{q_0^i} \quad \text{and} \quad g(t) = \sum_{i=0}^{n-1} b_i t^{q_0^i}.$$

We follow the argument at the end of Section 3 and dualise with respect to the bilinear form $\hat{b}$. A point $\langle x_0, x_1, x_2\rangle \in \mathcal{W}^*$ if and only if

$$Tr_{q\to q_0}(-4x_0f(t) + 4x_1t - 2x_2g(t)) = 0$$

for all $t \in GF(q)$ if and only if

$$Tr_{q\to q_0}((-4c_0x_0 + 4x_1 - 2b_0x_2)t + \sum_{i=1}^{n-1}(-4c_ix_0 - 2b_ix_2)t^{q_0^i}) = 0$$

if and only if

$$Tr_{q\to q_0}((-4c_0x_0 + 4x_1 - 2b_0x_2 + \sum_{i=1}^{n-1}(-4c_ix_0 - 2b_ix_2)^{q_0^{n-i}})t) = 0$$

for all $t \in GF(q)$. Hence

$$4x_1 = \sum_{i=0}^{n-1}(4c_ix_0 + 2b_ix_2)^{q_0^{n-i}}.$$

The set

$$\mathcal{W}^* = \{\langle x_0, \sum_{i=0}^{n-1}(c_ix_0 + \frac{1}{2}b_ix_2)^{q_0^{n-i}}, x_2\rangle \mid x_0, x_2 \in GF(q)\}.$$

Now if we were to cone $\mathcal{W}^*$ to the set $\langle W^*, P\rangle$ where $P$ is a point not on the hyperplane $X_3 = 0$ we would have $q^2$ points of an ovoid of the $\mathcal{T}_2(\mathcal{C})$ defined with conic $4X_0X_1 = X_2^2$. However we wish to have an ovoid of the $\mathcal{T}_2(\mathcal{C})$ defined by the conic $X_0X_1 = X_2^2$ and so we use the map $\psi$ that takes

$$X_0 \mapsto X_0$$
$$X_1 \mapsto X_1$$
$$X_2 \mapsto \tfrac{1}{2}X_2$$

and maps the subspace $\mathcal{W}^*$ to the subspace

$$\{\langle x, F(x,y), y\rangle \mid x, y \in GF(q)\}$$

where

$$F(x,y) = \sum_{i=0}^{n-1} (-c_i x + b_i y)^{q_0^{n-i}}.$$

We take the point $P$ to be the point $\langle 0,0,0,1\rangle$ so that the set

$$\{\langle x, F(x,y), y, 1\rangle \mid x, y \in GF(q)\}$$

is a set of $q^2$ points of an ovoid of $T_2(\mathcal{C})$. We apply the isomorphism $\phi^{-1}$ from the previous section to give the explicit points of an ovoid of $Q(4,q)$ that comes from the semifield flock defined by the functions $f$ and $g$,

$$\mathcal{O} = \{\langle x, F(x,y), y, 1, y^2 - xF(x,y) \mid x, y \in GF(q)\} \cup \{\langle 0,0,0,0,1\rangle\}.$$

## 6 Correspondence Between the Commutative Semifield and the Flock

In this section we look at the correspondence between the commutative semifields of rank 2 over their middle nucleus and the associated semifield flocks. This is a proof of Theorem 6.

Let $\mathcal{S}$ and $\hat{\mathcal{S}}$ be commutative semifields of rank 2 over their middle nucleus $GF(q)$, $q = p^n$, constructed from the pairs of functions $(f,g)$ and $(\hat{f},\hat{g})$ respectively. The functions $f$, $g$, $\hat{f}$, $\hat{g}$ are linear over $GF(p)$ so we can write them as

$$f(x) = \sum_{i=0}^{n-1} f_i x^{p^i}, \quad g(x) = \sum_{i=0}^{n-1} g_i x^{p^i},$$

$$\hat{f}(x) = \sum_{i=0}^{n-1} \hat{f}_i x^{p^i}, \quad \hat{g}(x) = \sum_{i=0}^{n-1} \hat{g}_i x^{p^i}.$$

Let us assume that there exists a one-to-one $GF(p)$-linear map $H$ from $\mathcal{S}$ to $\hat{\mathcal{S}}$ and a one-to-one linear map $F$ from $\mathcal{S}$ to $\hat{\mathcal{S}}$ such that

$$((x,y)(u,v))H = ((x,y)F).((u,v)F)$$

for all $(x,y)$ and $(u,v) \in \mathcal{S}$. Expanding the left-hand side we get

$$((x,y)(u,v))H = ((xv + yu + \hat{g}(ux), yv + \hat{f}(ux))H$$

$$= \left( \sum_{i=0}^{n-1} h_i(xv + yu + \hat{g}(ux))^{p^i} + \sum_{i=0}^{n-1} m_i(yv + \hat{f}(ux))^{p^i}, \right.$$
$$\left. \sum_{i=0}^{n-1} k_i(xv + yu + \hat{g}(ux))^{p^i} + \sum_{i=0}^{n-1} l_i(yv + \hat{f}(ux))^{p^i} \right),$$

for some $h_i$, $m_i$, $k_i$ and $l_i$. Expanding the right-hand side we get

$$((x,y)F).((u,v)F) = (\alpha_0 x + \alpha_1 y, \beta_0 x + \beta_1 y).(\alpha_0 u + \alpha_1 v, \beta_0 u + \beta_1 v) =$$
$$(2\alpha_0\beta_0 xu + 2\alpha_1\beta_1 yv + (\alpha_0\beta_1 + \alpha_1\beta_0)(xv + yu) +$$
$$g((\alpha_0 x + \alpha_1 y)(\alpha_0 u + \alpha_1 v)),$$
$$\beta_0^2 xu + \beta_1^2 yv + \beta_0\beta_1(xv + yu) + f((\alpha_0 x + \alpha_1 y)(\alpha_0 u + \alpha_1 v))\big),$$

for some $\alpha_0$, $\alpha_1$, $\beta_0$ and $\beta_1$. Equate the coefficient of $(yv)^{p^i}$ to get

$$(i > 0)\ \ m_i = \alpha_1^{2p^i} g_i \quad (i = 0)\ \ m_0 = 2\alpha_1\beta_1 + \alpha_1^2 g_0,$$

$$(i > 0)\ \ l_i = \alpha_1^{2p^i} f_i \quad (i = 0)\ \ l_0 = \beta_1^2 + \alpha_1^2 f_0.$$

Equate the coefficient of $(yu)^{p^i}$ to get

$$(i > 0)\ \ h_i = (\alpha_0\alpha_1)^{p^i} g_i \quad (i = 0)\ \ h_0 = \alpha_0\beta_1 + \alpha_1\beta_0 + \alpha_0\alpha_1 g_0,$$

$$(i > 0)\ \ k_i = (\alpha_0\alpha_1)^{p^i} f_i \quad (i = 0)\ \ k_0 = \beta_0\beta_1 + \alpha_0\alpha_1 f_0.$$

Equate the coefficient of $(xu)^{p^j}$ to get

$$(j > 0) \quad \sum_{i=0}^{n-1} h_i \hat{g}_{j-i} + \sum_{i=0}^{n-1} m_i \hat{f}_{j-i} = \alpha_0^{2p^j} g_j,$$

$$(j = 0) \quad \sum_{i=0}^{n-1} h_i \hat{g}_{n-i} + \sum_{i=0}^{n-1} m_i \hat{f}_{n-i} = 2\alpha_0\beta_0 + \alpha_0^2 g_0,$$

$$(j > 0) \quad \sum_{i=0}^{n-1} k_i \hat{g}_{j-i} + \sum_{i=0}^{n-1} l_i \hat{f}_{j-i} = \alpha_0^{2p^j} f_j,$$

$$(j = 0) \quad \sum_{i=0}^{n-1} k_i \hat{g}_{n-i} + \sum_{i=0}^{n-1} l_i \hat{f}_{n-i} = \beta_0^2 + \alpha_0^2 f_0,$$

where all indices are taken modulo $n$. Substitute the expressions for the $h_i$, $m_i$, $k_i$ and $l_i$ in the previous four equations and get the equations $A_j$ for $j = 1, \ldots, n-1$

$$\sum_{i=0}^{n-1} (\alpha_0\alpha_1)^{p^i} g_i \hat{g}_{j-i} + (\alpha_0\beta_1 + \alpha_1\beta_0)\hat{g}_j + \sum_{i=0}^{n-1} \alpha_1^{2p^i} g_i \hat{f}_{j-i} + 2\alpha_1\beta_1 \hat{f}_j = \alpha_0^{2p^j} g_j,$$

the equation $A_0$

$$\sum_{i=0}^{n-1}(\alpha_0\alpha_1)^{p^i} g_i\hat{g}_{n-i} + (\alpha_0\beta_1 + \alpha_1\beta_0)\hat{g}_0 + \sum_{i=0}^{n-1}\alpha_1^{2p^i} g_i\hat{f}_{n-i} + 2\alpha_1\beta_1\hat{f}_0 =$$

$$2\alpha_0\beta_0 + \alpha_0^2 g_0,$$

the equations $B_j$ for $j = 1, \ldots, n-1$

$$\sum_{i=0}^{n-1}(\alpha_0\alpha_1)^{p^i} f_i\hat{g}_{j-i} + \beta_0\beta_1\hat{g}_j + \sum_{i=0}^{n-1}\alpha_1^{2p^i} f_i\hat{f}_{j-i} + \beta_1^2\hat{f}_j = \alpha_0^{2p^j} f_j,$$

and the equation $B_0$

$$\sum_{i=0}^{n-1}(\alpha_0\alpha_1)^{p^i} f_i\hat{g}_{n-i} + \beta_0\beta_1\hat{g}_0 + \sum_{i=0}^{n-1}\alpha_1^{2p^i} f_i\hat{f}_{n-i} + \beta_1^2\hat{f}_0 = \beta_0^2 + \alpha_0^2 f_0.$$

Now the sums $\sum_{j=0}^{n-1} A_j t^{p^j}$ and $\sum_{j=0}^{n-1} B_j t^{p^j}$ give

$$g(\alpha_0\alpha_1\hat{g}(t)) + (\alpha_0\beta_1 + \beta_0\alpha_1)\hat{g}(t) + g(\alpha_1^2\hat{f}(t)) + 2\alpha_1\beta_1\hat{f}(t) = 2\alpha_0\beta_0 t + g(\alpha_0^2 t)$$

and

$$f(\alpha_0\alpha_1\hat{g}(t)) + \beta_0\beta_1\hat{g}(t) + f(\alpha_1^2\hat{f}(t)) + \beta_1^2\hat{f}(t) = \beta_0^2 t + f(\alpha_0^2 t).$$

The functions $f$ and $g$ are additive and so these equations can be written as

$$g(-\alpha_0^2 t + \alpha_1^2\hat{f}(t) + \alpha_1\alpha_0\hat{g}(t)) = 2\alpha_0\beta_0 t - 2\alpha_1\beta_1\hat{f}(t) - (\alpha_0\beta_1 + \beta_0\alpha_1)\hat{g}(t)$$

and

$$f(-\alpha_0^2 t + \alpha_1^2\hat{f}(t) + \alpha_1\alpha_0\hat{g}(t)) = \beta_0^2 t - \beta_1^2\hat{f}(t) - \beta_0\beta_1\hat{g}(t).$$

Put $u = -\alpha_0^2 t + \alpha_1^2\hat{f}(t) + \alpha_0\alpha_1\hat{g}(t)$ and rewrite the above equations in matrix form as

$$\begin{pmatrix} -\alpha_0^2 & -\alpha_1^2 & \alpha_0\alpha_1 & 0 \\ -\beta_0^2 & -\beta_1^2 & \beta_0\beta_1 & 0 \\ 2\alpha_0\beta_0 & 2\alpha_1\beta_1 & -(\alpha_0\beta_1 + \beta_0\alpha_1) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ -\hat{f}(t) \\ \hat{g}(t) \\ 1 \end{pmatrix} = \begin{pmatrix} u \\ -f(u) \\ g(u) \\ 1 \end{pmatrix}.$$

The matrix is an element of the stabiliser group of the quadratic cone defined by the equation $4X_0X_1 = X_2^2$ with vertex $\langle 0,0,0,1\rangle$. Dualising as in the previous section this implies that there is an element of the stabiliser group of the quadratic cone defined by the equation $X_0X_1 = X_2^2$ with vertex $\langle 0,0,0,1\rangle$ that maps the set of planes

$$\{tX_0 - \hat{f}(t)X_1 + \hat{g}(t)X_2 + X_3 = 0 \mid t \in GF(q)\}$$

to the planes

$$\{uX_0 - f(u)X_1 + g(u)X_2 + X_3 = 0 \mid u \in GF(q)\}.$$

The converse argument works following the above argument in reverse. Note that the determinant of the matrix is $-(\alpha_0\beta_1 - \alpha_1\beta_0)^3$ and the determinant of map $F$ is $\alpha_0\beta_1 - \alpha_1\beta_0$. Therefore $F$ will be a non-singular map and hence $H$ will be non-singular too.

## 7 $q$-clans and Translation Generalised Quadrangles

A *$q$-clan* is a set $\{A_t \mid t \in GF(q)\}$ of $q$ two by two matrices with entries from $GF(q)$ with the property that the difference of any two distinct matrices is anisotropic, i.e.

$$\alpha(A_t - A_s)\alpha^T = 0$$

$s \neq t$ implies $\alpha = (0,0)$. A $q$-clan is *additive* if $A_t + A_s = A_{t+s}$.

Consider the set of matrices

$$\left\{ \begin{pmatrix} t & g(t) \\ 0 & -f(t) \end{pmatrix} \mid t \in GF(q) \right\}$$

where $f$ and $g$ are linear over some subfield $GF(q_0)$. Let $(v,u)$ be such that $(v,u)(A_t - A_s)(v,u)^T = 0$, $s \neq t$. It follows that

$$(v,u) \begin{pmatrix} z & g(z) \\ 0 & -f(z) \end{pmatrix} (v,u)^T = 0$$

where $z = t - s$. This implies that $zv^2 + vug(z) - u^2 f(z) = 0$ and $z \neq 0$. If either $u = 0$ or $v = 0$ then $(u,v) = (0,0)$. If $u \neq 0$ then making the substitution $z = v/u$

$$zw^2 + wg(z) - f(z) = 0.$$

If this quadratic has no solutions for $w, z \in GF(q)$ and $z \neq 0$ this set of matrices is a $q$-clan. However this is the same condition as in Theorem 2 and so to a commutative semifield of rank 2 over its middle nucleus $GF(q)$ there is an associated additive $q$-clan. The following theorem is from [21]. For the definition of an egg see Section 3.

**Theorem 11.** *The set $\{A_t \mid t \in GF(q)\}$ of $2 \times 2$ matrices over $GF(q)$ is an additive $q$-clan if and only if the set $\mathcal{E} = \{E_\gamma \mid \gamma \in GF(q)^2 \cup \{\infty\}\}$, with*

$$E_\gamma = \{\langle t, -\gamma A_t \gamma^T, -\gamma(A_t + A_t^T)\rangle \mid t \in GF(q)\},$$

$$E_\infty = \{\langle 0, t, 0, 0\rangle \mid t \in GF(q)\},$$

*and tangent spaces $T_\mathcal{E} = \{T_{E_\gamma} \mid \gamma \in GF(q)^2 \cup \{\infty\}\}$,*

$$T_{E_\gamma} = \{\langle t, \beta\gamma^T + \gamma A_t^T \gamma^T, \beta\rangle \mid t \in GF(q),\ \beta \in GF(q)^2\},$$

$$T_{E_\infty} = \{\langle 0, t, \beta\rangle \mid t \in GF(q),\ \beta \in GF(q)^2\}$$

*is an egg of $PG(4n-1, q_0)$ where $q = q_0^n$.*

The construction of a generalized quadrangle $T(\mathcal{E})$ in Section 3 from an egg $\mathcal{E}$ implies that from a commutative semifields of rank 2 over its middle nucleus one can construct a generalized quadrangle of order $(q, q^2)$. This is a special case of a more general construction of generalized quadrangles due to Kantor [17]. If a generalized quadrangle $\mathcal{G}$ has an abelian collineation group that acts regularly on the points not collinear with a base point $P$ while fixing every line incident with $P$ then $\mathcal{G}$ is called a *translation generalized quadrangle*. The following theorem is from [23, (8.7.1)].

**Theorem 12.** *The incidence structure $T(\mathcal{E})$ is a translation generalized quadrangle of order $(q^n, q^m)$ with base point $(\infty)$ and conversely every translation generalized quadrangle is isomorphic to a $T(\mathcal{E})$ for some egg $\mathcal{E}$ of $PG(2n + m - 1, q)$.*

For more details and other results concerning eggs and translation generalized quadrangles refer to [20] or [21].

## 8 Concluding Remarks

It was the intention of this article to show how useful pairs of functions $f$ and $g$ from $GF(q) \to GF(q)$ linear over a subfield with the property that $g^2(x) + 4xf(x)$ is a non-square for all $x \in GF(q)^*$ are. Of course it would be of great interest to have more examples. The recent geometrical construction of the Penttila-Williams ovoid by Cardinali [9] from a Cohen-Ganley Thas-Payne flock and a Dickson Kantor Knuth flock gives hope that there may be a geometrical way to construct new examples.

The fact that the set $\mathcal{W}$ is a subspace of rank $n$ contained in the internal points of a conic is not necessarily required in the hypothesis of Theorem 8. The theorem only requires that $\mathcal{W}$ contains a subplane. One might expect that a much stronger bound should hold in Corollary 1 and Corollary 2 if one could utilise the fact that $\mathcal{W}$ is a much larger subspace for $n \geq 4$.

We have seen that the functions $f$ and $g$ allow us to construct not just a commutative semifield of rank 2 over its middle nucleus but other semifields as well. A geometrical explanation of these semifields (including the commutative semifield of rank 2 over its middle nucleus) will appear in [6].

## 9 Acknowledgements

We would like to thank Peter Cameron and Greg Stein for helpful discussions and to Dieter Jungnickel for useful remarks.

## References

1. Albert, A. A. (1952) On non-associative division algebras. Trans. Amer. Math. Soc. **72**, 296–309.

2. Bader L., Ghinelli D., Penttila T. (2001) On monomial flocks. European J. Combin. **22**, 454–474.
3. Bader, L., Lunardon, G., (1994) On non-hyperelliptic flocks. European J. Combin. **15**, 411–415.
4. Bader, L., Lunardon, G., Pinneri, I. (1999) A new semifield flock. J. Combin. Theory Ser. A **86**, 49–62.
5. Ball, S., Blokhuis, A., Lavrauw, M. On the classification of semifield flocks, preprint.
6. Ball, S., Brown, M. The six semifields associated with a semifield flock, preprint.
7. Bloemen, I., Thas, J. A., van Maldeghem, H. (1998) Translation ovoids and generalised quadrangles and hexagons. Geom. Dedicata **72**, 19–62.
8. Cameron, P. J. *Projective and Polar Spaces*, available from `http://www.maths.qmw.ac.uk/~pjc/pps/`
9. Cardinali, I., Polverino, O., Trombetti, R. On the sporadic semifield flock, preprint.
10. Cohen, S. D., Ganley, M. J. (1982) Commutative semifields two dimensional over their middle nuclei. J. Algebra **75**, 373–385.
11. Dembowski, P. *Finite Geometries*, Springer-Verlag, New York, 1968.
12. Dickson, L. E. (1906) Linear algebra in which division is always uniquely possible. Trans. Amer. Math. Soc. **7**, 370–390, 514–527.
13. Dickson, L. E. (1935) Linear algebras with associativity not assumed. Duke. Math. J. **1**, 113–125.
14. Gevaert, H., Johnson, N. L. (1988) Flocks of quadratic cones, generalized quadrangles and translation planes. Geom. Dedicata **27**, 301–317.
15. Hall Jr., M. *The Theory of Groups*, Macmillan, New York, pp. 346–420, 1959.
16. Hughes, D. R., Piper, F. *Projective Planes*, Springer-Verlag, New York, 1973.
17. Kantor, W. M. (1980) Generalized quadrangles associated with $G_2(q)$. J. Combin. Theory Ser. A **29**, 212–219.
18. Kantor, W. M. (1982) Ovoids and translation planes. Canad. J. Math. **34**, 1195–1207.
19. Knuth, D. E. (1965) Finite semifields and projective planes. J. Algebra **2**, 182–217.
20. Lavrauw, M., *Scattered subspaces with respect to spreads and eggs in finite projective spaces*, Ph. D. thesis, Technical University of Eindhoven, The Netherlands, 2001.
21. Lavrauw, M., Penttila, T. (2001) On eggs and translation generalized quadrangles. J. Combin. Theory Ser. A **96**, 303–315.
22. Lunardon, G. (1997) Flocks, ovoids of $Q(4, q)$ and designs. Geom. Dedicata **66**, 163–173.
23. Payne, S. E., Thas, J. A. *Finite generalized quadrangles*, Research Notes in Mathematics, 110. Pitman, Boston, 1984.
24. Penttila, T., Williams, B. (2000) Ovoids of parabolic spaces. Geom. Dedicata **82**, 1–19.
25. Prince, A. R. (2000) Two new families of commutative semifields. Bull. London Math. Soc. **32**, 547–550.
26. Thas, J. A. (1987) Generalized quadrangles and flocks of cones. European J. Combin. **8**, 441–452.
27. Thas, J. A. (1997) Generalized quadrangles of order $(s, s^2)$. II. J. Combin. Theory Ser. A **79**, 223–254.

28. Thas, J. A., Payne, S. E. (1994) Spreads and ovoids in finite generalized quadrangles. Geom. Dedicata **52**, 227–253.
29. Tits, J. (1962) Ovoides et Groupes de Suzuki. Arch. Math. Vol. XIII, 187–198.

# A Rao-Nam like Cryptosystem with Product Codes

Ángela I. Barbero[1] and Juan G. Tena[2]

[1] Dept. Matemática Aplicada a la Ingeniería (E.T.S. de Ingenieros Industriales)
Universidad de Valladolid, Valladolid, Spain.
[2] Dept. Álgebra, Geometría y Topología (Fac. de Ciencias)
Universidad de Valladolid, Valladolid, Spain

**Abstract.** The purpose of this paper is to design a private key cryptosystem that uses error correcting codes in an efficient way.

A secret key variant of the McEliece public key cryptosystem [5] was introduced by Rao and Nam in [7].

One of the practical drawbacks of the Rao-Nam system is that it needs to keep in memory the set of error vectors and syndromes in order to remove the errors in the decryption process.

Our proposal is to use the product of random error-correcting linear codes, and to take advantage of the product structure in order to tag the positions in error. The cryptosystem we present will have two main advantages:

- There are no memory requirements.
- The decryption process is easy.

## 1 Introduction

The well known McEliece Public Key Cryptosystem (PKC) (see [5], [1]) is based on error correcting codes. The private key of each user of the system is the generator matrix $G$ of a linear $[n, k, d]$ code with a good decoding algorithm (in the original proposal by McEliece it was a binary classic Goppa code), which is disguised as $G' = SGP$, where $S$ is an invertible matrix and $P$ a permutation matrix. $G'$, being the generator matrix of a linear code with the same parameters as the one generated by $G$, but supposedly hard to decode, is the public key of the user. The sender encrypts a $k$–bit message vector $m$ into an $n$–bit ciphertext vector $c$ as

$$c = mG' + e$$

where $e$ is a random $n$–bit error vector of weight less than or equal to the correcting capability $t$.

Jordan [4] and Rao [6] propose to use the same idea for a private key cryptosystem (by also keeping $G'$ secret). They at first thought that it would allow a drastic reduction in the size of the keys used. Nonetheless, as van Tilburg points out in [8], the small weight of the error-vectors permits, by

means of an easy *Majority Voting (MV) analysis*, the recovering of the $G'$-matrix (bringing the problem back to McEliece's PKC but with an easily breakable parameter size).

The only case in which the MV attack is not successful is when the average weight of the error vectors is $n/2$ in the binary case ($n(q-1)/q$ in the $q$-ary case). That kind of error vectors is beyond the correcting capability of any linear code. To overcome the problem, Rao and Nam in [7] propose choosing the error vectors among the elements of a (secret) set $\mathcal{Z}$ of predefined error vectors of average weight $n/2$ and each with a different syndrome.

The RN-ciphering is as follows: $y = (xSG + z)P$, $z \in \mathcal{Z}$, and $S$ and $P$ as above. At the deciphering end the user computes $y' = yP^{-1}$ and the syndrome of $y'$. Looking up in the syndrome-error table $(\mathcal{Z}, H\mathcal{Z}^T)$ he identifies $z$ and removes it.

The two main drawbacks of RN-system are the linearity of the whole process and the need to store in memory, as part of the key, the syndrome–error table, which in practice results in a limited size of such table.

Struik and van Tilburg, in [8], take advantage of both features pointed out before and cryptanalize the equivalent scheme $y = xG + z$ (let us note that the matrices $S, P$ do not play a relevant role in the RN-scheme because it does not make use of any actual decoding algorithm for $G$). That cryptanalysis relies strongly on the fact that in the binary case the sum of two encryptions $xG + z_1$, $xG + z_2$ of the same plaintext $x$ is an element $z_1 + z_2 \in \mathcal{Z} + \mathcal{Z}$, and this, together with the limited size of $\mathcal{Z}$ (limitation forced by the need to store $\mathcal{Z}$), allows the breaking of the system.

To avoid that kind of attack, which makes use of the linearity of the cryptosystem, Struik and van Tilburg propose a modification consisting of breaking the linearity by means of introducing invertible non-linear functions $f_s$ indexed by the syndrome, that scramble the space of message vectors.

Still that system has the other weakness, namely, the need to store the syndrome–error table. Making use of the limited size of the set of error vectors $\mathcal{Z}$, it can be broken as Barbero and Ytrehus show in [2].

Hence, in order to build a secure Rao-Nam-like cryptosystem one should fight against both characteristics at once.

Assuming that the problem with linearity is already solved in the way proposed by Struik and van Tilburg, we still need to design a system in which the errors used must have enormous average weight on the one hand, but on the other hand they must be in some sense correctable by the code so that there is no need to store them and, on a third hand, the total number of different errors used must be large enough to avoid attacks like that of Barbero and Ytrehus that make use of the feasibility of retrieving all cryptograms associated to a given plaintext.

All together it seems an impossible task, since random errors of enormous weight are non correctable by a linear code, but here we can make use of another characteristic of the system, that is, the errors do not appear randomly,

but are introduced on purpose by the sender, hence he can control the type of errors introduced. Now the idea is to use some family of codes that can correct errors of enormous weights provided they have a certain structure.

This paper presents a variation of the RN system that can solve the problem. We propose the use of some nice properties of the product codes in order to be able to avoid the need to store any set of errors. The idea is to make use of the structure of a product code in order to remove certain errors, that can be of considerably high weight, provided the errors satisfy certain conditions. The errors do not need to be stored since they can be generated randomly, just paying attention to satisfying the required conditions, and this can be done by a simple checking. Also the codes used as factors do not need to be of any special kind, since we do not need to use them for actual decoding. Hence random linear codes can be used, paying attention in principle only to their lengths and dimensions.

Besides, the total number of errors for suitable parameters can be made large enough so as to make computationally infeasible attacks like that of Barbero and Ytrehus, described in [2].

Regarding the problem with linearity, we follow the idea used by Struik and van Tilburg.

The only practical problem that affects this system is the fact that, as we will see later, the information rate can be poor, that is, the plaintext will in general experiment a large expansion before being sent through the channel.

The paper is organized as follows. In Section 2, we give the definition and some properties of product codes as well as some results, both of general linear codes and of product codes, that will be used in the design of our cryptosystem. Section 3 contains the description of our scheme together with a small example to illustrate it. Section 4 is for some considerations about the cryptosystem presented, specially regarding the set of errors used and also some suggestions about the choice of parameters in order to achieve good performance of the system. The last section is for the conclusions.

## 2 Product Codes

Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two linear codes over the finite field $\mathbf{F}_q$ with parameters $(n_1, k_1, d_1)$ and $(n_2, k_2, d_2)$ respectively. The definition of the product code is as follows:

**Definition 1.** The product code $\mathcal{C}$ of $\mathcal{C}_1$ and $\mathcal{C}_2$, denoted $\mathcal{C}_1 \otimes \mathcal{C}_2$, is the set of all the $n_1 \times n_2$ matrices over $\mathbf{F}_q$, whose columns are in $\mathcal{C}_1$ and whose rows are in $\mathcal{C}_2$.

The set so defined is itself a linear code over $\mathbf{F}_q$, with parameters $n = n_1 n_2$, $k = k_1 k_2$ and $d = d_1 d_2$.

If $G_1$ and $G_2$ are generator matrices of $\mathcal{C}_1$ and $\mathcal{C}_2$ respectively, then the Kronecker product matrix $G = G_1 \otimes G_2$ is a generator matrix of the product

$\mathcal{C}$. However, not every generator matrix of $G$ can be expressed in this way. It should also be remarked that the property of being susceptible to being expressed as a product code is not invariant under equivalence, that is to say, an equivalent code of a product code does not have to be a product code (See [3]).

Thus, a permutation $P$ of the columns of a generator matrix $G$ of a product code of sufficiently large length, plus a change of basis (i.e. taking $G' = SGP$) will make the recovering of the original factorization of the given code infeasible.

*Remark 1. For words of a product code of length $n = n_1 n_2$ we will denote the coordinates or positions with just one index $1 \leq j \leq n$, or with a pair $(i, j)$, $1 \leq i \leq n_1$, $1 \leq j \leq n_2$, depending on whether we are considering the word as a vector of length $n$ or as a matrix of size $n_1 \times n_2$. This will be clear from the context and we hope it will not lead to confusion.*

*Also, to pass from one notation to the other is immediate, since we will always consider the same deterministic way of passing from an $n-$vector to an $n_1 \times n_2$ matrix and viceversa. Namely, the vector will be split in $n_1$ consecutive portions of length $n_2$ and these will be placed one under another as the rows of an $n_1 \times n_2$ matrix. and reciprocally, the rows of an $n_1 \times n_2$ matrix will be placed consecutively one after another to form a vector of length $n$.*

As has already been said, the product code $\mathcal{C} = \mathcal{C}_1 \otimes \mathcal{C}_2$ is a linear code and consequently it has a correcting capability $t = [(d_1 d_2 - 1)/2]$. Nonetheless, its structure of product code enables the correction, and specially in the case that interests us, the detection and location of error configurations of weight, in most cases, much larger than its capability as a linear code.

In what follows $\mathrm{supp}(v)$ will denote the support of a vector $v$, that is the set of coordinates in which $v$ is not zero. In the same way we will denote the support of a matrix.

We will make use of the following result:

**Theorem 1.** *Let $Y = X + E$ be the received word (matrix), where $X \in \mathcal{C}$ is the sent word.*

*When the error matrix $E$ is such that none of the non zero columns of $E$ is a codeword of $\mathcal{C}_1$ and none of the non zero rows of $E$ is a codeword of $\mathcal{C}_2$, then the error $E$ can be located, in the sense that the minimum rectangle of positions $R$ such that $R \supseteq \mathrm{supp}(E)$ can be determined.*

*Proof.* Let $H_i$ be a parity check matrix for the code $\mathcal{C}_i$, $i = 1, 2$. Taking into account the hypotheses that the error matrix $E$ satisfies, we have that $H_1 Y$ shows the columns of $Y$ affected by errors, and analogously $Y H_2^T$ shows the rows of $Y$ that have errors. Hence we can determine the minimum rectangle

of positions, $R$, that contains the support of $E$ as follows,

$$R = \left\{ (i,j) \mid \begin{array}{l} i\text{–th row of } YH_2^T \text{ is non zero} \\ \text{and} \\ j\text{–th column of } H_1 Y \text{ is non zero} \end{array} \right\}$$

It is well known that a certain set of positions $J = \{j_1, \ldots, j_k\}$ that are information positions in a generator matrix $G$ of a linear $(n,k)$-code $\mathcal{C}$ ( i.e., such that the matrix $G_k(J)$, formed considering the columns corresponding to the positions in $J$, has rank $k$), can be taken as information positions in any other generator matrix $G'$ of the same code $\mathcal{C}$. Hence, one can talk about *information positions (or coordinates) in a code $\mathcal{C}$ as those that are information coordinates in any generator matrix of $\mathcal{C}$.*

Another well known fact regarding linear codes is that the information positions determine any codeword. Therefore, given a a linear $(n,k)$-code $\mathcal{C}$ over $\mathbf{F}_q$, if $e \in \mathbf{F}_q^n$, $e \neq \underline{0}$, is a vector such that $\{1, \ldots, n\} \backslash \mathrm{supp}(e)$ contains $k$ information positions in $\mathcal{C}$ then $e \notin \mathcal{C}$. This is clear because any codeword which has zeros in a set of information positions can only be the all-zero word.

Now, more specifically about product codes, let us recall some facts that we will use in the sequel.

We will call the next one lemma in order to refer to it more easily when we need it.

**Lemma 1.** *Let $G_{k_i}$ be an invertible matrix of size $k_i$, $i = 1, 2$. Then the Kronecker product $G_{k_1 \times k_2} = G_{k_1} \otimes G_{k_2}$ is invertible and $G_{k_1 \times k_2}^{-1} = G_{k_1}^{-1} \otimes G_{k_2}^{-1}$*

As a consequence of the previous result one can see that the product of information positions for the factor codes will give a rectangle of information positions for the product code.

## 3 The Proposed Scheme

Let $\mathcal{C}_i$, $i = 1, 2$, be linear codes of parameters $(n_i, k_i)$. These codes do not need to be "good" linear codes in the usual sense, in fact we are not going to use any of the main characteristics of good codes, like the minimum distance or the existence of a good decoding algorithm. In fact we can even generate these factor codes in a random way by simply taking two matrices $G_i$, $i = 1, 2$ with the corresponding sizes and ranks. Let $\mathcal{C}$ be the product code $\mathcal{C} = \mathcal{C}_1 \otimes \mathcal{C}_2$ with parameters $n = n_1 n_2$ and $k = k_1 k_2$ and let $G = G_1 \otimes G_2$ be a generator matrix of $\mathcal{C}$. We also choose matrices $S, P$ as in the McEliece PKC, that is, $S$ is an invertible scrambling matrix of size $k$ and $P$ a permutation matrix of size $n$. The role played by these matrices here is to hide the product structure of $G$. Finally let us denote by $H_i$ a parity check matrix of $\mathcal{C}_i$, for $i = 1, 2$.

Encryption: Let $m$ be the cleartext, a vector of length $k_1k_2$ over $\mathbf{F}_q$. Let $E$ be a $n_1 \times n_2$ error-matrix with support contained in a rectangle $R$ as in the setting of Theorem 1 and such that the complementary of the positions in $R$ contains $k_1 \times k_2$ information positions. This can be constructed as follows.

Supposing we first want to choose the rows with errors, then we can proceed in the following steps:

- Choose $k_1$ information positions $I = \{i_1, \ldots, i_{k_1}\}$ in $G_1$.
- Choose a vector $e_1$ of length $n_1$ and such that $\mathrm{supp}(e_1) \subseteq \{1, \ldots, n_1\} \backslash I$ (note that this guarantees that $e_1 H_1^T \neq \underline{0}$).
- Choose another vector $e_2$ of length $n_2$ such that $e_2 H_2^T \neq \underline{0}$.
- Construct $E = e_1^T \otimes e_2$

Analogously one could start by choosing the columns with errors and by exchanging the roles played by rows and columns in the description above.

Let us note that this gives us a very simple procedure to construct errors that verify the hypotheses needed in Theorem 1. But still many other error matrices that meet the required conditions can be constructed by slightly modifying the above procedure, that is, choosing, for instance, a set $I = \{i_1, \ldots, i_{k_1}\}$ of $k_1$ information positions in $\mathcal{C}_1$, and then placing random non zero values in any random set of positions $(i, j)$ chosen among those with $i \notin I$. Finally use $H_2$ to check that every non zero row of the error matrix is not a codeword of $\mathcal{C}_2$. In case some row turns out to be a codeword just make a new choice of values for the positions in that row until the resulting $n_2$-vector is not a codeword of $\mathcal{C}_2$.

$E$ can now be written as an $n_1 n_2$ vector $e$ just by putting the rows one after another.

The ciphertext will then be

$$c = (f_R(m)SG + e)P$$

where $f_R$ is a non linear invertible function $f_R : \mathbf{F}_q^{k_1 k_2} \longrightarrow \mathbf{F}_q^{k_1 k_2}$ indexed by $R$. Such a function can consist, for instance, of a permutation of the elements of $\mathbf{F}_q^{k_1 k_2}$ depending on the positions in $R$.

Decryption:

- The receiver starts by computing $c' = cP^{-1} = f_R(m)SG + e$.
- The next step is to write $c'$ as a matrix $C'$ of dimensions $n_1 \times n_2$ as explained in the remark in the previous section.
- Determine the rectangle of errors $R$ as in the proof of Theorem 1.
- Next choose $k_1 \times k_2$ information positions in $G$ outside $R$ and compute the inverse of the submatrix $G_{k_1 \times k_2}$ corresponding to the columns of $G$ in those positions. The way in which $E$ was generated, guarantees the existence of $k_1 \times k_2$ information positions outside $R$. It is also clear that $c'$ and $f_R(m)SG$ coincide in all the positions outside $R$, and hence coincide in the positions corresponding to $G_{k_1 \times k_2}$. Then

$f_R(m)SG_{k_1\times k_2}$ will be computed by simply selecting the coordinates in $c'$ in the corresponding positions.

- Compute $f_R(m)SG_{k_1\times k_2}G^{-1}_{k_1\times k_2}S^{-1} = f_R(m)$.
- Finally recover $m$ as $f_R^{-1}(f_R(m))$.

Let us point out now some characteristics of the design of the cryptosystem:

1. As we remarked previously, the factor codes $\mathcal{C}_1$ and $\mathcal{C}_2$ can be chosen randomly, just paying attention to their lengths and dimensions in order to achieve suitable parameters of the product code in the way we will precise below (see Section 4).
2. The private key the users need to exchange consists of the matrices $G_1$ and $G_2$, the matrix $S$ and the permutation $P$. The matrix $G$ and the parity check matrices $H_1$ and $H_2$ can be computed from $G_1$ and $G_2$. The two legal users do not even need to arrive to the same parity check matrices, since they only need to use them to characterize which vectors are in the corresponding factor codes, so any parity check matrix $H_i$ will work.
3. There is no need to store any set of errors. In the first procedure described, the 'factor' errors $e_1$ and $e_2$ can be generated randomly after selecting the set $I$, and then, checking whether they verify the required conditions or not can be done multiplying by the corresponding parity check matrix. In case the vector $e_i$ does not verify $e_iH_i^T \neq \underline{0}$ it can be rejected and another $e_i$ can be generated randomly until one is found that verifies the condition. In the more general procedure, described afterwards, the checking must be done over all the non zero rows (or columns).
4. The choice of $k_1 \times k_2$ information positions in $G$ and the computation of the inverse of $G_{k_1\times k_2}$ is done according to Lemma 1. Hence the computational complexity of those parts of the encryption and decryption processes depends on the parameters $n_1, n_2, k_1$ and $k_2$, and not on the parameters $n$ and $k$.

### 3.1 Example

Here we give a little example to illustrate the processes of encryption and decryption with our cryptosystem. The parameters used in the example are small to allow a full description of the whole system in a limited space, but do not pretend to be suggestions for the size of the parameters in a real application. For that we refer the reader to the next section.

In our example we will consider over the field $\mathbf{F}_2$ the codes $\mathcal{C}_1$ and $\mathcal{C}_2$ generated by the following matrices:

$$G_1 = \begin{bmatrix} 1\,1\,0\,1 \\ 0\,1\,1\,1 \end{bmatrix} \quad \text{and} \quad G_2 = \begin{bmatrix} 1\,1\,0\,0\,0 \\ 0\,1\,0\,1\,1 \\ 0\,0\,1\,0\,1 \end{bmatrix}$$

The product code $\mathcal{C}$ will have parameters $n = 20$ and $k = 6$ so the rate is aproximately $1/3$.

We will consider any couple of parity check matrices $H_1$ and $H_2$, for instance:

$$H_1 = \begin{bmatrix} 0\,1\,0\,1 \\ 1\,1\,1\,0 \end{bmatrix} \quad \text{and} \quad H_2 = \begin{bmatrix} 0\,0\,1\,1\,1 \\ 1\,1\,1\,0\,1 \end{bmatrix}$$

The generator matrix of the product code is $G = G_1 \otimes G_2$.

We will take also the scrambling matrix

$$S = \begin{bmatrix} 1\,1\,0\,0\,0\,0 \\ 0\,1\,1\,0\,0\,0 \\ 0\,0\,1\,0\,1\,0 \\ 0\,0\,0\,1\,0\,0 \\ 0\,0\,0\,0\,1\,1 \\ 0\,0\,0\,0\,0\,1 \end{bmatrix}$$

And a permutation

$$P = (2, 15, 12, 14, 5, 4, 3, 10, 8, 19, 16, 7, 6, 18, 17, 11, 1, 13, 20, 9)$$

Now recall that the actual private key is $G_1, G_2, S$ and $P$, hence the total size of the key is $8 + 15 + 36$ binary symbols to represent the matrices $G_1, G_2$ and $S$, and 69 bits to represent a permutation of order 20 (See [2]).

Now suppose Alice wants to send the message $m = (1, 1, 0, 1, 0, 1)$.

Encryption:

- First she chooses an error vector $e_1$ of length 4 whose support leaves out $k_1 = 2$ information positions of $\mathcal{C}_1$. Since in this case the code $\mathcal{C}_1$ is such that any two positions are information positions, that means she can choose any vector $e_1$ with $w(e_1) \leq 2$, for instance $e_1 = (0, 1, 0, 1)$.
  Now as $e_2$ she can choose any vector in $\mathbf{F}_2^5 \backslash \mathcal{C}_2$, for instance $e_2 = (1, 1, 1, 1, 1)$ $(H_2^T e_2 \neq \underline{0})$.
  The error matrix is

$$E = \begin{bmatrix} 0\,0\,0\,0\,0 \\ 1\,1\,1\,1\,1 \\ 0\,0\,0\,0\,0 \\ 1\,1\,1\,1\,1 \end{bmatrix}$$

  or, written as a vector of length 16,

$$e = (0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1).$$

  Let us note that $w(e) = 8 = n/2$.
  The minimum rectangle that contains the support of $E$ is $R = \{(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (4, 1), (4, 2), (4, 3), (4, 4), (4, 5)\}$.

- Now suppose that the non linear function $f_R : \mathbf{F}_2^6 \longrightarrow \mathbf{F}_2^6$ is defined as the $r$-th cyclic permutation acting over the numbers $0, 1, \ldots, 2^6 - 1$ (written as 6-uples of binary numbers), and where $r$ is the sum modulo $2^6$ of all the indices that appear in the positions in $R$ (let us remark that we have chosen this function here for the sake of simplicity in the example. In a real application a 'wiser' choice of $f_R$ should be made, considering the non linearity and also a reasonably even distribution of its outputs over the set $\mathbf{F}_q^{k_1 k_2}$).
  In our case $r = 2+1+2+2+\cdots+4+5 = 60$ and $f_R(x) = x+60 \mod 2^6$.

$$m = (1, 1, 0, 1, 0, 1) = 1 + 2 + 2^3 + 2^5 = 43,$$

  hence

$$f_R(m) = 43 + 60 \mod 2^6 = 39 = 1 + 2 + 2^2 + 2^5 = (1, 1, 1, 0, 0, 1)$$

- Finally the encrypted text will be

$$c = (f_R(m)SG + e)P = (1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 1, 0, 1, 1, 1, 0)$$

  At the other end Bob receives $c$ and proceeds to decrypt it.
  Decryption:

- $c' = cP^{-1} = f_R(m)SG+e = (1, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1)$.

$$C' = \begin{bmatrix} 1\ 1\ 0\ 0\ 0 \\ 0\ 1\ 0\ 0\ 1 \\ 0\ 1\ 1\ 1\ 0 \\ 0\ 1\ 0\ 0\ 1 \end{bmatrix}$$

  We will now denote by $C_i$ and $c_i$ the $i$-th column and row of $C'$ respectively.
- Now Bob will use the parity check matrices to compute the syndromes, by columns

$$C_1 H_1^T \neq \underline{0}, C_2 H_1^T \neq \underline{0}, C_3 H_1^T \neq \underline{0}, C_4 H_1^T \neq \underline{0}, C_5 H_1^T \neq \underline{0}$$

  and by rows

$$c_1 H_2^T = \underline{0}, c_2 H_2^T \neq \underline{0}, c_3 H_2^T = \underline{0}, c_4 H_2^T \neq \underline{0}$$

  Hence $R = \{(2,1), (2,2), (2,3), (2,4), (2,5), (4,1), (4,2), (4,3), (4,4), (4,5)\}$, and $f_R(x) = x + 60 \mod 2^6$ and so $f_R^{-1}(x) = x - 60 \mod 2^6$.
- The next step is to choose $k_1 k_2 = 6$ information positions of $\mathcal{C}$ outside $R$. To achieve that, Bob should first consider 2 information positions in $\mathcal{C}_1$ outside $\{2, 4\}$, so there is only one option, those positions must be $I = \{1, 3\}$. Then he should choose 3 information positions in $\mathcal{C}_2$. For instance $J = \{1, 2, 3\}$.

Hence

$$G_{k_1}(I) = \begin{bmatrix} 1\,0 \\ 0\,1 \end{bmatrix} \text{ and } G_{k_2}(J) = \begin{bmatrix} 1\,1\,0 \\ 0\,1\,0 \\ 0\,0\,1 \end{bmatrix}$$

$$G_{k_1 \times k_2}(I \times J)^{-1} = G_{k_1}(I)^{-1} \otimes G_{k_2}(J)^{-1} = \begin{bmatrix} 1\,1\,0\,0\,0\,0 \\ 0\,1\,0\,0\,0\,0 \\ 0\,0\,1\,0\,0\,0 \\ 0\,0\,0\,1\,1\,0 \\ 0\,0\,0\,0\,1\,0 \\ 0\,0\,0\,0\,0\,1 \end{bmatrix}$$

- The vector $f_R(m)SG_{k_1 \times k_2}(I \times J)$ is the vector of length $|I| \cdot |J|$ obtained from $c'$ by selecting only those positions corresponding to $I \times J$, so in this case
$$f_R(m)SG_{k_1 \times k_2}(I \times J) = (1, 1, 0, 0, 1, 1)$$
- $$f_R(m) = (1, 1, 0, 0, 1, 1)G_{k_1 \times k_2}(I \times J)^{-1}S^{-1} = (1, 1, 1, 0, 0, 1)$$
- And finally
$$m = f_R^{-1}(f_R(m)) = f_R^{-1}(39) = 39 - 60 \mod(2^6) = 43 = (1, 1, 0, 1, 0, 1)$$

## 4 About the Set of Errors. How to Choose Parameters

Let us call $\mathcal{E}$ the set of admissible error vectors in our cryptosystem, that is, the set of errors $e$ such that, once written as $n_1 \times n_2$ matrices $E$, and denoting by $R = I \times J$ the minimum rectangle of positions such that $R \supseteq \text{supp}(E)$, it verifies either

- $\{1, \ldots, n_1\} \backslash I$ contains $k_1$ information positions of $\mathcal{C}_1$ and no nonzero row of $E$ is in $\mathcal{C}_2$, or
- $\{1, \ldots, n_2\} \backslash J$ contains $k_2$ information positions of $\mathcal{C}_2$ and no nonzero column of $E$ is in $\mathcal{C}_1$.

The weight of the error vectors considered in $\mathcal{E}$ is $\leq max\{(n_1 - k_1)n_2, (n_2 - k_2)n_1\}$. Large $k_i$ will result in acceptable information rates, while small $k_i$ will allow the use of error vectors of high weight. The introduction of the non linear functions $f_R$ that scramble the messages fades the immediate threat of MV on each coordinate, since it is in principle impossible for the attacker to recover, or better, to recognize, encryptions that differ only in the error vector used, but still it would be safer to count on the possibility of using error vector of high weight in order to avoid other hypothetical attacks that

could take advantage of the weakness that error vectors of low weight might suppose.

Hence, to have a balance between both desirable characteristics one possibility would be to combine two different codes, one with low and the other with high rate.

Also the range of weights that the error vectors can have is wide, and this avoids the weakness of having error vectors of constant or almost constant weight, which could allow other kinds of attacks.

Besides, the next lemma guarantees that the error vectors are evenly distributed in the vector space $\mathbf{F}_q^{n_1 n_2}$, which in turn avoids attacks that could take advantage of the fact of the error vectors being concentrated in some particular coset of the code $\mathcal{C}$.

**Lemma 2.** *Two different error vectors in $\mathcal{E}$ will always belong to different cosets of $\mathcal{C}$*

*Proof.* Let us consider two error vectors $e_1, e_2 \in \mathcal{E}$, and let us consider the corresponding matrices $E_1$ and $E_2$. Let $R_i = I_i \times J_i$ be the minimum rectangle such that $R_i \supseteq \mathrm{supp}(E_i)$, for $i = 1, 2$. And suppose that, for instance, $I_1$ leaves out $k_1$ information positions of $\mathcal{C}_1$ (analogously for $I_2$).

Let $R = I \times J$ be the minimum rectangle such that $R \supseteq \mathrm{supp}(E_1 - E_2)$.

- If $\{1, \ldots, n_1\} \backslash I$ contains $k_1$ information positions. Each column of $E_1 - E_2$ will belong to $\mathcal{C}_1$ if and only if it is the allzero vector, that is, $E_1 - E_2$ will belong to $\mathcal{C}$ if and only if $E_1 = E_2$.
- On the contrary, if $\{1, \ldots, n_1\} \backslash I$ does not contain $k_1$ information positions of $\mathcal{C}_1$, then it is clear that there exists $i_0 \in I_2 \backslash I_1$. And then the $i_0$-th row of $E_1 - E_2$ will be the $i_0$-th row of $E_2$ multiplied by $-1$, and hence, not a vector in $\mathcal{C}_2$. Thus $E_1 - E_2 \notin \mathcal{C}$.

Regarding the cardinality of $\mathcal{E}$, the exact formula cannot be given in general since it depends strongly on the factor codes chosen in each particular case. But some bounds can be given specially in the case where we do not consider all the possible error vectors in $\mathcal{E}$, but only those that result from the construction first described in the encryption process in Section 3. We will denote by $\mathcal{E}'$ the set of those error vectors.

Let us continue with the same notations as before, and consider all the possible choices of $e_1$ such that the complementary of its support contains $k_1$ information positions in $\mathcal{C}_1$. How many different $e_1$ can be considered? It depends on the particular code $\mathcal{C}_1$ we are using; to be precise, it depends on how many ways there are to choose $k_1$ information positions in $\mathcal{C}_1$. One extreme case is that any $k_1$ positions in $\mathcal{C}_1$ are information. For instance, the code generated by $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ over $\mathbf{F}_2$ has that property. In that case, for $i \leq n_1 - k_1$, the number of error vectors $e_1$ of weight $i$ that can be taken is

$\binom{n_1}{i}(q-1)^i$, and the total number of vectors $e_1$ would be

$$\sum_{i=1}^{n_1-k_1}\binom{n_1}{i}(q-1)^i$$

and for each $e_1$ one can choose a vector $e_2$ with the only condition that $e_2 H_2^T \neq \underline{0}$, that is, there are $q^{n_2}-q^{k_2}$ possible choices for $e_2$.

The same can be done exchanging the roles played by $e_1$ and $e_2$.

Finally we have to discount the patterns that have been counted twice, that is, those of the form $e_1^T \otimes e_2$ where $e_i$ leaves out $k_i$ information positions for both $i=1$ and 2, so we have a total number of errors given by

$$\begin{aligned}
&U(n_1,k_1,n_2,k_2,q) = \\
&\left(\sum_{i=1}^{n_1-k_1}\binom{n_1}{i}(q-1)^i\right)(q^{n_2}-q^{k_2}) + \left(\sum_{i=1}^{n_2-k_2}\binom{n_2}{i}(q-1)^i\right)(q^{n_1}-q^{k_1}) - \\
&\left(\sum_{i=1}^{n_1-k_1}\binom{n_1}{i}(q-1)^i\right)\left(\sum_{i=1}^{n_2-k_2}\binom{n_2}{i}(q-1)^i\right) = \\
&= -\prod_{j=1}^{2}\left[\left(\sum_{i=1}^{n_j-k_j}\binom{n_j}{i}(q-1)^i\right) - (q^{n_j}-q^{k_j})\right] + \prod_{j=1}^{2}(q^{n_j}-q^{k_j})
\end{aligned}$$

On the other hand, the other extreme case would be that there is only one possible choice of $k_i$ information coordinates in $\mathcal{C}_i$, but this requires $\mathcal{C}_i$ to be degenerated, that is, to have some all-zero column in its generator matrix. Provided $\mathcal{C}_i$ is not degenerated, the worst case would give us $\binom{n_i-k_i+1}{j}(q-1)^j$ choices of a vector $e_i$ of weight $j$ whose support leaves out $k_i$ information positions. For instance, the code generated by $\begin{bmatrix}1&0&0&0\\0&1&1&1\end{bmatrix}$ corresponds to that worst case.

In this case the total number of possible error choices is

$$\begin{aligned}
&L(n_1,k_1,n_2,k_2,q) = \\
&-\prod_{j=i}^{2}\left[\left(\sum_{i=1}^{n_j-k_j}\binom{n_j-k_j+1}{i}(q-1)^i\right) - (q^{n_j}-q^{k_j})\right] + \prod_{j=1}^{2}(q^{n_j}-q^{k_j})
\end{aligned}$$

In a general case the cardinality of the set of possible error vectors $\mathcal{E}'$ will be

$$L(n_1,k_1,n_2,k_2,q) \leq |\mathcal{E}'| \leq U(n_1,k_1,n_2,k_2,q)$$

We should remark here that in the case of using arbitrarily chosen matrices, with no particular structure, the actual number of possible error vectors tends to be closer to the upper bound $U(n_1, k_1, n_2, k_2, q)$ than to the lower one $L(n_1, k_1, n_2, k_2, q)$. Just to show a little example, for the matrix $G_1 = \begin{bmatrix} 1\,1\,0\,1\,1\,0 \\ 0\,1\,0\,1\,0\,1 \\ 0\,0\,1\,0\,1\,1 \end{bmatrix}$ over $\mathbf{F}_2$ the actual number of choices of error vectors whose support leaves out $k_1 = 3$ information positions is 33, while the best case would give 41 different choices and the worst case only 14 different possibilities.

Nonetheless, the actual cardinality of $\mathcal{E}$ is in general much larger than the cardinality of $\mathcal{E}'$. Just to show a little unreal example, if we consider over $\mathbf{F}_2$ the codes $\mathcal{C}_1$ and $\mathcal{C}_2$, generated by $G_1 = G_2 = \begin{pmatrix} 1\,0\,1\,1 \\ 0\,1\,1\,0 \end{pmatrix}$, we have $|\mathcal{E}| = 1307$, while for $\mathcal{E}'$ we have $L(4, 2, 4, 2, 2) = 108 \leq \sharp(\mathcal{E}') \leq 140 = U(4, 2, 4, 2, 2)$.

Finally, to illustrate the fact that the set of error vectors can have enormous size, even for small parameters of the factor codes, let us consider

$$G_1 = G_2 = \begin{bmatrix} 1\,2\,5\,5\,1\,0 \\ 0\,1\,2\,5\,5\,1 \end{bmatrix}$$

over $\mathbf{F}_7$. In this case we have

$$|\mathcal{E}'| = U(6, 2, 6, 2, 7) = 5131586304$$

and the cardinality of $\mathcal{E}$ will be in fact much larger.

### 4.1 The Choice of Parameters

According to the remarks pointed out above, it would be wise to consider parameters that allow a large number or error vectors, but taking care not to reduce the rate of the code and, consequently, increase the costs too drastically. In order to achieve this we would suggest considering a balanced combination of field size and rates of the factor codes.

For instance, if one of the factor codes has parameters $(n_1, n_1 - 1)$, while the other has parameters $(n_2, n_2/T)$ we can achieve a code of rate approximately $1/T$, which will allow us the use of error vectors of weights up to $n(T-1)/T$ (in case $\mathbf{F}_q^{n_2} \backslash \mathcal{C}_1$ contains some vector of weight $n_2$). So, these parameters conveniently chosen, together with a suitable field size, can give us an enormous cardinality of the set of errors for each code used as private key, and also a really large number of different private keys to choose with the same parameters, which will avoid attacks based on exhaustive search of the key used.

As an example, if we consider over $\mathbf{F}_{16}$ factor codes of parameters $(10, 9)$ and $(10, 3)$ we will obtain a cryptosystem with rate approximately $1/3$. Even if we do not consider all the possible error vectors, but only those that can

be constructed with the first procedure described, we still have that $|\mathcal{E}'|$ is between $L(10, 9, 10, 3, 16)$ and $U(10, 9, 10, 3, 16)$, and a simple computation shows that these numbers are enormous. Also we remark that the actual number of admissible errors $|\mathcal{E}|$ will be much bigger. Hence an attack like that of Barbero and Ytrehus that needs the computation of all cryptotexts associated to one chosen plain text will be infeasible here.

Besides, provided $\mathbf{F}_{16}^{10} \backslash \mathcal{C}_1$ contains vectors of weight 10 this will give us error vectors of weights up to 70. Certainly the average weight will not be that high, but still this shows that we can use errors in a window of weights large enough to avoid attacks that could take advantage of the weakness of using errors of low weight, or errors of weight kept constant or almost constant. Even when the immediate threat of a majority voting attack is now less important since the non linearity of the functions used to scramble the space of messages protects against such attack, it should not be discarded that the use of errors of low weight can allow some other kind of attack.

The size of the private key for such a code will be: $10 \cdot 9 \cdot 4$ bits to store one of the generator matrices, $10 \cdot 3 \cdot 4$ bits to store the other generator matrix, $27 \cdot 27 \cdot 4$ bits to store the scrambling matrix $S$ and 473 bits to represent a permutation $P$ of order 100. Hence the total size of the key is 3869 bits. Less than 0.4 Kb to encrypt $16^{27}$ different messages.

And finally, the number of different private keys that correspond to that particular choice of parameters is the number of different matrices with 10 columns and 9 rows and rank 9 over $\mathbf{F}_{16}$ times the number of different matrices with 10 columns and 3 rows and rank 3 over the same field. Just computing the number of $9 \times 10$ matrices of rank 9 over $\mathbf{F}_{16}$ gives us $(\sum_{i=0}^{8}(16^9 - 16^i))(\sum_{i=0}^{9} 16^i)$ and this amounts to $0.5 \cdot 10^{23}$ different matrices with those characteristics, and still this number should be multiplied by the number of matrices $3 \times 10$ with rank 3 over $\mathbf{F}_{16}$. This simple example of a possible choice of parameters gives us the idea of the enormous size of the set of different private keys that can be chosen by the two parties. This is possible precisely because the system do not need of any particular kind of code to be used as component of the product. Even those with very poor performance as usual linear codes can be used here, since the process does not make use of any decoding algorithm in the usual sense.

## 5 Conclusions

We have presented a variation of the Rao-Nam cryptosystem that, with a reduced key size, is strong against known attacks that can break the Rao-Nam original cryptosystem and some of its modifications.

The construction is based on profitting from properties of the product codes.

The only drawback of our system when compared with the previous versions of private key cryptosystems that make use of error correcting codes

may be the low information rate that can be achieved in general with our system.

On the other hand, the possibility that product codes offer to locate errors very easily when they are placed in some particular configuration, together with the idea of leaving free of errors a sufficient number of information positions allows us to recover the original message with neither the need to use any actual decoding algorithm for the code $\mathcal{C}$ nor the obligation to store in memory large error-syndrome tables. Also the whole process of ciphering and deciphering requires low complexity, since it simply consists of matrix multiplications. And for the more complex tasks like inverting some matrices, this can be done by means of working with the factor components, so the complexity, even in those steps, depends only on the small parameters of the factor codes and not on the actual parameters of the product code.

Also in the last section we have shown how even reasonably small parameters will provide enough number of different private keys, and once a particular private key has been chosen, enough number of different error vectors to make our system resilient against attacks that need some kind of exhaustive search like those that have been designed against the original Rao–Nam system and some of its variations.

## References

1. C.M. Adams and H. Meijer. "Security–Related Comments Regarding McEliece's Public–Key Cryptosystem". *Lecture Notes on Computer Science* vol. 293. Eurocrypt 87. Springer Verlag 1987.
2. A.I.Barbero and Ø. Ytrehus "An attack on the Rao-Nam-Struik-van Tilburg Cryptosystem". *Manuscript sent to IEEE Trans. on Info. Theory.* August 2001.
3. A.I. Barbero. "An Algorithm for Characterizing Linear Bidimensional Product Codes". *Arithmetic. Geometry and Coding Theory. Proceedings AGCT–IV.* Luminy 1993. Walter de Gruyter, 1996.
4. J.P.Jordan. "A variant of a public-key cryptosystem based on Goppa codes". *Sigact news* 15, 1983, no. 1, pp. 61–66.
5. R.J. McEliece. "A Public–Key Cryptosystem Based on Algebraic Coding Theory". *DSN Progress Report.* 42–44. Jet Propulsion Laboratory 1978, California Inst. of Tech. Pasadena, Cal.
6. T.R.N. Rao. "Cryptosystem using algebraic codes". *Int. Conf. on Computer Systems&Signal Processing*, December 1984, Bangalore, India.
7. T.R.N. Rao and Nam. "Private–Key Algebraic–Coded Cryptosystems".*Lecture Notes on Computer Science.* vol.263, Springer Verlag 1987. pp.35–48.
8. J.van Tilburg. "Security–Analysis of a Class of Cryptosystems Based on Linear Error–Correcting Codes". *Royal PTT* 1994, The Nederlands.

# Pseudorandom Sequences from Elliptic Curves

P.H.T. Beelen and J.M. Doumen *

Department of Mathematics and Computer Science, Technische Universiteit Eindhoven, PO Box 513, 5600 MB Eindhoven, The Netherlands.
E-mail: p.h.t.beelen@tue.nl and doumen@win.tue.nl

**Abstract.** In this article we will generalize some known constructions to produce pseudorandom sequences with the aid of elliptic curves. We will make use of both additive and multiplicative characters on elliptic curves. Keywords are: Artin-Schreier extension, Kummer extension, elliptic curve, exponential sum, correlation, balance, linear recurrencies.

## 1 Introduction

Nowadays, many applications call for random numbers. One of the most preferable ways to generate those would be to take a monkey, give him a coin to flip, and write down the result of each coin flip. Unfortunately this process is quite slow, and we would like a faster way to generate random numbers. On second thought, a sequence of numbers that *appears* random would be just as good – who could tell the difference? We will call such a sequence pseudorandom.

Many people have constructed pseudorandom number generators using many, diverse methods (see for example Chap. 5 of [7] for an overview). The first study of using linear congruences on elliptic curves to generate pseudorandom sequences was done in [6]. Further results on these generators were obtained in [3,4,8,15]. We will generalize some of these constructions and introduce another construction using linear recurrence relations on elliptic curves. An instance of this last construction was investigated in [5].

## 2 Some Properties of Elliptic Curves

As we will use elliptic curves throughout this article, we will start by fixing some notation and giving some elementary properties of elliptic curves. We will denote the Galois field of $q$ elements by $\mathbb{F}_q$ and an elliptic curve by $\mathcal{E}$. The group of $\mathbb{F}_q$-rational points on the curve $\mathcal{E}$ will be denoted by $\mathcal{E}(\mathbb{F}_q)$ and the function field of an algebraic curve $\mathcal{C}$ by $\mathcal{F}(\mathcal{C})$, or by $\mathcal{F}_q(\mathcal{C})$ if we only want the functions with coefficients in $\mathbb{F}_q$. The algebraic closure of a field $F$ will be denoted by $\overline{F}$. Scalar multiplication of a point $P$ on an elliptic curve

* This author was supported by STW in the project *Strong Authentication Methods*, number EWI.4536.

by $n$ will be denoted by $[n]P$. We denote the $n$-torsion subgroup of $\mathcal{E}(\overline{\mathbb{F}}_q)$ by $\mathcal{E}[n]$.

For the following proposition, see p.145 of [11].

**Proposition 1.** *Let $\mathcal{E}$ be an elliptic curve defined over the finite field $\mathbb{F}_q$. There exist numbers $k$ and $l$ such that as abelian groups*

$$\mathcal{E}(\mathbb{F}_q) \cong \mathbb{Z}/k\mathbb{Z} \times \mathbb{Z}/kl\mathbb{Z}.$$

*Furthermore, $k$ divides $(q-1)$.*

We will denote the $k$ and $l$ from the above proposition respectively by $k(\mathcal{E})$ and $l(\mathcal{E})$ or, if we want to stress the field of definition, by $k(\mathcal{E}, \mathbb{F}_q)$ and $l(\mathcal{E}, \mathbb{F}_q)$ respectively.

Since $k = k(\mathcal{E}.\mathbb{F}_q)$ divides $q-1$, we see that the multiplication by $k$ map from $\mathcal{E}$ to $\mathcal{E}$ is unramified of degree $k^2$. Further note that $\mathcal{E}[k] \subset \mathcal{E}(\mathbb{F}_q)$.

## 3 Pseudorandom Sequences

In this section we will give some basic definitions concerning pseudorandom sequences.

**Definition 1.** Let $S = \{s(0), s(1), \ldots, s(N-1)\}$ be a sequence of elements of $\mathbb{F}_q$ and let $\alpha \in \mathbb{F}_q^*$. Denote the characteristic of $\mathbb{F}_q$ by $p$. We define the *balance with respect to* $\alpha$ in the following way:

$$B_S(\alpha) = \frac{1}{N} \sum_{i=0}^{N-1} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q|\mathbb{F}_p}(\alpha s(i))},$$

with $\zeta_p$ the $p^{\text{th}}$ root of unity $\exp(2\pi i/p)$. Further we define the *balance* to be

$$B_S = \max_{\alpha \in \mathbb{F}_q^*} \{|B_S(\alpha)|\}.$$

Now we introduce a similar concept for sequences defined over $\mathbb{Z}/m\mathbb{Z}$. We will assume that $m$ divides $q-1$. Then it is possible to identify $\mathbb{Z}/m\mathbb{Z}$ with $(\mathbb{F}_q^*)^{(q-1)/m}$. Thus there exists a surjective homomorphism of groups $\chi_m : \mathbb{F}_q^* \to \mathbb{Z}/m\mathbb{Z}$.

If $S = \{s(0), s(1), \ldots, s(N-1)\}$ is a sequence of elements of $(\mathbb{F}_q^*)^{(q-1)/m}$, then we define the balance with respect to $\alpha$ to be

$$B_S(\alpha) = \frac{1}{N} \sum_{i=0}^{N-1} \zeta_m^{\chi_m(\alpha s(i))},$$

with $\alpha \in \mathbb{F}_q^*$ and $\zeta_m = \exp(2\pi i/m)$.

We now define the autocorrelation of a sequence.

**Definition 2.** Let $\{s(0), s(1), \ldots, s(N-1)\}$ be a sequence $S$ of period $N$ defined over the finite field $\mathbb{F}_q$. Write $p$ for the characteristic of this field. Furthermore, let $\alpha, \beta \in \mathbb{F}_q^*$.
We define the *autocorrelation with respect to* $\alpha$ *and* $\beta$ of a sequence as follows:

$$C_S(d, \alpha, \beta) = \frac{1}{N} \sum_{i=0}^{N-1} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q|\mathbb{F}_p}(\alpha s(i+d) - \beta s(i))},$$

with $0 \leq d < N$ and $\zeta_p = \exp(2\pi i/p)$. For a sequence $S$ defined over $(\mathbb{F}_q^*)^{(q-1)/m}$ we define

$$C_S(d, \alpha, \beta) = \frac{1}{N} \sum_{i=0}^{N-1} \zeta_m^{\chi_m(\alpha s(i+d) - \beta s(i))},$$

with $\zeta_m = \exp(2\pi i/m)$.

Note that in the above definition $i + d$ should be read modulo $N$. Also note that for binary sequences this definition amounts to

$$C_S(d) = \frac{1}{N} \sum_{i=0}^{N-1} (-1)^{s(i+d)+s(i)},$$

which is the usual definition of the autocorrelation (see for example Chap. 5, Sect. 4 of [7]).

Another useful object is the crosscorrelation of two sequences. It is defined as follows:

**Definition 3.** Let $S = \{s(i)\}$ and $T = \{t(i)\}$ be two sequences defined over $\mathbb{F}_q$ having the same period $N$. Denote the characteristic of $\mathbb{F}_q$ by $p$ and let $\alpha, \beta \in \mathbb{F}_q^*$. We define the *crosscorrelation of* $S$ *and* $T$ *with respect to* $\alpha$ *and* $\beta$ by

$$C_{S,T}(d, \alpha, \beta) = \frac{1}{N} \sum_{i=0}^{N-1} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q|\mathbb{F}_p}(\alpha s(i+d) - \beta t(i))},$$

with $\zeta_p = \exp(2\pi i/p)$ and $0 \leq d < N$. For sequences $S$ and $T$ defined over $(\mathbb{F}_q^*)^{(q-1)/m}$ we define

$$C_{S,T}(d, \alpha, \beta) = \frac{1}{N} \sum_{i=0}^{N-1} \zeta_m^{\chi_m(\alpha s(i+d) - \beta t(i))},$$

with $\zeta_m = \exp(2\pi i/m)$.

The problem is to find a family of sequences $\Sigma = \{S_i | i \in I\}$ such that for all $i, j \in I$ the crosscorrelations $C_{S_i,S_j}(d, \alpha, \beta)$ are small.

## 4 Pseudorandom Sequences from Elliptic Curves Using Additive Characters

Some generalizations of known constructions of pseudorandom sequences from elliptic curves will be given in this section.

Let $\mathcal{E}$ be an elliptic curve defined over a finite field $\mathbb{F}_{q^e}$ of characteristic $p$. Suppose for now that this group is cyclic of order $N$ and has generator $P$. Let $f \in \mathcal{F}_{q^e}(\mathcal{E})$ be a function on $\mathcal{E}$ defined over $\mathbb{F}_{q^e}$. We can define a pseudorandom sequence $S = \{s(i)\}$ as follows:

$$s(i) = \mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f([i]P)),$$

with $0 \leq i < N$. Here $\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}$ denotes the trace map from $\mathbb{F}_{q^e}$ to $\mathbb{F}_q$ defined by

$$\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(x) = x + x^q + x^{q^2} + \cdots + x^{q^{e-1}}.$$

We see that the condition that $\mathcal{E}(\mathbb{F}_{q^e})$ is a cyclic group is a natural one, since we need an ordering of the points in $\mathcal{E}(\mathbb{F}_{q^e})$. In the literature this assumption is often made. Moreover, the field $\mathbb{F}_q$ is usually assumed to be $\mathbb{F}_p$. We will remove both restrictions. First we need some definitions.

**Definition 4.** Let $\mathcal{C}$ be an algebraic curve of genus $g$ defined over $\mathbb{F}_q$. Let $f \in \mathcal{F}_q(\mathcal{C})$ be a rational function on $\mathcal{C}$ defined over $\mathbb{F}_q$ as well. We define $\mathcal{C}^{\mathrm{AS}}(f, \mathbb{F}_q)$ to be the set of all $\mathbb{F}_q$-rational points $Q$ on $\mathcal{C}$ such that there exists a $g \in \mathcal{F}_q(\mathcal{C})$ (depending on $Q$) with the property that $f - g^p + g$ is defined at $Q$.

Note that for every function $f \in \mathcal{F}_q(\mathcal{C})$ and point $Q \in \mathcal{C}(\overline{\mathbb{F}_q})$ there exists a function $g \in \mathcal{F}_q(\mathcal{C})$ such that either $v_Q(f - g^p + g) \geq 0$ or $v_Q(f - g^p + g) < 0$ and $p \nmid v_Q(f - g^p + g)$. We define $m_Q = -1$ in the former case and $m_Q = -v_Q(f - g^p + g)$ in the latter. Of course $m_Q$ depends on $f$ as well. When we want to make this explicit we will write $m_Q(f)$ instead of $m_Q$. For more details see p.114 of [12].

Also observe that the quantity $\mathrm{Tr}_{\mathbb{F}_q|\mathbb{F}_p}((f - g^p + g)(Q))$ does not depend on $g$ as long as the function $f - g^p + g$ is defined at $Q$. This is why for $Q \in \mathcal{C}^{\mathrm{AS}}(f, \mathbb{F}_q)$ we will write $\mathrm{Tr}_{\mathbb{F}_q|\mathbb{F}_p}(f(Q))$ for this quantity even if $f$ itself is not defined in $Q$.

We will now define the sequence we want to study.

**Definition 5.** Let $\mathcal{E}$ be an elliptic curve defined over $\mathbb{F}_{q^e}$. Suppose that $P$ is a generator of the group $[k(\mathcal{E}, \mathbb{F}_{q^e})]\mathcal{E}(\mathbb{F}_{q^e})$ and denote its order by $N$. Let $f \in \mathcal{F}_{q^e}(\mathcal{E})$. We define the sequence $S^{\mathrm{AS}}(f, P) = \{s(i)\}_{0 \leq i < N}$ by

$$s(i) = \mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f([i]P)).$$

Here we use the convention that $\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f(Q)) = 0$ if $Q \notin \mathcal{E}^{\mathrm{AS}}(f, \mathbb{F}_{q^e})$. Of course this sequence depends on the elliptic curve as well, but we do not make this explicit in the notation.

Note that in the above notation $N = l(\mathcal{E}, \mathbb{F}_{q^e})$ and that if $\mathcal{E}(\mathbb{F}_{q^e})$ is cyclic, we are back in the situation that has already been studied in the literature [4,15].

From the point of view of coding theory we do not need an ordering of the points in $\mathcal{E}(\mathbb{F}_{q^e})$. In this case any change of ordering gives rise to an equivalent code. Indeed there is in that case no need of restricting oneself to elliptic curves. The resulting codes are called trace-codes. They have been studied in for example Chap. VIII of [12] and [14].

Before we give some estimates for the parameters of the pseudorandom sequences defined above, we need some theory. The key to the results is the proposition about the following exponential sum.

**Definition 6.** Let $\mathcal{C}$ be an algebraic curve defined over $\mathbb{F}_q$. Let $f \in \mathcal{F}_q(\mathcal{C})$. We define the following exponential sum:

$$ES^{\mathrm{AS}}(\mathcal{C}, f) = \sum_{P \in \mathcal{C}^{\mathrm{AS}}(f, \mathbb{F}_q)} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_q|\mathbb{F}_p}(f(P))},$$

with $\zeta_p = \exp(2\pi i/p)$.

We will now give a known upper bound for this exponential sum.

**Proposition 2.** *Let $\mathcal{C}$ be an algebraic curve of genus $g$ defined over $\mathbb{F}_q$. Let $f \in \mathcal{F}_q(\mathcal{C})$ and suppose that $f \neq z^p - z$ for all $z \in \overline{\mathcal{F}(\mathcal{C})}$. Then the following holds:*

$$\left|ES^{\mathrm{AS}}(\mathcal{C}, f)\right| \leq \left(2g - 2 + \sum_{P \in \mathcal{C}(\overline{\mathbb{F}}_q)} (m_P + 1)\right) \sqrt{q}.$$

*Proof.* This proposition was proven in [2,8,15]. We give the gist of the proof for the convenience of the reader. We can rephrase the proposition by considering the curve $\mathcal{D}$ defined over $\mathbb{F}_q$ whose function field is given by $\mathcal{F}_q(\mathcal{C})(z)$ with $z^p - z = f$. Denote its genus by $h$. The $L$-function of $\mathcal{D}$ is the product of the $L$-function of $\mathcal{C}$ with the following $p-1$ expressions $(1 \leq i \leq p-1)$:

$$\exp\left(\sum_{e \geq 1} \frac{T^e}{e} \sum_{P \in \mathcal{C}^{\mathrm{AS}}(f, \mathbb{F}_{q^e})} \zeta_p^{i\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_p}(P)}\right).$$

As a matter of fact the above expressions turn out to be polynomials. By Hasse-Weil's theorem these polynomials have roots of length $1/\sqrt{q}$ and hence we find for all $e \geq 1$

$$\left|\sum_{P \in \mathcal{C}^{\mathrm{AS}}(f, \mathbb{F}_{q^e})} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_p}(P)}\right| \leq \frac{2(h-g)}{p-1} \sqrt{q^e}.$$

Using the theory for Artin-Schreier extensions we can find an explicit expression for the genus $h$ (see for example Chap. III, Sect. 7 of [12]). This leads to the upper bound given in the proposition.

We will now apply the above result to give an estimate for the parameters of the sequences $S^{\mathrm{AS}}(f,P)$.

**Lemma 1.** *Let $\mathcal{E}$ be an elliptic curve defined over the finite field $\mathbb{F}_{q^e}$ of characteristic $p$. Set $k=k(\mathcal{E},\mathbb{F}_{q^e})$. Furthermore, let $f\in\mathcal{F}_{q^e}(\mathcal{E})$ be a function and $P$ be a generator of the group $[k]\mathcal{E}(\mathbb{F}_{q^e})$.*

*Suppose that for all $z\in\overline{\mathcal{F}}(\mathcal{E})$ the relation $z^p-z\neq f\circ[k]$ holds and that*

$$\mathcal{E}^{\mathrm{AS}}(f\circ[k],\mathbb{F}_{q^e})=[k]^{-1}\left(\mathcal{E}^{\mathrm{AS}}(f,\mathbb{F}_{q^e})\cap\langle P\rangle\right).$$

*Then $B_{S^{\mathrm{AS}}(f,P)}$ is bounded from above by*

$$\frac{1}{N}\left(N-\#(\mathcal{E}^{\mathrm{AS}}(f,\mathbb{F}_{q^e})\cap\langle P\rangle)+\frac{1}{k(\mathcal{E})^2}\sum_{Q}(m_Q(f\circ[k])+1)\sqrt{q^e}\right),$$

*with $N=\#\langle P\rangle$.*

*Proof.* Denote by $S$ the sequence $\{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f(Q))\}$ with $Q\in\mathcal{E}^{\mathrm{AS}}(f,\mathbb{F}_{q^e})\cap\langle P\rangle$. Further denote by $T$ the sequence $\{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f\circ[k](Q))\}$ with $Q\in\mathcal{E}^{\mathrm{AS}}(f\circ[k],\mathbb{F}_{q^e})$. We know that $\mathcal{E}^{\mathrm{AS}}(f\circ[k],\mathbb{F}_{q^e})=[k]^{-1}\left(\mathcal{E}^{\mathrm{AS}}(f,\mathbb{F}_{q^e})\cap\langle P\rangle\right)$. Hence, for each point $R$ in $\mathcal{E}^{\mathrm{AS}}(f,\mathbb{F}_{q^e})\cap\langle P\rangle$, there exist exactly $k^2$ points $Q$ in the set $\mathcal{E}^{\mathrm{AS}}(f\circ[k],\mathbb{F}_{q^e})$ such that $[k]Q=R$. Hence for any $\alpha\in\mathbb{F}_q^*$

$$\#\mathcal{E}^{\mathrm{AS}}(f\circ[k],\mathbb{F}_{q^e})B_T(\alpha)=k^2\#(\mathcal{E}^{\mathrm{AS}}(f,\mathbb{F}_{q^e})\cap\langle P\rangle)B_S(\alpha).$$

Hence

$$B_T(\alpha)=B_S(\alpha).$$

Since

$$N\cdot|B_{S^{\mathrm{AS}}(f,P)}(\alpha)|\leq N-\#(\mathcal{E}^{\mathrm{AS}}(f,\mathbb{F}_{q^e})\cap\langle P\rangle)+\#(\mathcal{E}^{\mathrm{AS}}(f,\mathbb{F}_{q^e})\cap\langle P\rangle)|B_S(\alpha)|,$$

we see that an upper bound for $B_T(\alpha)$ results in an upper bound for $B_{S^{\mathrm{AS}}(f,P)}$. However, since $\#\mathcal{E}^{\mathrm{AS}}(f\circ[k],\mathbb{F}_{q^e})B_T(\alpha)=ES^{\mathrm{AS}}(\mathcal{E},f\circ[k])$, such an upper bound is available from Proposition 2. This concludes the proof.

Note that the technical condition

$$\mathcal{E}^{\mathrm{AS}}(f\circ[k],\mathbb{F}_{q^e})=[k]^{-1}\left(\mathcal{E}^{\mathrm{AS}}(f,\mathbb{F}_{q^e})\cap\langle P\rangle\right)$$

is fulfilled if $k=1$. Also note that the righthandside set is always contained in the lefthandside set.

We consider a special case of the above lemma. Denote by $\mathrm{wdeg}(f(x,y))$ the weighted degree of a polynomial in two variables defined by $\mathrm{wdeg}(x)=2$ and $\mathrm{wdeg}(y)=3$.

**Theorem 1.** *Let $\mathcal{E}$ be an elliptic curve defined over the finite field $\mathbb{F}_{q^e}$ of characteristic $p$ given by a Weierstrass equation. Let $f$ be a polynomial in the coordinate functions $x$ and $y$ such that $\deg_y(f) \leq 1$. Further let $P$ be a generator of the group $[k(\mathcal{E})]\mathcal{E}(\mathbb{F}_{q^e})$ and define $N = \#\langle P\rangle$. Suppose that $p$ does not divide $\mathrm{wdeg}(f)$. Then we have*

$$B_{S^{\mathrm{AS}}(f,P)} \leq \frac{1}{N}\left(1 + (1 + \mathrm{wdeg}(f))\sqrt{q^e}\right).$$

Note that the above theorem also follows from Theorem 1 of [8] or the work of Bombieri [2]. Also note that the condition $\deg_y(f) \leq 1$ is not a real restriction, because we can use the Weierstrass equation to reduce this degree if $\deg_y(f) \geq 2$. Further note that if this condition is met, we have $v_{\mathcal{O}}(f) = -\mathrm{wdeg}(f)$ with $\mathcal{O}$ the point at infinity $[0:1:0]$. For the proof of the above theorem note that $\mathcal{E}^{\mathrm{AS}}(f, \mathbb{F}_{q^e}) = \mathcal{E}(\mathbb{F}_{q^e}) \setminus \{\mathcal{O}\}$ and that $\mathcal{E}^{\mathrm{AS}}(f \circ [k(\mathcal{E})], \mathbb{F}_{q^e}) = \mathcal{E}(\mathbb{F}_{q^e}) \setminus \mathcal{E}[k(\mathcal{E})]$.

In the same way we can investigate the autocorrelation of the sequences $S^{\mathrm{AS}}(f, P)$. We do this in the following theorem. First we state a lemma.

**Lemma 2.** *Let $\mathcal{E}$ be an elliptic curve defined over the field $\mathbb{F}_{q^e}$ of characteristic $p$. Let $f \in \mathcal{F}_{q^e}(\mathcal{E})$ and choose $\alpha, \beta \in \mathbb{F}_q^*$. Write $k = k(\mathcal{E}, \mathbb{F}_{q^e})$ and choose a generator $P$ of the group $[k]\mathcal{E}(\mathbb{F}_{q^e})$ and a number $d$ satisfying $1 \leq d < N$ with $N = \#\langle P\rangle$. Define $h \in \mathcal{F}_{q^e}(\mathcal{E})$ by*

$$h(X) = \alpha f(X \oplus [d]P) - \beta f(X).$$

*Denote by $S$ the sequence $\{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_p}(h \circ [k](Q))\}$ with $Q \in \mathcal{E}^{\mathrm{AS}}(h \circ [k], \mathbb{F}_{q^e})$. Finally suppose that $\mathcal{E}^{\mathrm{AS}}(h \circ [k], \mathbb{F}_{q^e}) = [k]^{-1}\left(\mathcal{E}^{\mathrm{AS}}(h, \mathbb{F}_{q^e}) \cap \langle P\rangle\right)$. We then have*

$$N \cdot C_{S^{\mathrm{AS}}(f,P)}(d, \alpha, \beta) = c + \#(\mathcal{E}^{\mathrm{AS}}(h, \mathbb{F}_{q^e}) \cap \langle P\rangle) B_S.$$

*Here*

$$c = \sum_{Q} \zeta_p^{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_p}(h(Q))},$$

*where the sum is over points $Q$ such that $Q \in \langle P\rangle$ and $Q \notin \mathcal{E}^{\mathrm{AS}}(h, \mathbb{F}_{q^e})$.*

*Proof.* Note that $C_{S^{\mathrm{AS}}(f,P)}(d, \alpha, \beta) = B_{S^{\mathrm{AS}}(h,P)}(1)$. Using similar tricks as in the proof of Lemma 1 we obtain the result.

Using an upper bound for exponential sums we can derive an upper bound for the autocorrelation, if some technical conditions are met. More explicitly we find the following theorem, in the case that $f$ is a polynomial in the coordinate functions.

**Theorem 2.** *Let $\mathcal{E}$ be an elliptic curve defined over the field $\mathbb{F}_{q^e}$ given by a Weierstrass equation. Let $f$ be a polynomial in the two coordinate functions $x$ and $y$, such that $\deg_y(f) \leq 1$. Choose $\alpha, \beta \in \mathbb{F}_q^*$. Further choose a generator*

*$P$ of the group $[k(\mathcal{E})]\mathcal{E}(\mathbb{F}_{q^e})$ and a number $d$ satisfying $1 \leq d < N$ with $N = \#\langle P\rangle$. Suppose that the characteristic does not divide* $\mathrm{wdeg}(f)$. *We then have*

$$|C_{S^{\mathrm{AS}}(f,P)}(d,\alpha,\beta)| \leq \frac{1}{N}\left(2 + 2(1+\mathrm{wdeg}(f))\sqrt{q^e}\right).$$

Analogous to the autocorrelation, we can derive properties about the crosscorrelations of sequences. We state some results in the following theorem.

**Theorem 3.** *Let $\mathcal{E}$ be an elliptic curve defined over the finite field $\mathbb{F}_{q^e}$ of characteristic $p$, given by a Weierstrass equation. Let $P$ be a generator of the group $[k(\mathcal{E})]\mathcal{E}(\mathbb{F}_{q^e})$ and write $N = \#\langle P\rangle$. Let $f_1$ and $f_2$ be two polynomials in the coordinate functions $x$ and $y$ such that $\deg_y(f_i) \leq 1$ for $i = 1, 2$, and such that for all $(\alpha,\beta) \in \mathbb{F}_{q^e}^2 \setminus \{(0,0)\}$ we have $p \nmid \mathrm{wdeg}(\alpha f_1 - \beta f_2)$. Write $S_1 = S^{\mathrm{AS}}(f_1, P)$ and $S_2 = S^{\mathrm{AS}}(f_2, P)$. For all $\alpha, \beta \in \mathbb{F}_q^*$ and $0 \leq d < N$ we have*

$$|C_{S_1,S_2}(d,\alpha,\beta)| \leq \frac{1}{N}\left(2 + (2 + \mathrm{wdeg}(f_1) + \mathrm{wdeg}(f_2))\sqrt{q^e}\right),$$

*unless $d = 0$ and $\alpha f_1 = \beta f_2$.*

*Proof.* This is a straightforward generalization of the proof of Theorem 2.

We now give an example of a family of sequences having good crosscorrelations. We assume in this example that the characteristic is 2, since this is the most interesting case for applications.

*Example 1.* Let $\mathcal{E}$ be an elliptic curve defined over the finite field $\mathbb{F}_{2^e}$. Denote by $P$ a generator of the group $[k(\mathcal{E})]\mathcal{E}(\mathbb{F}_{q^e})$ and write $N = \#\langle P\rangle$. Let $\mathbf{a}$ be defined by $\mathbf{a} = (a_0, \cdots, a_m) \in \mathbb{F}_{2^e}^{m+1}$ and let $S_{\mathbf{a}}$ be the binary sequence $S^{\mathrm{AS}}(f_{\mathbf{a}}, P)$ with defining function $f_{\mathbf{a}} = a_0 y + \cdots + a_m x^m y$.
For any number $0 \leq d < N$ and $\mathbf{a}, \mathbf{b} \in \mathbb{F}_{q^e}^{n+1} \setminus \{\mathbf{0}\}$ we have

$$\begin{aligned} |C_{S_{\mathbf{a}},S_{\mathbf{b}}}(d)| &\leq \tfrac{1}{N}\left(2 + (2 + \mathrm{wdeg}(f_{\mathbf{a}}) + \mathrm{wdeg}(f_{\mathbf{b}}))\sqrt{2^e}\right) \\ &\leq \tfrac{1}{N}\left(2 + (8 + 4m)\sqrt{2^e}\right), \end{aligned}$$

unless $d = 0$ and $\mathbf{a} = \mathbf{b}$.

## 5 Pseudorandom Sequences from Elliptic Curves Using Multiplicative Characters

We will now give results which are similar to those of the previous section, but depend on the use of multiplicative characters and Kummer extensions, instead of additive characters and Artin-Schreier extensions. Codes have been obtained using this approach in [9]. Sequences have been constructed in this way using the projective line in [1]. We will construct sequences using elliptic curves.

**Definition 7.** Let $\mathcal{C}$ be an algebraic curve defined over $\mathbb{F}_q$. Choose $1 < m < q-1$ a divisor of $q-1$ and let $f \in \mathcal{F}_q(\mathcal{C})$. Define $\mathcal{C}^K(f, \mathbb{F}_q)$ to be the set of $\mathbb{F}_q$-rational points on $\mathcal{C}$ such that there exists a $g \in \mathcal{F}_q(\mathcal{C})$ such that $v_P(f \cdot g^m) = 0$.

Note that if $\phi : \mathbb{F}_q^* \to \mathbb{Z}/m\mathbb{Z}$ is a homomorphism of groups, the quantity $\phi((f \cdot g^m)(Q))$ does not depend on $g$, as long as $v_Q(f \cdot g^m) = 0$. Hence for $Q \in \mathcal{C}^K(f, \mathbb{F}_q)$ we will write $\phi(f(Q))$, even if $v_Q(f) \neq 0$. In particular we see that the quantity $f(Q)^{(q-1)/m}$ is well-defined for $Q \in \mathcal{C}^K(f, \mathbb{F}_q)$. If $Q \notin \mathcal{C}^K(f, \mathbb{F}_q)$, we can always find $g \in \mathcal{F}_q(\mathcal{C})$ such that $(f \cdot g^m)(Q) = 0$. Hence we define $f(Q)^{(q-1)/m} = 0$ for $Q \notin \mathcal{C}^K(f, \mathbb{F}_q)$, even if $f$ has a pole in $Q$. In the same way we define $\phi(f(Q)) = 0$ for $Q \notin \mathcal{C}^K(f, \mathbb{F}_q)$.

**Definition 8.** Let $\mathcal{E}$ be an elliptic curve defined over $\mathbb{F}_q$. Fix a natural number $1 < m < q-1$ dividing $q-1$. Denote by $\chi_m : \mathbb{F}_q^* \to \mathbb{Z}/m\mathbb{Z}$ some fixed, surjective homomorphism of groups. Let $P \in \mathcal{E}(\mathbb{F}_q)$ and $f \in \mathcal{F}_q(\mathcal{E})$. Define $S^K(f, P) = \{s(i)\}$ by

$$s(i) = \chi_m(f([i]P)).$$

We need the homomorphism $\chi_m$ to define the balance and correlations of the sequence $S^K(f, P)$ (see Section 3). Note that $\chi_m(f(Q)) = 0$ if $Q \notin \mathcal{E}^K(f, \mathbb{F}_q)$.

*Example 2.* Let $\mathbb{F}_p$ be a prime field with $p$ odd. Let $\alpha$ be a generator of the multiplicative group $\mathbb{F}_p^*$. Let $f[X]$ be a polynomial in $\mathbb{F}_p[X]$ of degree $m$. By evaluating this polynomial in all elements $\alpha, \alpha^2, \ldots$ of $\mathbb{F}_p^*$ we obtain a codeword from a Reed-Solomon code (RS-code). We obtain a binary sequence from this codeword by applying coordinatewise the map $\chi_2 : \mathbb{F}_p \to \mathbb{Z}/2\mathbb{Z}$, defined by

$$\chi_2(a) = \begin{cases} 0 & \text{if } a = 0 \text{ or } \left(\frac{a}{p}\right) = 1, \\ 1 & \text{if } \left(\frac{a}{p}\right) = -1 \end{cases}$$

If we take for example $p = 13$, $f[X] = X^2 + X$, and $\alpha = 2$ we find the codeword

$$(2, 6, 7, 7, 12, 3, 0, 2, 12, 4, 6, 4)$$

and corresponding binary sequence

$$(1, 1, 1, 1, 0, 0, 0, 1, 0, 0, 1, 0).$$

As we said before it is possible to obtain codes using this construction. This was done in [9]. In that article non-linear codes were found and investigated. It is possible to find interesting linear codes as is shown in the following example. After the example we return to the study of sequences.

*Example 3.* Choose $\mathcal{C}$ to be the projective line defined over some finite field $\mathbb{F}_q$ of odd characteristic. For $\mathcal{M} \subset \mathbb{F}_q(x)/(\mathbb{F}_q(x))^2$ we choose the group generated by the residue classes of $x - \beta$ with $\beta$ in some non-empty subset $S$ of $\mathbb{F}_q$. Then every element of $\mathcal{M}$ has a representative of the form $\prod_{\beta \in S}(x-\beta)^{\epsilon_\beta}$ with $\epsilon_\beta \in \{0,1\}$. As a vector space over $\mathbb{F}_2$, the group $\mathcal{M}$ has dimension $\#S$. Define $\chi_2$ as in example 2. Evaluating $\chi_2 \circ f$ in the set $\mathbb{F}_q \setminus S$ for all functions $f \in \mathcal{M}$, yields a binary linear code $C$ of length $\#(\mathbb{F}_q \setminus S)$. Its dimension is less than or equal to $\min(\#S, \#(\mathbb{F}_q \setminus S))$.

Equality need not hold in this equation. Suppose for example that $q$ is a square. The evaluation of the polynomial $f(X) = X^{\sqrt{q}} + X$ in an element $a$ of $\mathbb{F}_q$ is either zero or a square. To see this note that for $a \in \mathbb{F}_q$ we have $f(a)^{\sqrt{q}} = f(a)$, and hence either $f(a) = 0$ or $f(a)^{(q-1)/2} = 1$. Hence if we choose $S = \{a \in \mathbb{F}_q \,|\, a^{\sqrt{q}} + a = 0\}$, then the polynomial $X^{\sqrt{q}} + x$ will correspond to the all-zero codeword. This means that in this case the dimension of the code cannot equal the cardinality of $S$.

Note that the curve given by the equation $Y^2 = X^{\sqrt{q}} + X$ has maximum number of $\mathbb{F}_q$-rational points for its genus. As a matter of fact the number of $\mathbb{F}_q$-rational points can be seen to be $2q - \sqrt{q} + 1$, while its genus equals $(\sqrt{q}-1)/2$. The fact that it is maximal also follows from the fact that it can be covered by the Hermitian curve which has equation $Y^{\sqrt{q}+1} = X^{\sqrt{q}} + X$. Using the Hasse-Weil bound and investigating the curve $Y^2 = f(X)$, one can show that for sets $S$ with cardinality strictly smaller than $\sqrt{q}$, only the zero-polynomial can give the all-zero codeword. This means that in this case the dimension of the resulting codes equals the cardinality of the set $S$.

Refining this argument, we see that for the minimum distance $d$ of these codes we have the statement

$$d \geq \frac{q - (\#S - 1)(\sqrt{q}-1)}{2} - \#S,$$

which is a non-trivial lower bound if $\#S < \sqrt{q}$.

In a similar way as in the previous section we give statements about the balance, autocorrelation and crosscorrelation of the sequences $S^K(f, P)$.

**Definition 9.** Let $\mathcal{C}$ be an algebraic curve defined over $\mathbb{F}_q$. Let $1 < m < q-1$ be a divisor of $q-1$ and denote by $\chi_m : \mathbb{F}_q^* \to \mathbb{Z}/m\mathbb{Z}$ some surjective homomorphism of groups. Let $f \in \mathcal{F}_q(\mathcal{C})$. We define the following exponential sum:

$$ES^K(\mathcal{C}, f) = \sum_{P \in \mathcal{C}^K(f, \mathbb{F}_q)} \zeta_m^{\chi_m(f(P))},$$

with $\zeta_m = exp(2\pi i/m)$.

For this exponential sum a bound exists similar to that of the exponential sum defined in Definition 6. See for example [9], where similar upper bounds are derived. The proof of the following proposition is analagous to

that of Proposition 2. Instead of Artin-Schreier extensions we use Kummer extensions (see for example Chap. III, Sect. 7 of [12]).

**Proposition 3.** *Let $\mathcal{C}$ be an algebraic curve of genus $g$ defined over $\mathbb{F}_q$. Let $f \in \mathcal{F}_q(\mathcal{C})$ and suppose that $f \neq z^l$ for all $z \in \overline{\mathcal{F}(\mathcal{C})}$ and all divisors $l > 1$ of $m$. Write $r_P = \gcd(m, v_P(f)) > 0$. Then the following holds:*

$$\left|ES^K(\mathcal{C}, f)\right| \leq \left(2g - 2 + \sum_{P \in \mathcal{E}(\overline{\mathbb{F}}_q)} (1 - \frac{r_P - 1}{m - 1})\right) \sqrt{q}.$$

The $r_P$ occurring in the above proposition are standard in the theory of Kummer extensions. When we want to stress the role of $f$, we will write $r_P(f)$.

**Theorem 4.** *Let $\mathcal{E}$ be an elliptic curve defined over the finite field $\mathbb{F}_q$ of characteristic $p$. Let $f \in \mathcal{F}_q(\mathcal{E})$ be a function and write $k = k(\mathcal{E}, \mathbb{F}_q)$. Let $P$ be a generator of the group $[k]\mathcal{E}(\mathbb{F}_q)$. Suppose that the polynomial $T^m - f \circ [k]$ is absolutely irreducible. Then we have*

$$B_{S^K(f,P)} \leq \frac{1}{N} \left( N - \#(\mathcal{E}^K(f, \mathbb{F}_q) \cap \langle P \rangle) + \sum_Q (1 - \frac{r_Q(f) - 1}{m - 1}) \sqrt{q} \right),$$

*with $N = \#\langle P \rangle$.*

*Proof.* Note that $v_Q(f \circ [k]) = v_{[k]Q}(f)$ and hence $r_Q(f \circ [k]) = r_{[k]Q}(f)$. Moreover, note that for $Q \in \mathcal{E}(\mathbb{F}_q)$ we have $r_Q = m$ if and only if $Q \in \mathcal{E}^K(f, \mathbb{F}_q)$. Hence we see that $\mathcal{E}^K(f \circ [k], \mathbb{F}_q) = [k]^{-1}\mathcal{E}^K(f, \mathbb{F}_q) \cap \langle P \rangle$. The rest of the proof is similar to that of Lemma 1.

In the following corollary we again use the weighted degree of a polynomial in two variables defined by $\mathrm{wdeg}(x) = 2$ and $\mathrm{wdeg}(y) = 3$.

**Corollary 1.** *Let the notation be as in the above theorem and suppose that $\mathcal{E}$ is given by a Weierstrass equation. Suppose that $f$ is a non-trivial polynomial of total degree $\Delta$ in the coordinate functions $x$ and $y$ satisfying $\deg_x(f) \leq 2$. Suppose that $\gcd(\mathrm{wdeg}(f), m) = 1$. Then we have*

$$B_{S^K(f,P)} \leq \frac{1}{N} \left( N - \#(\mathcal{E}^K(f, \mathbb{F}_q) \cap \langle P \rangle) + (3\Delta + 1)\sqrt{q} \right).$$

*If we additionally demand $(\langle P \rangle \setminus \{\mathcal{O}\}) \subset \mathcal{E}^K(f, \mathbb{F}_q)$, we find*

$$B_{S^K(f,P)} \leq \frac{1}{N} \left( 1 + (3\Delta + 1)\sqrt{q} \right).$$

*Proof.* This follows from the above theorem by remarking that by Bézout's theorem (see for example Sect. 83 of [13]) $f$ has at most $3\Delta$ zeros on $\mathcal{E}$. Further note that the point $\mathcal{O}$ is the only pole $f$ has on $\mathcal{E}$. These zeros and poles are the only points $Q$ for which it can happen that $r_Q < m$. Using that $r_Q \geq 1$ for these points $Q$, the result follows. Note that $r_{\mathcal{O}}(f \circ [k(\mathcal{E})]) = r_{\mathcal{O}}(f) = \gcd(\mathrm{wdeg}(f), m) = 1$. Hence the polynomial $T^m - f \circ [k(\mathcal{E})]$ is absolutely irreducible by the theory of Kummer extensions.

Note that it can be useful to rewrite $f$, using the equation of $\mathcal{E}$, in such a form that the total degree is minimal. This explains why we now assume $\deg_x(f) \leq 2$ instead of assuming $\deg_y(f) \leq 1$ as we did before.

We will now give some statements about the autocorrelation and cross-correlations of these sequences. We omit most of the proofs, since they are analogous to the proofs in the Artin-Schreier case.

**Theorem 5.** *Let $\mathcal{E}$ be an elliptic curve defined over the field $\mathbb{F}_q$ of characteristic $p$. Let $f \in \mathcal{F}_q(\mathcal{E})$ and choose $\alpha, \beta \in \mathbb{F}_q^*$. Write $k = k(\mathcal{E}, \mathbb{F}_q)$ and choose a generator $P$ of the group $[k]\mathcal{E}(\mathbb{F}_q)$ and a number $d$ satisfying $1 \leq d < N$ with $N = \#\langle P \rangle$. Define $h \in \mathcal{F}_q(\mathcal{E})$ by*

$$h(X) = \alpha f(X \oplus [d]P) - \beta f(X).$$

*Suppose that the polynomial $T^m - h \circ [k]$ is absolutely irreducible. We then have*

$$|C_{S^K(f,P)}(d, \alpha, \beta)| \leq \frac{1}{N}\left(N - \#(\mathcal{E}^K(h, \mathbb{F}_q) \cap \langle P \rangle) + \sum_Q (1 - \frac{r_Q(h) - 1}{m-1})\sqrt{q}\right).$$

**Corollary 2.** *Let the notation be the same as in the above theorem. Suppose that $\mathcal{E}$ is given by a Weierstrass equation. Further assume that $f$ is a non-trivial polynomial in the coordinate functions of total degree $\Delta$ satisfying $\deg_x(f) \leq 2$ and $\gcd(\mathrm{wdeg}(f), m) = 1$. Then we have*

$$C_{S^K(f,P)}(d, \alpha, \beta) \leq \frac{1}{N}\left(N - \#(\mathcal{E}^K(h, \mathbb{F}_q) \cap \langle P \rangle) + (3\Delta + 3\mathrm{wdeg}(f) + 2)\sqrt{q}\right),$$

*If we additionally demand $(\langle P \rangle \setminus \{\mathcal{O}, [-d]P\}) \subset \mathcal{E}^K(h, \mathbb{F}_q)$, we find*

$$C_{S^K(f,P)}(d, \alpha, \beta) \leq \frac{1}{N}\left(2 + (3\Delta + 3\mathrm{wdeg}(f) + 2)\sqrt{q}\right).$$

*Proof.* Again we want to use Bézout's theorem to estimate the number of zeros of the function $h$. Using the addition formula (see for example Chap. III, Sect. 2 of [11]) we find that $(x, y) \oplus (a, b)$ can be written as $(g_1(x, y)/(x - a)^2, g_2(x, y)/(x - a)^3)$ with $g_1$ (respectively $g_2$) a polynomial in $x$ and $y$ of total degree less than or equal to 2 (respectively 3). This means that $\alpha f((x, y) \oplus (a, b))$ can be written as $k(x, y)/(x - a)^{\mathrm{wdeg}(f)}$ with $k$ a polynomial

of total degree less than or equal to $\mathrm{wdeg}(f)$. Hence, after multiplying the rational function $h$ with $(x-a)^{\mathrm{wdeg}(f)}$. we get a polynomial of total degree less than or equal to $\Delta + \mathrm{wdeg}(f)$. This gives an upper bound for the total number of zeros of the function $h$ while its poles are $\mathcal{O}$ and $-[d]P$. The rest of the proof is analogous to the proof of Corollary 1.

**Theorem 6.** *Let $\mathcal{E}$ be an elliptic curve defined over the finite field $\mathbb{F}_q$ of characteristic $p$. Let $P$ be a generator of the group $[k(\mathcal{E})]\mathcal{E}(\mathbb{F}_q)$ and write $N = \#\langle P\rangle$. Let $f_1$ and $f_2$ be two functions and choose $\alpha, \beta \in \mathbb{F}_q^*$ as well as a natural number $0 \le d < N$. Write $S_1 = S^K(f_1, P)$ and $S_2 = S^K(f_2, P)$. Define $h \in \mathcal{F}_q(\mathcal{E})$ by*

$$h(X) = \alpha f_1(X \oplus [d]P) - \beta f_2(X)$$

*and suppose that the polynomial $T^m - h \circ [k(\mathcal{E})]$ is absolutely irreducible. We have*

$$|C_{S_1,S_2}(d,\alpha,\beta)| \le \frac{1}{N}\left(N - \#(\mathcal{E}^K(h,\mathbb{F}_q) \cap \langle P\rangle) + \sum_Q (1 - \frac{r_Q(h)-1}{m-1})\sqrt{q}\right).$$

**Corollary 3.** *Let the notation be the same as in the above theorem. Suppose that $\mathcal{E}$ is given by a Weierstrass equation. Further assume that $f_1$ and $f_2$ are non-trivial polynomials in the coordinate functions of total degree $\Delta_1$ and $\Delta_2$ with $\Delta_1 \ge \Delta_2$ and satisfying $\deg_x(f_i) \le 2$ with $i = 1, 2$. Further suppose that $\gcd(\mathrm{wdeg}(f_1), m) = 1$ if $d \ne 0$ and $\gcd(\mathrm{wdeg}(\alpha f_1 - \beta f_2), m) = 1$ if $d = 0$. We then have*

$$C_{S_1,S_2}(d,\alpha,\beta) \le \frac{1}{N}\left(N - \#(\mathcal{E}^K(h,\mathbb{F}_q) \cap \langle P\rangle) + (3\mathrm{wdeg}(f_1) + 3\Delta_2 + 2)\sqrt{q}\right).$$

*If we additionally demand $(\langle P\rangle \setminus \{\mathcal{O}, [-d]P\}) \subset \mathcal{E}^K(h,\mathbb{F}_q)$, we find*

$$C_{S_1,S_2}(d,\alpha,\beta) \le \frac{1}{N}\left(2 + (3\Delta_2 + 3\mathrm{wdeg}(f_1) + 2)\sqrt{q}\right).$$

## 6 Pseudorandom Sequences Using Linear Recurrence Relations on Elliptic Curves

In this section we will investigate the balance and the period of a family of sequences obtained by using linear recurrence relations on the points of $\mathcal{E}$.

Suppose that $G$ is a cyclic subgroup of $\mathcal{E}$ of order $N$ generated by a point $P \in \mathcal{E}$. In this section we will assume that $N$ is a prime number.

Let $r(X) = X^n + r_{n-1}X^{n-1} \cdots + r_0$ be a monic polynomial of degree $n > 1$ over $\mathbb{Z}/N\mathbb{Z}$ with $\gcd(r_0, N) = 1$ and let $\Omega(r, G)$ be the vector space over $\mathbb{Z}/N\mathbb{Z}$ of bi-infinite sequences of points in $G$ that satisfy the linear recurrence

relation with characteristic polynomial $r(X)$. This vector space has dimension $n$.

We suppose from now on that the characteristic polynomial $r(X)$ of the recursion is irreducible over $\mathbb{Z}/N\mathbb{Z}$. It is known from the theory of linear recurrencies (for example, Chap. 7 in [10]) that if $r(X)$ is irreducible, every sequence, apart from the zero sequence, has the same period $k(N, r)$. As a matter of fact $k(N, r)$ is the smallest positive integer $k$ such that for every root $\alpha$ of $r(X)$ we have $\alpha^k = 1$.

Define $\Psi(r, G)$ to be the set of sequences $\Omega(r, G)$ modulo cyclic shifts.

**Lemma 3.** *Every point $Q \in G$ occurs the same number of times in sequences in $\Psi(r, G)$, i.e. the number of pairs*

$$\#\left\{(i, \psi) \mid 0 \leq i < k(N, r); \psi(i) = Q; \psi \in \Psi(r, G)\right\}$$

*is independent of the choice of $Q$.*

*Proof.* Since we demanded that $\gcd(r_0, N) = 1$ in the choice of the recursion polynomial $r$ and this polynomial has degree $n$, each sequence in $\Psi(r, G)$ is uniquely determined by the choice of $n$ consecutive points. Conversely, each $n$-tuple of points occurs exactly once in $\Psi(r, G)$ (note that this is modulo cyclic shifts). Since each point $Q$ occurs equally often in the set of all $n$-tuples of points, we have that this is the case in $\Psi(r, G)$.

Let $f \in \mathcal{F}_{q^e}$ be a function on $\mathcal{E}$. Now look at the sequence $S^{\mathrm{AS}}(f, P)$ which was defined earlier by

$$S^{\mathrm{AS}}(f, P) = \left\{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}\left(f\left([i]P\right)\right)\right\}_{0 \leq i < N}.$$

Furthermore, define the set of sequences $\Psi_f(r, G)$ by applying the function $f$ to each point in each sequence in $\Psi(r, G)$, and then taking the trace to the ground field $\mathbb{F}_q$ of the result:

$$\Psi_f(r. G) = \left\{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}\left(f\left(\psi\right)\right) | \psi \in \Psi(r, G)\right\}.$$

Hence each sequence of points in $\Psi(r, G)$ corresponds with a sequence in $\Psi_f(r, G)$.

Here we use the same convention as before, namely that $\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f(Q)) = 0$ if $Q \notin \mathcal{E}^{\mathrm{AS}}(f, \mathbb{F}_{q^e})$.

**Theorem 7.** *Choose a point $P \in \mathcal{E}$. Let $G$ be the subgroup of $\mathcal{E}$ generated by $P$ and suppose that its order is a prime $N$. Furthermore, let $r$ be a monic recursion polynomial of degree $n$ whose tail coefficient is coprime to $N$. Then the average balance of a sequence in $\Psi_f(r, G)$ is the same as the balance of the sequence $S^{\mathrm{AS}}(f, P)$.*

*Proof.* Start by defining the sequence $T$ by concatenating all sequences of $\Psi_f(r, G)$. Since the order of points is not important in the definition of balance and because according to Lemma 3 each point occurs an equal number of times, we can reorder the points of $T$ such that we get a number of copies of the sequence $S^{\mathrm{AS}}(f, P)$. Of course, this is the same sequence as $S^{\mathrm{AS}}(f, P)$ itself. Thus the average balance of sequences in $\Psi_f(r, G)$ is equal to the balance of the sequence $S^{\mathrm{AS}}(f, P)$.

It is well known from the theory of linear recurrencies that the period of a sequence can be larger than the group order. So sequences defined in the above way can have a larger period than the sequences described in the previous sections. But this only applies to the sequences of points on $\mathcal{E}$. The next theorem links this period to the period of the generated pseudorandom sequence. We still suppose that $r(X)$ is irreducible.

**Theorem 8.** *Let $r$ and $G$ be defined as in the above theorem. Suppose that the order of $G$ is a prime $N$ and that the degree of the recursion polynomial is $n$. Denote by $T_f(a)$ the number of points $Q$ in $G$ for which $\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f(Q)) = a$. Suppose that all sequences in $\Psi_f(r, G)$ have period dividing $k(N, r)/d$. Then $d$ is a divisor of $\gcd(N - k(N, r), T_f(a), N^n - 1)$ for all $a \in \mathbb{F}_q \backslash \{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f(\mathcal{O}))\}$.*

*Proof.* We know that all non-zero sequences in $\Psi_f(r, G)$ have as period a divisor of $k(N, r)/d$. Hence the number of times $a$ occurs in the corresponding sequences is divisible by $d$. Write $b = \mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f(\mathcal{O}))$. Then $d$ divides $N^n T_f(a)$ with $a \in \mathbb{F}_q \setminus \{b\}$, and $d$ divides $N^n T_f(b) - k(N, r)$. Since $k(N, r)$ divides $N^n - 1$, we see that $d$ divides $\gcd(N^n T_f(b) - k(N, r), T_f(a), N^n - 1)$ for all $a \in \mathbb{F}_q \backslash \{b\}$. Using $\sum_{a \in \mathbb{F}_q} T_f(a) = N$, we find for all $a \in \mathbb{F}_q \backslash \{\mathrm{Tr}_{\mathbb{F}_{q^e}|\mathbb{F}_q}(f(\mathcal{O}))\}$, after eliminating $T_f(b)$, that $d$ divides

$$\gcd(N^{n+1} - k(N, r), T_f(a), N^n - 1).$$

The result follows directly from this.

*Example 4.* Let $\mathcal{E}$ be the elliptic curve defined over $\mathbb{F}_2$ given by the equation

$$Y^2 + Y = X^3 + X + 1.$$

Let $G$ be a prime-order subgroup of $\mathcal{E}(\mathbb{F}_q)$ with $q$ an odd power of 2. Using the addition formulas on $\mathcal{E}$, one can derive that for this curve $T_x(0) - 1 = T_x(1) = (N - 1)/2$. Hence, according to the above theorem, we find that $d$ divides $\gcd((N - 1)/2, k(N, r) - 1)$. Take for example $r(X) = X^2 - X - 1$, the Fibonacci-recursion. The polynomial $r(X)$ is irreducible if and only if $N \equiv 2, 3 \pmod 5$. Assuming this, we find that $k(N, r)$ divides $2N + 2$, since for any root $\rho$ of $r(X)$ we have $\rho^{2N+2} = (\rho \cdot \rho^N)^2 = (-1)^2 = 1$. Here we used that $r(X)$ is absolutely irreducible and hence that its roots are given by $\rho$ and $\rho^N$. Let us further assume that $k(N, r) = 2N + 2$. In this case we find that $d$ divides 3.

## References


1. A. Barg. *A large family of sequences with low periodic correlation.* Discrete Math. 176(1-3), pages 21-27, 1997.
2. E. Bombieri. *On exponential sums in finite fields.* Amer. J. Math. 88, pages 71-105, 1966.
3. E. El Mahassni and I.E. Shparlinski. *On the uniformity of distribution of congruential generators over elliptic curves.* Sequences and their Applications - SETA '01, pages 257-264. Springer, London, 2002.
4. G. Gong, T.A. Berson and D.R. Stinson. *Elliptic curve pseudorandom sequence generators.* Selected areas in cryptography (Kingston, ON, 1999), pages 34-48. Springer, Berlin, 2000.
5. G. Gong, C.C.Y. Lam. *Recursive sequences over elliptic curves.* Sequences and their Applications - SETA '01, pages 182-196. Springer, London, 2002.
6. S. Hallgren. *Linear congruential generators over elliptic curves.* Preprint CS-94-143, Dept. of Comp. Sci., Cornegie Mellon Univ., 1994.
7. A.J. Menezes, P.C. van Oorschot and S.A. Vanstone. *Handbook of applied cryptography.* CRC Press, Boca Raton, FL, 1997.
8. D.R. Kohel and I.E. Shparlinski. *Exponential sums and group generators for elliptic curves over finite fields.* Lecture Notes in Comp. Sci. 1838, pages 395-404. Springer, Berlin, 2000.
9. M. Perret. *Multiplicative character sums and nonlinear geometric codes.* Eurocode '90 (Udine, 1990), pages 158-165. Springer, Berlin, 1991.
10. I.E. Shparlinski. *Finite Fields: Theory and Computation.* Kluwer Academic Publishers, Dordrecht, 1999.
11. J.H. Silverman. *The arithmetic of elliptic curves.* Springer, Berlin, 1986.
12. H. Stichtenoth. *Algebraic function fields and codes.* Springer, Berlin, 1993.
13. B.L. van der Waerden. *Moderne Algebra.* Springer, Berlin, 1940.
14. C. Voss. *Abschätzungen der Parameter von Spurcodes mit Hilfe algebraischer Funktionenkörper.* Ph.D. Thesis, Universität Essen, 1993.
15. J.F. Voloch and J.L. Walker. *Euclidean weights of codes from elliptic curves over rings.* Trans. Amer. Math. Soc., 352(11), pages 5063-5076, 2000.

# On Cryptographic Complexity of Boolean Functions

Claude Carlet*

GREYC, University of Paris 8 and INRIA

**Abstract.** Cryptographic Boolean functions must be complex to satisfy Shannon's principle of confusion. Two main criteria evaluating, from crytpographic viewpoint, the complexity of Boolean functions on $F_2^n$ have been studied in the literature: the nonlinearity (the minimum Hamming distance to affine functions) and the algebraic degree. We consider two other criteria: the minimum number of terms in the algebraic normal forms of all affinely equivalent functions (we call it the algebraic thickness) and the non-normality. We show that, asymptotically, almost all Boolean functions have high algebraic degrees, high nonlinearities, high algebraic thicknesses and are highly non-normal.

## 1 Introduction

Let $n$ be any positive integer. We denote by $\mathcal{B}_n$ the set of all Boolean (i.e. $F_2$-valued) functions on $F_2^n$. We denote by $\oplus$ the additions in $F_2$, in $F_2^n$ and in $\mathcal{B}_n$. Every Boolean function $f$ admits a unique representation (cf. [11]) called its *algebraic normal form* (A.N.F.) as a polynomial over $F_2$ in $n$ binary variables of the form:

$$f(x_1, \cdots, x_n) = \bigoplus_{u \in {F_2}^n} a_u \left( \prod_{i=1}^{n} {x_i}^{u_i} \right) = \bigoplus_{u \in {F_2}^n} a_u \, x^u; \; a_u \in F_2.$$

The degree of the A.N.F. is called the *algebraic degree* of the function. It is an *affine invariant*: the degree of any function $f$ equals that of any *affinely equivalent* function $f \circ A$ ($A$ element of the general affine group). The Boolean functions whose algebraic degrees do not exceed 1 are called *affine*.

The Hamming weight of a Boolean function $f$ is the size of its support $\{x \in F_2^n;\ f(x) = 1\}$ and the Hamming distance between two functions $f$ and $g$ is the Hamming weight of the Boolean function $f \oplus g$. The *nonlinearity* $\mathcal{NL}(f)$ of a Boolean function $f$ is its minimum Hamming distance to affine functions. It is an affine invariant and can be expressed by means of the discrete Fourier-Walsh-Hadamard transform of the function. The *discrete Fourier-Walsh-Hadamard transform* of $f$ is by definition the integer-valued

* INRIA Projet CODES, Domaine de Voluceau, BP 105, 78153 Le Chesnay Cedex, France

function $\widehat{f}$ defined by

$$\forall u \in F_2^n, \quad \widehat{f}(u) = \sum_{x \in F_2^n} f(x)(-1)^{u \cdot x}$$

where $u \cdot x$ denotes the usual inner product $u \cdot x = u_1 x_1 \oplus \cdots \oplus u_n x_n$. Fourier-Walsh-Hadamard transform corresponds to the expression of $f$ in the orthogonal basis of the so called Walsh functions $x \mapsto (-1)^{u \cdot x}$. We will also have to consider the Fourier-Walsh-Hadamard transform of the "sign" function $\chi_f = (-1)^f$:

$$\widehat{\chi_f}(u) = \sum_{x \in F_2^n} (-1)^{f(x) \oplus u \cdot x} = 2^n \delta_0(u) - 2\widehat{f}(u))$$

where $\delta_0$ is the Dirac symbol ($\delta_0(u)$ equals 1 if $u = 0$ and 0 otherwise). We have:

$$\mathcal{NL}(f) = 2^{n-1} - \frac{1}{2} \max_{u \in F_2^n} |\widehat{\chi_f}(u)| \ . \tag{1}$$

Because of Parseval's relation:

$$\sum_{u \in F_2^n} \widehat{\chi_f}^{\,2}(u) = 2^{2n},$$

any Boolean function $f$ in $n$ variables satisfies $\mathcal{NL}(f) \leq 2^{n-1} - 2^{n/2-1}$. This upper bound can only be achieved for even values of $n$. The functions for which equality holds are called *bent functions*.

In this paper, we are interested in those cryptographic criteria on Boolean functions (used in conventional cryptosystems) which are related to Shannon's principle of *confusion*. This principle [22] has been introduced in 1949 (with another principle, called diffusion, that we do not study in this paper), and since then, its relevance to modern cryptography has always been verified. Concerning the Boolean functions involved in the cryptosystems (stream ciphers, block ciphers), this principle is related to the *complexity* of the functions. The complexity criteria and the corresponding complexity measures which are relevant to cryptography being related to the attacks on the cryptosystems where Boolean functions play a role, they are different from those used in circuit complexity (see e.g. [24]).

Nonlinearity is the most important of these criteria. It is related to attacks on stream ciphers (cf. [1]) and block ciphers as well (cf. the linear attack by Matsui [13]). Two other criteria play also important roles: the algebraic degree and the number of monomials in the A.N.F. (i.e. the number of nonzero $a_u$'s). The complexity of the "higher order differential attack" on block ciphers due to Knudsen and Lai [8,9] depends on the highest algebraic degree of the Boolean functions involved in the system. The linear complexity of a sequence generated by several Linear Feedback Shift Registers (LFSRs) combined by a nonlinear function depends on the degree of the function and on the number

of terms in its A.N.F. (these parameters condition therefore the resistance to Berlekamp-Massey algorithm, cf. [12,20]). The nonlinear functions selected for filtered LFSR's must also have high degrees and include many terms in their A.N.F.s (cf. [15] page 208).

As already pointed out by W. Meier and O. Staffelbach in [14], the general complexity criteria which are mostly interesting in cryptographic framework are affine invariants because the attacks on cryptosystems using Boolean functions (e.g. filtered Linear Feedback Shift Registers, block ciphers) often work with the same complexity when the functions are replaced by affinely equivalent ones. This is why we shall consider an affine invariant related to the number of terms, and not the number of terms itself.

Recall that, asymptotically, almost all Boolean functions have high circuit complexities. Lupanov [10] calls this the *Shannon effect*: Shannon [23] observed in 1949 that Boolean functions with high circuit complexity must exist because of the double-exponential increment of the number of all Boolean functions.

In this paper, we study to what extent this Shannon effect applies to cryptographic complexity criteria (the nonlinearity, the algebraic degree and the affine invariant, denoted by $\mathcal{T}(f)$ and called algebraic thickness, which is related to the number of terms). We also study another criterion, called normality, whose definition will be given later. We show that, asymptotically, almost all Boolean functions on $F_2^n$ have high degrees (greater than $\beta n$ where $\beta$ is any positive number smaller than 1), high nonlinearities (greater than $2^{n-1} - n^{\alpha} 2^{\frac{n}{2}-1}$ where $\alpha$ is any number greater than $\frac{1}{2}$), high algebraic thicknesses and are highly non-normal. We can also require that these functions admit no linear structure (i.e. that there does not exist $a \neq 0$ in $F_2^n$ and $\epsilon$ in $F_2$ such that, for every $x \in F_2^n$, $f(x \oplus a) = f(x) \oplus \epsilon$, cf. [6]). Our method is very similar to the methods used in circuit complexity: counting the functions which do not match the above constraints.

We finally generalize our study to $q$-ary functions (i.e. functions from $F_q^n$ to $F_q$ where $q$ is a power of a prime).

## 2 The Number of Terms in the A.N.F.s of Boolean Functions

The number of terms (i.e. of monomials with nonzero coefficients) in the A.N.F. of a Boolean function can obviously be any integer between 0 and $2^n$. But as explained by Meier and Staffelbach [14], a Boolean function having many terms in its A.N.F. can be nevertheless inadequate for cryptographic use if it is affinely equivalent to a function with few terms in its A.N.F.. They take the example of the function whose A.N.F. contains all monomials: this function is equal to $\prod_{i=1}^{n}(x_i \oplus 1)$.

This is why we are interested in an affine invariant related to this parameter of the function:

**Definition 1.** We call *algebraic thickness* of a Boolean function $f$ and we denote by $\mathcal{T}(f)$ the minimum number of terms (i.e. of monomials with nonzero coefficients) in the A.N.F.s of the functions $f \circ A$, where $A$ ranges over the set of all affine automorphisms of $F_2^n$.

**Examples**:

– For every nonzero affine function $f(x_1, \cdots, x_n) = a_1\, x_1 \oplus \cdots \oplus a_n\, x_n \oplus a_0$, where $a = (a_1, \cdots, a_n) \in {F_2}^n$ and $a_0 \in F_2$, we have $\mathcal{T}(f) = 1$, since the constant function 1 has one term in its A.N.F. and since every non-constant affine function is equivalent to $x_1$ (for instance).

– More generally, let $E$ be any flat (affine subspace) of $F_2^n$ and let $k$ be its dimension. $E$ is the intersection of $(n-k)$ independent affine hyperplanes. Thus, the indicator of $E$ (defined by $1_E(x) = 1$ if $x \in E$; 0 otherwise) is equivalent to $\prod_{i=1}^{n-k} x_i$ and thus $\mathcal{T}(1_E) = 1$.

– We know (cf. [11]) that every non-affine quadratic function (i.e. any function of degree 2) is equivalent to $x_1x_2 \oplus \cdots \oplus x_{2k-1}x_{2k} \oplus x_{2k+1}$ (where $2k+1 \leq n$) if the function is balanced (i.e. if its output is uniformly distributed) and to $x_1x_2 \oplus \cdots \oplus x_{2k-1}x_{2k}$ or to $x_1x_2 \oplus \cdots \oplus x_{2k-1}x_{2k} \oplus 1$ (where $2k \leq n$) otherwise. Thus, $\mathcal{T}(f) \leq \lfloor n/2 \rfloor + 1$, where $\lfloor\ \rfloor$ denotes the integer part. Moreover, two functions of two different forms, or of the same form but with different values of $k$ are affinely inequivalent. Thus the maximum of $\mathcal{T}(f)$ when $f$ ranges over the set of all quadratic functions equals $\lfloor n/2 \rfloor + 1$.

We see that classical Boolean functions have small algebraic thicknesses. The question addressed in this section is an approximation of the maximum possible value of $\mathcal{T}(f)$ when $f$ ranges over $\mathcal{B}_n$. Clearly, the number of terms in the A.N.F. of any Boolean function $f$, and *a fortiori* $\mathcal{T}(f)$, is smaller than or equal to $2^n$. But is $\max_{f \in \mathcal{B}_n} \mathcal{T}(f)$ polynomial or exponential in $n$? If it was polynomial, then this would indicate a potential weakness of many ciphers using Boolean functions.

### 2.1 A Lower Asymptotic Bound on $\max_{f \in \mathcal{B}_n}(\mathcal{T}(f))$

We show in the next proposition that $\max_{f \in \mathcal{B}_n}(\mathcal{T}(f))$ is exponentially large, since $\liminf_{n \to \infty} \frac{\max_{f \in \mathcal{B}_n}(\mathcal{T}(f))}{2^n} \geq \frac{1}{2}$. Moreover, for almost all Boolean functions $f$, the number $\mathcal{T}(f)$ is not substantially smaller than $2^{n-1}$.

**Theorem 1.** *For every number $\lambda < 1/2$, the density in $\mathcal{B}_n$ of the subset $\{f \in \mathcal{B}_n \mid \mathcal{T}(f) \geq \lambda\, 2^n\}$ is greater than $1 - 2^{2^n H_2(\lambda) - 2^n + n^2 + n}$ where $H_2(x) = -x \log_2(x) - (1-x)\log_2(1-x)$ is the entropy function. This density tends to 1 when $n$ tends to infinity. Thus, there exists $N$ such that for every $n \geq N$, a Boolean function $f$ such that $\mathcal{T}(f) \geq \lambda\, 2^n$ exists. We can take $N = 9$ for $\lambda = \frac{1}{4}$ and $N = 12$ for $\lambda = \frac{3}{8}$.*

*Proof.* Let $k$ be any positive integer. The number of Boolean functions on $F_2^n$ whose A.N.F.s have at most $k$ terms equals $1 + \binom{2^n}{1} + \cdots + \binom{2^n}{k}$.

The number of affine automorphisms on $F_2^n$ equals $(2^n-1)(2^n-2)(2^n-4)\cdots(2^n-2^{n-1})\ 2^n$.
Thus, the number of Boolean functions $f$ such that $\mathcal{T}(f) \leq k$ is smaller than or equal to

$$\left(1+\binom{2^n}{1}+\cdots+\binom{2^n}{k}\right)(2^n-1)(2^n-2)(2^n-4)\cdots(2^n-2^{n-1})\ 2^n,$$

which is smaller than:

$$N(n,k)=\left(1+\binom{2^n}{1}+\cdots+\binom{2^n}{k}\right)2^{\,n^2+n}.$$

We know (cf. for instance [11], page 310) that for every positive $\lambda < \frac{1}{2}$ and every positive integer $N$:

$$\sum_{0\leq i\leq \lambda N}\binom{N}{i} \leq 2^{\,N\,H_2(\lambda)}.$$

Thus the density of the set $\{f \in \mathcal{B}_n \mid \mathcal{T}(f) \geq \lambda 2^n\}$ is greater than $1-\frac{N(n,\lambda 2^n)}{2^{2^n}} \geq 1-2^{2^n\,(H_2(\lambda)-1)+n^2+n}$. The entropy function is strictly increasing on $[0;1/2]$ and its value at $1/2$ is 1. Thus for $\lambda < 1/2$, the expression $2^n\,(H_2(\lambda)-1)+n^2+n$ tends to $-\infty$ when $n$ tends to infinity and the density tends to 1.

We have checked these values $N=9$ and $N=12$ by computation.

**Remark**: $N(n,2^{n-1}) = \left(2^{n-1}+\frac{1}{2}\binom{2^n}{2^{n-1}}\right)2^{\,n^2+n}$. Thus, our method does not work for $\lambda = 1/2$. We do not know if there exist functions $f$ such that $\mathcal{T}(f) > 2^{n-1}$.

### 2.2 An Upper Bound

**Proposition 1.** *For every Boolean function $f$ in $\mathcal{B}_n$, $\mathcal{T}(f)$ is smaller than or equal to $\frac{2}{3}\ 2^n$.*

*Proof.* The proof is by induction on $n$. The assertion is clearly valid for $n=1$. Let $n$ be any integer greater than 1 and assume that the assertion is valid for $n-1$. Let $f$ be any Boolean function in $\mathcal{B}_n$ and let $f_0$ and $f_1$ be the Boolean functions on $F_2^{n-1}$ such that $f(x_1,\cdots,x_n) = f_0(x_1,\cdots,x_{n-1}) \oplus x_n f_1(x_1,\cdots,x_{n-1})$. We shall denote by $|f|$ the number of terms in the A.N.F. of $f$. We have $|f| = |f_0| + |f_1|$. By hypothesis, there exists an affine isomorphism $A$ on $F_2^{n-1}$ such that $|f_1 \circ A| \leq \frac{2}{3}\ 2^{n-1}$. Thus, we can assume without loss of generality that $|f_1| \leq \frac{2}{3}\ 2^{n-1}$. Assume that $|f| = |f_0| + |f_1|$ is greater than $\frac{2}{3}\ 2^n$. Let $r$ be the number of terms which are in both A.N.F.s of $f_0$ and $f_1$. We have $|f_0|+|f_1|-r \leq 2^{n-1}$, since $2^{n-1}$ is the total number of monomials $x^u$ ($x$, $u \in F_2^{n-1}$). Thus $r$ is greater than $\frac{2}{3}\ 2^n - 2^{n-1} = \frac{1}{3}\ 2^{n-1}$. Changing $x_n$

into $x_n \oplus 1$ in the A.N.F. of $f$ keeps $f_1$ unchanged and replaces $f_0$ by $f_0 \oplus f_1$. We have $|f_0 \oplus f_1| + |f_1| = |f_0| + 2|f_1| - 2r = (|f_0| + |f_1| - r) + |f_1| - r < 2^{n-1} + \frac{2}{3}\,2^{n-1} - \frac{1}{3}\,2^{n-1} = \frac{2}{3}\,2^n$.

**Remark**: All the affine isomorphisms we use in the proof above are in fact translations.

### 2.3 Relationship Between $\mathcal{T}$ and the Other Complexity Criteria

We shall say that a function $f$ has high thickness if $\mathcal{T}(f)$ equals $\lambda\,2^n$ where $\lambda$ is near 1/2: this is coherent with Theorem 1. We shall say that $f$ has high nonlinearity if $\mathcal{NL}(f)$ is greater than $2^{n-1} - \kappa_n\,2^{n/2-1}$ where the sequence $\kappa_n$ is under-linear in $n$: we know that for every $n$, there exist (e.g. quadratic) functions on $F_2^n$ with nonlinearity greater than or equal to $2^{n-1} - 2^{\lceil n/2 \rceil - 1}$, but the number of such functions seems very small (this has never been proved but it can be verified for small values of $n$ with computer help). We shall say that $f$ has high algebraic degree if its degree is greater than $\beta n$ where $\beta$ is near 1.

We also consider that $f$ has low thickness if $\mathcal{T}(f)$ is polynomial in $n$ (for some values of the exponent and of the coefficient), that $f$ has low nonlinearity if $\mathcal{NL}(f)$ is smaller than $2^n\lambda$ where $\lambda < 1/2$, and that $f$ has low algebraic degree if its degree is smaller than $\beta n$ where $\beta$ is small.

All those functions whose algebraic thicknesses are high have degrees not substantially smaller than $n/2$, since for every Boolean function of algebraic degree $d$ we have

$$\mathcal{T}(f) \leq \sum_{i=0}^{d} \binom{n}{i} \tag{2}$$

and $\sum_{i=0}^{d} \binom{n}{i}$ is polynomial. $\mathcal{T}(f)$ is small if $d$ is small, but the converse is false: recall that $\mathcal{T}(f)$ equals 1 when $f$ is the indicator of any flat. If this flat is a singleton, then the degree of $f$ equals $n$.

There exist functions with low algebraic thicknesses and with highest possible nonlinearity (e.g. quadratic bent functions). There also exist functions with high algebraic thicknesses and low nonlinearities, since there exist functions with high algebraic thicknesses and low weights: take $\lambda < \lambda' < 1/2$; the number of functions of weights smaller than or equal to $2^n\lambda'$ equals $\sum_{i=0}^{2^n\lambda'} \binom{2^n}{i} \geq \frac{2^{2^n H_2(\lambda')}}{\sqrt{2^{n+3}\lambda'(1-\lambda')}}$ (cf. [11], page 310) and we have seen above that the number of functions $f$ such that $\mathcal{T}(f) \leq 2^n\lambda$ is smaller than or equal to

$$\left(1 + \binom{2^n}{1} + \cdots + \binom{2^n}{k}\right) 2^{\,n^2+n} \leq 2^{2^n H_2(\lambda) + n^2 + n};$$

thus, the latter is asymptotically smaller than the former and there exist functions of weights smaller than or equal to $2^n\lambda'$ satisfying $\mathcal{T}(f) > 2^n\lambda$.

But the most interesting question is clearly: "does there exist functions with high degrees, high nonlinearities and high algebraic thicknesses?". The answer is yes, according to Theorem 1 and to the following results:

**Theorem 2.** *Let $\alpha$ and $\alpha'$ be two numbers such that $\frac{1}{2} < \alpha' < \alpha$. Then, asymptotically, the density of the set $\{f \in \mathcal{B}_n, \mathcal{NL}(f) \geq 2^{n-1} - n^{\alpha}\, 2^{\frac{n}{2}-1}\}$ is greater than $1 - 2^{-n^{2\alpha'}}$ and tends to 1 when $n$ tends to infinity. Thus, there exists $N$ such that for every $n \geq N$, a majority of Boolean functions $f$ are such that $\mathcal{NL}(f) \geq 2^{n-1} - n^{\alpha}\, 2^{\frac{n}{2}-1}$. For instance, for $\alpha = 0.55$ we can take $N = 11$.*

*Proof.* The number of affine functions is $2^{n+1}$. For a given affine function $l$, the number of Boolean functions $f$ such that $d_H(f,l) \leq 2^{n-1} - n^{\alpha}\, 2^{\frac{n}{2}-1}$ equals $\sum_{0 \leq i \leq 2^{n-1} - n^{\alpha}\, 2^{\frac{n}{2}-1}} \binom{2^n}{i} \leq 2^{2^n H_2(\frac{1}{2} - n^{\alpha}\, 2^{-\frac{n}{2}-1})}$. Thus, the number of those Boolean functions which have nonlinearities smaller than or equal to $2^{n-1} - n^{\alpha}\, 2^{\frac{n}{2}-1}$ is smaller than $2^{n+1+2^n H_2(\frac{1}{2} - n^{\alpha}\, 2^{-\frac{n}{2}-1})}$. Since $H_2(1/2) = 1$, $H_2'(1/2) = 0$, and $H_2''(1/2) = -\frac{4}{\ln 2}$, Taylor formula gives $H_2(\frac{1}{2} - n^{\alpha}\, 2^{-\frac{n}{2}-1}) = 1 - \frac{n^{2\alpha} 2^{-n}}{2 \ln 2} + o(n^{2\alpha} 2^{-n})$. Thus, asymptotically, we have $(n+1)2^{-n} + H_2(\frac{1}{2} - n^{\alpha}\, 2^{-\frac{n}{2}-1}) < 1 - n^{2\alpha'} 2^{-n}$ (indeed, take $\alpha' < \alpha'' < \alpha$; then $n+1$ is negligible with respect to $n^{2\alpha} - n^{2\alpha''}$, since $2\alpha > 2\alpha'' > 1$, and $\frac{n^{2\alpha''}}{2 \ln 2}$ is asymptotically greater than $n^{2\alpha'}$) and the density of the set of those Boolean functions which have nonlinearity greater than $2^{n-1} - n^{\alpha}\, 2^{\frac{n}{2}-1}$ is greater than $1 - 2^{2^n (1 - n^{2\alpha'} 2^{-n}) - 2^n} = 1 - 2^{-n^{2\alpha'}}$ and tends to 1 when $n$ tends to infinity. We have checked the last sentence by computation.

Since the density of the set of those functions whose nonlinearities exceed $2^{n-1} - n^{\alpha}\, 2^{\frac{n}{2}-1}$ and of the set of those functions whose algebraic thicknesses exceed $\lambda\, 2^n$ with $\lambda$ near $1/2$ both tend to 1, we deduce that the density of those functions which have both properties tends also to 1. We have seen that these functions cannot have low degrees since relation (2) and the fact that $\mathcal{T}(f)$ exceeds $\lambda\, 2^n$ imply that $d$ cannot be significantly smaller than $\frac{n}{2}$. But we can in fact require that these functions have high degrees:

**Theorem 3.** *Let $\beta$ and $\beta'$ be two positive numbers such that $\beta < \beta' < 1$. The density of the set of all functions on $F_2^n$ whose degrees are greater than $\beta n$ is asymptotically greater than $1 - 2^{-2^{n H_2(1-\beta')}}$ and tends to 1 when $n$ tends to infinity.*

*Proof.* The number of all functions on $F_2^n$ of degrees smaller than or equal to $\beta n$ equals $2^{\tau_n}$ where $\tau_n = \sum_{0 \leq i \leq \beta n} \binom{n}{i} = 2^n - \sum_{0 \leq i \leq n - \beta n - 1} \binom{n}{i}$. We have (cf. [11], page 310) $\tau_n \leq 2^n - \frac{2^{n H_2(1-\beta-1/n)}}{\sqrt{8n(1-\beta-1/n)(\beta+1/n)}}$. Thus, assuming without loss

of generality that $\beta > 1/2$, the density of the set of all functions on $F_2^n$ whose degrees are greater than $\beta n$ equals $1 - 2^{\tau_n - 2^n} \geq 1 - 2^{-\frac{2^n H_2(1-\beta-1/n)}{\sqrt{8n(1-\beta-1/n)(\beta+1/n)}}}$. It is asymptotically greater than $1 - 2^{-2^{n H_2(1-\beta')}}$ and tends to 1 when $n$ tends to infinity.

So, to conclude this section, we can say that, for every $\alpha$ such that $\frac{1}{2} < \alpha$ and every $\beta$ such that $0 < \beta < 1$, there exist functions with degrees greater than $\beta n$, nonlinearities greater than $2^{n-1} - n^{\alpha} 2^{\frac{n}{2}-1}$ and with $\mathcal{T}(f)$ near $2^{n-1}$.

## 3 Normality of Boolean Functions

The complexity criterion on Boolean functions we shall now consider is not yet related to explicit attacks on ciphers. But it is a natural criterion to consider. The situation of the three other criteria when they were first considered was similar. For instance, at the time D.E.S. was designed, only the differential attack was known (and kept secret), but the notion of nonlinearity had been already introduced by Rothaus [19]. The linear attack has been discovered sixteen years later (cf. [13]).
There is a relation (cf. Proposition 2) between the criterion we shall introduce and nonlinearity which shows that to have a chance to be highly nonlinear, a function must satisfy this criterion at a reasonably high level.

Hans Dobbertin has introduced in [4] the following notion: a Boolean function defined on $F_2^n$ ($n$ even) is normal if it is constant on at least one $n/2$-dimensional flat. We generalize this notion and extend it to any $n$:

**Definition 2.** Let $k \leq n$. A Boolean function $f$ on $F_2^n$ is called $k$-normal (resp. $k$-weakly-normal) if there exists a $k$-dimensional flat on which $f$ is constant (resp. affine).

The complexity criterion we are interested in is non-k-normality with small $k$. Philippe Langevin calls index of $f$ the maximum value of $k$ such that $f$ is $k$-normal. Clearly, $k$-normality implies $k$-weak-normality and $k$-weak-normality implies $(k-1)$-normality.

**Examples**:
– Every symmetric Boolean function (i.e. every function whose output is invariant under permutation of its input bits, i.e. whose output depends only on the weight of the input) is $\lfloor \frac{n}{2} \rfloor$-normal and $\lceil \frac{n}{2} \rceil$-weakly-normal since its restriction to the $\lceil \frac{n}{2} \rceil$-dimensional flat:

$$\{(x_1, \ldots x_n) \in F_2^n \mid x_{i+\frac{n}{2}} = x_i \oplus 1, \forall i \leq \frac{n}{2}\}$$

is constant if $n$ is even and affine if $n$ is odd. Indeed, if $n$ is even, all the elements of this flat have same weight $\lfloor \frac{n}{2} \rfloor$ and $f(x)$ takes therefore constant

value; if $n$ is odd, we have $f(x) = f(x_1, \cdots, x_{n-1}, 0) \oplus x_n[f(x_1, \cdots, x_{n-1}, 0) \oplus f(x_1, \cdots, x_{n-1}, 1)]$ where the functions $f(x_1, \cdots, x_{n-1}, 0)$ and $f(x_1, \cdots, x_{n-1}, 1)$ are constant on this flat.
– Every Boolean function on $F_2^n$ with $n \leq 7$ is $\lfloor \frac{n}{2} \rfloor$-normal, as proved by S. Dubuc [5].

There is a mutual upper bound on $k$ and on the nonlinearity of the function:

**Proposition 2.** *Let $f$ be a $k$-weakly-normal Boolean function on $F_2^n$. Denote by $\mathcal{NL}(f)$ its nonlinearity. Then*

$$\mathcal{NL}(f) \leq 2^{n-1} - 2^{k-1},$$

*or equivalently*

$$k \leq \log_2[2^{n-1} - \mathcal{NL}(f)] + 1.$$

Proposition 2 is a direct consequence of a known property of Fourier-Walsh-Hadamard transform. This property is interesting by itself:

**Proposition 3.** *Let $f$ be any Boolean function on $F_2^n$, $E$ any vector subspace of $F_2^n$ and $a$, $b$ two vectors of $F_2^n$. Then*

$$\sum_{u \in b \oplus E} (-1)^{a \cdot u} \widehat{\chi_f}(u) = |E| \ (-1)^{a \cdot b} \sum_{x \in a \oplus E^\perp} (-1)^{f(x) \oplus b \cdot x},$$

*where $|E|$ denotes the size of $E$ and $E^\perp = \{x \in F_2^n;\ \forall y \in E,\ x \cdot y = 0\}$ is its orthogonal. If $f \oplus b \cdot x$ is constant on the flat $a \oplus E^\perp$, then $\sum_{u \in b \oplus E} (-1)^{a \cdot u} \widehat{\chi_f}(u) = \pm 2^n$.*

*Proof.* Let $\varphi$ be any real-valued function on $F_2^n$ and $\widehat{\varphi}$ its Fourier-Walsh-Hadamard transform. We have:

$$\sum_{u \in b \oplus E} (-1)^{a \cdot u} \widehat{\varphi}(u) = \sum_{u \in b \oplus E; x \in F_2^n} \varphi(x) (-1)^{a \cdot u \oplus x \cdot u} =$$

$$\sum_{u \in E; x \in F_2^n} \varphi(x) (-1)^{(a \oplus x) \cdot (b \oplus u)}.$$

Since the sum $\sum_{u \in E} (-1)^{(a \oplus x) \cdot (b \oplus u)}$ is null for every $x \notin a \oplus E^\perp$, we deduce:

$$\sum_{u \in b \oplus E} (-1)^{a \cdot u} \widehat{\varphi}(u) = |E| \sum_{x \in a \oplus E^\perp} (-1)^{(a \oplus x) \cdot b} \varphi(x).$$

Applying this result to $\varphi = \chi_f$ implies

$$\sum_{u \in b \oplus E} (-1)^{a \cdot u} \widehat{\chi_f}(u) = |E| \ (-1)^{a \cdot b} \sum_{x \in a \oplus E^\perp} (-1)^{f(x) \oplus b \cdot x}.$$

If the restriction of $f$ to the flat $a \oplus E^\perp$ equals $b \cdot x \oplus \epsilon$ ($\epsilon \in F_2$), then applying this last equality gives $\sum_{u \in b \oplus E} (-1)^{a \cdot u} \widehat{\chi_f}(u) = \pm 2^n$.

Proposition 3 implies that, if $f \oplus b \cdot x$ is constant on the flat $a \oplus E^\perp$, then the mean of $(-1)^{a \cdot u}\widehat{\chi_f}(u)$ when $u$ ranges over $b \oplus E$ equals $\pm\frac{2^n}{|E|} = \pm|E^\perp|$. Thus the maximum magnitude of $\widehat{\chi_f}(u)$ is greater than or equal to $|E^\perp|$. This implies Proposition 2, according to relation (1).

**Remark**: Proposition 3 also implies a more general result due to Zhang, Zheng and Imai, but proved in a complex way in [25]: let $A$ be any $k$-dimensional flat ($k \leq n$). Let $f$ be a Boolean function on $F_2^n$ and $f'$ its restriction to $A$. Denote by $\mathcal{NL}(f')$ the nonlinearity of $f'$ (i.e. the minimum Hamming distance between $f'$ and any affine function on $A$). Then we have

$$\mathcal{NL}(f) - \mathcal{NL}(f') \leq 2^{n-1} - 2^{k-1}.$$

Indeed, according to Proposition 3 with $A = a \oplus E^\perp$, we have: $\max_{b \in F_2^k} |\widehat{\chi_{f'}}(b)| \leq \max_{(b,u) \in F_2^n} |\widehat{\chi_f}(b,u)|$ which completes the proof.

Let us see now the consequences of Proposition 2 on the properties of Boolean functions with specified degrees.

1. Every quadratic Boolean function $f$ on $F_2^n$ is $\frac{n}{2}$-normal if $n$ is even and $\frac{n+1}{2}$-weakly-normal if $n$ is odd, according to the properties of quadratic functions recalled in the previous section.

2. According to Proposition 2, this implies that the maximum possible nonlinearity of quadratic functions (known by coders as the covering radius of the Reed-Muller code $RM(1,n)$ in the Reed-Muller code $RM(2,n)$) is upper bounded by $2^{n-1} - 2^{\frac{n-2}{2}}$ if $n$ is even and by $2^{n-1} - 2^{\frac{n-1}{2}}$ if $n$ is odd (these values are in fact the exact ones).

We know that, at least for $n \geq 15$, $n$ odd, properties 1 and 2 above *do not generalize to all Boolean functions*. Indeed we know (cf. [16]) that for these values of $n$, there exist Boolean functions with nonlinearities greater than $2^{n-1} - 2^{\frac{n-1}{2}}$. According to Proposition 2, these functions cannot be $\frac{n+1}{2}$-weakly-normal (and *a fortiori* they cannot be $\frac{n+1}{2}$-normal). S. Blackburn and Hans Dobbertin have also shown in [4] that for every even $n \geq 12$, there exist non-$\frac{n}{2}$-normal Boolean functions on $F_2^n$. We investigate now the values of $k$ (depending on $n$, whatever is its parity) for which these results extend.

**Theorem 4.** *Let $\alpha$ be greater than 1. Let $(k_n)_{n \in \mathbf{N}^*}$ be a sequence of positive integers such that $\alpha \log_2 n \leq k_n \leq n$. The density in $\mathcal{B}_n$ of the set of all Boolean functions on $F_2^n$ which are not $k_n$-weakly-normal (and thus which are not $k_n$-normal) is greater than $1 - 2^{n(k_n+1)-2^{k_n}}$. This density tends to 1 when $n$ tends to infinity. Therefore, there exists a positive integer $N$ such that, for every $n \geq N$, non-$k_n$-normal functions exist. For $k_n = \lfloor \frac{n}{2} \rfloor$ we can take $N = 12$.*

*Proof.* Let $\lambda_n$ be the number of $k_n$-dimensional flats in $F_2^n$. Fix such a flat $A$. Let $\mu_n$ be the number of Boolean functions whose restrictions to $A$ are

affine (clearly, this number does not depend on the choice of $A$). The number of $k_n$-weakly-normal functions on $F_2^n$ is smaller than or equal to $\lambda_n \mu_n$.

The number of $k_n$-dimensional vector subspaces of $F_2^n$ equals (cf. [11]) :

$$\begin{bmatrix} n \\ k_n \end{bmatrix} = \frac{(2^n-1)(2^n-2)(2^n-2^2)\cdots(2^n-2^{k_n-1})}{(2^{k_n}-1)(2^{k_n}-2)(2^{k_n}-2^2)\cdots(2^{k_n}-2^{k_n-1})}$$

and the number of $k_n$-dimensional flats in $F_2^n$ is:

$$\lambda_n = 2^{n-k_n} \begin{bmatrix} n \\ k_n \end{bmatrix}.$$

We choose now as particular $k_n$-dimensional flat the set $F_2^{k_n} \times \{(0,\cdots,0)\}$. The restriction to $F_2^{k_n} \times \{(0,\cdots,0)\}$ of a Boolean function on $F_2^n$ is affine if and only if the algebraic normal form of the function contains no monomial of degree at least 2 involving the coordinates $x_1,\cdots,x_{k_n}$ only. The number of such functions is $\mu_n = 2^{\nu_n}$, where $\nu_n = 2^n - 2^{k_n} + 1 + k_n$. The number of $k_n$-weakly-normal functions on $F_2^n$ is smaller than or equal to $2^{n-k_n} \begin{bmatrix} n \\ k_n \end{bmatrix} 2^{\nu_n}$. The number of Boolean functions on $F_2^n$ being equal to $2^{2^n}$, the density of the subset $\mathcal{A}_n$ of $\mathcal{B}_n$ containing all Boolean functions on $F_2^n$ which are not $k_n$-weakly-normal is greater than or equal to

$$1 - 2^{n-k_n} \begin{bmatrix} n \\ k_n \end{bmatrix} 2^{\nu_n - 2^n}.$$

We have $\begin{bmatrix} n \\ k_n \end{bmatrix} < 2^{nk_n - {k_n}^2 + k_n}$, since every factor in the numerator of $\begin{bmatrix} n \\ k_n \end{bmatrix}$ is smaller than $2^n$ and every factor in its denominator is greater than or equal to $2^{k_n-1}$. Thus, the density of $\mathcal{A}_n$ is greater than or equal to

$$1 - 2^{n(k_n+1)+k_n+1-{k_n}^2-2^{k_n}} > 1 - 2^{n(k_n+1)-2^{k_n}}.$$

The exponent $n(k_n+1) - 2^{k_n}$ is smaller than or equal to $2^{k_n/\alpha}(k_n+1) - 2^{k_n}$ and thus tends to $-\infty$ when $n$ tends to $+\infty$.

The last sentence of the proposition can be checked by computation (the sequences $1 - 2^{n-k_n} \begin{bmatrix} n \\ k_n \end{bmatrix} 2^{\nu_n - 2^n}$, $n$ even and $n$ odd are increasing and positive respectively for $n \geq 12$ and $n \geq 13$).

**Remark**: Theorem 4 remains valid if we only assume that $\frac{2^{k_n}}{nk_n}$ tends to infinity. It also remains valid (except "$N = 12$") if, in the definition of $k$-weakly-normal functions, we replace "there exists a $k$-dimensional flat on which it is affine" by "there exists a $k$-dimensional flat such that the restriction of

the function to this flat has degree $\leq l$", where $l$ is some fixed positive integer: the value of $\nu_n$ has then to be changed into $2^n - 2^{k_n} + 1 + \binom{k_n}{1} + \ldots + \binom{k_n}{l}$.

X.-D. Hou has shown in [7] that, for any odd $n \leq 13$, the maximum nonlinearity of all functions of degree 3 is the same as for quadratic functions: $2^{n-1} - 2^{\frac{n-1}{2}}$. So we could hope that Boolean functions of degree 3 behave for every $n$ as quadratic functions with respect to nonlinearity or to normality. For nonlinearity, this is an open problem. But for normality, we will show the existence of non-$k_n$-normal Boolean functions of degree 3, where $k_n$ is negligible with respect to $n$. This confirms the feeling that these functions behave merely as general functions, considering their combinatorial properties (cf. [2]).

**Proposition 4.** *Let $\lambda > \frac{1}{2}$. Let $(k_n)_{n \in \mathbf{N}^*}$ be a sequence of positive integers such that $n^\lambda \leq k_n \leq n$. The density of the set of all Boolean functions of degrees at most 3 on $F_2^n$ which are not $k_n$-weakly-normal (and thus which are not $k_n$-normal) in the set of all Boolean functions of degrees at most 3 is greater than or equal to $1 - 2^{n(k_n+1)-{k_n}^2-\binom{k_n}{2}-\binom{k_n}{3}}$. This density tends to 1 when $n$ tends to infinity. Therefore, there exists a positive integer $N$ such that, for every $n \geq N$, such functions exist. For $k_n = \lceil \frac{n}{2} \rceil$ we can take $N = 15$.*

*Proof.* Let $\mu'_n$ be the number of Boolean functions of degrees at most 3 whose restrictions to $F_2^{k_n} \times \{(0, \cdots, 0)\}$ are affine. The number of $k_n$-weakly-normal functions of degrees at most 3 on $F_2^n$ is smaller than or equal to $\lambda_n \mu'_n$, where $\lambda_n$ is the number of $k_n$-dimensional flats in $F_2^n$.

The number of functions whose restrictions to $F_2^{k_n} \times \{(0, \cdots, 0)\}$ are affine equals $\mu'_n = 2^{\nu'_n}$, where $\nu'_n = 1 + n + \binom{n}{2} + \binom{n}{3} - \binom{k_n}{2} - \binom{k_n}{3}$. The number of $k_n$-weakly-normal functions of degree at most 3 on $F_2^n$ is smaller than or equal to $2^{n-k_n} \begin{bmatrix} n \\ k_n \end{bmatrix} 2^{\nu'_n}$ and the density of this set in the set of all Boolean functions of degree at most 3 is greater than or equal to

$$1 - 2^{n-k_n} \begin{bmatrix} n \\ k_n \end{bmatrix} 2^{\nu'_n - \kappa_n}$$

where $\kappa_n = 1 + n + \binom{n}{2} + \binom{n}{3}$ and is therefore greater than

$$1 - 2^{n + nk_n - {k_n}^2 - \binom{k_n}{2} - \binom{k_n}{3}}.$$

Since the binomial coefficient $\binom{k_n}{3}$ has degree 3 with respect to $k_n$ and since the sequence $\frac{{k_n}^2}{n}$ tends to infinity, the exponent $n + nk_n - {k_n}^2 - \binom{k_n}{2} - \binom{k_n}{3}$ tends to $-\infty$ when $n$ tends to $\infty$.

For $k_n = \lceil \frac{n}{2} \rceil$, it is a simple matter to check that, for $n \geq 16$, we have $n + \frac{n^2}{4} < \binom{n/2}{2} + \binom{n/2}{3}$ ($n$ even) and $n + \frac{n^2-1}{4} < \binom{(n+1)/2}{2} + \binom{(n+1)/2}{3}$ ($n$ odd). For $n = 15$ (and $n = 13$) we checked that $1 - 2^{n-k_n} \begin{bmatrix} n \\ k_n \end{bmatrix} 2^{\nu'_n - \kappa_n} > 0$.

**Remark**: Proposition 4 remains valid if we assume only that the sequence $\frac{{k_n}^2}{n}$ tends to infinity. It also remains valid (except "$N = 15$") if, in the definition of $k$-weakly-normal functions, we replace "there exists a $k$-dimensional flat on which it is affine" by "there exists a $k$-dimensional flat on which it is quadratic" (we then just have to withdraw the term $\binom{k_n}{2}$ from the proof). Obviously, Proposition 4 can also be generalized to all other fixed degrees.

### 3.1 Relationship Between Normality and the Other Complexity Criteria

We have seen in Proposition 2 that if $f$ is $k$-(weakly)-normal, then $\mathcal{NL}(f) \leq 2^{n-1} - 2^{k-1}$. Notice that, since every Boolean function has nonlinearity upper bounded by $2^{n-1} - 2^{n/2-1}$, this gives no information if $k \leq n/2$. But the high nonlinearity of bent functions ($2^{n-1} - 2^{n/2-1}$) implies that they cannot be $(\frac{n}{2} + 1)$-weakly-normal.

Anyway, $k$-normality with large $k$ implies low nonlinearity, but we do not know whether the converse is true or not.

Low degree of Boolean functions does not imply their normality: we have seen in Proposition 4 that there exist functions of degree 3 which are non-k-normal with low $k$.

$k$-normality does not imply either low degree: take a function of high degree on $F_2^{n-1}$ (considered as a subspace of $F_2^n$) and complete it by 0 to obtain a function on $F_2^n$.

There exist functions $f$ with low algebraic thicknesses (e.g. functions of degree 3) which are non-$k$-normal with small $k$ according to Proposition 4; and there exist functions with high algebraic thicknesses which are k-normal with large $k$: take a function $g$ on $F_2^{n-1}$ with high $\mathcal{T}(g)$ and complete it by 0 to obtain a function $f$ on $F_2^n$; it is a simple matter to check that $\mathcal{T}(f) \geq \mathcal{T}(g)$.

The most interesting point is that *almost all functions have high degrees, high nonlinearities, high algebraic thicknesses and are non-k-normal with small k's*, since we have seen at subsection 2.3 that the density of those functions which have high degrees (greater than $\beta n$ where $\beta < 1$), high nonlinearities (greater than $2^{n-1} - n^{\alpha}\, 2^{\frac{n}{2}-1}$ where $\frac{1}{2} < \alpha$), high algebraic thicknesses (more than approximately $2^{n-1}$) tends to 1 and we know that the density of the set of those functions which are non-$k$-normal with $k$ logarithmic in $n$ tends also to 1.

**Remark**: We can also require that these functions admit no linear structure. This can be necessary because, if the function used in a cipher admits a linear structure, the complexity of an exhaustive search of the key may be reduced [6]. Thus the non-existence of a linear structure can also be considered as a complexity criterion (of a different kind, since functions either satisfy this criterion or do not satisfy it, while all the other criteria are satisfied at levels quantified by numbers). A linear structure of a Boolean function $f$ is any nonzero word $a \in F_2^n$ such that the function $D_a f(x) = f(x \oplus a) \oplus f(x)$

is constant. The existence of a linear structure for a function $f$ is equivalent to the existence of a Boolean function $g$ on $F_2^{n-1}$ and of an affine function $l$ on $F_2^n$ such that $f(x_1, \ldots, x_n)$ is affinely equivalent to the function $g(x_1, \ldots, x_{n-1}) \oplus l(x_1, \ldots, x_n)$ (cf. for instance [5]). Thus, the number of functions admitting linear structures is smaller than or equal to $2^{2^{n-1}}$, times the number of affine automorphisms, times the number of affine functions, and is therefore smaller than $2^{2^{n-1}+n^2+2n+1}$. The density of the set of functions admitting no linear structure is then greater than or equal to $1 - 2^{n^2+2n+1-2^{n-1}}$ and tends to 1.

Meier and Staffelbach introduced in [14] a complexity criterion (in their paper, they wrote "nonlinearity criterion") satisfied at levels quantified by numbers and related to this notion: a Boolean function on $F_2^n$ being given, its "distance to linear structures" is its distance to the set of all Boolean functions admitting linear structures (among which we have all affine functions). Let $\rho$ be a positive number smaller than 1/2. The number of Boolean functions on $F_2^n$ which lie at distance smaller than or equal to $\rho\, 2^n$ from this set is smaller than or equal to $2^{2^{n-1}+n^2+2n+1} \sum_{i=0}^{\rho\, 2^n} \binom{2^n}{i} \leq 2^{2^{n-1}+n^2+2n+1+2^n H_2(\rho)}$. Thus, this number is negligible with respect to $2^{2^n}$ if $H_2(\rho) < 1/2$ and, asymptotically, almost all functions lie then at distance greater than $\rho\, 2^n$ to the set of all Boolean functions admitting linear structures.

**General remark**:

1. We see that all complexity citeria studied in this paper are not contradictory to each others. This is different when we also consider criteria more related to the principle of diffusion, such as correlation immunity, resilience, cf. [3]. However, *all the results above are essentially valid if we restrict ourselves to balanced functions.* Indeed, the number of balanced functions on $F_2^n$ equals $\binom{2^n}{2^{n-1}} = \Theta(2^{2^n - n/2})$ (cf. [11], page 309) and all our arguments can be used , replacing the number of functions, $2^{2^n}$, by $\binom{2^n}{2^{n-1}}$.

2. There is the same "Shannon effect" with (linear) error correcting codes, Boolean functions from circuit viewpoint, and Boolean functions from cryptographic viewpoint: we know by combinatorial arguments that, asymptotically, good codes (resp. good functions) exist; moreover, we know that, for sufficiently large values of their lengths (resp. of their numbers of variables) almost all of them are good, but the values of the length (resp. of the number of variables) for which we can assert that many are good make impossible the verification of the quality of such codes (resp. functions) chosen at random. We know very few examples of non-$\lceil \frac{n}{2} \rceil$-normal functions: the functions obtained by Patterson and Wiedmann [16] and a few functions obtained for $n$ even $\geq 8$ by Sylvie Dubuc in her thesis [5].

## 4 On $q$-ary Functions

Since the motivation of this paper was cryptographic, we have focused on Boolean functions. But our main results generalize to functions from $F_q^n$ to $F_q$, where $q$ is any power of a prime. We denote by $\mathcal{B}_n^q$ the set of such $q$-ary functions. The algebraic normal form of any function in $\mathcal{B}_n^q$ has the form:

$$f(x_1, \cdots, x_n) = \sum_{u \in {Z_q}^n} a_u \left( \prod_{i=1}^{n} {x_i}^{u_i} \right), a_u \in F_q$$

where $Z_q = \{0, \ldots, q-1\}$. It exists and is unique (see [18]).

We denote by $\mathcal{T}(f)$ again the minimum number of terms (i.e. of monomials with nonzero coefficients) in the A.N.F.s of the functions $f \circ A$, where $A$ ranges over the set of all affine automorphisms of $F_q^n$. A Boolean function $f$ on $F_q^n$ is called $k$-normal (resp. $k$-weakly-normal) if there exists a $k$-dimensional flat on which $f$ is constant (resp. has degree at most 1).

**Theorem 5.** *For every $\lambda$ such that $\frac{H_2(\lambda)}{\log_2 q} + \lambda < 1$, the density of the subset of $\mathcal{B}_n^q$ which contains all functions such that $\mathcal{T}(f) \geq \lambda\, q^n$ is greater than $1 - 2^{q^n H_2(\lambda)}\, q^{\lambda q^n + n^2 + n - q^n}$. This density tends to 1 when $n$ tends to infinity.*

*Proof.* Let $k$ be any positive integer. The number of functions in $\mathcal{B}_n^q$ whose A.N.F.s have at most $k$ terms equals $1 + \binom{q^n}{1}(q-1) + \cdots + \binom{q^n}{k}(q-1)^k$.
The number of affine automorphisms on $F_q^n$ equals $(q^n - 1)(q^n - q)(q^n - q^2) \cdots (q^n - q^{n-1})\, q^n$.
Thus, the number of functions $f$ such that $\mathcal{T}(f) \leq k$ is smaller than or equal to

$$\left(1 + \binom{q^n}{1}(q-1) + \cdots + \binom{q^n}{k}(q-1)^k\right)(q^n-1)(q^n-q)(q^n-q^2)\cdots(q^n-q^{n-1})\, q^n$$

$$< \left(1 + \binom{q^n}{1} + \cdots + \binom{q^n}{k}\right) q^{\,k+n^2+n}.$$

Thus the density of the set $\{f \in \mathcal{B}_n^q \mid \mathcal{T}(f) \geq \lambda q^n\}$ in $\mathcal{B}_n^q$ is greater than $1 - 2^{q^n H_2(\lambda)}\, q^{\lambda q^n + n^2 + n - q^n} = 1 - q^{q^n \left(\frac{H_2(\lambda)}{\log_2 q} + \lambda q^n - q^n + n^2 + n\right)}$ and tends to 1 since $\frac{H_2(\lambda)}{\log_2 q} + \lambda < 1$.

**Theorem 6.** *Let $\alpha$ be greater than 1. Let $(k_n)_{n \in \mathbf{N}^*}$ be a sequence of positive integers such that $\alpha \log_q n \leq k_n \leq n$. The density of the subset of $\mathcal{B}_n^q$ containing those Boolean functions on $F_q^n$ which are not $k_n$-weakly-normal (and thus which are not $k_n$-normal) is greater than $1 - q^{n(k_n+1)+k_n-k_n^2+1-q^{k_n}}$. This density tends to 1 when $n$ tends to infinity.*

*Proof.* The number of $k_n$-dimensional flats of $F_q^n$ equals:

$$q^{n-k_n}\frac{(q^n-1)(q^n-q)(q^n-q^2)\cdots(q^n-q^{k_n-1})}{(q^{k_n}-1)(q^{k_n}-q)(q^{k_n}-q^2)\cdots(q^{k_n}-q^{k_n-1})} \leq q^{n-k_n+nk_n-k_n(k_n-1)}.$$

The number of $q$-ary functions whose restrictions to a given $k_n$-dimensional flat (e.g. $F_q^{k_n} \times \{(0,\cdots,0)\}$) have degrees at most 1 equals $q^{q^n-q^{k_n}+1+k_n}$. Thus, the number of $k_n$-weakly-normal functions on $F_q^n$ is smaller than or equal to $q^{q^n-q^{k_n}+n(k_n+1)+k_n-k_n^2+1}$ and the density of the subset of $\mathcal{B}_n^q$ containing all Boolean functions on $F_q^n$ which are not $k_n$-weakly-normal is greater than or equal to

$$1-q^{n(k_n+1)+k_n-k_n^2+1-q^{k_n}}$$

and tends to 1.

## Acknowledgement

The author wishes to thank Anne Canteaut, Gérard Cohen, Grisha Kabatianski, Horacio Tapia Recillas and Gerardo Vega for useful discussions.

## References

1. A. Canteaut and M. Trabbia. *Improved fast correlation attacks using parity check equations of weight 4 and 5.* Advances in Cryptology – EUROCRYPT 2000, number 1805 in Lecture Notes in Computer Science, pp. 573-588. Springer-Verlag, 2000.
2. C. Carlet *A transformation on Boolean functions, its consequences on some problems related to Reed-Muller codes*, actes de EUROCODES' 90, Lecture Notes in Computer Sciences n° 514, pp. 42-50, Springer-Verlag (1991)
3. C. Carlet and P. Sarkar. *Spectral Domain Analysis of Correlation Immune and Resilient Boolean Functions*, Finite Fields Appl. To appear (2001).
4. H. Dobbertin. *Construction of bent functions and balanced Boolean functions with high nonlinearity.* Fast Software Encryption (Proceedings of the 1994 Leuven Workshop on Cryptographic Algorithms), Lecture Notes in Computer Science 1008, Springer Verlag, pp. 61-74, 1995.
5. S. Dubuc. *Etude des propriétés de dégérescence et de normalité des fonctions booléennes et construction de fonctions q-aires parfaitement non linéaires.* PhD thesis, University of Caen, 2001.
6. J. H. Evertse, *Linear structures in block ciphers*, Advances in Cryptology, EUROCRYPT' 87, Lecture Notes in Computer Science 304, pp. 249-266, Springer Verlag, 1988.
7. X.-D. Hou, *On the covering radius of* $r(1,m)$ *into* $r(3,m)$. IEEE Transactions on Information Theory, vol 42 n°3, pp. 1035-1037, 1996.
8. L.R. Knudsen. *Truncated and higher order differentials.* Fast Software Encryption, Second International Workshop, Lecture Notes in Computer Science, n 1008. pp. 196–211. – Springer-Verlag, 1995.

9. X. Lai. *Higher order derivatives and differential cryptanalysis.* Proc. "Symposium on Communication, Coding and Cryptography", in honor of J. L. Massey on the occasion of his 60'th birthday. 1994.
10. Lupanov. *On circuits of functional elements with delay.* Probl. Kibern. 23, pp. 43-81 (1970).
11. F.J. MacWilliams and N.J.A. Sloane. *The theory of error-correcting codes.* North-Holland, 1977.
12. J.L. Massey. *Shift-register synthesis and BCH decoding.* IEEE Transactions on Information Theory, vol. 15, pp. 122–127, 1969.
13. M. Matsui. *Linear cryptanalysis method for DES cipher.* Advances in Cryptology - EUROCRYPT'93, number 765 in Lecture Notes in Computer Science. Springer-Verlag, 1994.
14. W. Meier and O. Staffelbach, *Nonlinearity Criteria for Cryptographic Functions*, Advances in Cryptology, EUROCRYPT' 89, Lecture Notes in Computer Science 434, pp. 549–562, Springer Verlag (1990)
15. A. Menezes, P. van Oorschot, and S. Vanstone. *Handbook of Applied Cryptography.* CRC Press Series on Discrete Mathematics and Its Applications, 1996.
16. N.J. Patterson and D.H. Wiedemann. *The covering radius of the* $[2^{15}, 16]$ *Reed-Muller code is at least 16276.* IEEE Trans. Inform. Theory, IT-29, pp. 354–356, 1983.
17. N.J. Patterson and D.H. Wiedemann. *Correction to [16].* IEEE Trans. Inform. Theory, IT-36(2), pp. 443, 1990.
18. V. S. Pless, W. C. Huffman, Eds, R. A. Brualdi, *Handbook of Coding Theory*, Amsterdam, the Netherlands: Elsevier, 1998.
19. O. S. Rothaus. On bent functions, *J. Comb. Theory*, 20A, pp. 300-305, 1976.
20. R. A. Rueppel *Analysis and design of stream ciphers* Com. and Contr. Eng. Series, Berlin, Heidelberg, NY, London, Paris, Tokyo 1986
21. C.E. Shannon. *A mathematical theory of communication.* Bell system technical journal, 27, pp. 379–423, 1948.
22. C.E. Shannon. *Communication theory of secrecy systems.* Bell system technical journal, 28, pp. 656–715, 1949.
23. C.E. Shannon. *The synthesis of two-terminal switching circuits.* Bell system technical journal, 28, pp. 59–98, 1949.
24. I. Wegener. *The complexity of Boolean functions.* B.G. Teubner, Stuttgart. John Wiley and sons. 1987.
25. Y. Zheng, X.-M. Zhang and H. Imai. *Restriction, terms and nonlinearity of Boolean functions.* Theoretical Computer Science, 226(1-2), pp. 207–223, 1999.

# On Divisibility of Exponential Sums over Finite Fields of Characteristic 2

F.N. Castro and O. Moreno

Department of Mathematics, University of Puerto Rico, PO Box 23355, SJ, PR, 00931-3355

**Abstract.** Moreno-Moreno's result and the covering method give estimates on the 2-divisibility of the number of solutions of a system of polynomial equations. Let us call $\mu$ the order of the 2-divisibility given by Moreno-Moreno's result and $\mu'$ that of the covering method. In 2000 we proved that $\mu' \geq \mu$. In this present paper we give general conditions under which $\mu' > \mu$. We also give some results concerning the tightness of the covering method.

## 1 Intoduction

In [12] and [8], Moreno et al. introduced the covering method. This is a combinatorial technique that give a lower bound for the 2-divisibility of exponential sums over finite fields of characteristic two. In [12] and [8], they proved that the covering method improves the binary Ax's theorem. In [9], we introduce a generalization to the covering method, and prove the Moreno-Moreno's theorem for finite fields of characteristic two. In [10], we give a survey of the best methods to obtain divisibility results. Let $\mu$ be the order of the 2-divisibility given by Moreno-Moreno's result and $\mu'$ be the order of the 2-divisibility given by the covering method. In [9], we proved that $\mu' \geq \mu$. In the present paper we give general conditions under which $\mu' > \mu$. Another theorem of this paper is a tight lower bound on the 2-divisibility of the number of solutions of a system of polynomial equations over $\mathbb{F}_2$. Our bound is tight in the following sense: if the $\epsilon_1 x_1^{e_{11}} \cdots x_n^{e_{1n}}, \ldots, \epsilon_N x_1^{e_{N1}} \cdots x_n^{e_{Nn}}$ monomials are fixed and the coefficients $\epsilon_i$ varied, then there is a choice of coefficients which yield a polynomial for which the lower bound is tight. Since a polynomial over $\mathbb{F}_2$ is a Boolean function, our result can be stated as follows: Consider the following set $\mathcal{G} = \{\epsilon_1 x_1^{e_{11}} \cdots x_n^{e_{1n}} + \cdots + \epsilon_N x_1^{e_{N1}} \cdots x_n^{e_{Nn}} : \epsilon_i \in \mathbb{F}_2\}$ of Boolean functions, if the minimum number of monomials that cover all the variables $x_1, \ldots x_n$ is $r$, then there is a Boolean function $F \in \mathcal{G}$ such that the exact 2-divisibility of the number $|\{(a_1, \ldots, a_n) : F(a_1, \ldots, a_n) = 0\}|$ is $2^{r-1}$.

Finally we want to point out that the divisibility of exponential sums of characteristic two is very important and has been used many times in coding theory(see [12], and [5]).

## 2 On the Divisibility of the Number of Rational Points

In this section, we are going to state some definitions and results about the divisibility of the number of solutions of a system of polynomial equations over $\mathbb{F}_q$. Let $F_1(x_1,\ldots,x_n),\ldots,F_t(x_1,\ldots,x_n)$ be polynomials over $\mathbb{F}_q$, and let $N(F_1,\ldots,F_t)$ be the number of solutions of the system of polynomial equations: $F_1 = 0,\ldots,F_t = 0$. Without of lost of generality, we assume throughout the rest of the paper, that the $F_i$'s are not polynomials in some proper subset of the variables $\{x_1,\ldots,x_n\}$.

In 1935, Chevalley proved a conjectured by E. Artin.

**Theorem 1 (Chevalley).** *If $F(x_1,\ldots,x_n)$ is a homogeneous polynomial of total degree $d$ over a finite field $\mathbb{F}_q$ having $q = p^f$ elements and $n > d$, then $F$ has a nontrivial zero.*

Ax obtained an improvement of the Chevalley's theorem(see [2]). Now we state the Ax's theorem.

**Theorem 2 (Ax).** *With the notation of Theorem 1. If $\mu$ is equal to $\lceil n/d\rceil - 1$, where $\lceil a\rceil$ is the smallest integer larger or equal to $a$, then the number of zeros of $F$ is divisible by $q^\mu$.*

In 1971, N. Katz improved Ax's theorem(see [4]). Now we state the Katz's theorem:

**Theorem 3 (Katz).** *Let $F_1,\ldots,F_t$ be polynomials in $n$ variables with coefficients in $F_q$ of total degrees $d_1,\ldots,d_t$, respectively. Let $\mu$ be the least integer that satisfies*

$$\mu \geq \frac{n - \sum_{i=1}^{t} d_i}{max_i\, d_i}\,.$$

*Then $q^\mu$ divides $N(F_1,\ldots,F_t)$.*

*Example 1.* Let $F_1(x_1,\ldots,x_5) = x_1^3 + \cdots + x_5^3$, and $F_2(x_1,\ldots,x_5) = x_1 + \cdots + x_5$ be polynomials over $\mathbb{F}_q$. Applying Katz's theorem, we have that $q$ divides $N(F_1,F_2)$.

Moreno and Moreno gave in [11] an improvement to the Ax-Katz's theorem. Before we state the Moreno and Moreno result, we need to give a definition.

**Definition 1.** For each integer $n$ with $p$-expansion

$$n = a_0 + a_1 p\cdots + a_s p^s \text{ where } 0 \leq a_i < p\,,$$

we denote its $p$-weight by $\sigma_p(n) = \sum_{i=0}^{s} a_i$. The $p$-weight degree of a monomial $x^d = x_1^{d_1}\cdots x_n^{d_n}$ is $w_p(x^d) = \sigma_p(d_1) + \cdots + \sigma_p(d_n)$. The $p$-weight degree of a polynomial $F(x_1,\ldots,x_n) = \sum_d a_d x^d$ is $w_p(F) = max_{x^d,\, a_d\neq 0} w_p(x^d)$

The Moreno-Moreno's result is the following:

**Theorem 4 (Moreno-Moreno).** *Let $F_1, \ldots, F_t$ be polynomials in $n$ variables with coefficients in $\mathbb{F}_q$, a finite field with $q = p^f$ elements. Let $w_p(F_i)$ be the $p$-weight degree of $F_i$ and let $\mu$ be the smallest integer such that*

$$\mu \geq f\left(\frac{n - \sum_{i=1}^{t} w_p(F_i)}{max_i\, w_p(F_i)}\right).$$

*Then $p^\mu$ divides $N(F_1 \ldots, F_t)$.*

*Example 2.* Let $F_1(x_1, \ldots, x_4) = x_1^3 + \cdots + x_4^3$, and $F_2(x_1, \ldots, x_4) = x_1 + \cdots + x_4$ be polynomials over $\mathbb{F}_{2^f}$. Applying Moreno-Moreno's theorem, we have that $2^{\lceil f/2 \rceil}$ divides $N(F_1, F_2)$.

*Remark 1.* Theorem 3 improves Theorem 4 whenever the characteristic is small with comparison to the degrees, i.e., we need, say $p < \max_i d_i$ in order for an improvement to occur.

Now we are going to describe the Adolphson-Sperber's method(see [1]). Let $F(x_1, \ldots, x_n) = \sum_{(d_1,\ldots,d_n)\in D} a_{d_1,\ldots,d_n} x_1^{d_1} \cdots x_n^{d_n}$ be a polynomial over $\mathbb{F}_q$. The Newton polyhedron $\Delta(F)$ is defined to be convex hull in $\mathbf{R}^n$ of the set $D \bigcup \{(0, \ldots, 0)\}$. Let $\omega(F)$ be the smallest positive rational number such that $\omega(F)\Delta(F)$ contains at least one point with positive integral coordinates. Now, we state the Adolphson and Sperber's theorem:

**Theorem 5 (Adolphson-Sperber).** *With the above notation and assumptions, we have*

$$q^\mu \text{ divides } N(F_1, \ldots, F_t).$$

*where $\mu$ is the smallest integer greater than or equal to $\omega(\sum_{i=1}^{t} y_i F_i) - t$.*

Let us illustrate Theorem 6, computing $\omega(vF)$ for a polynomial.

*Example 3.* Let $F(x, y, z) = x^3 + y^5 + z^7 + xy + xz + yz$ be a polynomial over $\mathbb{F}_q$. The convex hull associated to $F$ is the convex set generated by the following points:

$$v_1 = (3, 0, 0, 1),\ v_2 = (0, 5, 0, 1),\ v_3 = (0, 0, 7, 1),\ v_4 = (1, 1, 0, 1),$$
$$v_5 = (1.0, 1, 1),\ v_6 = (0, 1, 1, 1),\ v_7 = (0, 0, 0, 0).$$

Then $\omega(F) = 2$. Hence $q$ divides $N(x^3 + y^5 + z^7 + xy + xz + yz)$.

## 3 On 2-Divisibility of Exponential Sums; Method of Covering

Let $\mathbb{F}_2$ be the binary field and $F(x_1, \ldots, x_n)$ be a polynomial in $n$ variables over $\mathbb{F}_2$. Let $\mathcal{C}(F)$ be a minimal set of the monomials of $F$ covering all variables, that is, every variable $x_i$ is in at least one monomial in $\mathcal{C}(F)$ and $\mathcal{C}(F)$ is minimal with that property. We call this set $\mathcal{C}(F)$ a *minimal covering of the variables of* $F$ and we assume that its cardinality is $r$.

*Example 4.* Let $F(x_1, \ldots, x_n) = x_1x_2x_3 + x_4x_5 + x_1 + \cdots + x_5$ be a polynomial over $\mathbb{F}_2$. In this case $\mathcal{C}(F) = \{ x_1x_2x_3, x_4x_5 \}$.

The following lemma establishes divisibility properties of an exponential sum which will determine the divisibility of the number of zeros of a system of polynomial equations(see [8], [12]). Lemma 1 is an improvement to the binary Ax's theorem.

**Lemma 1 (Moreno et al).** *Let $F(x) = F(x_1, \ldots, x_n)$ be a polynomial in $n$ variables over $\mathbb{F}_2$. Let $\mathcal{C}(F)$ be a minimal set of monomials of $F(x)$ covering all the variables and $r$ be the cardinality of $\mathcal{C}(F)$. Then*

$$2^r \text{ divides } S(F) = \sum_{x_1, \ldots, x_n \in \mathbb{F}_2} (-1)^{F(x_1, \ldots, x_n)}.$$

For the proof see [8], [12].

*Example 5.* Let $F(x_1, \ldots, x_6) = x_1x_2 + x_2x_3 + x_3x_4 + x_4x_5 + x_5x_6 + x_1x_6 + x_1 \cdots + x_6$ be a polynomial over $\mathbb{F}_2$. We have $\mathcal{C}(F) = \{x_1x_2, x_3x_4, x_5x_6\}$, hence 8 divides $S(F)$.

In [9], we proved that Lemma 1 is an improvement to the Adolphson-Sperber's theorem for polynomial equations over $\mathbb{F}_2$, i.e., $r \geq \omega(F)$.

Now, we state an immediate consequence of Lemma 1:

**Theorem 6.** *Let $F(x_1, \ldots, x_n)$ be a polynomial of degree $d$ over $\mathbb{F}_2$. If there exists a variable $x_j$ that does not appear in the leader monomials of $F$, then*

$$2^{[n/d]+1} \text{ divides } S(F) = \sum_{x_1, \ldots, x_n \in \mathbb{F}_2} (-1)^{F(x_1, \ldots, x_n)}.$$

*Remark 2.* If $d$ divides $n$, Theorem 6 gives an extra two factor when it is compared to the binary Ax's theorem.

Let $F_1(x_1, \ldots, x_2), \ldots, F_t(x_1, \ldots, x_n)$ be polynomials over $\mathbb{F}_2$. We denote the number solutions of the system of polynomial equations $F_1(x_1, \ldots, x_n) = 0, \ldots, F_t(x_1, \ldots, x_n) = 0$ by $N(F_1, \ldots, F_t)$. Now we state the identity that associated exponential sums and the number of solutions of a system of polynomial equations:

$$N(F_1, \ldots, F_t) = \frac{1}{2^{ft}} \sum_{\substack{x_1, \ldots, x_n \in \mathbb{F}_2 \\ y_1, \ldots, y_t \in \mathbb{F}_2}} \psi(\sum_{i=1}^{t} y_i F_i(x_1, \ldots, x_n)),$$

where $\psi$ is an additive character.

Combining Lemma 1 and the above identity, we have the following theorem:

**Theorem 7.** *Let $F_1(x_1,\ldots,x_n),\ldots,F_t(x_1,\ldots,x_n)$ be polynomials over $\mathbb{F}_2$. Let $\mathcal{C}(F_1,\ldots,F_t)$ be a minimal set of monomials of $\sum_{i=1}^{t} y_iF_i(x)$ covering all the variables and $r$ be the cardinality of $\mathcal{C}(F_1,\ldots,F_t)$. Then*

$$2^{r-t} \text{ divides } N(F_1,\ldots,F_t)$$

*Example 6.* Let $F_1(x_1,\ldots,x_5) = x_1x_2 + x_1 + \cdots + x_5$, and $F_2(x_1,\ldots,x_5) = x_1x_2x_3 + x_4x_5 + x_1x_6$ be polynomials over $\mathbb{F}_2$. In this case, we have that $|\mathcal{C}(F_1,F_2)| = 3$, hence 2 divides $N(F_1,F_2)$. In particular, the system $F_1(x_1,\ldots,x_5) = 0$ and $F_2(x_1,\ldots,x_5) = 0$ has a nontrivial solution. Note that Theorem 3 and 4 do not give any information about $N(F,F_2)$.

As preparation for the statement and the proof of the main result of this section, we note the following lemma.

**Lemma 2.** *Let $\epsilon_1 x_1^{e_{11}}\cdots x_n^{e_{1n}}, \ldots, \epsilon_N x_1^{e_{N1}}\cdots x_n^{e_{Nn}}$ be monomials over $\mathbb{F}_2$, and let $r$ be the minimal number of monomials of $\{\epsilon_1 x_1^{e_{11}}\cdots x_n^{e_{1n}}, \ldots, \epsilon_N x_1^{e_{N1}}\cdots x_n^{e_{Nn}}\}$ that covers $x_1,\ldots,x_n$. If $\mathcal{G}$ is the class of polynomials generated by the monomials*

$$\epsilon_1 x_1^{e_{11}}\cdots x_n^{e_{1n}}, \ldots, \epsilon_N x_1^{e_{N1}}\cdots x_n^{e_{Nn}},$$

*where $\epsilon_i \in \{0,1\}$.*

$$\mathcal{G} = \{\epsilon_1 x_1^{e_{11}}\cdots x_n^{e_{1n}} + \cdots + \epsilon_N x_1^{e_{N1}}\cdots x_n^{e_{Nn}} \,:\, \epsilon_i \in \{0,1\}\}.$$

*Then there is one polynomial $F'$ in $\mathcal{G}$ such that $N(F')$ is divisible by $2^{r-1}$ but not divisible $2^r$.*

*Proof.* Let $yF(x_1,\ldots,x_n) = \sum_{i=1}^{N} \epsilon_i y x_1^{e_{i1}}\cdots x_n^{e_{in}}$. We are going to use the following identities $(-1)^{\epsilon_i y x_1^{e_{i1}}\cdots x_n^{e_{in}}} = 1 - 2\epsilon_i y x_1^{e_{i1}}\cdots x_n^{e_{in}}$ and $x_i^l = x_i$ for $l > 0$, throughout the proof.

$$2N(F) = \sum_{x_1,\ldots,x_n,y\in\mathbb{F}_2} \prod_{(e_{i1},\ldots,e_{in})} (1 - 2\epsilon_i y x_1^{e_{i1}}\cdots x_n^{e_{in}}).$$

If we expand this equation, we get

$$2N(F) = 2^n + \sum_{\lambda} 2^{n(\lambda)} \sum_{(x_1,\ldots,x_n,y)\in\mathbb{F}_2} g_\lambda(\epsilon_1,\ldots,\epsilon_N,x_1,\ldots,x_n,y), \qquad (1)$$

where the $g_\lambda$'s are monomials. In [12] and [8], Moreno et al. proved that

$$\min_{\lambda} \operatorname{ord}_2\Big(2^{n(\lambda)} \sum_{x_1,\ldots,x_n,y\in\mathbb{F}_2} g_\lambda(\epsilon_1,\ldots,\epsilon_N,x_1,\ldots,x_n,y)\Big) = r. \qquad (2)$$

Recall that if $a$ is a positive integer. then $\mathrm{ord}_2(a) = r \Leftrightarrow 2^r \mid a$ but $2^{r+1} \nmid a$. Note that

$$\frac{2^n + \sum_\lambda 2^{n(\lambda)} \sum_{(x_1,\ldots,x_n,y)\in\mathbb{F}_2} g_\lambda(\epsilon_1,\ldots,\epsilon_N,x_1,\ldots,x_n,y)}{2^r} \tag{3}$$

is an integer.Moreover

$$2^{n(\lambda)-r} \sum_{(x_1,\ldots,x_n,y)\in\mathbb{F}_2} g_\lambda(\epsilon_1,\ldots,\epsilon_N,x_1,\ldots,x_n,y)$$

is an integer. If we apply reduction modulo 2 to (3), yields a polynomial in the variables $\bar{\epsilon}_1,\ldots.\bar{\epsilon}_N$ ($\bar{\epsilon}_i = \epsilon \bmod 2$). Let

$$P(\bar{\epsilon}_1,\ldots,\bar{\epsilon}_N) = 2^{n-r} + \sum_\lambda 2^{n(\lambda)-r} \sum_{(x_1,\ldots,x_n,y)\in\mathbb{F}_2} g_\lambda(\bar{\epsilon}_1,\ldots,\bar{\epsilon}_N,x_1,\ldots,x_n,y).$$

Note that if $\epsilon_{i_1} x_1^{e_{i_1 1}} \cdots x_n^{e_{i_1 n}}$, $\ldots$, $\epsilon_{i_r} x_1^{e_{i_r 1}} \cdots x_n^{e_{i_r n}}$ is a minimal covering, then

$$2^{r-r} \sum_{x_1,\ldots,x_n,y\in\mathbb{F}_2} \bar{\epsilon}_{i_1} x_1^{e_{i_1 1}} \cdots x_n^{e_{i_1 n}} \times \cdots \bar{\epsilon}_{i_r} x_1^{e_{i_r 1}} \cdots x_n^{e_{i_r n}} = \bar{\epsilon}_{i_1} \cdots \bar{\epsilon}_{i_r}$$

is one of the term of $P(\bar{\epsilon}_1,\ldots,\bar{\epsilon}_N)$(Note that (2) is attained when $g_\lambda(\epsilon_1,\ldots,\epsilon_N,x_1,\ldots,x_n,y) = \epsilon_{i_1}\cdots\epsilon_{i_r}$). Hence there is a term $\bar{\epsilon}_{i_1}\cdots\bar{\epsilon}_{i_r}$ whose coefficient is 1. Therefore there is at least a nonzero coefficient in $P(\bar{\epsilon}_1,\ldots,\bar{\epsilon}_N)$. Hence $P(\bar{\epsilon}_1,\ldots,\bar{\epsilon}_N) \not\equiv 0$. Recall that the degree of each $\bar{\epsilon}_i$ is less than or equal to 1. If $P(\bar{\epsilon}_1,\ldots,\bar{\epsilon}_N) = 0$ for all $(\bar{\epsilon}_1,\ldots,\bar{\epsilon}_N) \in \mathbb{F}_2^n$, then $P$ has $2^n$ solutions. By lemma 10.2 in [3], we have that $P(\bar{\epsilon}_1,\ldots,\bar{\epsilon}_N)$ is the zero polynomial. This is a contradiction. Therefore there is a $N$-tuple $\boldsymbol{\epsilon} \neq \mathbf{0}$ of zeros and ones such that $P(\boldsymbol{\epsilon}) \neq 0$. Then $\mathrm{ord}_2(N(F_{\boldsymbol{\epsilon}})) = r - 1$ for the polynomial $F_{\boldsymbol{\epsilon}}$ that has coefficients given by $\boldsymbol{\epsilon}$. This completes the proof of Lemma 2.

*Example 7.* Let

$$\mathcal{G} = \{\, \epsilon_1 x_1 x_2 x_3 x_4 x_5 + \epsilon_2 x_1 x_2 x_3 + \sum_{i,j} \epsilon_{ij} x_i^{e_i} x_j^{e_j} \;:\; \epsilon_1, \epsilon_2, \epsilon_{ij} \in \mathbb{F}_2 \text{ and } i,j \leq 10 \,\}.$$

In this case $r = 4$. Hence there exists a polynomial $F \in \mathcal{G}$ such that $\mathrm{ord}_2(N(F)) = 3$. Taking $F(x_1,\ldots,x_{10}) = x_1\cdots x_5 + x_1x_2x_3 + x_1x_2 + x_2x_3 + \cdots + x_9x_{10} + x_1x_{10}$, we obtain $\mathrm{ord}_2(N(F)) = 3$, i.e., $N(F) = 8 \times 67$.

Let $F_i(x_1,\ldots,x_n) = \sum_{j=1}^{N_i} \epsilon_{ij} x_1^{e_{ij1}} \cdots x_n^{e_{ijn}}$ be a polynomial over $\mathbb{F}_2$ for $i = 1,\ldots,t$. Let $\mathcal{G}$ be the class of polynomial generated by monomials

$$\epsilon_{11} x_1^{e_{111}} \cdots x_n^{e_{11n}}, \ldots, \epsilon_{1N_1} x_1^{e_{1N_1 1}} \cdots x_n^{1N_1 n},$$
$$\cdots, \epsilon_{t1} x_1^{e_{t11}} \cdots x_n^{e_{t1n}} \ldots, \epsilon_{tN_t} x_1^{e_{tN_t 1}} \cdots x_n^{e_{tN_t n}}.$$

Note that these are the monomials of $F_i$ for $i = 1,\ldots,t$.

Now we state the main theorem of this section.

**Theorem 8.** *With the above notation. If $\mathcal{G}$ is the class of polynomial generated by the monomials $\epsilon_{ij}x_1^{e_{ij1}}\cdots x_n^{e_{ijn}}$ for $i=1,\ldots,t$ and $j=1,\ldots,N_i$, then there are polynomials $F_1',\ldots,F_t'$ in $\mathcal{G}$ such that $N(F_1',\ldots,F_t')$ is divisible by $2^{r-t}$ but not divisible by $2^{r+1-t}$.*

*Proof.* The proof follows taking $F(x_1,\ldots,x_n)=\sum_{i=1}^{t} y_iF_i(x_1,\ldots,x_n)$ and repeating the same argument of Lemma 2 with the new $F$.

Theorem 8 is equivalent to Theorem 8.1 in [10] for the case $\mathbb{F}_2$. For polynomials of one variable Theorem 8.1 in [10] is equivalent to the theorem of McEliece(see [7]).

## 4 Application of the Covering Method

In this section we will combine the covering method and the reduction to the ground field method to obtain improvements to the Moreno-Moreno's theorem. For details of the reduction to the ground field method see [8], [10] and [11].

In [9], combining Lemma 1 and the reduction to the ground field method we obtained the Moreno-Moreno's result for finite fields of even characteristic. We denotes by $X_i$ the variable taking value in $\mathbb{F}_{2^f}$, and $x_i$ the variable taking value in $\mathbb{F}_2$. Let $F(X_1,\ldots,X_n)$ be a polynomial over $\mathbb{F}_{2^f}$, and let $F'(x_{11},\ldots,x_{1f},\ldots,x_{nf})$ be the polynomial over $\mathbb{F}_2$ associated to $F(X_1,\ldots,X_n)$, i.e.,

$$\sum_{X_1,\ldots,X_n\in\mathbb{F}_{2^f}}(-1)^{Tr(F(X_1,\ldots,X_n))}=\sum_{x_{11},\ldots,x_{nf}\in\mathbb{F}_2}(-1)^{F'(x_{11},\ldots,x_{nf})}.$$

Recall that $F'(x_{11},\ldots,x_{nf})$ is a polynomial in $nf$-variables over $\mathbb{F}_2$.

The following theorems give improvements to the Moreno-Moreno's result for finite fields of even characteristic.

**Theorem 9.** *Let $F(X_1,\ldots,X_n)$ be a polynomial over $\mathbb{F}_{2^f}$, and let $w_2(F)=l$. If there exists a variable $X_j$ that does not appear in the monomials of 2-weight degree $l$. then*

$$2^{\mu} \text{ divides } N(F),$$

*where*

$$\mu > f(\frac{n}{l}-1).$$

*Moreover if there are $k$ variables do not appear in the monomials of $F$ of 2-weight degree $l$. then*

$$2^{\mu} \text{ divides } N(F),$$

*where*

$$\mu > f(\frac{n-k}{l}+\frac{k}{l-1}-1).$$

*Proof.* Let $\alpha_1, \ldots, \alpha_f$ be a basis of $\mathbb{F}_{2^f}$ over $\mathbb{F}_2$. Then $X_i = \sum_{j=1}^{f} x_{ij}\alpha_i$, and $Y = \sum_{j=1}^{f} y_j\alpha_j$. Now we are going to apply the reduction to the ground field to $F(X_1, \ldots, X_n)$.

$$\begin{aligned} N(F) &= \frac{1}{2^f} \sum_{\substack{X_1,\ldots,X_n \in \mathbb{F}_{2^f} \\ Y \in \mathbb{F}_{2^f}}} (-1)^{Tr(F(X_1,\ldots,X_n))} \\ &= \frac{1}{2^f} \sum_{\substack{x_{11},\ldots,x_{nf} \in \mathbb{F}_2 \\ y_1,\ldots,y_f \in \mathbb{F}_2}} (-1)^{Tr\left(\left(\sum_{j=1}^{f} y_j\alpha_j\right)\left(F\left(\sum_{j=1}^{f} x_{1j}\alpha_j, \ldots, \sum_{j=1}^{f} x_{nj}\alpha_j\right)\right)\right)} \\ &= \frac{1}{2^f} \sum_{\substack{x_{11},\ldots,x_{nf} \in \mathbb{F}_2 \\ y_1,\ldots,y_f \in \mathbb{F}_{2^f}}} (-1)^{F'(x_{11},\ldots,x_{nf})}, \end{aligned}$$

where

$$F'(x_{11}, \ldots, x_{nf}) = \sum_{i=1}^{f} y_i P_i'(x_{11}, \ldots, x_{nf}).$$

Note that $w_2(P_i') \leq l$ for $i = 1, \ldots, f$. We have that $|\mathcal{C}(F')| > \frac{nf}{l}$ by Theorem 6 since the variables $x_{j1}, \ldots, x_{jf}$ do not appear in the monomials of 2-weight degree $l$ of $F'$. Hence $2^{[nf/l]}$ divides $N(F)$.

For the second part of Theorem 9, we assume that the variables $X_1, \ldots, X_{n-k}$ appear in the monomials of $F$ of 2-weight degree $l$, and $X_{n+1-k}, \ldots, X_n$ do not appear in the monomials of $F$ of 2-weight degree $l$. Let $r$ be the cardinality of the minimal covering $\mathcal{C}(F')$ of the variables of $F'$. Then

$$r \geq \frac{n-k}{l} + \frac{kf}{l-1},$$

since the polynomial $F'$ can be written

$$F'(x_{11}, \ldots, x_{nf}) = \sum_{i} y_i \Big( P_i'(x_{11}, \ldots, x_{n-k,f}) + P_i''(x_{11}, \ldots, x_{nf}) \Big),$$

where $w_2(P_i') \leq l.$ and $w_2(P_i'') \leq l-1$ for every $i$. Note that $P_i'$ is a polynomial in the variables $x_{11}, \ldots, x_{1f}, \ldots, x_{n-k,f}$. Recall that $x_{11}, \ldots, x_{n-k,f}$ are variables over $\mathbb{F}_2$ corresponding to $X_1, \ldots, X_{n-k}$. In the above argument, we cover the first $(n-k)f$ variables with monomials of 2-weight degree $l$, and the other $k$-variables are cover with monomials of 2-weight degree $< l$. This completes the proof.

*Example 8.* Let $F(X_1, X_2, X_3) = X_1^7 + X_2^7 + X_3^9 + G(X_1, X_2, X_3)$ be a polynomial over $\mathbb{F}_{2^f}$, where $w_2(G) \leq 2$, and $G(0,0,0) =$. By Theorem 9, we have that $F$ has a nontrivial solution (2 divides $N(F)$).

*Example 9.* Let $F(X_1, X_2, X_3, X_4) = X_1^7 + X_2^7 + X_3^9 + X_4^3 + X_1X_2 + X_3X_4$ be a polynomial over $\mathbb{F}_{2^f}$. By Theorem 9, we have that $2^\mu$ divides $N(F)$, where $\mu \geq \frac{2f}{3} + \frac{2f}{2} - f = \frac{2f}{3}$. Moreno-Moreno implies that $\mu \geq \frac{f}{3}$.

**Theorem 10.** *Let $F_i(X_1, \ldots, X_n)$ be a polynomial over $\mathbb{F}_{2^f}$, and let $w_2(F_i) = l$ for $i = 1, \ldots, t$. If there exists a variable $X_j$ that does not appear in the monomials of $F_i$ of 2-weight degree $l$ for $i = 1, \ldots, t$, then*

$$2^\mu \text{ divides } N(F),$$

*where*

$$\mu > \frac{f(n - tl)}{l}.$$

*Proof.* The proof of Theorem 10 is similar to the proof of Theorem 9. Applying the reduction to the ground field method to $\sum_{i=1}^{t} Y_i F_i(X_1, \ldots, X_n)$, we obtain a polynomial in $(n+t)f$ variables over $\mathbb{F}_2$. Then applying Theorem 6, we obtain the desired result.

Note that Theorem 10 implies Theorem 9. Combining Theorem 7 of [9] and Theorem 10, we obtain the following result:

**Theorem 11.** *Let $F_1(X_1, \ldots, X_n), \ldots, \ldots, F_t(X, \ldots, X_n)$ be polynomials in $n$-variables with coefficients in $\mathbb{F}_{2^f}$, and let $l_i$ be the 2-weight degree of $F_i(X_1, \ldots, X_n)$. Let $F_i'(x_{11}, \ldots, x_{nf})$ be the polynomial over $\mathbb{F}_2$ associated to $F_i(X_1, \ldots, X_n)$ for $i = 1, \ldots, t$. Let $\mathcal{C}(F_1', \ldots, F_t')$ be a minimal set of monomials of $\sum_{i=1}^{t} y_i F_i'(x_{11}, \ldots, x_{nf})$ that covers all the variables*

$$x_{11}, \ldots, x_{1f}, \ldots, x_{n1}, \ldots, x_{nf}.$$

*If $r$ is the cardinality of $\mathcal{C}(F_1', \ldots, F_t')$, then*

$$r - tf \geq \frac{f(n - \sum_{i=1}^{t} l_i)}{\max_i l_i}.$$

*Moreover, if $l = l_1 = \cdots = l_t$, and a variable $X_j$ does not appear in the monomials of $F_i$ of 2-weight degree $l$ for $i = 1, \ldots, t$, then*

$$r - tf > \frac{f(n - \sum_{i=1}^{t} l_i)}{l}.$$

*Example 10.* Let

$$F_1(x_1, \ldots, x_{10}) = x_1x_2x_3 + G_1(x_1, \ldots, x_{10}),$$

$$F_2(x_1, \ldots, x_{10}) = x_3x_4x_5 + G_2(x_1, \ldots, x_{10})$$

and

$$F_3(x_1, \ldots, x_{10}) = x_6x_7x_8x_9 + G_3(x_1, \ldots, x_{10})$$

be polynomials over $\mathbb{F}_{2^f}$, where $w_2(G_1) < 3$, $w_2(G_2) < 3$ and $w_2(G_3) < 3$ and $G_i(0, \ldots, 0) = 0$ for $i = 1, 2, 3$. Applying Theorem 6 and the reduction to the ground field method, we obtain that the system of polynomial equation has nontrivial solution. Note that the Ax-Katz's and Moreno-Moreno's theorems do not give any information about $N(F_1, F_2, F_3)$.

## References


1. A. Adolphson and S Sperber, $p$-adic Estimates for Exponential Sums and the Theorem of Chevalley-Warning, *Ann. Sci. École Norm. Super*, **20**, no 4, pp 545-556,1987.
2. J. Ax, Zeros of Polynomials over Finite Fields, *Am. J. of Math.*, **86**, pp 255-261, 1964.
3. K. Ireland and R. Rosen, *A Classical Introduction to Modern Number Theory*, GTM, Springer, 1990.
4. N. M. Katz, On a Theorem of Ax, *Am. J. Math.*, **93**, pp 485-499, 1971.
5. G. Cohen, L. Honkala. S. Litsyn, and A. Lobstein, *Covering Radius*, **54**, North-Holland Mathematical.
6. S. Litsyn, C.J. Moreno, and O. Moreno, Divisibility Properties and New Bounds for Exponential Sums in One Variable and Several Variables, *AAECC J.*, 1994.
7. R. McEliece, On Periodic Sequences from GF(q), *Journal of Comb. Theory*, **10**, pp 80-91, 1971.
8. O. Moreno, A. Cáceres, and M. Alonso, An Improved and Simplified Binary Ax's Theorem, *Proceeding 1994 IEEE International Symposium on Information Theory*, Trondheim, Norway 1994.
9. O. Moreno and F.N. Castro, and A. Cáceres, Exponential Sums in Several Variables over Finite Fields, *Coding Theory, Cryptography and Related Areas*, pp 209-220, Springer, 2000.
10. O. Moreno, and F. N. Castro, Comparison of Thechniques on Divisibility Properties of Exponential Sums and Applications, *Finite Fields, and Applications*, pp 363-389, Springer, 2001.
11. O. Moreno, and C. J. Moreno, Improvement of Chevalley-Warning and Ax-Katz Theorems, *Am. J. Math.*, **117**, no. 1 pp 241-244, 1995.
12. O. Moreno, and C. J. Moreno, The MacWilliams-Sloane Conjecture on the Tightness of Carlitz-Uchiyama Bound and the Weights of Duals of BCH Codes, *IEEE Trasactions on Information Theory*, **40**, no. 6, 1994.

# Value Sets of Polynomials over Finite Fields

Pinaki Das and Gary L. Mullen

Department of Mathematics
Pennsylvania State University
University Park, PA 16802
E-Mail: das@math.psu.edu and mullen@math.psu.edu

**Abstract.** We provide a lower bound for the cardinality of the value set of a polynomial over a finite field which improves upon several earlier bounds.

## 1 Introduction

Let $f(x)$ be a polynomial over $F_q$, the finite field of order $q$ and characteristic $p$. Numerous papers have been written concerning the cardinality $|V_f|$ of the *value set* $V_f = \{f(a) : a \in F_q\}$ of $f(x)$. If $f(x)$ has degree $n$, since a polynomial cannot have more than $n$ roots in any field, it is easy to see that

$$\lfloor \frac{q-1}{n} \rfloor + 1 \leq |V_f| \leq q.$$

Polynomials achieving the lower bound are said to be *minimal value set polynomials* while those with value sets achieving the upper bound $q$ are known as *permutation polynomials*, see Chapter 7 of Lidl and Niederreiter [2].

In general it is very difficult to determine the cardinality $|V_f|$ of the value set $V_f$ of a polynomial $f(x)$ over the finite field $F_q$. This has only been done for several classes of polynomials, including the power polynomials $x^n$ whose value set has cardinality $|V_{x^n}| = (q-1)/d + 1$, where $d = (q-1, n)$, Dickson polynomials $D_n(x, a)$ see [1], linearized polynomials all of whose nonzero terms involve exponents which are powers of the characteristic $p$ of $F_q$, and a few other small classes, see for example [3] and [4].

## 2 Main Result

In Theorem 2.1 of [4] it was shown that if $u_p(f)$ is the smallest positive integer $i$ such that $\sum_{a \in F_q} (f(a))^i \neq 0$, then $|V_f| \geq u_p(f) + 1$.

**Notation:** Let $f(x)$ be a polynomial of degree $n$ over $F_q$. Since $f(x)$ and $f(x) + a$ both have values sets of the same cardinality, we may assume that $f(0) = 0$. By the Lagrange Interpolation Formula [2] page 369, we can assume that $n \leq q-1$. Thus for each $i = 1, \ldots, q-1$, we may write $(f(x))^i = \sum_{j=0}^{q-1} a_{ij} x^j \bmod (x^q - x)$.

Let $A_f$ be the matrix $A_f = (a_{ij}^{q-1})$, for $1 \leq i, j \leq q-1$ so that the entries of $A_f$ consist only of 0 and 1. In particular, the $(i,j)$ entry of $A_f$ is 1 if the coefficient of $x^j$ in $(f(x))^i$ mod $(x^q - x)$ is nonzero. If $f(x)$ is not the zero polynomial, the matrix $A_f$ has at least one nonzero column. If the $j$-th column of $A_f$ consists entirely of 0's or entirely of 1's, set $l_j = 0$. Otherwise for a nonzero $j$-th column, arrange the entries in a circle and define $l_j$ to be the maximum number of consecutive zeros appearing in this circular arrangement. Let $L_f$ be the maximum of the values of $l_j$, where the maximum is taken over all of the $q-1$ columns of the matrix $A_f$.

For example, if $f(x) = x^3 + 6x$ over $F_7$, then

$$A_f = \begin{pmatrix} 1\,0\,1\,0\,0\,0 \\ 0\,1\,0\,1\,0\,1 \\ 1\,0\,0\,0\,1\,0 \\ 0\,1\,0\,1\,0\,1 \\ 0\,0\,1\,0\,1\,0 \\ 0\,1\,0\,1\,0\,1 \end{pmatrix},$$

and so $l_1 = l_3 = l_5 = 3, l_2 = l_4 = l_6 = 1$. Note $l_5 = 3$ because of our circular arrangement. In the fifth column of $A_f$ the zero in the last row and the two zeros in the first two rows are counted as consecutive zeros. Hence $L_f = 3$, and as we will show in Theorem 1 below, $|V_f| \geq L_f + 2 = 5$. For this particular example, Theorem 2.1 of [4] shows that $|V_f| \geq 3$. In fact $|V_f| = 5$.

**Theorem 1:** With notation as above, we have:

$$|V_f| \geq L_f + 2.$$

**Proof:** First observe that

$$F_q[x]/(x^q - x) = \bigoplus_{\alpha_k \in F_q} F_q[x]/(x - \alpha_k).$$

This follows from the Chinese Remainder Theorem since over $F_q$, $x^q - x = \prod_{\alpha_k \in F_q}(x - \alpha_k)$. It follows that for each $i$ with $1 \leq i \leq q-1$, $(f(x))^i$ mod $(x^q - x)$ is the unique polynomial of degree at most $q-1$ which represents the function $\phi_i : F_q \to F_q$ given by $\phi_i(\alpha_k) = (f(\alpha_k))^i$.

Assume that $l = l_j \neq 0$ for the $j$-th column of the matrix $A_f$. Then there are three possibilities:

1) Some $a_{kj} = 1$ and $a_{k-l,j} = \cdots = a_{k-1,j} = 0$,

or

2) some $a_{kj} = 1$ and $a_{k+1,j} = \cdots = a_{k+l,j} = 0$,

or

3) $a_{k+1,j} = \cdots = a_{q-1,j} = 0$ for $r$ consecutive zeros, and $a_{1,j} = \cdots = a_{s,j} = 0$ for $s$ consecutive zeros with $r + s = l$, and $a_{s+1,j} \neq 0$.

Without loss of generality, we first assume that $a_{k-l,j} = \cdots = a_{k-1,j} = 0$ and $a_{kj} = 1$.

Then $(f(x))^k \bmod (x^q - x)$ has a nonzero coefficient of $x^j$ and the coefficient of $x^j$ is zero in $(f(x))^i \bmod (x^q - x)$ for the previous $l$ values, $i = k - l, \ldots, k - 1$.

If the set of nonzero elements of $F_q$ is denoted by $\{\alpha_1, \ldots, \alpha_{q-1}\}$, the polynomial $(f(x))^k \bmod (x^q - x)$ is the interpolating polynomial for the set of values $\{(f(\alpha_1))^k, \ldots, (f(\alpha_{q-1}))^k\}$. Hence we have

$$\begin{aligned}
&(f(x))^k \bmod (x^q - x)\\
&= \textstyle\sum_{\alpha_h \in F_q} (f(\alpha_h))^k (1 - (x - \alpha_h)^{q-1})\\
&= \textstyle\sum_{\alpha_h \neq 0 \in F_q} (f(\alpha_h))^k (1 - (x - \alpha_h)^{q-1})\\
&= \textstyle\sum_{\alpha_h \neq 0 \in F_q} (f(\alpha_h))^k \Big(1 - \sum_{n=0}^{q-1} \binom{q-1}{n} x^n (-\alpha_h)^{q-1-n}\Big).
\end{aligned}$$

Since $a_{kj} = 1$ we have

$$\sum_{\alpha_h \neq 0 \in F_q} (f(\alpha_h))^k \alpha_h^{q-1-j} \neq 0. \tag{1}$$

Moreover, $a_{k-l,j} = \cdots = a_{k-1,j} = 0$ implies that

$$\sum_{\alpha_h \neq 0 \in F_q} (f(\alpha_h))^s \alpha_h^{q-1-j} = 0, s = k - l, \ldots, k - 1. \tag{2}$$

Assume the number of distinct nonzero elements in the image of $f(x)$ is $m \leq l$. Then the image of $f(x)$ is $f(F_q) = \{0, \beta_1, \ldots, \beta_m\}$, where $\beta_i \neq \beta_j$. Then (1) implies that $c_1\beta_1^k + \cdots + c_m\beta_m^k \neq 0$, where $c_i \in F_q$, and not all $c_i = 0$.

From (2) we have the system of equations

$$\begin{aligned}
c_1\beta_1^{k-m} + \cdots + c_m\beta_m^{k-m} &= 0\\
\vdots \qquad\qquad \vdots \qquad\quad &\vdots\\
c_1\beta_1^{k-1} + \cdots + c_m\beta_m^{k-1} &= 0
\end{aligned} \tag{3}$$

Since by assumption $m \leq l$, and $c_1, \ldots, c_m$ are not all 0, we have

$$\det \begin{pmatrix} \beta_1^{k-m} & \ldots & \beta_m^{k-m} \\ \vdots & & \vdots \\ \beta_1^{k-1} & \ldots & \beta_m^{k-1} \end{pmatrix} = 0.$$

Hence

$$(\beta_1 \ldots \beta_m)^{(k-m)m} \det \begin{pmatrix} 1 & \ldots & 1 \\ \beta_1 & \ldots & \beta_m \\ \vdots & & \vdots \\ \beta_1^{m-1} & \ldots & \beta_m^{m-1} \end{pmatrix} = 0\,.$$

But this is impossible since the $\beta_i$'s are nonzero and distinct and the Vandermonde determinant can be evaluated as $\prod_{1 \le i < j \le m} (\beta_i - \beta_j)$.

Hence the number of distinct nonzero elements in the image of $f(x)$ must be at least $l + 1$. Since 0 lies in the image of $f(x)$ we have $|V_f| \ge l + 2$. Similarly this is true for the maximum value $L_f$ of $l$.

A similar argument works in the case when some $a_{kj} = 1$ and $a_{k+1,j} = \cdots = a_{k+l,j} = 0$.

In the case when we have $r$ consecutive zeros in the last $r$ rows and $s$ consecutive zeros in the top $s$ rows in the $j$-th column of the matrix $A_f$, if $m > r$, then we have a system like (3), except the exponents on the $\beta_i$ run through the values $k + 1, k + 2, \ldots, q - 1, 1, \ldots, k + m - q + 1$. In this case also the argument is the same as above, and so the proof is complete.

**Remark 1)** We note that $|V_f| \ge l_j + 2$ for each value $j = 1, \ldots, q - 1$. In particular, this holds for the last column when $j = q - 1$ and hence as a corollary, we have Theorem 2.1 from [4].

**Remark 2)** We also note that our result extends the Hermite/Dickson criterion for permutation polynomials, see [2, Thm. 7.4, page 349]. This is due to the fact that the Hermite/Dickson criterion is essentially equivalent to the first $q - 2$ consecutive elements of the last column of the matrix $A_f$ being 0, with the last element of that column equal to 1. In particular, $f(x)$ is a permutation polynomial if and only if $L_f = q - 2$.

**Example 1)** We now provide a larger example which shows that our method yields a better bound than the Wan, Shiue, and Chen bound from [4]. Consider the polynomial $f(x) = x^7 + x$ over $F_{19}$. After forming the matrix $A_f$ we find that the first six entries of the last column of $A_f$ contain 6 consecutive zeros, with a 1 in row 7 so that the bound from [4] is $|V_f| \ge 6 + 1 = 7$, while for $j = 3$, we have a string of 11 consecutive zeros in the third column of $A_f$ and so from Theorem 1, we have $|V_f| \ge 11 + 2 = 13$. In fact, our bound is sharp since $|V_f| = 13$.

**Example 2)** When considering polynomials $f(x)$ over extension fields $F_q$ of prime power order $q = p^e$ with $e > 1$, there are examples where there is no $i$ such that $\sum_{a \in F_q} (f(a))^i \ne 0$, and so in these cases the bound $|V_f| \ge i + 1$, from [4] cannot be applied. However in these cases, our bound from Theorem

1 still applies. For example, if $f(x) = x^2 + x$ over the field $F_8$, then

$$A_f = \begin{pmatrix} 1\,1\,0\,0\,0\,0\,0 \\ 0\,1\,0\,1\,0\,0\,0 \\ 0\,0\,1\,1\,1\,1\,0 \\ 1\,0\,0\,1\,0\,0\,0 \\ 0\,1\,1\,1\,1\,0\,0 \\ 1\,0\,1\,0\,1\,1\,0 \\ 1\,1\,1\,1\,1\,1\,0 \end{pmatrix},$$

and hence $L_f = 2$, and from Theorem 1, $|V_f| \geq 4$. In this case we actually have $|V_f| = 4$.

**Example 3)** The following example shows that our method, while providing an improvement of Theorem 2.1 of [4], is still not best possible, i.e. that the condition is not necessary and sufficient for a polynomial $f(x)$ to have a value set of cardinality $L_f + 2$. Consider the polynomial $f(x) = x^9 + x^5 + 7x^2 + x$ over the field $F_{19}$. In this case $L_f = 3$ and so Theorem 1 predicts a value set of cardinality $|V_f| \geq 5$, while in reality, we have $|V_f| = 10$.

**Example 4)** We now include a table which shows that our Theorem 1 improves the values cited in Table 1 of [4] for many values of $a$.

Table 1: $f(x) = x^7 + ax$ over $F_{19}$

| $a$ | $\lfloor \frac{q-1}{deg(f)} \rfloor + 1$ | $u_p(f) + 1$ | $L_f + 2$ | $\lvert V_f \rvert$ |
|---|---|---|---|---|
| 1 | 3 | 7 | 13 | 13 |
| 2 | 3 | 7 | 13 | 13 |
| 3 | 3 | 7 | 13 | 13 |
| 4 | 3 | 7 | 7 | 7 |
| 6 | 3 | 7 | 7 | 7 |
| 7 | 3 | 7 | 13 | 13 |
| 8 | 3 | 7 | 13 | 13 |
| 9 | 3 | 7 | 7 | 7 |
| 10 | 3 | 7 | 13 | 13 |
| 11 | 3 | 7 | 13 | 13 |
| 12 | 3 | 7 | 13 | 13 |
| 13 | 3 | 7 | 13 | 13 |
| 14 | 3 | 7 | 13 | 13 |
| 15 | 3 | 7 | 13 | 13 |
| 18 | 3 | 7 | 13 | 13 |

We note that our Theorem 1 improves upon 12 of the values in Table 1 of [4]. We also point out that for $a = 6, 8, 9$, the cardinalities of the value sets are given in [4] as $13, 7, 13$, when they should have been given as $7, 13, 7$ as indicated above.

## References

1. W.-S. Chou, J. Gomez-Calderon, and G.L. Mullen, *Value sets of Dickson polynomials over finite fields*, J. Number Theory. 30(1988), 334–344.
2. R. Lidl and H. Niederreiter, *Finite Fields*, Encyclo. Math. Appl., Vol. 20, Cambridge Univ. Press, 1997.
3. D. Wan and R. Lidl, *Permutation polynomials of the form $x^r f(x^{(q-1)/d})$ and their group structure*, Monatsh. Math. 112(1991), 149–163.
4. D. Wan, P.J.-S. Shiue, and C.-S. Chen, *Value sets of polynomials over finite fields*, Proc. Amer. Math. Soc. 119(1993), 711–717.

# Bounds for Completely Decomposable Jacobians

Iwan Duursma[1] and Jean-Yves Enjalbert[2]

[1] University of Illinois at U-C, Urbana IL 61801, USA
[2] Université de Limoges, F-87060 Limoges Cedex, France

**Abstract.** A curve over the field of two elements with completely decomposable Jacobian is shown to have at most six rational points and genus at most 26. The bounds are sharp. The previous upper bound for the genus was 145. We also show that a curve over the field of $q$ elements with more than $q^{m/2}+1$ rational points has at least one Frobenius angle in the open interval $(\pi/m, 3\pi/m)$. The proofs make use of the explicit formula method.

## 1 Introduction

The Jacobian of an (absolutely irreducible, projective, non-singular) algebraic curve is said to be completely decomposable if it is isogenous over the base field to a product of elliptic curves. Many examples are known of curves with completely decomposable Jacobian [ES93], both in characteristic zero and in finite characteristic. For a curve over a finite field $F_q$, the genus of a curve with completely decomposable Jacobian is bounded [TV97], [Ser97]. For $q = 2$, Serre [Ser97] gives a first order estimate $g < 146$. We use the explicit formula method developed in [Ser83] to obtain $g \leq 26$. The upper bound is sharp and is attained by the modular curve $X(11)$ for which Hecke showed that the Jacobian decomposes as $E_1^5 \times E_2^{10} \times E_3^{11}$ [Lig77].

For an algebraic curve (absolutely irreducible, projective, non-singular) of genus $g$ over a finite field of $q$ elements, the Hasse-Weil bound gives that the number of rational points $N$ does not exceed $q + 1 + 2g\sqrt{q}$. For the explicit formula method, the number of rational points is expressed in terms of the Frobenius eigenvalues as

$$N = q + 1 - \sum_{j=1}^{g} (\alpha_j + \bar{\alpha}_j).$$

By Weil's theorem, we may write $\alpha_j = \sqrt{q}e^{i\theta_j}$, for elements $\theta_j$ in $[0, \pi]$ for all $j$. The $\theta_j$ are called the *Frobenius angles*. Over an extension field of size $q^m$, the number of rational points $N_m$ is given by

$$N_m = q^m + 1 - \sum_{j=1}^{g} (\alpha_j^m + \bar{\alpha}_j^m) = q^m + 1 - r^m \sum_{j=1}^{g} 2\cos m\theta_j,$$

where $r = \sqrt{q}$. For curves of large genus, the distribution of the Frobenius angles is restricted by the constraints $N_{dm} \geq N_m$, for all $m, d$. This allows one to obtain upper bounds of the form $N \leq ag + b$ for the number of rational points that are better than the Hasse-Weil bound when the genus is large [Iha81], [Ser83]. Asymptotically, the Drinfeld-Vladuts bound gives $\limsup_{g\to\infty} N/g \leq \sqrt{q} - 1$ [VD83], where the limit is over an infinite family of curves of increasing genus. In Section 2.1, we recall the main steps of the explicit formula method.

Tsfasman-Vladuts [TV97] and Serre [Ser97] study the distribution of the Frobenius angles for families of curves of increasing genus. It is easy to see that any infinite family of curves of increasing genus contains a subfamily for which $N_m/g$ approaches a limit, for each $m$, when the genus increases. Such subfamilies are called asymptotically exact in [TV97]. For curves in an asymptotically exact family, the distribution of the Frobenius angles approaches a limit distribution that is given by a continuous measure on $[0, \pi]$. In particular, the Frobenius angles in an asymptotically exact family are dense in $[0, \pi]$. This shows that any family of curves for which the Frobenius angles are not dense in $[0, \pi]$ is finite. We consider the following problem.

**(Problem 1)** Given a discrete subset $\Theta$ of $[0, \pi]$, maximize $N$ and $g$ for a curve over $F_q$ with all Frobenius angles in $\Theta$.

The elliptic curves over the field of two elements have Frobenius angle $\theta$ such that $2\sqrt{2}\cos\theta \in \{-2, -1, 0, 1, 2\}$. The corresponding Frobenius eigenvalues are of degree at most two. As a special case of the previous problem we have

**(Problem 2)** Maximize $N$ and $g$ for a curve over $F_q$ with all Frobenius eigenvalues of bounded degree at most $d$.

The case $d = 2$ corresponds to curves with completely decomposable Jacobian. In Section 2.3 and Section 2.4, respectively, we show that a curve over $F_2$ with completely decomposable Jacobian has $N \leq 6$ and $g \leq 26$, respectively. Similarly, the family of curves with no Frobenius angle in a given interval is finite. And we can ask for the largest number of rational points or the largest genus for curves in the family.

**(Problem 3)** Given a (small) subset $I$ of $[0, \pi]$, maximize $N$ and $g$ for a curve over $F_q$ with all Frobenius angles outside $I$.

In Section 3, we prove that any curve over $F_q$ with $N > q^{m/2} + 1$ has a Frobenius angle in the open interval $(\pi/m, 3\pi/m)$. We formulate one other problem along the same lines. It will not be considered in this paper however.

**(Problem 4)** Given $\delta$, maximize $N$ and $g$ for a curve over $F_q$ such that $[0, \pi] \not\subset \cup_j(\theta_j - \delta, \theta_j + \delta)$.

## 2 The Explicit Formula Method

We first recall the explicit formula method and its use in obtaining general upper bounds for the number of rational points on a curve [Ser83], [Han95]. Then we present three variations of the method that yield better bounds for curves whose Frobenius angles are restricted to a subset $\Theta$ of $[0, \pi]$. In particular, curves that exceed one of the latter bounds, necessarily have at least one Frobenius angle outside $\Theta$.

### 2.1 General Upper Bounds for the Number of Rational Points

For an algebraic curve $X$ of genus $g$ over the finite field $F_q$ of $q$ elements, let the Frobenius angles be $\theta_1, \theta_2, \ldots, \theta_g$. So that the number of rational points $N_n$ over $F_{q^n}$ satisfies

$$N_n = q^n + 1 - q^{n/2} \sum_{j=1}^{g} 2 \cos n\theta_j.$$

With $r = \sqrt{q}$, we rewrite the equation as

$$N_1 r^{-n} + (N_n - N_1) r^{-n} = \quad r^n + r^{-n} - \sum_{j=1}^{g} 2 \cos n\theta_j. \tag{1}$$

Let $f$ be an auxiliary cosine polynomial with real coefficients $u_n$,

$$f(\theta) = u_0 + \sum_{n \geq 1} u_n \cos n\theta. \tag{2}$$

Define

$$\psi(x) = \sum_{n \geq 1} u_n x^n. \tag{3}$$

The equations (1) scaled by $u_n$, for $n = 1, 2, \ldots$, add up to

$$N_1 \psi(r^{-1}) + \sum_{n \geq 2} u_n (N_n - N_1) r^{-n} =$$

$$= 2u_0 g + \psi(r) + \psi(r^{-1}) - 2 \sum_{j=1}^{g} f(\theta_j). \tag{4}$$

The equation (4) leads to upper bounds for the number of points. As in [Ser83], choose $\{u_n\}$ such that $u_0 = 1$, and

**(a)** $u_n \geq 0, \forall n \geq 1$
**(b)** $f(\theta) \geq 0$, for all $\theta \in [0, \pi]$.

Then Equation (4) yields

$$N\psi(r^{-1}) \leq 2g + \psi(r^{-1}) + \psi(r).$$

As an example, the choice

$$\begin{aligned} f(\theta) &= \cos^2\theta(1 - \cos\theta/\cos(\frac{5\pi}{6}))^2 \\ &= 1 + \sqrt{3}\cos\theta + \frac{7}{6}\cos 2\theta + \frac{\sqrt{3}}{3}\cos 3\theta + \frac{1}{6}\cos 4\theta \end{aligned}$$

gives, for $q = 3$, the upper bound

$$N \leq \frac{54}{41}(g - 15) + 28 < 1.317\, g + 8.244.$$

This is better than the Hasse-Weil bound $N \leq 2\sqrt{3}\, g + 4$ for all $g \geq 2$. A curve attains the upper bound above only if $N_1 = N_2 = N_3 = N_4$ and if all its Frobenius angles are among $\{\pi/2, 5\pi/6\}$. The unique such curve is the Deligne-Lusztig curve associated to ${}^2G_2(3)$ [HP93]. The curve is of genus $g = 15$ and has $N = 28$. Its zeta function $Z(T) = P(T)/(1 - T)(1 - 3T)$ has numerator $P(T) = (1 + 3T^2)^7(1 + 3T + 3T^2)^8$.

### 2.2 Restricted Upper Bounds for the Number of Rational Points ($u_0 = 1$)

The upper bound in the previous subsection generalizes as follows. Choose $\{u_n\}$ in Equation (4) such that

**(a)** $u_0 = 1$ and $u_n \geq 0, \forall n \geq 2$.
**(b)** $f(\theta) \geq 0$, for all $\theta \in \Theta \subset [0, \pi]$.

Then, for a curve that has all its Frobenius angles contained in $\Theta$,

$$N\psi(r^{-1}) \leq 2g + \psi(r^{-1}) + \psi(r).$$

The converse yields that a curve with

$$N\psi(r^{-1}) > 2g + \psi(r^{-1}) + \psi(r).$$

has a Frobenius angle outside $\Theta$. For $0 < \alpha < \beta < \pi$, let

$$\begin{aligned} f_2(\theta) &= (\cos\theta - \cos\alpha)\,(\cos\theta - \cos\beta)\,, \\ &= \frac{1}{2} + \cos\alpha\cos\beta - (\cos\alpha + \cos\beta)\cos\theta + \frac{1}{2}\cos 2\theta. \end{aligned}$$

Then $f_2(\theta)$ is non-negative on $\Theta = [0, \pi]\backslash(\alpha, \beta)$. For $q = 2$, and for $\alpha = \pi/3$ and $\beta = 3\pi/4$, we obtain

$$N > \frac{8 - 2\sqrt{2}}{7}(g - 1) + 5 \quad \Rightarrow \quad \exists\theta_j \in (\frac{\pi}{3}, \frac{3\pi}{4}).$$

The inequality on the left applies in the range $2 \leq g \leq 38$. In that range the inequality holds for a curve that meets the Oesterlé upper bound for the number of points. For another example, let

$$\begin{aligned} f(\theta) &= (1 + \sqrt{2}\cos\theta)(1 - 2\sqrt{2}\cos\theta)^2, \\ &= 1 + 3\sqrt{2}\cos\theta + 2\sqrt{2}\cos 3\theta. \end{aligned}$$

We obtain, for a curve over $F_2$,

$$N > \frac{1}{2}(g-1) + 5 \quad \Rightarrow \quad \exists \theta_j \in (\frac{3\pi}{4}, \pi].$$

### 2.3 Uniform Upper Bounds for the Number of Rational Points ($u_0 = 0$)

By choosing $u_0 = 0$, we obtain upper bounds for the number of rational points that are independent of the genus $g$. Choose $\{u_n\}$ in Equation (4) such that

**(a)** $u_0 = 0$ and $u_n \geq 0, \forall n \geq 2$.
**(b)** $f(\theta) \geq 0$, for all $\theta \in \Theta \subset [0, \pi]$.

Then the number $N$ of rational points on a curve with all Frobenius angles contained in $\Theta$ satisfies

$$N\psi(r^{-1}) \leq \psi(r^{-1}) + \psi(r).$$

If, moreover, the coefficients $u_n$ have the following symmetry property, for some positive integer $m$ with $m > \deg(\psi)$,

**(c)** $u_n = u_{m-n}$, for $n = 0, 1, \ldots, m$,

then the upper bound becomes

$$N \leq 1 + \frac{\sum_{n=0}^{m} u_n r^n}{\sum_{n=0}^{m} u_{m-n} r^{n-m}} = r^m + 1.$$

The function

$$\begin{aligned} f(\theta) &= \frac{\sqrt{2}}{5}\cos\theta(1 - 2\cos^2\theta)(1 - 8\cos^2\theta) \\ &= \frac{7}{10}\sqrt{2}\cos\theta + \frac{1}{2}\sqrt{2}\cos 3\theta + \frac{1}{5}\sqrt{2}\cos 5\theta \end{aligned}$$

cancels at the Frobenius angles of the five different elliptic curves over $F_2$. It leads to the bound $N \leq 6$ for any curve $X$ over $F_2$ with completely decomposable Jacobian. The bound is tight only when $N_1 = N_3 = N_5$. The

smallest feasible zeta function is of genus 3 with uniquely determined zeta polynomial $P(T) = (1 + 2T + 2T^2)^2(1 - T + 2T^2)$. It is realized by the curve

$$y^2 + y = \frac{x^2 + x}{(x^2 + x + 1)^3}.$$

We give two examples that use Condition (c). The choice $f(\theta) = \cos\theta$ yields that a curve with $N > r^2 + 1$ has a Frobenius angle in $(\pi/2, 3\pi/2)$ (indeed the Frobenius trace can only be negative if at least one Frobenius angle has $\cos\theta < 0$). The choice $f(\theta) = \cos\theta + \cos 2\theta$ yields that a curve with $N > r^3 + 1$ has a Frobenius angle in $(\pi/3, \pi)$. In both cases, the bound on $N$ is sharp. The projective line with $N = r^2 + 1$ has no Frobenius angle in $(\pi/2, 3\pi/2)$, and the Hermitian curve (see [RS94]) over $F_{r^2}$ with $N = r^3 + 1$ has no Frobenius angle in $(\pi/3, \pi)$. The latter example confirms that the Hasse-Weil bound is not sharp for curves with $N > r^3 + 1$. In Section 3, we show more generally that a curve with $N > r^m + 1$ has a Frobenius angle in $(\pi/m, 3\pi/m)$.

### 2.4 Uniform Upper Bounds for the Genus ($u_0 = -1$)

By choosing $u_0 = -1$, we obtain upper bounds for the genus $g$. Choose $\{u_n\}$ in Equation (4) such that

**(a)** $u_0 = -1$ and $u_n \geq 0, \forall n \geq 2$.
**(b)** $f(\theta) \geq 0$, for all $\theta \in \Theta \subset [0, \pi]$.

Then the genus of a curve with all Frobenius angles contained in $\Theta$ satisfies

$$N\psi(r^{-1}) + 2g \leq \psi(r) + \psi(r^{-1}).$$

If, moreover, the coefficients $u_n$ satisfy

**(d)** $\psi(r^{-1}) = 0$,

then the upper bound becomes

$$2g \leq \psi(r).$$

The function

$$f(\theta) = -1 - \frac{4}{3}\cos\theta + \frac{7}{9}\cos 2\theta + \frac{26}{9}\cos 3\theta + \frac{16}{9}\cos 4\theta$$

is of minimal degree such that it cancels at the Frobenius angles of the three elliptic curves over $F_4$ that are defined over $F_2$ and such that Condition (d) holds. It leads to the bound $2g \leq 52$ for any curve $X$ over $F_2$ with completely decomposable Jacobian. A previous estimate showed that $g \leq 145$ [Ser97]. The bound is tight only when $N_1 = N_2 = N_3 = N_4$ for the base field $F_4$. It is attained by the modular curve $X(11)$, which has $g = 26$, $N = 55$ over $F_4$, and zeta polynomial $P(T) = (1 + 4T + 4T^2)^5(1 + 3T + 4T^2)^{10}(1 + 4T^2)^{11}$.

## 3 An Asymptotic Example

Let $m \geq 4$ and let $\alpha = \pi/m$. Conditions (a)-(c) in Section 2.3 hold with $\Theta = [0, \pi] \backslash (\pi/m, 3\pi/m)$ for coefficients $\{u_n\}$ that are defined by

$$f(\theta) = \frac{1 + \cos m\theta}{4(\cos\theta - \cos\alpha)(\cos\theta - \cos 3\alpha)} = \sum_{n=2}^{m-2} u_n \cos n\theta. \tag{5}$$

So that

$$u_n = \frac{\sin(n-1)\alpha \sin n\alpha \sin(n+1)\alpha}{\sin\alpha \sin 2\alpha \sin 3\alpha}, \quad n = 0, 1, \ldots, m. \tag{6}$$

Thus, a curve with number of rational points $N > r^m + 1$, for $m \geq 4$, has at least one Frobenius angle in the open interval $(\pi/m, 3\pi/m)$. For $f(\theta)$ we may write

$$f(\theta) = 2^{m-3} \prod_{k=2}^{m-1} \big(\cos\theta - \cos(2k+1)\alpha\big),$$

which justifies writing the right hand side of (5) as a cosine polynomial. To see that the coefficients of the cosine polynomial are those given by (6), we use a generating function for gaussian polynomials [And98]

$$\frac{1}{(1-T)(1-yT)(1-y^2T)(1-y^3T)} = \sum_{i \geq 0} \begin{bmatrix} i+3 \\ 3 \end{bmatrix} T^i,$$

where

$$\begin{bmatrix} i+3 \\ 3 \end{bmatrix} = \frac{(y^{i+3}-1)(y^{i+2}-1)(y^{i+1}-1)}{(y^3-1)(y^2-1)(y-1)}.$$

For $y$ with $y^m = 1$, the right hand side is periodic and, for $n = i + 2$,

$$\frac{T^2(1-T^m)}{(1-T)(1-yT)(1-y^2T)(1-y^3T)} = \sum_{n=2}^{m-2} \frac{(y^{n+1}-1)(y^n-1)(y^{n-1}-1)}{(y^3-1)(y^2-1)(y-1)} T^n.$$

Let $x = e^{i\alpha}$, so that $x^m = -1$. With $y = x^2$ and $t = x^3T$, we obtain

$$\frac{(1+t^m)}{(t+t^{-1}-2\cos\alpha)(t+t^{-1}-2\cos 3\alpha)} = \sum_{n=2}^{m-2} \frac{\sin(n-1)\alpha \sin n\alpha \sin(n+1)\alpha}{\sin\alpha \sin 2\alpha \sin 3\alpha} t^n.$$

Now sum the two equations with $t = e^{i\theta}$ and $t = e^{-i\theta}$, respectively, and divide by 2.

The cases $m = 2$ and $m = 3$ were considered in Section 2.3, so that the claim extends to all $m \geq 2$. For $m = 4$ and $m = 6$ the bounds are sharp, as can be seen by considering curves of Suzuky type or Ree type, respectively. The Suzuki curve over $F_8$ has $N = 65$ but has no Frobenius angle in $(\pi/4, 3\pi/4)$. The Ree curve over $F_3$ has $N = 28$ but has no Frobenius angle in $(\pi/6, \pi/2)$.

## 4 Conclusion

Results by Tsfasman-Vladuts and Serre led us to consider Problems (1)-(4) in the Introduction. For Problems (1)-(3) we have given methods that yield partial results. One result is a sharp upper bound for the number of points ($N \leq 6$) or the genus ($g \leq 26$) for a curve over $F_2$ with completely decomposable Jacobian. We also showed that a curve over $F_q$ with $N > q^{m/2} + 1$ has at least one Frobenius angle in the interval $(\pi/m, 3\pi/m)$. No results were obtained towards Problem (4).

## References


[And98] George E. Andrews. *The theory of partitions.* Cambridge University Press, Cambridge, 1998. Reprint of the 1976 original.

[ES93] Torsten Ekedahl and Jean-Pierre Serre. Exemples de courbes algébriques à jacobienne complètement décomposable. *C. R. Acad. Sci. Paris Sér. I Math.*, 317(5):509–513, 1993.

[Han95] Søren Have Hansen. *Rational points on curves over finite fields.* Aarhus Universitet Matematisk Institut, Aarhus, 1995.

[HP93] Johan P. Hansen and Jens Peter Pedersen. Automorphism groups of Ree type, Deligne-Lusztig curves and function fields. *J. Reine Angew. Math.*, 440:99–109, 1993.

[Iha81] Yasutaka Ihara. Some remarks on the number of rational points of algebraic curves over finite fields. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 28(3):721–724 (1982), 1981.

[Lig77] Gérard Ligozat. Courbes modulaires de niveau 11. In *Modular functions of one variable, V (Proc. Second Internat. Conf., Univ. Bonn, Bonn, 1976)*, pages 149–237. Lecture Notes in Math., Vol. 601. Springer, Berlin, 1977.

[RS94] Hans-Georg Rück and Henning Stichtenoth. A characterization of Hermitian function fields over finite fields. *J. Reine Angew. Math.*, 457:185–188, 1994.

[Ser83] Jean-Pierre Serre. Sur le nombre des points rationnels d'une courbe algébrique sur un corps fini. *C. R. Acad. Sci. Paris Sér. I Math.*, 296(9):397–402, 1983.

[Ser97] Jean-Pierre Serre. Répartition asymptotique des valeurs propres de l'opérateur de Hecke $T_p$. *J. Amer. Math. Soc.*, 10(1):75–102, 1997.

[Tsf92] Michael A. Tsfasman. Some remarks on the asymptotic number of points. In *Coding theory and algebraic geometry (Luminy, 1991)*, pages 178–192. Springer, Berlin, 1992.

[TV97] M. A. Tsfasman and S. G. Vlăduţ. Asymptotic properties of zeta-functions. *J. Math. Sci. (New York)*, 84(5):1445–1467, 1997. Algebraic geometry, 7.

[VD83] S. G. Vlăduţ and V. G. Drinfelʹd. The number of points of an algebraic curve. *Funktsional. Anal. i Prilozhen.*, 17(1):68–69, 1983.

# Twin Irreducible Polynomials over Finite Fields

Gove W. Effinger[1], Kenneth H. Hicks[2], and Gary L. Mullen[3]

[1] Department of Mathematics & Computer Science, Skidmore College, Saratoga Springs, NY 12866, *email: effinger@skidmore.edu*,
[2] Department of Physics and Astronomy, Ohio University, Athens OH 45701, *email: hicks@ohio.edu*
[3] Department of Mathematics, The Pennsylvania State University, University Park PA 16802, *email: mullen@math.psu.edu*

**Abstract.** We discuss a finite field polynomial analogue of the twin primes conjecture.

## 1 Introduction

One of the most famous conjectures of classical number theory is the Twin Primes Conjecture, which asserts that there are infinitely many pairs $(p, p+2)$ of primes. For example, Hardy and Wright [3] conjecture a specific formula for the number of such pairs below an integer $x$ as $x$ goes to infinity. In this paper we discuss an analogue to the Twin Primes Conjecture in the domain of monic polynomials (in one variable) over the finite field $\mathbf{F}_q$. We first (in Section 2) define the notion of twin irreducible polynomials and consider the case (which is *not* the true analogue of the Twin Primes Conjecture) of a fixed degree $r$ but growing field order $q$, showing that in this case there are indeed arbitrarily many pairs of twins (indeed also triples, quadruples, and so on) as $q$ goes to infinity. However, again, we observe that the true polynomial analogue of the famous conjecture is rather in the case where the order $q$ of the base field is fixed and the degree $r$ goes to infinity. This much more challenging case is studied analytically in Section 3, resulting in a specific formula (Conjecture 2) for the expected number of pairs of twin irreducibles of degree $r$ over $\mathbf{F}_q$. We then turn to a sieving technique, which supplies us both with an alternate theoretical framework in which to predict the numbers of pairs of twin irreducibles (Section 4), and also with actual counts of such pairs for specific small values of $q$ and $r$ (Section 5). We shall see that both frameworks (Sections 4 and 5) give excellent predictions of the computed actual counts of pairs of twins. Though proof of the classical Twin Primes Conjecture has eluded mathematicians to this point, perhaps this paper can point us toward a solution to our analogous Twin Irreducibles Conjecture (Conjecture 1 of Section 3).

## 2 Definitions and the Fixed Degree Case

For $q$ a prime power, let $\mathbf{F}_q$ denote the finite field of order $q$. All polynomials over $\mathbf{F}_q$ considered herein are *monic*, for these are the correct polynomial analogues to *positive* integers. Among the positive integers, two primes are called *twins* if they differ by as little as possible, meaning (except for 2 and 3) that they differ by 2. Hence it make sense to call two irreducible polynomials "twins" provided that they differ by as little as possible. Consider first polynomials over the field $\mathbf{F}_2$. By how little can they differ? We observe that a polynomial $Y(x)$ over $\mathbf{F}_2$ is divisible by $x$ if its constant term is 0 and is divisible by $x+1$ if it has an even number of nonzero terms. Hence if $P$ is irreducible over $\mathbf{F}_2$ (of degree $\geq 3$), then none of $P(x)+1$, $P(x)+x$, $P(x)+x+1$, $P(x)+x^2$, or $P(x)+x^2+1$ can be irreducible. Hence the smallest possible gap between two irreducibles of degree $\geq 3$ over $\mathbf{F}_2$ is $x^2+x$. However, if $q > 2$, then two (or more) irreducibles can differ by just a constant, so th! at is the smallest gap possible. For example, for $p$ a prime, the polynomials $x^p - x - a$ with $a = 1, \ldots, p-1$, provide a set of $p-1$ consecutive nonlinear irreducibles over the field $\mathbf{F}_p$, see Corollary 3.79 of [8]. We now formalize these observations regarding twin irreducibles in the definitions which follow.

**Definition 1.** The *absolute value* of a polynomial $Y$ of degree $r$ over the finite field $\mathbf{F}_q$ of $q$ elements, denoted $|Y|$, is $q^r$.

**Definition 2.** Two irreducible polynomials $P_1$ and $P_2$, both of degree $r$ over $\mathbf{F}_q$, are said to be *twin irreducible polynomials*, or simply *twin irreducibles*, provided that $|P_2 - P_1| = 4$ if $q = 2$ or $|P_2 - P_1| = 1$ otherwise. More generally, the members of a collection of $k$ distinct irreducible polynomials $P_1, P_2, \ldots, P_k$, all of degree $r$ over $\mathbf{F}_q$, are said to be a *k-tuplet of twin irreducibles* provided that each pair of them are twin irreducibles.

Again, in the cases when $q > 2$, this means that $P_i$ and $P_j$, $i \neq j$, differ only in their constant terms, whereas in the single case $q = 2$, they differ only in their linear and quadratic terms.

Throughout this paper the reader will notice many analogues in the polynomial ring over the finite field $\mathbf{F}_q$ of well studied ideas and results in the ring of integers. While we would like to continually remind the reader of these incredible similarities, we would also like to avoid any possible confusion and so we will refer to twins in the integer setting as twin primes, while in the polynomial setting we will refer to twins as twin irreducibles.

*Example 1.* Over $\mathbf{F}_2$, $x^3+x+1$ and $x^3+x^2+1$ are twin irreducibles. Over $\mathbf{F}_5$, the polynomials $x^2+2$ and $x^2+3$ are twin irreducibles. Over $\mathbf{F}_7$, $x^2+1$, $x^2+2$, and $x^2+4$ form a 3-tuplet (i.e., triplet) of twin irreducibles.

**Definition 3.** Let $N_q(r)$ denote the number of monic irreducible polynomials of degree $r$ over $\mathbf{F}_q$ so that

$$N_q(r) = \frac{1}{r} \sum_{d|r} \mu(d) q^{r/d},$$

where $\mu$ is the Möbius function,

see Theorem 3.25 of [8].

We can now immediately obtain a result on the existence of $k$-tuplets of twin irreducible polynomials provided that we fix the polynomial degree $r$. It comes as no surprise that for fixed $r$, as we increase $q$, we increase the odds of finding $k$-tuplets of twin irreducibles. The result is as follows:

**Proposition 1.** *For every degree $r \geq 2$ and every $k \geq 2$ there exists at least one $k$-tuplet of twin irreducible polynomials of degree $r$ over $\mathbf{F}_q$ provided that $q \geq 2(k-1)r$.*

Proof: We observe that if $q \geq 2(k-1)r \geq 2r$, then

$$(k-1)rq^{r-1} \leq \frac{q^r}{2} = q^r(1 - \frac{1}{2}) \leq q^r(1 - \frac{1}{2q^{\frac{r}{2}-1}})$$
$$= q^r(1 - \frac{q}{2}(\frac{1}{q^{r/2}})) \leq q^r(1 - r(\frac{1}{q^{r/2}})) < \sum_{d|r} \mu(d) q^{r/d}.$$

Dividing through by $r$, we obtain that if $q \geq 2(k-1)r$, then

$$(k-1)q^{r-1} < N_q(r).$$

But now just suppose that among irreducible polynomials of degree $r$ over $\mathbf{F}_q$ there exist only $(k-1)$-tuplets, then for each of the $q^{r-1}$ combinations of coefficients of all but the constant term, there could be at most $k-1$ irreducibles, so the total number of irreducibles $N_q(r)$ would be less than or equal to $(k-1)q^{r-1}$, contradicting the above result. Hence if $q \geq 2(k-1)r$, we are guaranteed at least one k-tuplet of twin irreducibles of degree $r$ over $\mathbf{F}_q$.

**Corollary 1.** *In the collection of all polynomials over all finite fields, there exist infinitely many $k$-tuplets of twin irreducible polynomials for every $k \geq 2$.*

Impressive as this corollary may sound, it stems simply from the fact that if you have enough constant coefficients, you will be able to obtain $k$-tuplet twin irreducibles. This however, is *not* the true analogue of the classical Twin Primes Conjecture that there are infinitely many primes $p$ with the property that $p+2$ is also prime; the correct analogue is the case where the *base field* is fixed. We now consider this much more problematic case.

## 3 The Fixed Base Field Case

We start by stating the desired result:

*Conjecture 1.* For every finite field $\mathbf{F}_q$, there exist infinitely many twin irreducible polynomials over $\mathbf{F}_q$.

One can see immediately that this result is much more difficult than that of the previous section. For with the base field fixed, the number of constant coefficients stays fixed as the degree $r$ goes to infinity, and so the density of twin irreducibles (and of course $k$-tuplets of twin irreducibles for all $k$) decreases rapidly.

**Definition 4.** Let $N_{2,q}(r)$ denote the number of twin irreducible polynomials of degree $r$ over $\mathbf{F}_q$.

In what follows we shall attempt to make an argument, following one made for the classical case in, for example, [3], for the plausibility of an affirmative answer to our conjecture. Specifically, we put forward the following conjecture.

*Conjecture 2.* As the degree $r$ goes to infinity, we have

$$N_{2,q}(r) \sim \delta \left(\frac{q-1}{2}\right) \frac{q^r}{r^2} \prod_P \left(1 - \frac{1}{(|P|-1)^2}\right),$$

where by $\sim$ we mean that the ratio of the quantities on the two sides approaches 1, where $\delta = 4$ if $q = 2$ and 1 otherwise, and where the product is over all irreducible polynomials $P$ over $\mathbf{F}_q$ provided that $q > 2$, but over all irreducibles of degree $\geq 2$ when $q = 2$.

To obtain this conjecture, we mimic the argument presented in Section 22.20 of [3], adapting it to the polynomial setting as needed. In particular, we require a polynomial version of Mertens' Theorem (Theorem 429 in [3]), which states that as $x \to \infty$,

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right) \sim \frac{e^{-\gamma}}{\log x},$$

where, as throughout the remainder of this paper, log denotes the natural logarithm.

We start then with the analogue of this result.

**Theorem 1.** *(Mertens' Theorem for Polynomials) As $n \to \infty$,*

$$\prod_{\deg P \leq n} \left(1 - \frac{1}{|P|}\right) \sim \frac{e^{-\gamma}}{n},$$

*where $P$ runs over monic irreducible polynomials and $\gamma$ is Euler's Constant.*

The following proof was shown to us by K. Conrad.

Proof: We consider the reciprocals of the quantities in the statement, *i.e.*, we show that as $n \to \infty$,

$$\prod_{\deg P \leq n} \frac{1}{1 - 1/|P|} \sim e^{\gamma} n. \tag{1}$$

The left side of (1) is a finite product. Taking logarithms of both sides and expanding the log, the asymptotic (1) is equivalent to

$$\sum_{\deg P \leq n} \sum_{k \geq 1} \frac{1}{k|P|^k} = \log n + \gamma + o(1) \tag{2}$$

with $o(1)$ an expression tending to 0 as $n \to \infty$.

We will prove (2), thus obtaining (1) by exponentiation.

The key idea in proving (2) is to observe that on the left side we can replace the double sum, whose inner sum has infinitely many terms, with a single finite sum over prime powers $P^k$ of degree up to $n$. The change from $\deg P \leq n$ and all $k$ to $\deg P^k \leq n$ is a negligible change in the following sense:

$$\sum_{\deg P \leq n} \sum_{k \geq 1} \frac{1}{k|P|^k} = \sum_{\deg P^k \leq n} \frac{1}{k|P|^k} + o(1). \tag{3}$$

To see why (3) holds, subtract the sum on the right from the sum on the left. Every term in the right sum is a term in the left, so after subtraction we are left with

$$\begin{aligned}
\sum_{\deg P \leq n} \sum_{k > \frac{n}{\deg P}} \frac{1}{k|P|^k} &\leq \sum_{\deg P \leq n} \sum_{k > \frac{n}{\deg P}} \frac{\deg P}{n|P|^k} \\
&= \sum_{\deg P \leq n} \frac{\deg P}{n} \sum_{k > \frac{n}{\deg P}} \frac{1}{|P|^k} \\
&\leq \sum_{\deg P \leq n} \frac{\deg P}{n} \left(\frac{1}{|P|^{n/\deg P}}\right)\left(\frac{1}{1 - 1/|P|}\right) \\
&\leq \sum_{\deg P \leq n} \frac{\deg P}{n} \frac{1}{q^n}\left(\frac{1}{1 - 1/q}\right) \\
&= \frac{1}{q^n}\left(\frac{1}{1 - 1/q}\right) \sum_{\deg P \leq n} \frac{\deg P}{n} \\
&\leq \frac{1}{q^n}\left(\frac{1}{1 - 1/q}\right) \#\{P : \deg P \leq n\} \\
&\leq \frac{1}{q^n}\left(\frac{1}{1 - 1/q}\right) O\left(\frac{q^n}{n}\right)
\end{aligned}$$

which evidently goes to 0 as $n \to \infty$.

This verifies (3), which means the desired identity (2) is equivalent to

$$\sum_{\deg P^k \le n} \frac{1}{k|P|^k} = \log n + \gamma + o(1). \tag{4}$$

To show (4), we write the left side as

$$\sum_{\deg P^k \le n} \frac{1}{k|P|^k} = \sum_{m=1}^{n} \sum_{\deg P^k = m} \frac{1}{kq^m} = \sum_{m=1}^{n} \frac{b_m}{q^m},$$

where

$$b_m = \sum_{\deg P^k = m} \frac{1}{k} = \sum_{d|m} \frac{1}{d} N_q\left(\frac{m}{d}\right) = \sum_{d|m} \frac{d}{m} N_q(d) = \frac{1}{m} \sum_{d|m} d N_q(d) = \frac{q^m}{m}.$$

Thus

$$\sum_{\deg P^k \le n} \frac{1}{k|P|^k} = \sum_{m=1}^{n} \frac{1}{m} = \log n + \gamma + o(1),$$

by the definition of $\gamma$. This establishes (4), so we are done.

Remark. This argument, as Conrad first found it, applies to any function field $K$ over a finite field $\mathbf{F}_q$, as follows. For any place $v$ on $K$, let its residue field have size $\mathrm{N}v = q^{\deg v}$. Then as $n \to \infty$,

$$\prod_{\deg v \le n} \frac{1}{1 - 1/\mathrm{N}v} \sim \frac{L(1/q)}{1 - 1/q} e^{\gamma} n, \tag{5}$$

where the product runs over all places of $K$ and where $L(T)$ is the numerator of the zeta function of $K$. The analogue of (5) for number fields has a zeta residue in place of $L(1/q)/(1 - 1/q)$. For a discussion of Mertens' Theorem in both the number field and function field cases, see [11].

We are now in a position to develop Conjecture 2 by employing an appropriate translation of the argument made for the classical case in [3], Section 22.20.

Fix a degree $r$. Define

$$M = \prod_{\deg P \le r/2} P$$

where the $P$ are irreducible. (One could call $M$ a "primorial" polynomial since it is the product of some initial segment of irreducible polynomials with degree $\le (r/2)$. We note that $M$ is not the same as $M_k$ given in section 4 of this paper.)

Let us denote the degree of $M$ as $m$. Following [3], we call a polynomial $Y$ *special* if it is relatively prime to $M$. For any degree $k$, let $S(k)$ denote the number of special polynomials of degree $k$. Then by equation (4.30) of [2] and by our Theorem 1, we have

$$S(m) = \Phi_q(M) = |M| \prod_{\deg P \le r/2} \left(1 - \frac{1}{|P|}\right) \sim |M| \frac{2e^{-\gamma}}{r},$$

where $\Phi_q$ is the function defined for nonzero polynomials $f$ in $\mathbf{F}_q[x]$ which counts the number of polynomials in $\mathbf{F}_q[x]$ that are of smaller degree than the degree of $f$ and which are relatively prime to $f$. Lemma 3.69 of [8] shows that the function $\Phi_q$ is multiplicative and if $f = P_1^{e_1} \dots P_r^{e_r}$ where each $P_i$ is irreducible of degree $n_i$, then

$$\Phi_q(f) = \prod_{i=1}^{r} (q^{n_i e_i} - q^{n_i(e_i-1)}).$$

Now, the total number of monic polynomials whose degree is $m$ is $q^m = |M|$, so the *proportion* of special polynomials of degree $m$ is of order

$$\frac{2e^{-\gamma}}{r}.$$

Now we consider $S(r)$, noting that $r$ is much smaller than $m$. In fact, by the definition of $M$, we see that $S(r)$ is just the number of irreducible polynomials of degree equal to $r$. Hence

$$S(r) = N_q(r) \sim \frac{q^r}{r}.$$

Because the total number of polynomials having degree $r$ is $q^r$, the proportion of special polynomials of degree $r$ is then of order $\frac{1}{r}$.

Let us denote by $R$ the ratio of the calculated proportions of special polynomials of degree $r$ and $m$ respectively. We obtain

$$R \sim \frac{1}{2e^{-\gamma}}.$$

Now we turn our attention to twin irreducible polynomials. It is reasonable to *conjecture* that the ratio $R_2$ of special *pair* proportions of degree $r$ and $m$ respectively should in fact be $R^2$, i.e.,

$$R_2 = R^2 \sim \frac{1}{4e^{-2\gamma}}.$$

This is reasonable because if the probability that a polynomial $Y(x)$ (of degree either m or r) and $Y(x) + \alpha$ (or $Y(x) + x^2 + x$ when $q = 2$) are

irreducible is assumed be independent, then the probability that both are irreducible is just the product of the separate probabilities.

We observe that this assertion about $R_2$ is the only point at which are unable to provide proof, exactly as with the classical case as argued in [3].

Continuing, recall that we seek an asymptotic formula for $N_{2,q}(r)$, the number of twin irreducible polynomial pairs of degree $r$. But given our conjecture on $R_2$, we can now obtain our goal by finding the proportion of special pairs of the large degree $m$. Let us first assume that $q > 2$. Consider a special pair $Y$ and $Y + \alpha$. How many of these are there of degree $m$? For each $P$ (irreducible) of degree less than or equal to $r/2$, we must have $Y \not\equiv 0 \pmod{P}$ and $Y \not\equiv -\alpha \pmod{P}$, so we get $|P| - 2$ residue classes for each $P$, giving a count of

$$\prod_{\deg P \le r/2} (|P| - 2) = |M| \prod_{\deg P \le r/2} \left(1 - \frac{2}{|P|}\right)$$

such pairs of degree $m$ for a given $\alpha$.

Suppose first that $q$ is odd, so that $\alpha$ and $-\alpha$ are distinct. We observe that if $Y$ and $Y + \alpha$ are a special pair, then so are the exact same pair $Y + \alpha$ and $Y + \alpha - \alpha$. Hence we can obtain *distinct* special pairs by using half of the non-zero elements of $\mathbf{F}_q$ when $q$ is odd. On the other hand, if $q$ is even (and greater than 2), then each non-zero $\alpha$ has the property that if $Y$ and $Y + \alpha$ are a special pair, then $Y + \alpha$ and $Y + \alpha + \alpha$ are the same special pair. Hence again we obtain a factor of $(q-1)/2$. We conclude then that the total number of special pairs of degree $m$ for $q > 2$ is

$$\frac{q-1}{2} |M| \prod_{\deg P \le r/2} \left(1 - \frac{2}{|P|}\right).$$

The case $q = 2$ is somewhat different. Recall that here twins differ by $x^2 + x$ rather than by a constant. If $Y$ (of degree $m$) satisfies $Y \equiv 1 \pmod{x}$ or $Y \equiv 1 \pmod{x+1}$, then $Y + x^2 + x$ satisfies these same conditions. Now if $P_3 = x^2 + x + 1$ is the unique quadratic irreducible ($P_3$ is the "third" irreducible over $\mathbf{F}_2$), then it's easy to check that if $Y \equiv x \pmod{P_3}$ or $Y \equiv x + 1 \pmod{P_3}$ then $Y + x^2 + x$ is also relatively prime to $P_3$, but that if $Y \equiv 1 \pmod{P_3}$, then $P_3$ divides $Y + x^2 + x$. Thus $P_3$ provides 2 $(= |P_3| - 2)$ residue classes producing special pairs. Now for all irreducibles $P$ of degree 3 or greater, as in the $q > 2$ case, we require that $Y \not\equiv 0 \pmod{P}$ and $Y \not\equiv x^2 + x \pmod{P}$, so we get $|P| - 2$ residue classes for each $P$. Finally, we note that the special pair $Y$ and $Y + x^2 + x$ is identical to the special pair $Y + !x^2 + x$ and $(Y + x^2 + x) + (x^2 + x)$, so we must divide our count by 2 to eliminate this duplication. We obtain then in the case $q = 2$ a count of special pairs of degree $m$ of

$$\frac{1}{2} \prod_{2 \le \deg P \le r/2} (|P| - 2) = \frac{1}{8} |M| \prod_{2 \le \deg P \le r/2} \left(1 - \frac{2}{|P|}\right),$$

where the extra factor of 4 in the denominator on the right occurs because the two linear irreducibles (each of absolute value 2) are missing from the product as each $|P|$ is factored out.

We now have all the information we need to obtain our desired asymptotic formula. The basic equation is

$$R_2 = \frac{\text{proportion of special pairs of degree } r}{\text{proportion of special pairs of degree } m}$$

and so

$$N_{2,q}(r) \sim \frac{R_2(\text{total of degree } r)(\text{number of special pairs of degree } m)}{\text{total of degree } m}.$$

Recall that by Merten's Theorem, $2e^{-\gamma}/r \sim \prod_{P,\deg P \le r/2}(1 - \frac{1}{|P|})$. Hence we can compute

$$\begin{aligned} N_{2,q}(r) &\sim \frac{1}{4e^{-2\gamma}} \frac{q^r}{q^m} q^m \left(\frac{q-1}{2\beta}\right) \prod_{\lambda \le \deg P \le r/2} \left(1 - \frac{2}{|P|}\right) \\ &\sim \frac{1}{r^2(4e^{-2\gamma}/r^2)} q^r \left(\frac{q-1}{2\beta}\right) \prod_{\lambda \le \deg P \le r/2} \left(1 - \frac{2}{|P|}\right) \\ &\sim \left(\frac{q-1}{2\beta}\right) \frac{q^r}{r^2} \frac{\prod_{P, \lambda \le \deg P \le r/2}(1 - \frac{2}{|P|})}{(\prod_{P,\deg P \le r/2}(1 - \frac{1}{|P|}))^2} \\ &\sim \delta\left(\frac{q-1}{2}\right) \frac{q^r}{r^2} \prod_{\deg P \ge \lambda} \left(1 - \frac{1}{(|P|-1)^2}\right) \end{aligned}$$

where $\beta$ is 4 if $q = 2$ and is 1 otherwise, $\lambda$ is 2 if $q = 2$ and is 1 otherwise, and $\delta$ is 4 if $q = 2$ and is 1 otherwise. $\delta = 4$ arises in the $q = 2$ case because we must remove the factors of $(\prod_{P,\deg P \le r/2}(1 - \frac{1}{|P|}))^2$ corresponding to the two linear irreducibles over $\mathbf{F}_2$ from the product, and each contributes $1/4$ to the denominator. Then one of those factors of 4 (in the numerator) cancels with $\beta = 4$.

This completes the argument for Conjecture 2, which has been the goal of this section. We now turn in a different direction, using a sieving technique both to provide an alternate framework to understand (and predict counts of) twin irreducibles and to obtain exact counts (Section 5).

## 4 The Polynomial Wheel Sieve

The wheel sieve for integers was first described by Pritchard [9] as a fast algorithm for computer prime number sieve routines. In [5], this technique was used to study the distribution of primes in sets of arithmetic progressions of the form $a+nm_k$, where the multiplier $m_k$ is the $k$-th *primorial* number $p_1 \cdot$

$p_2 \dots p_k$ and $a < m_{k-1}$ is any number relatively prime to $m_k$. The heuristics in [5] suggest that the primes are distributed binomially among the arithmetic progressions $a + nm_k$, using a binomial probability given by the asymptotic value from Dirichlet's Theorem. Similarly, the heuristics from [5] suggest that twin primes in pairs of arithmetic progressions $a + nm_k$ and $(a + 2) + nm_k$ are also distributed binomially, with a binomial probability given from the twin prime conjecture [3].

An analogue of Dirichlet's theorem on the distribution of primes in an arithmetic progression also holds for polynomials over the finite field $\mathbf{F}_q$, see [7]. In [6], the authors considered a polynomial version of the wheel sieve, and discussed the distribution of irreducibles over the field $\mathbf{F}_2$. In particular, they discussed the distribution of irreducibles in arithmetic progressions, and made several conjectures, the most important of which (Conjecture 1) postulates that the irreducible polynomials in the progressions from the wheel sieve are distributed so as to asymptotically approach a binomial distribution using a binomial probability given by the asymptotic value from a theorem of Artin [8].

For the sake of completeness, we now briefly describe the polynomial wheel sieve and provide an example for purposes of illustration. For an integer $k \geq 1$, let $M_k(x) = P_1(x) \cdots P_k(x)$ be the product of the first $k$ monic irreducibles in $\mathbf{F}_p[x]$. The polynomial $M_k(x)$ corresponds to the $k$-th primorial number $p_1 \cdots p_k$, and will be called the *k-th primorial polynomial*. For each value of $k \geq 1$, the wheel sieve generates a sequence of polynomials, using an iterative process with polynomials from the previous cycle as seeds.

**Definition 5.** Over $\mathbf{F}_q$, let $W_1 = \mathbf{F}_q^* \cup \{x\}$. Given $W_k$ for $k > 0$, let $S_k = \{S \in W_k | P_k(x) \not| S\}$ be the set after sieving the set $W_k$ by the irreducible $P_k$. Let $W_{k+1} = \{S(x) + N(x)M_k(x) | S(x) \in S_k\}$, where $N(x)$ varies over all polynomials of degree less than the degree of $P_k(x)$.

Note if $q$ is a prime, $W_1$ may be taken as $1, 2, \dots, q-1, x$.

Let $(w_{ij}^k)$ be the matrix containing the set $W_k$. The first column is the set $S_{k-1}$, and the remaining columns as we move from left to right, contain successive multiples of the primorial polynomial $M_{k-1}(x)$. The set $S_k$, by construction, is pre-sieved for the first $k$ irreducibles. This reduces the work necessary to sieve using the remaining irreducibles of degree $(\deg(M_k)+1)/2$. (The addition of 1 is necessary only if $\deg(M_k)$ is odd.) After sieving $S_k$ by these irreducibles, the remaining set is examined for twin irreducibles. This procedure is carried out on a computer, using a program written in the C language.

We now illustrate the wheel sieve in the case when $q = 3$. We note that an example for $q = 2$ is given in [6].

*Example 2.* The first four irreducible polynomials over $\mathbf{F}_3$ are $P_1(x) = x$, $P_2(x) = x + 1$, $P_3(x) = x + 2$, $P_4(x) = x^2 + 1$, and the first three primorial

polynomials are $M_1(x) = x$, $M_2(x) = x^2 + x$, $M_3(x) = x^3 + 2x$. Then we have

$$W_1 = \{1, 2, x\}, S_1 = \{1, 2\},$$

$$W_2 = \left\{ \begin{matrix} 1 & x+1 & 2x+1 \\ 2 & x+2 & 2x+2 \end{matrix} \right\}, \quad S_2 = \left\{ \begin{matrix} 1 & & 2x+1 \\ 2 & x+2 & \end{matrix} \right\},$$

or using more compact notation, where the polynomial $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ is abbreviated in the form $a_n a_{n-1} \cdots a_0$, we have $S_2 = \{1, 2, 12, 21\}$.

For the next case,

$$W_3 = \left\{ \begin{matrix} 1 & 111 & 221 \\ 2 & 112 & 222 \\ 12 & 122 & 222 \\ 21 & 101 & 211 \end{matrix} \right\}, \quad S_3 = \left\{ \begin{matrix} 1 & & 221 \\ 2 & 112 & \\ & 122 & 222 \\ & 101 & 211 \end{matrix} \right\}.$$

Note that exactly 1 polynomial is removed per row, because each row spans a complete set of residues when sieved by $P_k(x)$.

The number of polynomials in $S_k$ is easily shown to be

$$\Phi_q(M_k(x)) = \prod_{i=1}^{k} (|P_i| - 1)$$

because these polynomials are all relatively prime to the first $k$ irreducibles. ($\Phi_q$ was defined in the previous section, but the form given here is specific to operation on $M_k(x)$, and emphasizes the similarity with the Euler $\phi$-function from number theory.) Note that the polynomials have, by construction, degree less than that of $M_k(x)$ and so there will be $\deg(M_k) - \deg(M_{k-1}) = \deg(P_k)$ different degrees in $S_k$ beyond those degrees found in $S_{k-1}$. This leads to an awkward counting of monic irreducibles in $S_k$, since we are interested in these counts for a particular degree $r$. It will be useful to define the number of monic polynomials of degree $r \geq m$ in $W_k$, which is denoted by

$$\Phi_q(M_{k-1}, r) = q^{r-m} \prod_{i=a}^{k-1} (|P_i| - 1)$$

where $m = \deg(M_{k-1}(x))$, and $a = 3$ when $q = 2$ and $a = 1$ otherwise. When $r = m$, this reduces to $\Phi(M_{k-1}(x))$, which is the number of elements in $S_{k-1}$ and hence the number of rows in the matrix representation of $W_k$. For example, when $q = 3$ and $k = 3$ then $\Phi_3(M_3) = 8$, which is the number of elements in $S_3$ (see the above example). The number of monic polynomials in $W_4$ of degrees $r = 3$ and $r = 4$ are respectively 8 and $3 \cdot 8 = 24$.

In analogy, the number of monic "special" pair polynomials in $S_k$ is given by

$$\Phi_{2,q}(M_k(x)) = \delta \left( \frac{q-1}{2} \right) \prod_{i=a}^{k} (|P_i| - 2)$$

where $a = 1$ when $q > 2$ and $a = 3$ when $q = 2$ (to avoid zeros in the product when $q = 2$). Our use of the term "special "in this context means the polynomials are relatively prime to $M_k$. The function $\Phi_{2,q}(M_k)$ plays a role in finite fields similar to the integer function $\phi_2$ given in [5]. The function $\Phi_{2,q}(M_k(x))$ counts the number of twin irrducible pairs that are relatively prime to $M_k(x)$, and the factor in front, $\delta(q-1)/2$ comes about from the same counting arguments as given in Section 3 (which were applied to the product $M$ but are equally valid with regard to the product $M_k$ above).

Similarly, the number of *monic* special pairs of degree $r \geq m$ in $W_k$ is

$$\Phi_{2,q}(M_{k-1}, r) = \delta \left( \frac{q-1}{2} \right) q^{r-m} \prod_{i=a}^{k-1} (|P_i| - 2)$$

where $m$ and $a$ are the same as above and $\delta$ was defined in Conjecture 2. As an example, over $\mathbf{F}_3$, $\delta = 1$ and $(q-1)/2 = 1$ so the number of monic special pairs in $W_4$ of degrees 3 and 4 are respectively 1 and $3 \cdot 1$.

The following is a generalization of Conjecture 1 in [6]:

*Conjecture 3.* The monic irreducible polynomials of degree $r$ in the progressions of the wheel sieve are distributed so as to asymptotically approach a binomial distribution in the parameter $p = (\Phi_q(M_{k-1}(x), r))^{-1}(q^r/r)$, for $r$ in the range $\deg(M_{k-1}) \leq r < \deg(M_k)$.

Analogously we make

*Conjecture 4.* The monic twin irreducible polynomials of degree $r$ in the progressions of the wheel sieve are distributed so as to asymptotically approach a binomial distribution in the parameter $p = ((\Phi_q(M_{k-1}(x), r))^{-1}(q^r/r))^2$, for $r$ in the range $\deg(M_{k-1}) \leq r < \deg(M_k)$.

Numerical calculations support Conjectures 3 and 4. In particular, Figures 1 and 2 from [6] provide numerical evidence for the $q = 2$ case. These conjectures are formulated in a similar way to Conjecture 2. The probability of obtaining two irreducibles with a minimum gap next to each other is *assumed independent*, and hence this twin probability is just the square of the probability to find a single irreducible in $S_k$ (compare the binomial parameters from Conjecture 3 and 4). This heuristic shows that the distribution of irreducibles in the rows of the sieved matrix are binomial, as shown in Figure 1 of [6], which is reminiscent of Bernoulli trials. It is now a small step to show that this conjecture is equivalent to the Conjecture 2 in Section 3.

The analog of the twin prime conjecture is obtained by combining Conjectures 3 and 4. The average probability that a given element in $W_k$ of degree $r$ is an irreducible is just the number $N_q(r)$ of irreducibles, divided by the number $\Phi_q(M_{k-1}, r)$ of polynomials of degree $r$ in the set $W_k$. Similarly, the average probability that a special pair in $W_k$ is a twin irreducible pair is given by the number $N_{2,q}(r)$ of twin irreducibles, divided by the number

$\Phi_{2,q}(M_{k-1}, r)$ of special pairs in the set $W_k$. Under the assumption that the probabilities are *uncorrelated*, the latter probability is equal to the square of the former, and so we obtain

$$\frac{N_{2,q}(r)}{\Phi_{2,q}(M_{k-1}, r)} \sim \left(\frac{N_q(r)}{\Phi_q(M_{k-1}, r)}\right)^2$$

where again $\deg(M_{k-1}) \leq r < \deg(M_k)$. Solving for $N_{2,q}(r)$ and using the definitions for $N_q(r)$, $\Phi_q$ and $\Phi_{2,q}$, one finds a similar asymptotic form as Conjecture 2. In particular,

$$N_{2,q}(r) \sim \delta\left(\frac{q-1}{2}\right)\left(q^{r-m}\prod_{i=a}^{k-1}(|P_i|-2)\right)\left(\frac{q^r}{r}\right)^2\left(q^{r-m}\prod_{i=a}^{k-1}(|P_i|-1)\right)^{-2}$$

and upon expanding the products and using $q^m = |M_{k-1}|$, we obtain

$$W_{predict} \sim \delta\left(\frac{q-1}{2}\right)\frac{q^r}{r^2}\prod_{i=a}^{k-1}\left(1-\frac{1}{(|P_i|-1)^2}\right)$$

as the prediction for $N_{2,q}(r)$, where $a = 3$ when $q = 2$ and $a = 1$ otherwise.

It is no surprise that Conjecture 2 and $W_{predict}$ have similar asymptotic forms (differing only in where the product is truncated), since they were formulated with similar assumptions. The predictions from both conjectures will be compared with numerical data in the next section.

We note that the product has a very similar form to the integer twin prime constant, $C_2$. In analogy, we define the constant

$$C_{2,q} = \delta\left(\frac{q-1}{2}\right)\prod_{i=a}^{k-1}\left(1-\frac{1}{(|P_i|-1)^2}\right)$$

as the twin irreducibles constant, where as before, $a = 3$ when $q = 2$ and $a = 1$ otherwise. This constant converges quickly, and for example when $q = 2$ it has the value $C_{2,2} = 0.8328783\ldots$.

Contrasting Sections 3 and 4, Section 3 used analytic techniques whereas Section 4 used probabilistic estimates based on the binomial distribution. We leave it to the reader to determine whether either approach will eventually yield a proof rather than a conjecture.

## 5 Numerical Data

In this section we provide some data related to the distribution of twin irreducibles over finite fields of small orders. In particular, we compare estimates given from the analytic theory developed in Section 3 and estimates arising

from the wheel sieve theory described in Section 4. We compare these estimates with actual counts of the number of twin irreducibles over fields of orders $2, 3, 4, 5$ and 7.

In the following tables,

$N_q(r)$ denotes the number of monic irreducibles of degree $r$ over $\mathbf{F}_q$,

$N_{2,q}(r)$ denotes the number of pairs of monic twin irreducibles of degree $r$ over $\mathbf{F}_q$,

$\Phi_q(M_{k-1}, r)$, abbreviated as $\Phi_q(k, r)$, denotes the number of monic polynomials of degree $r$ over $\mathbf{F}_q$ which are relatively prime to the first $k-1$ irreducibles,

$\Phi_{2,q}(M_{k-1}, r)$, abbreviated as $\Phi_{2,q}(k, r)$, denotes the number of special pairs of monic polynomials in $S_k$ of degree $r$ over $\mathbf{F}_q$ with minimal gap,

$A_{predict}$ denotes the predicted number of twin irreducibles of degree $r$ over $\mathbf{F}_q$ obtained from Conjecture 2 in Section 3,

$W_{predict}$ denotes the predicted number of twin irreducibles of degree $r$ over $\mathbf{F}_q$ obtained from Conjecture 4 in Section 4.

In Tables 1, 2, 3, 4, and 5, we provide numerical data on the distribution of twin irreducibles over fields of orders 2, 3, 4, 5, and 7. Thus in the table for the field $\mathbf{F}_q$, one can compare the actual number $N_{2,q}(r)$ of twin irreducibles of degree $r$ obtained by machine calculation with the estimate $A_{predict}$ from the analytic method by comparing the values in columns 3 and 6 for a given $r$. Similarly, one can compare the actual number $N_{2,q}(r)$ of twin irreducibles of degree $r$ with the estimate $W_{predict}$ from the wheel sieve method by comparing columns 3 and 7.

Armed with the conjectures, the analog of Brun's constant (the reciprocal sum of the twin primes) is easily calculated. The calculation is slightly different for $q = 2$, since the smallest "gap" is $x^2 + x$, except for the first 2 irreducibles, $P_1 = x$ and $P_2 = x + 1$. We decide that $P_1$ is the analogue of "2" in the integer case, and should not be included in the sum. Hence, for $q = 2$, the first irreducible pair $(x^3 + x + 1,\ x^3 + x^2 + 1)$ has degree 3 and contributes $1/8 + 1/8 = 0.25$ to the sum. Continuing in this manner, the wheel sieve can be used to calculate the exact value of $B_{2,2}$ up to degree 26, giving an intermediate sum of 0.9350. Using Conjecture 4 to estimate the sum up to degree 65190, the sum becomes 1.0585. Using analytic estimates to extrapolate as the degree goes to infinity, we obtain $B_{2,2} = 1.0591\ldots$ for the analog of Brun's constant over $\mathbf{F}_2$. The calculation of $B_{2,q}$ for $q > 2$ can be done in a similar manner. In the cases for $q = 3, 4, 5, 7$, we obtain $B_{2,3} = 2.2724\ldots$, $B_{2,4} = 4.0647\ldots$, $B_{2,5} = 5.5058\ldots$, $B_{2,7} = 8.7025\ldots$.

We note that while the two predictors $A_{predict}$ and $W_{predict}$ are based upon quite different methods (one analytic and the other probabilistic), they give very similar estimates for the number $N_{2,q}(r)$ of twin irreducibles of degree $r$ over $\mathbf{F}_q$.

**Table 1.** Distribution of twin irreducibles over $\mathbf{F}_2$ for degree $r$.

| $r$ | $N_2(r)$ | $N_{2,2}(r)$ | $\Phi_2(k,r)$ | $\Phi_{2,2}(k,r)$ | $A_{predict}$ | $W_{predict}$ |
|---|---|---|---|---|---|---|
| 2 | 1 | **0** | 0 | 0 | **0** | **0** |
| 3 | 2 | **1** | 2 | 1 | **2** | **1** |
| 4 | 3 | **1** | 3 | 1 | **2** | **1** |
| 5 | 6 | **2** | 6 | 2 | **2** | **1** |
| 6 | 9 | **2** | 12 | 3 | **3** | **2** |
| 7 | 18 | **4** | 21 | 6 | **4** | **4** |
| 8 | 30 | **7** | 42 | 12 | **7** | **6** |
| 9 | 56 | **8** | 84 | 24 | **11** | **10** |
| 10 | 99 | **16** | 147 | 36 | **17** | **16** |
| 11 | 186 | **28** | 294 | 72 | **28** | **29** |
| 12 | 335 | **55** | 588 | 144 | **47** | **47** |
| 13 | 630 | **76** | 1176 | 288 | **81** | **83** |
| 14 | 1161 | **142** | 2205 | 504 | **139** | **140** |
| 15 | 2182 | **224** | 4410 | 1008 | **243** | **247** |
| 16 | 4080 | **414** | 8820 | 2016 | **427** | **431** |
| 17 | 7710 | **758** | 17640 | 4032 | **756** | **770** |
| 18 | 14532 | **1340** | 33075 | 7056 | **1348** | **1362** |
| 19 | 27594 | **2456** | 66150 | 14112 | **2420** | **2456** |
| 20 | 52377 | **4436** | 132300 | 28224 | **4367** | **4424** |
| 21 | 99858 | **7926** | 264600 | 56448 | **7922** | **8040** |
| 22 | 190557 | **14362** | 496125 | 98784 | **14436** | **14573** |
| 23 | 364722 | **26638** | 992250 | 197568 | **26416** | **26693** |
| 24 | 698870 | **48358** | 1984500 | 395136 | **48520** | **49005** |
| 25 | 1342176 | **89048** | 3969000 | 790272 | **89431** | **90372** |
| 26 | 2580795 | **165368** | 7938000 | 1580544 | **165367** | **167067** |

As mentioned earlier, both approaches are based on the idea of special polynomials, however in Section 3 the estimates are based on $M$ (a primorial product up to $\deg P \leq (r/2)$) and in Section 4 the estimates are based on $M_k$ (a primorial product up to the $k$-th irreducible). Perhaps a deeper examination of these two methods will help elucidate the ideas behind the twin prime conjecture for finite fields.

There is a remarkably similar comparison between the forms of primes (over $\mathbf{Z}$) and irreducibles (over $\mathbf{F}_q$). This is best illustrated by putting these side-by-side. For the primes and irreducibles,

$$\pi(x) \sim \frac{x}{\log x} \; ; \; N_q(r) \sim \frac{q^r}{r}$$

and for the twin primes and twin irreducibles,

$$\pi_2(x) \sim 2C_2 \frac{x}{(\log x)^2} \; ; \; N_{2,q}(r) \sim C_{2,q} \frac{q^r}{r^2}$$

**Table 2.** Distribution of twin irreducibles over $\mathbf{F}_3$ for degree $r$.

| $r$ | $N_3(r)$ | $N_{2,3}(r)$ | $\Phi_3(k,r)$ | $\Phi_{2,3}(k,r)$ | $A_{predict}$ | $W_{predict}$ |
|---|---|---|---|---|---|---|
| 2 | 3 | **0** | 4 | 1 | **1** | **0** |
| 3 | 8 | **1** | 8 | 1 | **1** | **0** |
| 4 | 18 | **0** | 24 | 3 | **2** | **1** |
| 5 | 48 | **6** | 64 | 7 | **4** | **1** |
| 6 | 116 | **6** | 192 | 21 | **8** | **2** |
| 7 | 312 | **15** | 512 | 49 | **18** | **18** |
| 8 | 810 | **36** | 1536 | 147 | **41** | **41** |
| 9 | 2184 | **105** | 4096 | 343 | **96** | **97** |
| 10 | 5880 | **216** | 12288 | 1029 | **234** | **236** |
| 11 | 16104 | **585** | 36864 | 3087 | **580** | **589** |
| 12 | 44220 | **1506** | 106496 | 8575 | **1462** | **1478** |
| 13 | 122640 | **3747** | 319488 | 25725 | **3737** | **3791** |
| 14 | 341484 | **9510** | 958464 | 77175 | **9666** | **9796** |
| 15 | 956576 | **25555** | 2768896 | 214375 | **25261** | **25586** |
| 16 | 2690010 | **66606** | 8306688 | 643125 | **66606** | **67445** |
| 17 | 7596480 | **177561** | 24920064 | 1929375 | **177001** | **179285** |

**Table 3.** Distribution of twin irreducibles over $\mathbf{F}_4$ for degree $r$.

| $r$ | $N_4(r)$ | $N_{2,4}(r)$ | $\Phi_4(k,r)$ | $\Phi_{2,4}(k,r)$ | $A_{predict}$ | $W_{predict}$ |
|---|---|---|---|---|---|---|
| 2 | 6 | **3** | 9 | 6 | **4** | **3** |
| 3 | 20 | **4** | 27 | 12 | **7** | **6** |
| 4 | 60 | **18** | 81 | 24 | **15** | **13** |
| 5 | 204 | **36** | 324 | 96 | **37** | **38** |
| 6 | 670 | **130** | 1215 | 336 | **103** | **102** |
| 7 | 2340 | **312** | 4860 | 1344 | **303** | **311** |
| 8 | 8160 | **1008** | 18225 | 4704 | **928** | **943** |
| 9 | 29120 | **2836** | 72900 | 18816 | **2933** | **3002** |
| 10 | 104754 | **10158** | 273375 | 65856 | **9502** | **9670** |
| 11 | 381300 | **31116** | 1093500 | 263424 | **31410** | **32030** |

**Table 4.** Distribution of twin irreducibles over $\mathbf{F}_5$ for degree $r$.

| $r$ | $N_5(r)$ | $N_{2,5}(r)$ | $\Phi_5(k,r)$ | $\Phi_{2,5}(k,r)$ | $A_{predict}$ | $W_{predict}$ |
|---|---|---|---|---|---|---|
| 2 | 10 | **5** | 16 | 18 | **7** | **7** |
| 3 | 40 | **20** | 64 | 54 | **20** | **21** |
| 4 | 150 | **45** | 256 | 162 | **56** | **56** |
| 5 | 624 | **196** | 1024 | 486 | **178** | **180** |
| 6 | 2580 | **520** | 5120 | 2430 | **616** | **617** |
| 7 | 11,160 | **2280** | 24576 | 11178 | **2264** | **2305** |
| 8 | 48,750 | **8825** | 122880 | 55890 | **8662** | **8797** |
| 9 | 217000 | **34530** | 589824 | 128547 | **34221** | **34799** |
| 10 | 976248 | **138394** | 2949120 | 642735 | **138585** | **140863** |

**Table 5.** Distribution of twin irreducibles over $\mathbf{F}_7$ for degree $r$.

| $r$ | $N_7(r)$ | $N_{2,7}(r)$ | $\Phi_7(k,r)$ | $\Phi_{2,7}(k,r)$ | $A_{predict}$ | $W_{predict}$ |
|---|---|---|---|---|---|---|
| 2 | 21 | **21** | 36 | 75 | **30** | **26** |
| 3 | 112 | **84** | 216 | 375 | **94** | **101** |
| 4 | 588 | **336** | 1296 | 1875 | **366** | **386** |
| 5 | 3360 | **1680** | 7776 | 9375 | **1641** | **1750** |
| 6 | 19544 | **7770** | 46656 | 46875 | **7969** | **8225** |
| 7 | 117648 | **41847** | 279936 | 234375 | **40982** | **41396** |

where $C_{2,q}$ and the twin prime constant $C_2$ are also similar, both being convergent products. In particular

$$C_2 = \prod_p \left(1 - \frac{1}{(p-1)^2}\right); \qquad C_{2,q} = \delta\left(\frac{q-1}{2}\right) \prod_P \left(1 - \frac{1}{(|P|-1)^2}\right)$$

where for $C_2$ the product is over all primes $p \geq 3$, and for $C_{2,q}$ it is over all monic irreducibles of degree two or greater for $q = 2$ but over all monic irreducibles for $q > 2$.

## 6 Generalizations and Extensions

In this final section we briefly discuss several extensions and generalizations of the previous work. We say that two polynomials $f, g$ over $\mathbf{F}_q$ have *rank* $q^m$, $m \geq 0$, if $f - g$ is a polynomial of degree $m$. Thus for $q = 2$, we obtain the twin irreducible case when $m = 2$, while if $q > 2$, when $m = 0$ we obtain our earlier notion of twin irreducibles.

We begin by noting that in the fixed degree case, Proposition 1 can be generalized from twins to rank $q^m$.

**Proposition 2.** *Let $m \geq 0$. For every degree $r \geq 2$ and every $k \geq 2$, there exists at least one $k$-tuplet of twin irreducible polynomials of degree $r$ over $\mathbf{F}_q$ of rank $q^m$, provided that $q \geq (2(k-1)r)^{1/(m+1)}$.*

We also then have

**Corollary 2.** *In the collection of all polynomials over all finite fields, there exist infinitely many $k$-tuplets of twin irreducible polynomials of rank $q^m$ for every $k \geq 2$.*

As indicated in Section 3, the case of a fixed base field $\mathbf{F}_q$ is much more difficult. In this case we propose

*Conjecture 5.* For each $m \geq 0$ and for every finite field $\mathbf{F}_q$, there exist infinitely many irreducible polynomials over $\mathbf{F}_q$ of rank $q^m$.

Conjecture 2 postulates an asymptotic form for $N_{2,q}(r)$, the number of twin irreducibles of degree $r$ over $\mathbf{F}_q$. It would be of interest to have an asymptotic formula for the number of rank $q^m$ irreducibles over $\mathbf{F}_q$ of degree $r$. For $q > 2$ this appears to be

$$\left(\frac{q^{m+1}-1}{2}\right)\frac{q^r}{r^2}\prod_P\left(1-\frac{1}{(|P|-1)^2}\right).$$

We close by asking, more generally, can one obtain an asymptotic formula for the number of $k$-tuplets of twin irreducibles among rank $q^m$ irreducibles over $\mathbf{F}_q$ of degree $r$?

## References


1. G. Effinger, *A Goldbach Theorem for Polynomials of Low Degree over Odd Finite Fields*, Acta Arithmetica, 42:329–365, 1983.
2. G. Effinger and D. Hayes, Additive Number Theory of Polynomials over a Finite Field, Oxford University Press, Oxford, 1991.
3. G.H. Hardy and E.M. Wright, An Introduction to the Theory of Numbers, 5th Edition, Oxford Science Publications, Oxford, 1979.
4. D.R. Hayes, *The distribution of irreducibles in $GF[q,x]$*, Trans. Amer. Math. Soc. 117(1965), 101–127.
5. K. Hicks and I. Sato, *Heuristics of arithmetic progressions in the framework of the wheel sieve*, submitted.
6. K.H. Hicks, G.L. Mullen, and I. Sato, *Distribution of irreducible polynomials over finite fields*, published in these proceedings.
7. H. Kornblum, *Über die Primfunktionen in einer arithmetischen Progression*, Math. Z. 5(1919), 100–111.
8. R. Lidl and H. Niederreiter, Finite Fields, Cambridge Univ. Press, 1997.
9. P. Pritchard, *Explaining the wheel sieve*, Acta Informat. 17(1982), 477–485.
10. P. Ribenboim, The New Book of Prime Number Records Springer-Verlag, New York, 1995.
11. M. Rosen *A Generalization of Mertens' Theorem*, J. Ramanujan Math. Soc. 14(1999), 1–19.

# Invariants of Finite Groups over Finite Fields: Recent Progress and New Conjectures

Peter Fleischmann

Institute of Mathematics and Statistics,
University of Kent at Canterbury,
Canterbury, CT2 7NF, England
email: pf10@ukc.ac.uk

**Abstract.** Let $G$ be a finite group acting on a polynomial ring $A := \mathbb{F}[x_1, \ldots, x_n]$ by graded algebra automorphisms. If $\mathbb{F}$ is a field of characteristic zero, then due to classical results of Emmy Noether one knows that the invariant ring $A^G$ can be generated in degrees less or equal to $|G|$. If $\mathbb{F}$ is a field of positive characteristic $p$ dividing the group order $|G|$, this is no longer true. The situation in characteristic $p$ not dividing $|G|$ has been clarified recently after being open for several decades. This paper presents an account on these developments, including some related questions and conjectures dealing with constructive and structural properties of modular invariant rings.

## 1 Introduction

Let $\mathbb{F}$ be a field and let $\mathbb{F}[x_1, x_2, \cdots, x_n]$ be a polynomial algebra on which the group $G$ acts by graded algebra automorphisms. The *ring of invariants* consists of all polynomials unchanged by elements of $G$ and is the major object of study in classical invariant theory. One major motivation comes from algebraic geometry, where more general 'affine' algebras and group actions occur. Assume for the moment that $\mathbb{F}$ is algebraically closed and let $X$ be an affine algebraic variety with $A := \mathcal{O}(X)$ the corresponding algebra of regular functions, such that $X$ can be identified with the set of maximal ideals in $A$. For a group $G$ of automorphisms of $X$ let $X//G$ denote the *categorical quotient*. Then a natural question is, whether $X//G$ is again an affine algebraic variety. A necessary condition for this is, that the invariant ring

$$A^G := \{a \in A \mid g(a) = a \ \forall \ g \in G\}$$

is a finitely generated $\mathbb{F}$ - algebra. In this case the maximal ideal spectrum $\max - \operatorname{spec} A^G$ of the invariant ring is a natural candidate for $X//G$ and one might hope that $X//G$ coincides with the set - theoretic quotient, namely the orbit space $X/G$. In general neither does $A^G$ have to be finitely generated, nor does the categorical quotient $X//G$ have to coincide with the 'geometric quotient' $X/G$, but if $G$ is a finite group, then due to a fundamental theorem of Emmy Noether, the situation is much better (see e.g. [1] 1.4.4).

**Theorem 1.** *(Emmy Noether (1926)[14]) If $|G| < \infty$, then $A^G$ is finitely generated and $X//G \cong \max - \operatorname{spec} A^G$ is in bijection with the orbit - space $X/G$.*

Therefore the natural questions arise

- how can $A^G$ be constructed ?
- what are the geometrical properties of $A^G$?

When invariant theory was created in the middle of the 19'th century, the main interest was focused on the situation where $\mathbb{F}$ is the field of complex numbers and $G$ is a classical (infinite) group. Therefore invariant theory 'in characteristic zero' is highly developed. Recent applications in geometry, algebraic topology and cohomology theory ask for results in the situation where $\mathbb{F}$ is of positive characteristic, in particular a finite field, and $G$ is a finite group (see [21]). As I pointed out in my talk at the Fq4 - conference in Waterloo ([4]), much less is known in this situation, because many of the methods developed for characteristic zero do not carry over, and some results on 'classical invariant rings' are known to be false in positive characteristic. Since 1997 considerable progress has been achieved, in particular with respect to constructive methods. For example the problem of Emmy Noether's degree bound which was addressed in [4], has in the meantime been resolved to full satisfaction ([6], [8]). The experience gained in this process has led to various new questions and conjectures which also deal with geometrical and structural properties of invariant rings. In this paper I want to give an account on these developments and will present an admittedly subjective outlook into the nearer future.

**Notation**: Throughout the paper $\mathbb{N}_0$ denotes the set of nonnegative integers; for $n \in \mathbb{N}$, $\underline{n}$ denotes the set $\{1, 2, \cdots, n\}$ and $\mathbb{N}_0^{\underline{n}}$ denotes the set of functions from $\underline{n}$ to $\mathbb{N}_0$.

## 2 Constructive Aspects

Noether's proof of Theorem 1 was one of the first major applications of her newly developed theory of rings and modules with ascending chain condition. The price one has to pay for the generality of Theorem 1 is, that its proof is not constructive.

**Definition 1.** Let $R$ be a commutative ring, $A := R[a_1, a_2, \cdots, a_k]$ a finitely generated $R$ - algebra with generators $a_1, \cdots, a_k$ and $G$ a finite group acting on $A$ as $R$ - algebra automorphisms, stabilizing the $R$ - module $\sum_{i=1}^{k} Ra_i$. For $\alpha \in \mathbb{N}_0^{\underline{k}}$ we denote with $\underline{a}^\alpha$ the power product $a_1^{\alpha_1} a_2^{\alpha_2} \ldots a_k^{\alpha_k} \in A$ and with $\mathcal{M}_\ell((\underline{a_i}))$ we denote the $R$ - module spanned by

$$\{\underline{a}^\alpha \mid |\alpha| := \sum_{i=1}^{k} \alpha_i \leq \ell\} \subseteq A.$$

For a subalgebra or ideal $B$ of $A$ we define the **Noether number**

$$\beta(B, (a_1, \ldots, a_k)) \in \mathbb{N} \cup \{\infty\}$$

to be the infimum of the set of all $\ell \in \mathbb{N}$ such that $\{b_1, \ldots, b_m\} \subseteq \mathcal{M}_\ell(\underline{(a_i)})$, satisfying $B = R[b_1, \ldots, b_m]$ if $B$ is a subalgebra, or $B = (b_1, \ldots, b_m)A$ if $B$ is an ideal. If there is an obvious set of chosen generators $(a_i)$ or the choice does not matter, we will simply write $\beta(B)$. For example, if $A = R[x_1, \cdots, x_k]$ is a polynomial ring generated by variables $x_i$ of degree 1 and $B$ is a finitely generated graded subring or a homogeneous ideal, then $\beta(B) := \beta(B, \underline{(x_i)})$ will be the minimal number $k$ such that $B$ is generated by homogeneous elements in $A$ of degree less or equal to $k$.

*Remark 1.* By Theorem 1, $\beta := \beta(A^G, (a_1, \ldots, a_k)) < \infty$, if $G$ is finite and $R$ is Noetherian, but the proof does not give any bound for $\beta$.

Obviously the Noether number $\beta(B)$ is a measure for the *algorithmic complexity* of a subring or ideal. In particular if $R$ is a field and $B := A^G$, then invariants in $\mathcal{M}_\ell(\underline{(a_i)})$ can be computed by solving linear equations of size increasing with $\ell$. Since $\beta(A^G)$ is an upper bound for the $\ell$ needed to generate $A^G$, it also bounds the overall size of linear systems to solve.

In 1916 Emmy Noether gave two different proofs for the fact that

$$\beta(\mathbb{F}[x_1, \cdots, x_k]^G) \leq |G|,$$

if $G$ is a finite group and $R = \mathbb{F}$ is a field of characteristic zero. This is usually referred to as the 'Noether bound' in invariant theory. Both of these proofs fail in positive characteristics. On the other hand it is known from [16], that $\beta(A^G)$ can be arbitrarily large, if the characteristic of $\mathbb{F}$ divides $|G|$. After Emmy Noether's general result (Theorem 1), the question of degree bounds somewhat fell into oblivion, until it resurrected again with the upcoming of constructive invariant theory in conjunction with new powerful methods in computer algebra. In particular the so called 'Noether gap', i.e. the conjecture that the Noether bound $\beta(\mathbb{F}[x_1, \cdots, x_k]^G) \leq |G|$ holds if char $\mathbb{F}$ does not divide $|G|$, has de facto been open since 1916, but was considered more seriously during the past decade (see e.g. [19], [15], where the conjecture was shown to hold for solvable groups). This question has recently been answered to the affirmative by the author and J Fogarty independently and with slightly different approaches (see [6], [8]). In the following I will present a 'combined version' of these proofs, incorporating an essential observation by D Benson which makes the combinatorics in [6] more transparent.

The fact that Noether's degree bound does not hold in general can be seen already in the simple example $A := \mathbb{F}_2[x_1, \cdots, x_k, y_1, \cdots, y_k]$ with $G = \Sigma_2 = \langle g \rangle$ acting by swapping the 'variable types' $x_i \leftrightarrow y_i$. It is an easy exercise to show that the invariant $\mathfrak{r} := (x_1 \cdots x_k)^+ := x_1 \cdots x_k + y_1 \cdots y_k$ is

indecomposable, i.e. cannot be written as a sum of products of invariants of smaller degree. Therefore

$$\beta(A^G, (x_1, \cdots, y_k)) \geq k \to \infty \text{ if } k \to \infty.$$

It is interesting, though, to observe that

$$\mathfrak{x} = (x_2 \cdots x_k)^+ x_1 + (x_1 x_3 \cdots x_k)^+ y_2 - (x_3 \cdots x_k)^+ x_1 y_2,$$

i.e. $\mathfrak{x}$ decomposes in the *Hilbert - ideal* $A^{G,+}A$, generated *in* $A$ by all invariants of positive degree. Hence one can easily see that $\beta(A^{G,+}A) = 2$ for all $k \geq 2$. In fact a generalization of this observation led to the proof of Noether's bound in [6].

Let $H \leq G$ be a subgroup of index $n$ with coset - decomposition $G := \uplus_{i:=1}^{n} g_i H$. The homomorphism of $A^G$ - modules given by

$$t_H^G : \ A^H \to A^G, \ a \mapsto \sum_{g \in G:H} g(a)$$

is called the *relative transfer map* with respect to $H$; its image is an ideal in $A^G$, called the *relative transfer ideal* (w.r.t. $H$). The following lemma gives a useful decomposition in $A$ of high degree relative transfer elements:

**Lemma 1.** *For* $b, b_1, b_2, \cdots, b_n \in A^H$ *we have*

$$t_H^G(bb_1 \cdots b_n) = \sum_{I \subset \underline{n}, \ I \neq \underline{n}} (-1)^{n-|I|+1} \ t_H^G(b \prod_{j \in I} b_j) \prod_{j \notin I} g_j(b_j).$$

*Proof.* By Benson's trick we have for each fixed $i$ the obvious equality:

$$\prod_{j=1}^{n} ( \ g_i(b_j) - g_j(b_j) \ ) = 0.$$

Expansion and multiplication with $g_i(b)$ for fixed $i$ gives:

$$0 = \sum_{I \subseteq \underline{n}} (-1)^{|I|} \prod_{j \notin I} g_j(b_j) \cdot (\prod_{j \in I} g_i(b_j)) \cdot g_i(b).$$

Now summation over $i \in \underline{n}$ yields the claimed identity.

**Theorem 2.** *Let* $A$ *be as in Definition 1 and* $H$ *a subgroup of* $G$ *such that either* $|G|$ *invertible in* $R$ *or* $H \lhd G$ *a normal subgroup with index* $[G : H]$ *invertible in* $R$. *Then*

$$\beta(A^G) \leq \beta(A^H) \cdot [G : H].$$

*In particular, if* $|G| \in \mathbb{F}^*$, *then Noether's degree bound holds, i.e.*

$$\beta(\mathbb{F}[x_1, \cdots, x_k]^G) \leq |G|.$$

*Proof.* Under both assumptions on $H$, the relative transfer map $t_H^G$ is surjective. Also note that the elements $g_j(b_j)$ appearing in the previous lemma are in $A^H$, if $H$ is normal in $G$. Now suppose that $\beta := \beta(A^H) = \beta(A^H, (a_\ell))$ with $A := R[a_1, \cdots, a_\ell]$ and $A^H = R[b_1, \cdots, b_k]$ with $b_i \in A^H \cap \mathcal{M}_\beta(\underline{(a_\ell)})$. If $H \leq G$ and $|G|$ invertible, we have $t_H^G(bb_1 \cdots b_n) = \frac{1}{|G|} t_1^G(t_H^G(bb_1 \cdots b_n)) =$

$$\sum_{I \subset \underline{n},\ I \neq \underline{n}} (-1)^{n-|I|+1}\ t_H^G(b \prod_{j \in I} b_j) \frac{1}{|G|}\ t_1^G(\prod_{j \notin I}\ g_j(b_j)) \in R[A^G \cap \mathcal{M}_{n\beta}(\underline{(a_\ell)})].$$

If $H \lhd G$ and $|G/H|$ invertible, we replace $|G|$ by $|G/H|$ and $t_1^G$ by $t_H^G$ to conclude in a similar way that $t_H^G(bb_1 \cdots b_n) \in R[A^G \cap \mathcal{M}_{n \cdot \beta}(\underline{(a_\ell)})]$. Now an iterative application of this result finishes the proof.

One might hope to remove the requirement $H \lhd G$ for subgroups of invertible index:

*Conjecture 1.* : If $H \leq G$ with index $[G : H]$ invertible in $R$, then

$$\beta(A^G) \leq \beta(A^H) \cdot [G : H].$$

If $A = \mathbb{F}[x_1, \cdots, x_k]$, the formula in Lemma 1 describes a decomposition of relative transfer elements in the Hilbert - ideal $A^{G,+}A$. Extensive sample calculations done by Harm Derksen and Gregor Kemper led them to the following far reaching conjecture:

*Conjecture 2.* [Noether bound for Hilbert ideals] (H. Derksen / G. Kemper): Let $G$ be a finite group, $\mathbb{F}$ a field and $A := \mathbb{F}[x_1, \cdots, x_k]$ a polynomial ring, such that $G$ acts by graded algebra automorphisms. Then

$$\beta(A^{G,+}A) \leq |G|.$$

In the next section we will show that in special cases this conjecture can be verified using the techniques of the proof of Theorem 2, which we are now going to refine.

From now on for the rest of the paper let $R := \mathbb{F}$ be a field of characteristic $p \mid |G|$ and $V$ a finitely generated $\mathbb{F}G$ - module. We consider the symmetric algebra $A := \mathrm{Sym}(V^*)$ of the dual module $V^*$. Choosing a basis $\{x_1, \cdots, x_d\}$ for $V^*$, the ring $A$ can be viewed as a polynomial ring $\mathbb{F}[x_1, \cdots, x_d]$ with induced graded $G$ -action. Let $P$ be a fixed Sylow $p$ - group of $G$ with normalizer $N_G(P)$ and $N := N_G(P)/P$. For subgroups $U \leq H \leq G$ we define the (homogeneous) *relative transfer ideal*

$$\mathcal{I}_{<U}^H := \sum_{Y < U} t_Y^H(A^Y) \lhd A^H.$$

**Lemma 2.**

$$\beta(A^{G,+} \cdot A) \leq \max\{\beta(A^Q/\mathcal{I}_{<Q}^Q) \cdot [G : Q] \mid Q \leq P\}.$$

*Proof.* Let $f \in A^{G,+}$ be indecomposable in $A^{G,+}A$. Since $n := [G : P]$ is invertible, $f$ is of the form $t_P^G(h)$ for some $h \in A^{P,+}$, which itself can be written as

$$h = \sum_{Q \leq P} t_Q^P(b_{Q,1} b_{Q,2} \cdots b_{Q,\ell_Q})$$

with $b_{Q,j} + \mathcal{I}_{<Q}^Q \in A^Q/\mathcal{I}_{<Q}^Q$ and $b_{Q,j}$ homogeneous of positive degree $\leq \beta_Q := \beta(A^Q/\mathcal{I}_{<Q}^Q)$. Moreover we can assume that every nonzero transfer element

$$t_P^G(t_Q^P(b_{Q,1} b_{Q,2} \cdots b_{Q,\ell_Q})) = t_Q^G(b_{Q,1} b_{Q,2} \cdots b_{Q,\ell_Q})$$

is indecomposable in $A^{G,+}A$ as well. But from lemma 1 we see that this requires $\ell_Q \leq [G : Q]$, hence

$$f \in \sum_Q A^{G,+}_{\leq \beta_Q [G:Q]} \cdot A.$$

Hence Noether's bound in case of relative transfer quotients for $p$ - groups implies Conjecture 2:

**Corollary 1.** *If $\beta(A^Q/\mathcal{I}_{<Q}^Q) \leq |Q|$ for all $Q \leq P \in \mathrm{Syl}_p(G)$, then conjecture 2 holds, i.e. $\beta(A^{G,+}A) \leq |G|$.*

To obtain degree bounds for $A^G$ rather than $A^{G,+}A$ one can make use of the *Brauer homomorphism* from representation theory, i.e. is the canonical homomorphism $A^G \to \overline{A^G} := A^G/\mathcal{I}_{<P}^G$. Using Mackey's formula for the relative transfer, we get

$$t_P^G(b) = \sum_{g \in P\backslash G/P} t^P_{P \cap {}^gP}({}^g b) \equiv \sum_{g \in N/P} {}^g b \equiv t_1^N(b) \mod \mathcal{I}_{<P}^P,$$

where $P\backslash G/P$ denotes a chosen system of double cosets of $P$ in $G$. It is easy to see that $\mathcal{I}_{<P}^P \cap A^G = \mathcal{I}_{<P}^G$, hence we get

$$\overline{A^G} = (A^G + \mathcal{I}_{<P}^P)/\mathcal{I}_{<P}^P = t_1^N(A^P/\mathcal{I}_{<P}^P) \cong (A^P/\mathcal{I}_{<P}^P)^N.$$

Since $p$ does not divide $|N|$, Theorem 2 gives

**Lemma 3.**

$$\beta(\overline{A^G}) \leq \beta(\overline{A^P}) \cdot |N|.$$

$$\beta(A^G) \leq \max\{\beta(\mathcal{I}_{<P}^G),\ \beta(\overline{A^P}) \cdot |N|\}.$$

It has been conjectured by several experts that

$$\beta(\mathrm{Sym}(V^*)^G) \leq \max\{|G|, \dim V \cdot (|G| - 1)\}$$

is a 'natural degree bound' for modular invariant rings of type $\mathrm{Sym}(V^*)^G$. Using the above technique this certainly would follow from the next two slightly sharper conjectures:

*Conjecture 3.* Let $P$ be a Sylow $p$ - group of $G$. Then

1. $\beta(A^G/\mathcal{I}^G_{<P}) \leq |N_G(P)|$.
2. If $A = \mathrm{Sym}(V^*)$ with $\mathbb{F}G$ - module $V$, then

$$\beta(\mathcal{I}^G_{<P}) \leq \max\{|G|, \dim V \cdot (|G| - 1)\}.$$

## 3 *p*-permutation Modules

In this section we present some evidence for the previous conjectures, based on the analysis of a special type of invariant rings. An $\mathbb{F}G$ - module $V$ in characteristic $p$ is called a *$p$ - permutation module* or *trivial source module*, if its restriction to any Sylow $p$ - subgroup is an ordinary permutation module, or in other words, if a basis $\mathfrak{b}$ of $V$ can be found, which is permuted by some Sylow $p$ - group. Note that, since any two Sylow $p$ - groups are conjugate, the property of being a $p$ - permutation module does not depend on the choice of the Sylow group. Note also that $V$ is a $p$ - permutation module if and only if so is the dual $V^*$. The following lemma is a known fact in modular representation theory (e.g. see [18]):

**Lemma 4.** *The $\mathbb{F}G$ - module $V$ is a $p$ - permutation module if and only if it is a direct summand of a permutation module for $G$.*

Now choose a Sylow $p$ - group $P$ and assume that $V_{|P}$ is a permutation module of dimension $d$. Then for any subgroup $Q \leq P$ the restricted module $V^*_{|Q}$ is also a permutation module. Moreover for every integer $m$, the homogeneous component $A_m = \mathrm{Sym}(V^*)_m$ of degree $m$ is generated as a vector space by power products $\mathbf{x}^{\mathbf{a}} := x_1^{a_1} x_2^{a_2} \cdots x_d^{a_d}$ of exponent sum $m$, which themselves are permuted by $Q$. In particular an element

$$\mathfrak{x} := \sum_{\mathbf{a} \in \mathbb{N}_0^{\underline{d}}} c_{\mathbf{a}} \mathbf{x}^{\mathbf{a}}$$

is $Q$ - invariant, if and only if $c_{\mathbf{a}}$ is constant on the orbit $\mathbf{a}^Q := \{\mathbf{a} \circ g \mid g \in Q\}$, where $Q$ is viewed as permuting the set $\underline{d}$ and therefore acting naturally on the set of functions $\mathbb{N}_0^{\underline{d}}$. In other words, each $Q$ - invariant is a linear combination of *orbit - sums* of the form

$$(\mathbf{x}^{\mathbf{a}})^+ := \sum_{g \in Q/Q_{\mathbf{a}}} \mathbf{x}^{\mathbf{a} \circ g},$$

where $Q/Q_{\mathbf{a}}$ is a chosen cross section for the cosets of the stabilizer subgroup $Q_{\mathbf{a}}$ in $Q$. In particular $(\mathbf{x}^{\mathbf{a}})^+ \in \mathcal{I}^Q_{<Q}$ if and only if $Q_{\mathbf{a}} < Q$. On the other hand, the $Q$ - stable power products are products of 'norm - like elements' of the form $\mathfrak{n}_i := \prod_{x_i \in \mathcal{O}_i} x_i$, where $\{\mathcal{O}_i \mid i = 1, \cdots, s\}$ is the set of all $Q$ -

orbits on a chosen permutation basis $\{x_1, \cdots, x_d\}$ of $V$. From this it is easy to see that

$$A^Q = \mathcal{I}^Q_{<Q} \oplus \mathbb{F}[\mathfrak{n}_1, \cdots, \mathfrak{n}_s] \qquad 1.$$

In particular $A^Q/\mathcal{I}^Q_{<Q} \cong \mathbb{F}[\mathfrak{n}_1, \cdots, \mathfrak{n}_s]$ is a polynomial ring generated by elements of degree $\leq |Q|$. This together with Lemma 2 gives:

**Proposition 1.** *If $V$ is a $p$ - permutation module and $P$ is a Sylow $p$ - group of $G$, then $\beta(A^Q/\mathcal{I}^Q_{<Q}) \leq |Q|$ for each subgroup $Q \leq P$. In particular conjecture 2 holds for $A = \mathrm{Sym}(V^*)$.*

To show that the second degree bound of conjecture 3 holds in this case, we can argue in a similar way as in the proof of Lemma 2: Let $f \in \mathcal{I}^{G,+}_{<}$ be indecomposable; by the transitivity of relative transfers ($t^H_{U\,|A^U} = (t^H_Y \circ t^Y_U)_{|A^U}$ for $U \leq Y \leq H$) we can in fact assume that $f$ is of the form $t^G_Q(h)$ with $h$ being a power product in the $\mathfrak{n}_i$'s. Each polynomial $\prod_{g \in G/Q} (T - g(\mathfrak{n}_i)) \in A^G[T]$ has $\mathfrak{n}_i$ as a zero, showing that $\mathfrak{n}_i^{|G:Q|} \in \sum_{0 \leq j < |G:Q|} A^G \mathfrak{n}_i^j$. This allows for reductions of exponents in $h$ and since the operator $t^G_Q$ is $A^G$ - linear, we can assume that these exponents do not exceed $|G:Q|-1$. Hence the total degree of $f$ can be assumed to be less or equal to $s \cdot (|G:Q|-1)$. Since $s \leq d = \dim V$ and $\deg \mathfrak{n}_i \leq |Q|$ we conclude that $\beta(\mathcal{I}^G_{<P}) \leq \max\{|G|, d \cdot (|G|-1)\}$. Hence

**Proposition 2.** *If the $\mathbb{F}G$ - module $V$ is a direct summand of a permutation module, then*

$$\beta(\mathrm{Sym}(V^*)^G) \leq \max\{|G|, \dim\, V \cdot (|G|-1)\} \qquad 2.$$

Note that the modular group algebra $\mathbb{F}P$ in characteristic $p$ is a local Frobenius - algebra and therefore finitely generated projective, injective and free $\mathbb{F}P$ - modules coincide. Since the restriction $V_{|P}$ of any f.g. projective or injective $\mathbb{F}G$ - module is free and hence a permutation module (viz. a sum of copies of the regular module), any such module is a $p$ - permutation module. Moreover every $\mathbb{F}G$ - module appears as submodule and factor module of a suitable projective one. This adds to the evidence for (2) to be a natural modular degree bound, even though no general result seems to exist about Noether numbers for invariant rings of sub representations or quotients. If the group $G$ is cyclic, then $W \leq V$ implies $\beta(\mathrm{Sym}(W^*)^G) \leq \beta(\mathrm{Sym}(V^*)^G)$, due to a recent result of R J Shank and D Wehlau [17]. If moreover $G$ is of order $p$ then the bound (**) has been proved by D Hughes and G Kemper [9].

## 4 Structural Aspects

Let $B$ be a positively graded $\mathbb{F}$ - algebra with $B_0 \cong \mathbb{F}$, $M$ a finitely generated $B$ - module and $I \lhd B$ an ideal. Recall that a sequence $a_1, a_2, \cdots, a_r$ of homogeneous elements in $I$ is called a *regular $M$ - sequence* if $B(a_1, \cdots, a_r)M < M$

and for every $1 \leq i \leq r$ the multiplication with $a_i$ is an injective operator on the quotient $M/(a_1, a_2, \cdots, a_{i-1})M$. It is a known fact that all maximal regular $M$ - sequences in $I$ have the same length, which is called the *grade* of $I$ on $M$. The grade of $B$ on $M$ is called the *depth* of $M$ and the module $M$ is called *Cohen - Macaulay*, if its depth coincides with its Krull - dimension Dim $M$ = Dim $(B/\text{Ann}_B(M))$. Here $\text{Ann}_B(M) := \{a \in B \mid aM = 0\}$ is the annihilator of $M$ in $B$. The ring $B$ is called Cohen - Macaulay, if the regular module ${}_BB$ is Cohen - Macaulay. It can be shown that $B$ is Cohen - Macaulay if and only if $B$ is a finitely generated free module over some polynomial subring of $B$ (see [1] 4.3 or [2]).

It has been known for quite some time that rings of polynomial invariants of the form $\text{Sym}(V^*)^G$ are Cohen - Macaulay, if $G$ is a finite group whose order is coprime to the characteristic of $\mathbb{F}$. This is known to be false in general, once $p = \text{char } \mathbb{F}$ divides $|G|$ (see [10] and the references there). The degree to which it fails is measured by the *defect* def $A^G :=$ Dim $A^G -$ depth $A^G$. Clearly if $A = \text{Sym}(V^*)$ and $G$ is finite, then Dim $A =$ Dim $A^G = \dim_{\mathbb{F}} V$, because $A$ is a finite extension of $A^G$.

In 1980 G. Ellingsrud and T. Skjelbred proved the celebrated result that, if $P$ is a Sylow $p$ - group $P$ of $G$ with fixed point space $V^P$, one has

$$\text{depth } \text{Sym}(V^*)^G \geq 2 + \dim V^P \qquad 3$$

if $\dim V \geq \dim V^P + 2$, with equality if $G$ is a cyclic $p$ - group (see [3]). For almost two decades this has been the only general result on the depth of modular invariant rings, which remains to be one of their most interesting, but difficult to determine parameters. In particular the classification of modular Cohen - Macaulay - invariant rings is an open problem. The result of Ellingsrud and Skjelbred was achieved using homological algebra, in particular a Grothendieck spectral sequence. During the last five years or so, these techniques have been revitalized (see e.g. [10], [22], [12]), most notably by Gregor Kemper who was able to classify all groups, whose modular regular representation has a Cohen - Macaulay ring of invariants.

With regard to the techniques laid out in the previous sections of this paper, some recent results show that the relative transfer ideal $\mathcal{I}^G_{<P}$ and its radical $\sqrt{\mathcal{I}^G_{<P}}$ can shed some new light on the problem of determining the depth of modular invariant rings. Let $P \leq G$ be a chosen Sylow $p$ - group of $G$ and assume for technical reasons, that $\mathbb{F}$ is algebraically closed of characteristic $p > 0$. This allows us to consider $A := \text{Sym}(V^*)$ as the algebra of polynomial functions on $V$, and $A^G$ as the algebra of polynomial functions on the orbit space $V/G$. Hence for an ideal $\mathcal{I} \lhd A^G$, the variety $\mathcal{V}(\mathcal{I})$ consists of all orbits $v^G$ in $V$ such that $f(x) = 0$ for every $f \in \mathcal{I}$ and $x \in v^G$. On the other hand for each subset $S \subseteq V$ there is the ideal $\mathcal{I}(S) := \{f \in A \mid f(s) = 0 \ \forall s \in S\}$ and $\mathcal{I}^G(S) := \mathcal{I}(S) \cap A^G$. In [5] relative transfer ideals have been investigated geometrically, which led to the following description of $\sqrt{\mathcal{I}^G_{<P}}$ in terms of its variety in the orbit space $V/G$:

**Theorem 3.** *(P. Fl. [5])*

1. $\mathcal{V}(\mathcal{I}^G_{<P}) = \{v^G \in V/G \mid p \not| \, |v^G|\}$,
2. $\sqrt{\mathcal{I}^G_{<P}} = \mathcal{I}^G(V^P)$, *where* $V^P$ *denotes the space of* $P$ *- fixed points in* $V$.
3. *If the action of* $P$ *on* $V$ *is defined over* $\mathbb{F}_p$, *then*

$$A^G/\sqrt{\mathcal{I}^G_{<P}} \text{ is Cohen} - \text{Macaulay of Krull dimension } \dim_{\mathbb{F}} V^P.$$

Note that for a permutation module of a $p$ - group $Q$, the previous result

$$A^Q = \mathcal{I}^Q_{<Q} \oplus \mathbb{F}[\mathfrak{n}_1, \cdots, \mathfrak{n}_s]$$

implies that $\sqrt{\mathcal{I}^Q_{<Q}} = \mathcal{I}^Q_{<Q}$ with

$$A^Q/\mathcal{I}^Q_{<Q} \cong \mathbb{F}[\mathfrak{n}_1, \cdots, \mathfrak{n}_s]$$

being even a polynomial ring. Hence Theorem 3 3. is a natural generalization, which indicated that the ideal $\sqrt{\mathcal{I}^G_{<P}}$ might 'measure the depth' of $\mathrm{Sym}(V^*)^G$: the formula (3) of Ellingsrud - Skjelbred shows, that dim $V^P$ is a lower bound for the depth of $A^G$, so it was conceivable that the 'missing part' of the depth is provided by regular elements in the ideal $\mathcal{I}^G_{<P}$[1]. In the case of $p$ - groups this follows from a result of G. Kemper (see [11], Theorem 1.5) and for general groups it is a consequence of the following

**Theorem 4.** *(P. Fl., R.J. Shank, [7])*

$$\text{depth } A^G = \text{grade}(\mathcal{I}^G_{<P}, A^G) + \dim V^P.$$

Moreover, if $V$ is defined over $\mathbb{F}_q$, one can at least in principle use the 'Dickson invariants' $d_i \in \mathrm{Sym}(V^*)^{\mathrm{GL}_n(q)} \leq \mathrm{Sym}(V^*)^G$ to determine the grade of $\mathcal{I}^G_{<P}$ on $A^G$.

Theorem 4, in connection with Lemma and Conjecture 3, shows that the relative transfer ideal $\mathcal{I}^G_{<P}$ contains the clues for some of the most important structural and constructive properties of modular invariant rings. Therefore an efficient algorithm to find minimal generating sets for the ideal $\sqrt{\mathcal{I}^G_{<P}}$ is very much needed as an important step to determine its grade and henceforth the depth of $A^G$.

[1] note that the grade of an ideal always coincides with the grade of its radical.

## References


1. Benson D., *Polynomial Invariants of Finite Groups*, Cambridge Univ. Press 1993.
2. Bruns W., Herzog J., *Cohen-Macaulay Rings*, Cambridge Univ. Press 1993.
3. Ellingsrud G., Skjelbred T., *Profondeur d'anneaux d'invariants en caractéristique p*, Compos. Math. **41**, 233–244, (1980).
4. Fleischmann P., Lempken W. *On Degree Bounds for Invariant Rings of Finite Groups Over Finite Fields*, Cont. Math. **225**, 33–41, (1997).
5. Fleischmann P., Relative trace ideals and Cohen - Macaulay quotients of modular invariant rings in 'Computational Methods for Representations of Groups and Algebras', Dräxler P., Michler G.., Ringel C. M. Ed., Birkhäuser, Basel (1999).
6. Fleischmann P., *The Noether Bound in Invariant Theory of Finite Groups*, Adv. in Mathematics, **156**, 23–32, (2000).
7. Fleischmann P., Shank R. J., *Depth and the Relative Trace Ideal*, to appear in Arch. Math.
8. Fogarty J., *On Noether's Bound for Polynomial Invariants of a Finite Group.*, Elec. Res. Ann. of the AMS (7), 5–7, (2001).
9. Kemper G., Hughes I., *Symmetric Powers of Modular Representations, Hilbert Series and Degree Bounds*, Comm. in Algebra 28, 2059–2088, (2000).
10. Kemper G., *On the Cohen- Macaulay property of modular invariant rings*, J. of Algebra **215**, 330–351, (1999).
11. Kemper G., *The depth of invariant rings and cohomology*, with an appendix by K. Magaard, preprint (1999), Universität Heidelberg, to appear in the J. of Algebra.
12. Lorenz M., Pathak J., *On Cohen-Macaulay Rings of Invariants*, preprint (2001), Temple University.
13. Noether E., *Der Endlichkeitssatz der Invarianten endlicher Gruppen*, Math. Ann. **77**, 89–92, (1916).
14. Noether E., *Der Endlichkeitssatz der Invarianten endlicher linearer Gruppen der Charakteristik p*, Nachr. Ges. Wiss. Göttingen (1926), 28–35, in 'Collected Papers', pp. 485–492, Springer Verlag, Berlin (1983).
15. Richman D., *Explicit generators of the invariants of finite groups*, Adv. Math. **124**, 49–76, (1996).
16. Richman D., *Invariants of finite groups over fields of characteristic p*, Adv. Math. **124** (1996), 25–48.
17. Shank R. J., Wehlau D., *Noether numbers for sub representations of cyclic groups of prime order* , to appear in the Bulletin of the London Mathematical Society.
18. Landrock P., *Finite Group Algebras and their Modules*, LMS Lecture Note Series, Cambridge University Press (1984).
19. Smith L., *E. Noether's bound in the invariant theory of finite groups*, Arch. Math. **66**, 89–92, (1995).
20. Smith L., *Polynomial Invariants of Finite Groups*, A.K. Peters Ltd., (1995).
21. Smith L., *Polynomial invariants of finite groups, a survey of recent developments*, Bull. Amer. Math. Soc. (1997), 211–250.
22. Smith L., *Homological Codimension of Modular Rings of Invariants and the Koszul Complex*, preprint (1997).

# The Group Law on Elliptic Curves on Hesse form

Hege Reithe Frium

HQDC Norway
PB 14
N-1306 Baerum postterminal
Norway

**Abstract.** In this paper I will give an introduction to elliptic curves on *Hesse form*. The embedding of these curves in the projective plane make their symmetries especially nice. If we pick a point $p$ in the projective plane s.t. $p$ is not a 3-torsion point, $p$ is the parametrization of the curve that contains $p$. We will also see that the division polynomials are independent of chosen elliptic curve on Hesse form.

## 1 Introduction

The study of elliptic curves, elliptic integrals and elliptic functions were one of the great topics in the nineteenth century mathematics. Where the Norwegian mathematician Niels Henrik Abel were one of the masters together with Gauss, Jacobi and Legendre.

In this paper I will give an introduction to a certain family of elliptic curves, the elliptic curves on *Hesse form*. We will see that, among other properties on the elliptic curves on Hesse form, the division polynomials are independent of the chosen curve.

I was first introduced to these curves by my advisor Professor Kristian Ranestad as a topic for my master thesis in algebraic geometry. In cooperation with the Headquarters Defence Command, Norway and Thales Communications he started a seminar series on elliptic curves and cryptography at the Department of Mathematics at University of Oslo in fall 1998. This made us interested in studying the elliptic curves on Hesse form to see if there are some advantages compared with curves in the Weierstrass family.

## 2 Elliptic Curves on Hesse form

Let $k$ be a field and $K = \bar{k}$ its algebraic closure. And let $\mathbb{P}^2_k$ denote the projective plane over the field $k$ with projective coordinates $x_0, x_1$ and $x_2$. Elliptic curves are algebraic curves of genus 1 defined over the field $k$. Every elliptic curve can be embedded as a curve given by a smooth cubic equation in the projective plane $\mathbb{P}^2$.

A curve $F$ in $\mathbb{P}^2$ is smooth or nonsingular at a point $p$ if the three partial derivatives $\frac{\partial F}{\partial x_i}(p)$ are not all zero. The curve is smooth or nonsingular if it is nonsingular at every point $p \in F$.

### 2.1 Hessians and the Hesse pencil, $\mathcal{H}$

An important property of most plane nonsingular, irreducible cubics over the field $k$ is the existence of flexes. A flex is a nonsingular point, $p$, of a curve such that the curve intersects the tangent at $p$ with multiplicity at least three at p. If it intersects exactly three times it is called an ordinary flex else a higher flex. Nonsingular cubics over a finite field $k$ contain either 0, 1, 3 or 9 $k$-rational flexes.

In 1842 Ludwig Otto Hesse (1811-74) constructed a determinant, called a Hessian, that characterizes the flexes of a curve of degree at least 3. The Hessian of a plane projective curve, $F$, of degree $d$, is defined by

$$H(F) = \begin{vmatrix} F_{x_0x_0} & F_{x_0x_1} & F_{x_0x_2} \\ F_{x_0x_1} & F_{x_1x_1} & F_{x_1x_2} \\ F_{x_0x_2} & F_{x_1x_2} & F_{x_2x_2} \end{vmatrix}$$

where $F_{x_ix_j}$ is the second partial derivative of the polynomial $F$ with respect to $x_i$ and $x_j$. When $F$ is irreducible $H(F)$ is a form of degree $3(d-2)$. The following theorem gives us the relationship between $H(F)$ and the flexes of $F$.

**Theorem 1.** *Let $F$ be a curve of degree $d > 1$ in $\mathbb{P}^2$. Let $char(k) = 0$ or $char(k) \geq d$ and let $p$ be a $k$-rational point on the curve $F$. Then $p \in H(F) \cap F$ if and only if $p$ is either a flex or a singular point of $F$. The multiplicity of $p$ in $H(F) \cap F$ equals 1 if and only if $p$ is an ordinary flex.*

*Proof.* For a proof of this theorem, see either [2], [7] or [5].

When the characteristic of a finite field is less than the degree of the curve we will need to use the Hasse derivative instead of the usual derivative. The Hasse derivative, $D_x$, acts on $F(x) = \sum a_i x^i$ in $k[x]$ as $D_x^{(r)}(\sum a_i x^i) = \sum \binom{i}{r} a_i x^{i-r}$, where the binomial coefficient is taken modulo the prime characteristic. Write $i$ and $r$ as $p$-ary expansions, $i = i_0 + i_1 p + \cdots + n_e p^e$ and $r = r_0 + r_1 p + \cdots + r_e p^e$ with $0 \leq i_j < p$ and $0 \leq r_j < p$ for $0 \leq j \leq e$, then $\binom{i}{r} \equiv \binom{i_0}{r_0}\binom{i_1}{r_1} \cdots \binom{i_e}{r_e} \pmod{p}$. And $\binom{i}{r} \equiv 0 \pmod{p}$ if and only if $r_j > i_j$ for some $j$.

**Lemma 1.** *Let $F$ be a cubic defined over a field of characteristic $\geq 2$. A nonsingular point $p$ on $F$ is a flex point if $p$ lies in the intersection of $F$, $K^{(0)}$, $K^{(1)}$ and $K^{(2)}$, where $K^{(i)} = (D_{(j)})^2 D^2_{(kk)} + (D_{(k)})^2 D^2_{(jj)} - D_{(j)} D_{(k)} D^2_{(jk)}$ and $D_{(j)} = D_{x_j} F$, $D^2_{(jk)} = D^2_{x_j x_k} F$ and $\{i, j, k\} = \{0, 1, 2\}$.*

*Proof.* See [7].

We will need the following well known theorem, named Bezout's Theorem.

**Theorem 2.** *(Bezout) Let $F$ and $G$ be complex projective curves of degrees $m$ and $n$ such that $F$ and $G$ have no common factors of positive degree. Then $F$ and $G$ intersect exactly $mn$ times, counting multiplicities, in the complex projective plane.*

*Proof.* See f.ex. [2], [5] or [6].

This gives the next corollary.

**Corollary 1.** *A nonsingular, irreducible plane projective cubic over* $\mathbb{C}$ *has nine flexes, all ordinary.*

These nine flexes lie by threes on twelve lines. This is the classical configuration, $(9_4, 12_3)$, of flexes of a plane cubic.

A nonsingular , irreducible plane projective cubic curve over $\mathbb{R}$ has at least one flex. For nonsingular cubics over a finite field, $F_q$, the following theorem lists the possible number of flexes.

**Theorem 3.** *The number of rational flexes of a nonsingular cubic over* $F_q$ *is zero, one, three or nine. The possibilities are as follows:*

$$\begin{aligned} q &\equiv 0 \pmod 3 : \ 0, 1, 3; \\ q &\equiv 2 \pmod 3 : \ 0, 1, 3; \\ q &\equiv 1 \pmod 3 : 0, 1, 3, 9. \end{aligned}$$

*Proof.* See [7].

**Lemma 2.** *There exists a nonsingular plane cubic curve over* $F_q$ *with nine* $F_q$*-rational flexes if and only if* $q \equiv 1 \pmod 3$. *In this case the cubic has the canonical form* $E_{(a,b)} = ax_0x_1x_2 + b(x_0^3 + x_1^3 + x_2^3)$.

*Proof.* See [7].

The $(9_4, 12_3)$ configuration exists in $\mathbb{P}^2_{F_q}$ for nonsingular cubics if and only if $q \equiv 1 \pmod 3$. When we take the configuration in canonical form we get a pencil of cubic curves containing the nine points. This pencil is called the Hesse pencil.

**Definition 1.** The family of curves in $\mathbb{P}^2$, over the field $k$, generated by the two cubics $x_0x_1x_2 = 0$ and $x_0^3 + x_1^3 + x_2^3 = 0$

$$E_{(a,b)} : ax_0x_1x_2 + b(x_0^3 + x_1^3 + x_2^3) = 0, \quad (a, b) \in \mathbb{P}^1.$$

is called the Hesse pencil, $\mathcal{H}$.

$\mathcal{H}$ is an 1-dimensional linear subspace of $\mathbb{P}^9$, the space of cubics in $\mathbb{P}^2$. The name comes from the fact that the Hessian of a curve in the Hesse pencil is itself a curve in the Hesse pencil. The Hessian of $(a, b)$ is $(s, t)$ where $s = 216b^3 + 2a^3$ and $t = -6a^2b$. $H(E_{(a,b)})$ is a new curve different from $E_{(a,b)}$ in $\mathcal{H}$ if and only if $E_{(a,b)}$ is a nonsingular curve in $\mathcal{H}$.

**Definition 2.** We say that an elliptic curve in $\mathbb{P}^2$ is on Hesse form if it is a smooth curve in $\mathcal{H}$.

A curve in $\mathcal{H}$ is smooth if and only if $b \neq 0$ and $a^3+27b^3 \neq 0$. This means there are exactly four singular curves in $\mathcal{H}$ over $\mathbb{C}$ and over $F_q$ where $q \equiv 1 \pmod 3$. In $F_q$ the equation $x^n = 1$ has $d$ solutions, $x = 1, \omega^r, \omega^{2r}, \ldots, \omega^{(d-1)r}$, where $d = (n, q-1)$, $r = (q-1)/d$ and $\omega$ is a primitive root of $F_q$, i.e. $\omega$ is such that $F_q = \{0, 1, \omega, \ldots, \omega^{q-2} \mid \omega^{q-1} = 1\}$. So the polynomial $x^2 + x + 1$ has two distinct roots in $F_q$ if and only if $q \equiv 1 \pmod 3$. The four singular curves are the four triangles

$$T_\infty = E_{(1,0)} \quad T_{-3} = E_{(-3,1)}$$
$$T_{-3\epsilon} = E_{(-3\epsilon,1)} \quad T_{-3\epsilon^2} = E_{(-3\epsilon^2,1)}$$

where $\epsilon$ is a primitive $3th$ root of unity in $\mathbb{C}$ or in $F_q$ where $q \equiv 1 \pmod 3$. These four triangles are the four triples of lines containing the nine flexes mentioned above. We will later refer to the points on these four triangles as $T(= \cup T_\lambda)$.

Since the Hessian of a curve in $\mathcal{H}$ is itself a curve in $\mathcal{H}$ we have that the nine (three) common intersections of the curves in $\mathcal{H}$ are the nine (three) flexes on the elliptic curves, $E_{(a,b)}$, over $\mathbb{C}$ and over $F_q$ where $q \equiv 1 \pmod 3$ (over $\mathbb{R}$ and over $F_q$ where $q \equiv -1 \pmod 3$). The nine common intersections of the curves in $\mathcal{H}$ in $\mathbf{P}^2_{\mathbb{C}}$ is the set

$$\begin{aligned} U_{\mathbb{C}} = \{&(0,1,-1), (0,1,-\epsilon), (0,1,-\epsilon^2), \\ &(1,0,-1), (1,0,-\epsilon), (1,0,-\epsilon^2), \\ &(1,-1,0), (1,-\epsilon,0), (1,-\epsilon^2,0)\} \end{aligned}$$

It is easy to see that these nine points lie on a plane projective cubic if and only if the cubic is a curve in $\mathcal{H}$. The set $U_{\mathbb{R}}$ is the three points

$$U_{\mathbb{R}} = \{(0,1,-1), (1,0,-1), (1,-1,0)\}.$$

From Bezout's theorem it then follows that through a point $P \in \mathbb{P}^2 \setminus U_k$ there is exactly one curve in $\mathcal{H}$. So the curves in $\mathcal{H}$ spans the projective plane $\mathbb{P}^2$.

## 2.2 The Group Law

As a consequence of the Riemann-Roch theorem ([6], [10]) the set of points on an elliptic curve, $E$, over a field $k$ form an abelian group. We can form a group structure on $E$ by fixing any point on $E$ as the identity element, $O$. To define the "ordinary" group law on elliptic curves, however, we need the identity element to be a flex.

We say that points are collinear if they all lie on the same line. The ordinary group law on elliptic curves are characterized by two equivalent properties, one is that the identity element is a flex and the other is the collinearity condition, i.e. three points $P, Q, R$ on $E$ are collinear if and only if $P + Q + R = O$ in the group structure. For an introduction to the ordinary group law see f.ex. [2], [11], [10] or [9].

Now we want to define a group law on the elliptic curves on Hesse form. Let $E$ be an elliptic curve in $\mathcal{H}$. We fix the flex $(0, 1, -1)$, that lies on every curve in $\mathcal{H}$, as the identity element, $O$.

**Definition 3.** A point $x = (x_0, x_1, x_2) \in \mathbb{P}^2$ over the field $k$ is called general if $x \notin T$.

The matrix

$$M_{x,y} = (x_{i+j}y_{i-j})_{(mod3)} = \begin{pmatrix} x_0y_0 & x_1y_2 & x_2y_1 \\ x_1y_1 & x_2y_0 & x_0y_2 \\ x_2y_2 & x_0y_1 & x_1y_0 \end{pmatrix}$$

is called a *Moore matrix*. The determinant of $M_{x,y}$ is given by $detM_{x,y} = (y_0^3 + y_1^3 + y_2^3)x_0x_1x_2 - y_0y_1y_2(x_0^3 + x_1^3 + x_2^3)$, i.e. $detM_{x,y}$ is a curve $E_{(a,b)}$ in $\mathcal{H}$, where $(a, b) = (y_0^3 + y_1^3 + y_2^3, -y_0y_1y_2) \in \mathbb{P}^1$. Note that if $y \in U_k$ then $detM_{x,y} \equiv 0$.

**Lemma 3.** *If $x, y \in E$, then $rkM_{x,y} = 2$.*

*Proof.* We will divide the proof in three parts:

1. When both $x$ and $y \in U_k$ it is easy to see that $rkM_{x,y} = 2$. F.ex. if $x = (1, 0, -\epsilon)$ and $y = (1, -1, 0)$, then

$$M_{x,y} = \begin{pmatrix} 1 & 0 & \epsilon \\ 0 & -\epsilon & 0 \\ 0 & -1 & 0 \end{pmatrix}$$

   has rank 2.
2. When $y \in U_k$ and $x$ is general, then none of the coordinates of $x$ is zero so we can say $x_0 = 1$ and it is also easy to see that $rkM_{x,y} = 2$ in this case. F.ex. let $y = (1, -\epsilon, 0)$ and $x = (1, x_1, x_2)$, then

$$M_{x,y} = \begin{pmatrix} 1 & 0 & -\epsilon x_2 \\ -\epsilon x_1 & x_2 & 0 \\ 0 & -\epsilon & x_1 \end{pmatrix}$$

   and $(1, 0, -\epsilon x_2) = a(-\epsilon x_1, x_2, 0) + b(0, -\epsilon, x_1)$ where $a = -1/\epsilon x_1$ and $b = -\epsilon x_2/x_1$, so $M_{x,y}$ has rank 2.
3. Now let $x$ and $y$ be general, then none of the coordinates are zero so we can say $x_0 = y_0 = 1$, then

$$M_{x,y} = \begin{pmatrix} 1 & x_1y_2 & x_2y_1 \\ x_1y_1 & x_2 & y_2 \\ x_2y_2 & y_1 & x_1 \end{pmatrix}$$

Suppose $M_{x,y}$ has rank 1, then

$$\begin{aligned} &i) \quad x_1 - x_2^2 y_1 y_2 = 0 \\ &ii) \quad x_2 - x_1^2 y_1 y_2 = 0 \\ &iii) \; y_1 - x_1 x_2 y_2^2 = 0 \\ &iv) \; y_2 - x_1 x_2 y_1^2 = 0 \end{aligned}$$

By solving these equations we get $y_1^3 = y_2^3 = 1$. Then $y_0^3 + y_1^3 + y_2^3 = 3$ and $y_0 y_1 y_2 = 1, \epsilon$ or $\epsilon^2$ and $E$ is one of the triangles $T_{-3}$, $T_{-3\epsilon}$ or $T_{-3\epsilon^2}$, a contradiction.

Let $\iota$ be the involution $\iota : x_i \mapsto x_{-i}$, $i \pmod 3$. The unique solution of $M_{\iota(x),y} \cdot z = 0$, when $x, y \in E$ defines a group law on the points of $E$ with the flex $(0, 1, -1)$ as identity. The solution is given as the $2 \times 2$-minors of the matrix after removing one row.

$$M_{\iota(x),y} = \begin{pmatrix} x_0 y_0 & x_2 y_2 & x_1 y_1 \\ x_2 y_1 & x_1 y_0 & x_0 y_2 \\ x_1 y_2 & x_0 y_1 & x_2 y_0 \end{pmatrix}$$

**Theorem 4.** *Let $x = (x_0, x_1, x_2)$ and $y = (y_0, y_1, y_2)$ be two points in $\mathbb{P}_k^2$ on an elliptic curve $E$ in $\mathcal{H}$. The following equations defines a group law on $E$ over $k$:*

1. *let $x, y \in E$ and $x \neq y$, then*

$$x + y = (x_1 x_2 y_0^2 - x_0^2 y_1 y_2, x_0 x_1 y_2^2 - x_2^2 y_0 y_1, x_0 x_2 y_1^2 - x_1^2 y_0 y_2)$$

2. *when $x = y \in E$, then*

$$2x = (x_0 x_2^3 - x_0 x_1^3, x_1^3 x_2 - x_0^3 x_2, x_0^3 x_1 - x_1 x_2^3)$$

*Proof.* 1. When $x$ and $y \in E$ there are at least one nonvanishing equation given by the $2 \times 2$-minors of $M_{\iota(x),y}$. For the operation $x + y$, when $x \neq y$, we choose the symmetric equation given by the $2 \times 2$-minors of $M_{\iota(x),y}$ after eliminating the first row of the matrix. To see that the operation '+' is a group operation we check the group axioms:

(a) The operation '+' is closed on the curve $E$. Let $E_x$ denote the polynomial $ax_0x_1x_2 + b(x_0^3 + x_1^3 + x_2^3)$, where $(a, b) \in \mathbb{P}^1$ is the parametrization of $E$. If $x$ and $y$ are two points on $E$, then $E_x = ax_0x_1x_2 + b(x_0^3 + x_1^3 + x_2^3) = 0$ and $E_y = ay_0y_1y_2 + b(y_0^3 + y_1^3 + y_2^3) = 0$, i.e. $E_x$ and $E_y \in I(E)$, the ideal of $E$. When we factorize $E_{x+y}$ with respect to $E_x$ and $E_y$ we get either $E_x$ or $E_y$ as a factor in each term, so $E_{x+y} \in I(E)$ and $x + y$ is a point on the curve $E$.

(b) $(0, 1, -1)$ is the identity element:

$$\begin{aligned} x + (0, 1, -1) &= (x_0^2, x_0x_1, x_0x_2) \sim (x_0, x_1, x_2) = x \text{ and} \\ (0, 1, -1) + x &= (-x_0^2, -x_0x_1, -x_0x_2) \sim (x_0, x_1, x_2) = x \end{aligned}$$

(c) The involution $\iota(x) = (x_0, x_2. x_1)$ defines the inverse of an element $x = (x_0, x_1, x_2) \in E$:

$$\begin{aligned} x + \iota(x) &= (x_0^2 x_1 x_2 - x_0^2 x_1 x_2, x_0 x_1^3 - x_0 x_2^3, x_0 x_2^3 - x_0 x_1^3) \\ &\sim (0, 1, -1) \end{aligned}$$

(d) The operation '+' is commutative on the curve $E$:

$$\begin{aligned} x + y = (x_1 x_2 y_0^2 - x_0^2 y_1 y_2, x_0 x_1 y_2^2 - x_2^2 y_0 y_1, x_0 x_2 y_1^2 - x_1^2 y_0 y_2) \quad \sim \\ (y_1 y_2 x_0^2 - y_0^2 x_1 x_2, y_0 y_1 x_2^2 - y_2^2 x_0 x_1, y_0 y_2 x_1^2 - y_1^2 x_0 x_2) \quad = \quad y + x \end{aligned}$$

(e) The operation '+' is assosiative on the curve $E$. Let $x, y, z$ be three points on $E$ and let $p_1 = (x + y) + z$ and $p_2 = x + (y + z)$. The two points $p_1$ and $p_2$ are equal on $E$ if and only if $x, y, z \in E$. This will be shown in appendix A.

2. When $x = y$ the equation we chose in 1 is vanishing so we choose the equation we get by eliminating the second or third row in the matrix $M_{\iota(x),x}$.

Let $E[n] = \{p \in E \mid np = 0\}$ denote the $n$-torsion subgroup of $E$.

**Theorem 5.** *Let $k = K$, then for every $n \geq 2$ $E[n]$ is isomorphic to $\mathbb{Z}_n \times \mathbb{Z}_n$.*

*Proof.* See f.ex. [6].

**Lemma 4.** *A point $p \in E$ is a flex if and only if $3p = 0$.*

*Proof.* If $p$ is a flex, then $p \in U_k$ and it is easy to check that the points of $U_k$ are three torsion points on $E$. We have $U_k \simeq \mathbb{Z}_3$ if $k = \mathbb{R}$ or $k = F_q$ with $q \equiv 2 \pmod 3$ and $U_k \simeq \mathbb{Z}_3 \times \mathbb{Z}_3$ if $k = \mathbb{C}$ or $k = F_q$, $q \equiv 1 \pmod 3$. We then know from theorem 5 that the elements of $U_k$ are all the 3-torsion points on $E$.

### Collinearity

An elliptic curve $E$ in $\mathcal{H}$ satisfy the collinearity condition, i.e. three points on $E$ are collinear if and only if their sum equals zero on $E$.

If we have three collinear points, $x = (x_0, x_1, x_2)$, $y = (y_0, y_1, y_2)$ and $z = (z_0, z_1, z_2) \in \mathbb{P}^2$ then the matrix

$$M = \begin{pmatrix} x_0 & x_1 & x_2 \\ y_0 & y_1 & y_2 \\ z_0 & z_1 & z_2 \end{pmatrix}$$

has rank 2 and $detM = 0$.

**Proposition 1.** *Three points $x, y, z \in E$ are collinear if and only if $x + y + z = (0, 1, -1) \in E$.*

*Proof.* Suppose first that $x, y, z \in E$ are collinear. When $x, y \in E$, $x+y$ is the solution of $M_{\iota(x),y} \cdot a = 0$, $a = (a_0, a_1, a_2) \in E$. Let $L_i$ be the three equations from $M_{\iota(x),y} \cdot a = 0$:

$$\begin{aligned} L_1 &: x_0y_0a_0 + x_2y_2a_1 + x_1y_1a_2 = 0 \\ L_2 &: x_2y_1a_0 + x_1y_0a_1 + x_0y_2a_2 = 0 \\ L_3 &: x_1y_2a_0 + x_0y_1a_1 + x_2y_0a_2 = 0 \end{aligned}$$

then

$$L_3 - L_2 = (x_1y_2 - x_2y_1)a_0 + (x_0y_1 - x_1y_0)a_1 + (x_2y_0 - x_0y_2)a_2 = 0$$

and since $detM = 0$ we get

$$(x_1y_2 - x_2y_1)z_0 + (x_2y_0 - x_0y_2)z_1 + (x_0y_1 - x_1y_0)z_2 = 0$$

Bezout's theorem (2) tells us we can't have more than three points in the intersection $E \cap detM$, so we have to have $a = (a_0, a_1, a_2) = (z_0, z_2, z_1) = -z$ and $x + y = -z$.

Now suppose $x+y+z = (0, 1, -1) \in E$ then $L_i(x, \iota(y), z) = 0$, $(i = 1, 2, 3)$, where $L_i$ are the three equations from $M_{x,\iota(y)} \cdot z = 0$

$$\begin{aligned} L_1 &: x_0y_0z_0 + x_1y_1z_1 + x_2y_2z_2 = 0 \\ L_2 &: x_0y_1z_2 + x_1y_2z_0 + x_2y_0z_1 = 0 \\ L_3 &: x_0y_2z_1 + x_1y_0z_2 + x_2y_1z_0 = 0 \end{aligned}$$

$0 = L_2 - L_3 = detM$ and we have that $x, y, z$ are collinear.

**The group law on the singular curves in $\mathcal{H}$**

A curve $F$ in $\mathbb{P}^2$ is singular at a point $p$ if the three partial derivatives $(\partial F/\partial x_i)(p)$ $(i = 0, 1, 2)$ are all zero. The partial derivatives of a curve $E_{(a,b)} \in \mathcal{H}$ is $\partial E_{(a,b)}/\partial x_i = ax_jx_k + 3bx_i^2$ so $E_{(a,b)}$ is singular if $a^3 + 27b^3 = 0$ or if $b = 0$.

We have either 2 or 4 singular curves in $\mathcal{H}$ dependent of the characteristic of the field $k$. If $k = F_q$, $q \equiv 0 \pmod 3$ we have two singular curves, the triangle $T_\infty$ and the triple line $(x_0 + x_1 + x_2)^3$. If $k = \mathbb{R}$ or $k = F_q$, $q \equiv 2 \pmod 3$ we have two singular curves, the triangle $T_\infty$ and the curve $(x_0 + x_1 + x_2)(x_0^2 + x_1^2 + x_2^2 - x_0x_1 - x_0x_2 - x_1x_2)$. And when $k = \mathbb{C}$ or $k = F_q$, $q \equiv 1 \pmod 3$ we have four singular curves, the four triangles:

$$\begin{aligned} T_\infty &: x_0x_1x_2 = 0 \\ T_{-3} &: (x_0 + x_1 + x_2)(\epsilon x_0 + \epsilon^2 x_1 + x_2)(\epsilon^2 x_0 + \epsilon x_1 + x_2) = 0 \\ T_{-3\epsilon} &: (x_0 + \epsilon x_1 + x_2)(\epsilon x_0 + x_1 + x_2)(\epsilon^2 x_0 + \epsilon^2 x_1 + x_2) = 0 \\ T_{-3\epsilon^2} &: (x_0 + \epsilon^2 x_1 + x_2)(\epsilon x_0 + \epsilon x_1 + x_2)(\epsilon^2 x_0 + x_1 + x_2) = 0 \end{aligned}$$

On all these singular curves in $\mathcal{H}$, if we exclude the singular points, we have the same group law as we have on the elliptic curves in $\mathcal{H}$. We can extend the group operation on the elliptic curves in $\mathcal{H}$:

$$\begin{array}{ccc} E \times E & \longrightarrow & E \\ (x, y) & \mapsto & x + y \end{array}$$

to a rational map:

$$\begin{array}{ccc} \mathbf{P}^2 \times \mathbf{P}^2 & \dashrightarrow & \mathbf{P}^2 \\ (x, y) & \mapsto & x + y \end{array}$$

where $x + y$ is defined as in theorem 4.

### 2.3 The $j$-invariant

Algebraic curves over $K$ are classified by the discrete invariant genus, $g$, and by a point on the variety of moduli of curves of genus $g$, $\mathcal{M}_g$, which is a continuous invariant. $\mathcal{M}_g$ is an irreducible algebraic variety of dimension 1 if $g = 1$ or dimension $3g - 3$ if $g \geq 2$. Algebraic curves with $g = 1$ are called elliptic curves. For elliptic curves the point on $\mathcal{M}_g$ is called the $j$-invariant and it classifies elliptic curves up to isomorphism.

**Theorem 6.** *Let $k = K$ and $char(k) \neq 2$. Two elliptic curves $X$ and $X'$ over $k$ are isomorphic if and only if $j(X) = j(X')$.*

*Proof.* See [6].

**Proposition 2.** *Let $c = \frac{a}{b}, b \neq 0$. The $j$-invariant of an elliptic curve on Hesse form, $E_{(a,b)} = E_c : cx_0x_1x_2 + x_0^3 + x_1^3 + x_2^3$, over $K$ is given by*

$$j(E_c) = -\frac{c^3(c^3 - 216)^3}{c^9 + 81c^6 + 2187c^3 + 19683} = -\frac{c^3(c^3 - 216)^3}{(c+3)^3(c+3\epsilon)^3(c+3\epsilon^2)^3}.$$

*Proof.* We can transform $E_c$ into classic Weierstrass form $F_{(g_2,g_3)} : y^2z = 4x^3 - g_2xz^2 - g_3z^3$ using f.ex. Nagell's algorithm [3]. It follows that $E_c \simeq F_{(g_2,g_3)}$ where $(g_2, g_3) = (\frac{1}{12}c^4 - 18c, 27 - \frac{1}{216}c^6 - \frac{5}{2}c^3)$. The $j$-invariant of $E_c$ is then given by theorem 6 and $j(F_{(g_2,g_3)}) = 1728\frac{g_2^3}{g_2^3 - 27g_3^2}$.

**Proposition 3.** *Let $V = \{-3, -3\epsilon, -3\epsilon^2\}$. The transformation*

$$\begin{array}{ccc} \mathbf{A}^1 \setminus V & \longrightarrow & K \\ c & \mapsto & j(E_c) \end{array}$$

*is surjective and $12 : 1$, except over $j = 0$ where it is $4 : 1$.*

*Proof.* We will prove this proposition in Sect. 2.4.

**Corollary 2.** *1. Every elliptic curve $E$ over $K$ is isomorphic to a smooth curve in $\mathcal{H}$,*

$$E_c : cx_0x_1x_2 + x_0^3 + x_1^3 + x_2^3 = 0$$

*for $c \in \mathbf{A}^1 \setminus V$, where $V = \{-3, -3\epsilon, -3\epsilon^2\}$.*

*2. If we have an elliptic curve on classic Weierstrass form*

$$F_{(g_2,g_3)} : y^2z = 4x^3 - g_2xz^2 - g_3z^3,$$

*we can transform it into Hesse form, $E_c$, by solving*

$$j(F_{(g_2,g_3)}) = j(E_c)$$

*with respect to $c$.*

*Proof.*

1. $\{j(E_c)\} = K$, so there exists a curve in every isomorphism class of elliptic curves in $\mathcal{H}$.
2. Follows from theorem 6.

### 2.4 Symmetries of Curves in $\mathcal{H}$

The group $SL(3,k)$ acts on points in $\mathbf{P}_k^2$, if $\theta \in SL(3,k)$ then $\theta((x_0,x_1,x_2)) = (\theta(x_0),\theta(x_1),\theta(x_2))$.

#### The Heisenberg group of dimension 3, $H_3$

Let $\sigma$ and $\tau$ be two elements in $SL(3,K)$ such that $\sigma(x_i) = x_{i+1}$ and $\tau(x_i) = \epsilon^i x_i$. As matrices

$$\sigma = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad \tau = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon^2 \end{pmatrix},$$

where $\epsilon$ is a primitive third root of unity in $K$. The group generated by the matrices $\sigma$ and $\tau$ is called the *Heisenberg group of dimension 3* and is denoted $H_3$. $H_3$ is a finite nonabelian subgroup of $SL(3,K)$ with $[\sigma,\tau] = \epsilon id$. The order of $H_3$ is 27 [1].

**Proposition 4.** *$H_3$ leaves the curves in $\mathcal{H}$ invariant and operates on the points on an elliptic curve in $\mathcal{H}$ by translation by 3-torsion points.*

*Proof.* By looking at the generators for $\mathcal{H}$ we see that $H_3$ acts trivially on the elements of $\mathcal{H}$.

The generators of $H_3$ acts on points in $E$ as translation with 3-torsion points. Let $x \in E$, $x = (x_0,x_1,x_2)$, then $\sigma(x) = (x_1,x_2,x_0) = (x_0,x_1,x_2) + (1,-1,0)$ and $\tau(x) = (x_0,\epsilon x_1,\epsilon^2 x_2) = (x_0,x_1,x_2) + (0,1,-\epsilon)$. Further, the center of $H_3$ is $Z(H_3) = \{id, \epsilon id, \epsilon^2 id\}$ and $H_3$ is a central extension

$$1 \longrightarrow Z(H_3) \longrightarrow H_3 \longrightarrow \mathbb{Z}_3 \times \mathbb{Z}_3 \longrightarrow 1$$

where $\sigma \mapsto (1,0)$ and $\tau \mapsto (0,1)$.

### The normalizer of $H_3$ in $SL(3,K)$, $N_3$

The *normalizer of* $H_3$ *in* $SL(3,K)$ is the group that consists of those $\rho \in SL(3,K)$ such that $\rho H_3 = H_3\rho$ and is denoted $N_3$. $N_3$ is generated by the matrices $\sigma$, $\tau$, $\delta$ and $\nu$ where

$$\sigma = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \tau = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon^2 \end{pmatrix},$$
$$\delta = k_\delta \begin{pmatrix} 1 & 1 & 1 \\ 1 & \epsilon & \epsilon^2 \\ 1 & \epsilon^2 & \epsilon \end{pmatrix} \text{ and } \nu = \epsilon^{\frac{2}{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \epsilon^2 & 0 \\ 0 & 0 & \epsilon^2 \end{pmatrix}.$$

This group is a finite subgroup of $SL(3,K)$. We have $|N_3| = 648$. $k_\delta$ is a constant which make $det(\delta) = -3k_\delta^3(2\epsilon - 1)$ equal to 1 over the field in which we are working. For example in $char(k) = 2$ we have $k_\delta = 1$, in $char(k) = 5$ we have $k_\delta = (4 + 3\epsilon)^{\frac{1}{3}}$ and when $k = \mathbb{C}$ we have $k_\delta = -\frac{\sqrt{3}}{3}i$.

Remark; $N_3$ is not defined over fields of characteristic 3, because the determinant of $\delta$ then is zero.

### The group $G = N_3/H_3 \simeq SL(2,\mathbf{Z}_3)$

The factor group $G = N_3/H_3 \simeq SL(2,\mathbf{Z}_3)$ is generated by $\bar{\delta}$ and $\bar{\nu}$:

$$\begin{array}{ccccc} H_3 & \longrightarrow & N_3 & \longrightarrow & N_3/H_3 \\ \downarrow & & \downarrow & & \wr \\ H_3 & \longrightarrow & N_3 & \longrightarrow & SL(2,\mathbf{Z}_3) \\ & & \sigma, \tau & \mapsto & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ & & \delta & \mapsto & \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix} \\ & & \nu & \mapsto & \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \end{array}$$

$$SL(2,\mathbf{Z}_3) = (\bar{\delta} = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}, \bar{\nu} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix})$$

### The group $H \simeq A_4$

The group $H = G/ < \bar{\delta}^2 > \simeq A_4$ is generated by $\tilde{\delta}$ and $\tilde{\nu}$:

$$\begin{array}{ccccccccc} 1 & \longrightarrow & < \bar{\delta}^2 > & \longrightarrow & SL(2,\mathbf{Z}_3) & \longrightarrow & A_4 & \longrightarrow & 1 \\ & & & & \bar{\delta} & \mapsto & \tilde{\delta} & & \\ & & & & \bar{\nu} & \mapsto & \tilde{\nu} & & \end{array}$$

$H_3$ acts trivially on $\mathcal{H}$. This induces an action of $SL(2, \mathbf{Z}_3)$ on $\mathcal{H}$, where the orbit consists of isomorphic curves. Further

$$\bar{\delta}^2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} \in G$$

also acts trivially on $\mathcal{H}$, so we get an action of $H$ on $\mathcal{H}$.

Now we are ready to prove proposition 3 from p. 131:

*Proof.* Proof of proposition 3: There exists a $c \in K$ s.t. $j(E_c) \neq 0$, so the transformation is surjective. If $j(E_c) = 0$, then $c = 0$ or $c = \sqrt[3]{216}$, and therefore the isomorphism class of $E_0 : x_0^3 + x_1^3 + x_2^3 = 0$ consists of 4 curves in $\mathcal{H}$. When $j(E_c) \neq 0$ the isomorphism classes of $E_c$ concists of 12 curves, this follows from the formula for $j(E_c)$ and from the action of the group $H$ on the curves in $\mathcal{H}$.

## 3 Torsion Points and Division Polynomials

Let $E$ be an elliptic curve. For $n \in \mathbb{Z}$ let $\Phi_n$ be the *multiplication-by-n* map:

$$\begin{array}{rccc} \Phi_n : E & \xrightarrow{n\cdot} & E \\ p & \mapsto & np = \underbrace{p + \cdots + p}_{n\ times} \end{array}$$

The image of $\Phi_n$ is a subgroup $nE$ of $E$ and the kernel of $\Phi_n$ is the $n$-torsion subgroup of $E$, $E[n] = \{p \in E \mid np = 0\} = \Phi_n^{-1}(0)$.

**Proposition 5.** *Let $E$ be an elliptic curve and $n \in \mathbb{Z}$, $n \neq 0$.*

1. *$deg\Phi_n = n^2$.*
2. *If $char(k) = 0$ or if $n$ is prime to $char(k)$, then*

$$E[n] \simeq \mathbb{Z}_n \times \mathbb{Z}_n \ .$$

3. *If $char(k) = p$, then either*

$$E[p^r] \simeq 0 \quad \text{for all } r = 1, 2, \ldots; \text{ or}$$
$$E[p^r] \simeq \mathbb{Z}_{p^r} \quad \text{for all } r = 1, 2, \ldots$$

*Proof.* See [10].

If $E[p^r] = 0$ then $E$ is said to be *supersingular* else $E$ is said to be *ordinary*. Other equivalent definitions of supersingular elliptic curves can be found in [8].

In the next section we will take a closer look at the division polynomials on $E \in \mathcal{H}$. The results in Sect. 3.1 are taken from my master thesis [4].

If we only want to find the $n$-torsion points on $E$ we will describe a computational easier way in Sect. 3.2.

### 3.1 A First Algorithm for Finding the $n$-torsion Points on $E \in \mathcal{H}$

In this section we will work over the field of complex numbers, $\mathbb{C}$.

Let $E$ be an elliptic curve in $\mathcal{H}$. The formulas for addition and doubling on $E$ defined in theorem 4 are independent of chosen elliptic curve $E \in \mathcal{H}$. We can therefore extend $\Phi_n$ to a rational map:

$$\begin{aligned} \Phi_n : \mathbf{P}^2 &\dashrightarrow \mathbf{P}^2 \\ p &\mapsto np \end{aligned}$$

$\Phi_n$ maps a point $p = (x_0, x_1, x_2) \in \mathbb{P}^2$ to the point $np = (n-m)p + mp$ for a $m \in \mathbb{Z}$. Common factors in the polynomials representing $np$ are contained in the base locus and can be removed. By proposition 5 we expect to find that the polynomials representing $np$ after removing common factors are of degree $n^2$. We will show this for $n$ up to 10 in the next sections.

Let $x \in E$ and $nx = (F_0, F_1, F_2)$, where the $F_i$'s are polynomials of degree $n^2$ representing $nx$. $x$ is a $n$-torsion point on $E$ if $nx \sim (0, 1, -1)$ i.e. if $x \in Z(F_0) \cap Z(F_1 + F_2)$.

**Definition 4.** The set of $n$-torsion points on $E$ s.t. $mx \neq 0$ whenever $m \mid n$ and $m < n$ are called the primitive $n$-torsion points of $E$.

**Lemma 5.** *The number of primitive $n$-torsion points of $E$, $\mathfrak{a}_n$, can be found by executing the following recursive Maple-procedure:*

```
a := proc(n::posint)
  local m, j; option remember;
    if n = 1 then 1
    else
      j := n^2 - 1;
      for m in divisors(n) do if m ≠ n and m ≠ 1 then j := j - a(m)
      fi od;
      RETURN(j)
    fi
  end
```

We will state some claims and show them for $n$ up to 10 when $3 \nmid n$ in the following sections. The case with $3n$-torsion points will be discussed in Sect. 3.1.

*Claim.* If $n \neq 3$ then for all $m$ s.t. $m \mid n$ and $3 \nmid m$ $F_0$ and $F_1 + F_2$ have a common factor $P_m$ of degree $\frac{\mathfrak{a}_m}{3}$. The $P_m$'s are irreducible polynomials which are $SL(2, \mathbb{Z}_3)$-invariant.

When $3 \nmid n$ we can write the polynomials representing $nx$ as products of $P_m$'s:

$$nx = (x_0 \prod_{\substack{m,\\ m|n}} (P_m)(\tau P_m)(\tau^2 P_m),\ \sigma^n(x_0) \prod_{\substack{m,\\ m|n}} (\sigma^n P_m)(\tau\sigma^n P_m)(\tau^2\sigma^n P_m),$$
$$\sigma^{-n}(x_0) \prod_{\substack{m,\\ m|n}} (\sigma^{-n} P_m)(\tau\sigma^{-n} P_m)(\tau^2\sigma^{-n} P_m)),$$

where $\sigma$ and $\tau$ are the two generators of $H_3$ (2.4).

Let $X_m$ be the curve defined by $Z(P_m)$.

*Claim.* The intersection of $X_m$ and $E$ is exactly the primitive $m$-torsion points of $E$.

*Claim.* If $X_m$ is singular then the singularities are exactly the twelve vertices of the four triangles in $\mathcal{H}$. So we have no singular intersection points with $E$.

Since the $P_m$'s are $SL(2, \mathbb{Z}_3)$-invariant $X_m$ intersect isomorphic with each curve in the same isomorphism class of curves in $\mathcal{H}$. So $X_m$ intersects isomorphic with each of the 4 triangles in $\mathcal{H}$, and we need only to check the triangle $T_\infty : x_0x_1x_2 = 0$ when we want to study the intersection between $X_m$ and the singular curves in $\mathcal{H}$.

**Definition 5.** The number of branches of the curve $X_m$ at a point $q \in X_m$ is the number of locally irreducible components of $X_m$ in a sufficient small neighbourhood of $q$.

Denote the singularities of $X_m$: $[(x, y, z), a, b, c]$, where $(x, y, z) = q$ is the singular point, $a$ is the multiplicity of $q$, $b$ is the intersection multiplicity between $X_m$ and one of the triangles in $\mathcal{H}$ at $q$ and $c$ is the number of branches of $X_m$ at $q$.

*Claim.* The intersection multiplicity is 1 for all points $p \in X_m \cap E$.

To prove this when $X_m$ is nonsingular we use Hurwitz formula (7) on the morphism

$$f : X_m \xrightarrow[\mathfrak{a}_m:1]{(x_0x_1x_2, x_0^3+x_1^3+x_2^3)} \mathbf{P}^1$$

and show that $f$ is only ramified at points $(a, b) \in \mathbf{P}^1$ that correspond to singular curves in $\mathcal{H}$. (This morphism is defined for all $p \in X_m$ since $3 \nmid m$.)

When $X_m$ is singular we use Hurwitz formula on the minimal desingularisation $X_{m_{blown-up}}$ of $X_m$ as described in lemma 6.

**Theorem 7.** *(Hurwitz). Let $f : X \to Y$ be a finite separable morphism of complete, nonsingular curves. Let $n = degf$ and let $R$ be the ramification divisor of $f$. Then*

$$2g(X) - 2 = n \cdot (2g(Y) - 2) + degR,$$

*where*

$$degR = \sum_{P \in X} (e_p - 1).$$

*Proof.* See [6].

The ramification index, $e_p$, is the intersection multiplicity between $X_m$ and a curve in $\mathcal{H}$ at the point $p$.

**Definition 6.** A morphism $f : X \to Y$, defined as in theorem 7, is said to be *unramified at a point $y \in Y$ if the number of inverse images of $y$ equal $degf$, else $f$ is said to be ramified at $y \in Y$.*

**Lemma 6.** *Let* $f : X_{m_{blown-up}} \xrightarrow[\mathfrak{a}_m:1]{(x_0x_1x_2, x_0^3+x_1^3+x_2^3)} \mathbf{P}^1$ *be a finite separable morphism between complete nonsingular curves, where $X_{m_{blown-up}}$ is the minimal desingularisation of $X_m$. The following are equivalent:*

*i. computing the degree of the ramification divisor of $f$, $degR$, using Hurwitz formula on $f$.*
*ii. computing $degR = \sum_{q \in X_m}(I_q - r)$, where $I_q$ is the intersection multiplisity between $X_m$ and a curve in $\mathcal{H}$ at $q$ and $r$ is the number of branches of $X_m$ at the point $q$.*

*Proof.*

$$\begin{array}{ccc} p_1,\ldots,p_r \in X_{m_{blown-up}} & & \\ \downarrow & \searrow^{f} & \\ q \in X_m & \longrightarrow & \mathbf{P}^1 \end{array}$$

We blow up $X_m$ at the singular point $q$ with $r$ branches s.t. $q$ splits into $r$ points $p_1, ..., p_r \in X_{m_{blown-up}}$. Locally at $q \in X_m$ we can write $X_m$ as an union of irreducible components $Z_1 \cup \cdots \cup Z_r$. Then $I_q = \sum_{i=1}^{r}(I_{Z_i,q}) = \sum_{i=1}^{r}(I_{p_i})$, where $\sum_{i=1}^{r}(I_{Z_i,q})$ is the sum of the intersection multiplicities between $Z_i$ and a curve in $\mathcal{H}$ at the point $q$. And $\sum_{i=1}^{r}(I_{p_i})$ is the sum of the intersection multiplicities between $X_{m_{blown-up}}$ and a curve in $\mathcal{H}$ at the points $p_i$, $(i = 1 \ldots r)$.

**Algorithm 1** *An algorithm for finding the n-torsion points on $E$ when $3 \nmid n$ and to prove the claims:*

1. *Compute $nx$.*
2. *Compute the crossproduct $\overrightarrow{nx} \times \overrightarrow{(0,1,-1)}$ and find the common factors $P_m$ of $F_0$ and $F_1 + F_2$.*
3. *Check if $X_m$ is a singular curve.*
   *(Use f.ex. the Maple-procedure* algcurves[singularities]*.)*
4. *Check that $X_m$ intersects $E$ in $\frac{\mathfrak{a}_m}{3}$ different points.*

### $3n$-torsion points

Let $E$ be an elliptic curve in $\mathcal{H}$ and let $p, q$ be two points on $E$ where $q$ is given by an element $\alpha \in H_3$:

$$\begin{array}{ccc} \alpha : E & \longrightarrow & E \\ p & \mapsto & p+q \end{array}$$

This map is the *translation-by-q* map and it is an isomorphism with inverse element $\alpha^{-1} \in H_3$, (but it is not a group homomorphism). We have seen in Sect. 2.4 that $H_3$ operates on an elliptic curve in $\mathcal{H}$ by translation by 3-torsion points, so $q$ is an element of $U_k$. For $\alpha \in H_3$ consider the following composite map:

$$\begin{array}{ccccc} E & \xrightarrow{\alpha} & E & \xrightarrow{3n\cdot} & E \\ p & \mapsto & p+q & \mapsto & 3n(p+q) \\ & & & & = 3np + 3nq \end{array}$$

If $p \in ker(\Phi_n)$ then $3n(p+q) = 3np + 3nq = 0$. So we find the $3n$-torsion points on $E$ when $n \neq 3$ as the product of the orbit of $H_3$ on $P_n$, i.e. the $3n$-torsion points on $E$ are exactly the intersection between $E$ and the curve given by the product of the polynomials representing $nx$.

We find the 3-torsion points on $E$ by intersecting $E$ with any other curve in $\mathcal{H}$. But $H_3$ acts trivially on $\mathcal{H}$ so we have to consider the case with the 9-torsion points on $E$ as a special case.

Let $V_d$ denote the set of homogeneous elements of degree $d$ in the graded polynomial ring $K[x_0, x_1, x_2] = \bigoplus_{d \geq 0} V_d$. We have 8 cubics beside $\mathcal{H}$ in $V_3$ that are invariant as curves under the action of $H_3$:

$$B_1 : x_0^3 + \epsilon x_1^3 + \epsilon^2 x_2^3 \qquad B_2 : x_0^3 + \epsilon^2 x_1^3 + \epsilon x_2^3$$

$$B_3 : x_0^2 x_1 + x_1^2 x_2 + x_2^2 x_0 \qquad B_4 : x_0^2 x_1 + \epsilon x_1^2 x_2 + \epsilon^2 x_2^2 x_0$$

$$B_5 : x_0^2 x_1 + \epsilon^2 x_1^2 x_2 + \epsilon x_2^2 x_0 \quad B_6 : x_0^2 x_2 + x_1^2 x_0 + x_2^2 x_1$$

$$B_7 : x_0^2 x_2 + \epsilon x_1^2 x_0 + \epsilon^2 x_2^2 x_1 \quad B_8 : x_0^2 x_2 + \epsilon^2 x_1^2 x_0 + \epsilon x_2^2 x_1$$

$B_1, \ldots, B_8$ together with $x_0 x_1 x_2$ are the polynomials that represents $3x$ on $E$, and therefore $B_1, \ldots, B_8$ intersected with $E$ give the primitive 9-torsion points on $E$.

**Lemma 7.** *If $n \neq 3$ the curve associated with the polynomial $P_{3n}$ intersected with $E$ gives the $3n$-torsion points on $E$:*

$$P_{3n} = \prod_{\alpha \in H_3} \alpha(P_n),$$

*where $P_n$ is the polynomial which curve intersected with $E$ gives the $n$-torsion points on $E$. If $n = 3$ then*

$$P_9 = (x_0x_1x_2)\prod_{i=1}^{8} B_i,$$

*where $B_i$,$(i = 1..8)$, are the 8 $H_3$-invariant cubics mentioned above.*

In the next sections we will find the $n$-torsion points on $E$ for $n$ up to 10.

### 2-torsion points on $E$

| The 2-torsion points on $E$ consist of: |
|---|
| 1 origin |
| 3 primitive 2-torsion points |

We follow algorithm 1:

1. Compute $2x$:

$$\begin{aligned} 2x = (F_0, F_1, F_2) &= (x_0x_2^3 - x_0x_1^3, \quad x_1^3x_2 - x_0^3x_2, \quad x_0^3x_1 - x_1x_2^3) \\ &= (x_0(x_2 - x_1)(x_2 - \epsilon x_1)(x_2 - \epsilon^2 x_1), \\ &\quad x_2(x_1 - x_0)(x_1 - \epsilon x_0)(x_1 - \epsilon^2 x_0), \\ &\quad x_1(x_0 - x_2)(x_0 - \epsilon x_2)(x_0 - \epsilon^2 x_2)) \\ &= (x_0(P_2)(\tau P_2)(\tau^2 P_2), \quad x_2(\sigma^2 P_2)(\tau\sigma^2 P_2)(\tau^2\sigma^2 P_2), \\ &\quad x_1(\sigma P_2)(\tau\sigma P_2)(\tau^2\sigma P_2)) \end{aligned}$$

   The curve given by the product of these 12 polynomials intersected with $E$ gives the 6-torsion points on $E$, see Sect. 3.1 and 3.1.
2. Compute the crossproduct $\overrightarrow{nx} \times \overrightarrow{(0,1,-1)}$ and find the common factors $P_m$ of $F_0$ and $F_1 + F_2$:
   $F_0 : x_0x_2^3 - x_0x_1^3 = x_0(x_2 - x_1)(x_2 - \epsilon x_1)(x_2 - \epsilon^2 x_1) = 0$
   and
   $F_1 + F_2 : x_0^3x_2 - x_1^3x_2 - x_0^3x_1 + x_1x_2^3 = (x_2 - x_1)(x_1x_2^2 + x_1^2x_2 + x_0^3) = 0$
   Common factors of $F_0$ and $F_1 + F_2$ is the line

$$X_2 : x_2 - x_1 = 0$$

3. Check if $X_m$ is a singular curve:
   $X_2$ is a nonsingular curve.
4. Check that $X_m$ intersects $E$ in $\frac{a_m}{3}$ different points:
   $X_2$ intersects $E$ at 3 different points, this is because the ramification of the morphism

$$f : X_2 \xrightarrow[3:1]{(x_0x_1x_2, x_0^3 + x_1^3 + x_2^3)} \mathbf{P}^1$$

is a point of ramification index 2 on each of the four triangles. Hurwitz theorem says $degR = 4$, so these are all the ramifications. The four points (vertices) are

$$(1,0,0), \quad (1,1,1), \quad (1,\epsilon,\epsilon) \quad \text{and} \quad (1,\epsilon^2,\epsilon^2).$$

Conclusion:
The 2-torsion points on $E$ are the origin and the three intersection points between $E$ and the line $X_2$.

The three primitive 2-torsion points, $p_1, p_2, p_3$ are collinear so $p_1 + p_2 = -p_3$.

### 3-torsion points on $E$

We have seen in Sect. 2.1 that the set of 3-torsion points on $E$ are the set $U_k$. The primitive 3-torsion points are the set $U_k \setminus (0,1,-1)$.

$3x$ can be written as:

$$\begin{aligned}
&3x = (x_0^4x_1^4x_2 + x_0^4x_1x_2^4 + x_0x_1^4x_2^4 - x_0^7x_1x_2 - x_0x_1^7x_2 - x_0x_1x_2^7,\\
&3x_0^3x_1^3x_2^3 - x_0^6x_1^3 - x_1^6x_2^3 - x_0^3x_2^6, \quad 3x_0^3x_1^3x_2^3 - x_0^6x_2^3 - x_0^3x_1^6 - x_1^3x_2^6)\\
&= (-x_0x_1x_2(x_0^3 + \epsilon x_1^3 + \epsilon^2x_2^3)(x_0^3 + \epsilon^2x_1^3 + \epsilon x_2^3),\\
&-(x_0^2x_1 + x_1^2x_2 + x_2^2x_0)(x_0^2x_1 + \epsilon x_1^2x_2 + \epsilon^2x_2^2x_0)(x_0^2x_1 + \epsilon^2x_1^2x_2 + \epsilon x_2^2x_0),\\
&-(x_0^2x_2 + x_1^2x_0 + x_2^2x_1)(x_0^2x_2 + \epsilon x_1^2x_0 + \epsilon^2x_2^2x_1)(x_0^2x_2 + \epsilon^2x_1^2x_0 + \epsilon x_2^2x_1))\\
&= (-x_0x_1x_2(B_1)(B_2), \quad -(B_3)(B_4)(B_5), \quad -(B_6)(B_7)(B_8))
\end{aligned}$$

where $B_i \in V_3$ are the 8 cubics invariant as curves under the action of $H_3$, see p. 138. These nine cubics intersected with $E$ give the 9-torsion points on $E$, see Sect. 3.1 and 3.1.

### 4-torsion points on $E$

| The 4-torsion points on $E$ consist of: |
|---|
| 1 origin |
| 3 primitive 2-torsion points |
| 12 primitive 4-torsion points |

1. Compute $4x$:

$$\begin{aligned}4x = &(x_0(x_2-x_1)(x_2-\epsilon x_1)(x_2-\epsilon^2 x_1)(x_0^3x_1+x_0^3x_2-x_1^3x_2-x_2^3x_1)\\ &(x_0^3x_1+\epsilon x_0^3x_2-\epsilon x_1^3x_2-x_2^3x_1)(x_0^3x_1+\epsilon^2 x_0^3x_2-\epsilon^2 x_1^3x_2-x_2^3x_1),\\ &x_1(x_0-x_2)(x_0-\epsilon x_2)(x_0-\epsilon^2 x_2)(x_1^3x_0+x_1^3x_2-x_0^3x_2-x_2^3x_0)\\ &(x_1^3x_0+\epsilon x_1^3x_2-\epsilon x_0^3x_2-x_2^3x_0)(x_1^3x_0+\epsilon^2 x_1^3x_2-\epsilon^2 x_0^3x_2-x_2^3x_0),\\ &x_2(x_1-x_0)(x_1-\epsilon x_0)(x_1-\epsilon^2 x_0)(x_2^3x_1+x_2^3x_0-x_1^3x_0-x_0^3x_1)\\ &(x_2^3x_1+\epsilon x_2^3x_0-\epsilon x_1^3x_0-x_0^3x_1)(x_2^3x_1+\epsilon^2 x_2^3x_0-\epsilon^2 x_1^3x_0-x_0^3x_1)\\ =&(x_0(P_2)(\tau P_2)(\tau^2 P_2)(P_4)(\tau P_4)(\tau^2 P_4),\\ &x_1(\sigma P_2)(\tau\sigma P_2)(\tau^2\sigma P_2)(\sigma P_4)(\tau\sigma P_4)(\tau^2\sigma P_4),\\ &x_2(\sigma^2 P_2)(\tau\sigma^2 P_2)(\tau^2\sigma^2 P_2)(\sigma^2 P_4)(\tau\sigma^2 P_4)(\tau^2\sigma^2 P_4))\end{aligned}$$

   The curve given by the product of these polynomials representing $4x$ intersected with $E$ gives the 12-torsion points on $E$, see Sect. 3.1.
2. Compute the crossproduct $\overrightarrow{nx} \times \overrightarrow{(0,1,-1)}$ and find the common factors $P_m$ of $F_0$ and $F_1 + F_2$:

$$\begin{aligned}F_0 : &x_0(x_2-x_1)(x_2-\epsilon x_1)(x_2-\epsilon^2 x_1)(x_0^3x_1+x_0^3x_2-x_1^3x_2-x_2^3x_1)\\ &(x_0^3x_1+\epsilon x_0^3x_2-\epsilon x_1^3x_2-x_2^3x_1)(x_0^3x_1+\epsilon^2 x_0^3x_2-\epsilon^2 x_1^3x_2-x_2^3x_1)\\ &=0\end{aligned}$$

   and

$$\begin{aligned}F_1+F_2 : &(x_2-x_1)(x_0^3x_1+x_0^3x_2-x_1^3x_2-x_2^3x_1)(x_0^3x_1^8+x_0^3x_1^7x_2\\ &+x_1^6x_2^5+x_1^5x_2^6-3x_0^3x_1^5x_2^3-x_0^3x_1^4x_2^4-x_0^6x_1^4x_2+x_0^6x_1^3x_2^2\\ &-3x_0^3x_1^3x_2^5+x_0^6x_1^2x_2^3+x_0^3x_1x_2^7-x_0^6x_1x_2^4+x_0^9x_1x_2+x_0^3x_2^8)\\ &=0\end{aligned}$$

   Common factors of $F_0$ and $F_1 + F_2$ are the line

$$X_2 : x_2 - x_1 = 0$$

   and the curve

$$X_4 : x_0^3x_1 + x_0^3x_2 - x_1^3x_2 - x_2^3x_1 = 0$$

3. $X_4$ is a nonsingular curve.
4. The curve $X_4$ intersects $E$ in 12 different points, from Hurwitz formula the morphism

$$f : X_4 \xrightarrow[12:1]{(x_0x_1x_2, x_0^3+x_1^3+x_2^3)} \mathbf{P}^1$$

   has $degR = 28$. These ramification points lie on the singular curves in $\mathcal{H}$ with $degR = 7$ on each. The curve $X_4$ intersects $T_\infty : x_0x_1x_2 = 0$ with

multiplicity 4 at $(0,0,1)$ and $(0,1,0)$, with multiplicity 2 at $(1,0,0)$ and with multiplicity 1 at $(0,1,i)$ and $(0,1,-i)$. $\sum_{p\in(X_4\cap x_0x_1x_2)}(e_p-1) = 3+3+1 = 7$ and $X_4$ intersects $E$ at 12 different points, the primitive 4-torsion points on $E$.

Conclusion:
The 4-torsion points on $E$ are the origin, $E\cap X_2$ and $E\cap X_4$.

### 5-torsion points on $E$

| The 5-torsion points on $E$ consist of: |
|---|
| 1 origin |
| 24 primitive 5-torsion points |

1. Compute $5x$:

$$5x = (x_0(P_5)(\tau P_5)(\tau^2 P_5),\quad x_2(\sigma^2 P_5)(\tau\sigma^2 P_5)(\tau^2\sigma^2 P_5),\\ x_1(\sigma P_5)(\tau\sigma P_5)(\tau^2\sigma P_5))$$

   The curve given by the product of these polynomials representing $5x$ intersected with $E$ gives the 15-torsion points on $E$, see Sect. 3.1.
2. Common factors for $F_0$ and $F_1+F_2$ is the curve

$$X_5:\quad x_0^6x_1x_2 - x_0^3x_1^4x_2 - x_0^3x_1x_2^4 - x_0^3x_1^3x_2^2 - x_0^3x_1^2x_2^3\\ +x_0^3x_1^5 + x_0^3x_2^5 + x_1^6x_2^2 - x_1^5x_2^3 + x_1^4x_2^4 - x_1^3x_2^5 + x_1^2x_2^6 = 0.$$

3. $X_5$ is singular. The singularities are the 12 vertices of the triangles in $\mathcal{H}$. The singularities of $X_5$ on $T_\infty$ are:

$$[(1,0,0),2,10,2]\ [(0,1,0),2,5,1]\ [(0,0,1),2,5,1].$$

4. The morphism

$$f: X_{5_{blown-up}} \xrightarrow[24:1]{(x_0x_1x_2,x_0^3+x_1^3+x_2^3)} \mathbf{P}^1$$

   has $degR = 64$. $\sum_{q\in X_5}(I_q - r) = (10-2)+(5-1)+(5-1) = 16 = \frac{64}{4}$. So $X_5\cap E$ are 24 different points.

Conclusion:
The 5-torsion points on $E$ are the origin and $E\cap X_5$.

### 6-torsion points on $E$

| The 6-torsion points on $E$ consist of: |
|---|
| 1 origin |
| 3 primitive 2-torsion points |
| 8 primitive 3-torsion points |
| 24 primitive 6-torsion points |

From Sect. 3.1 we know we find the 6-torsion points on $E$ as the intersection between $E$ and the curve given by the product of the polynomials representing $2x$.

The primitive 6-torsion points on $E$ are given by the 8 lines:

$$\begin{array}{ll} Y_{6_1} (= Z(\epsilon^2\tau^2 P_2)) & : x_2 - \epsilon x_1 = 0 \\ Y_{6_2} (= Z(\epsilon\tau P_2)) & : x_2 - \epsilon^2 x_1 = 0 \\ Y_{6_3} (= Z(\sigma^2 P_2)) & : x_1 - x_0 = 0 \\ Y_{6_4} (= Z(\epsilon\tau^2\sigma^2 P_2)) & : x_1 - \epsilon x_0 = 0 \\ Y_{6_5} (= Z(\epsilon^2\tau\sigma^2 P_2)) & : x_1 - \epsilon^2 x_0 = 0 \\ Y_{6_6} (= Z(\sigma P_2)) & : x_0 - x_2 = 0 \\ Y_{6_7} (= Z(\tau^2\sigma P_2)) & : x_0 - \epsilon x_2 = 0 \\ Y_{6_8} (= Z(\tau\sigma P_2)) & : x_0 - \epsilon^2 x_2 = 0 \end{array}$$

We can write $6x$ as:

$$\begin{aligned} 6x = & (-(x_0x_1x_2)(B_1)(B_2)(P_2)(\tau P_2)(\tau^2 P_2)(\sigma^2 P_2)(\tau\sigma^2 P_2)(\tau^2\sigma^2 P_2) \\ & (\sigma P_2)(\tau\sigma P_2)(\tau^2\sigma P_2)((3\epsilon^2 - 3)b_1 - \epsilon^2 b_4 + b_5)((3\epsilon - 3)b_1 - \epsilon b_4 + b_5), \\ & -(B_6)(B_7)(B_8)(3b_1 + b_3 - b_4 - b_7)(3\epsilon b_1 + b_3 - \epsilon b_4 - b_7) \\ & (3\epsilon^2 b_1 + b_3 - \epsilon^2 b_4 - b_7), \\ & -(B_3)(B_4)(B_5)(3b_1 + b_3 - b_5 - b_7)(3\epsilon b_1 + b_3 - \epsilon b_5 - b_7) \\ & (3\epsilon^2 b_1 + b_3 - \epsilon^2 b_5 - b_7) \end{aligned}$$

where $B_i \in V_3$ are the 8 cubics invariant as curves under the action of $H_3$, see p. 138, and $b_i \in V_9$, $(i = 1, \ldots, 7)$ are the base elements of the polynomials in $V_9$ invariant under the action of $H_3$:

$$b_1 : x_0^3x_1^3x_2^3 \qquad b_2 : x_0^9 + x_1^9 + x_2^9$$

$$b_3 : x_0^7x_1x_2 + x_1^7x_2x_0 + x_2^7x_0x_1 \quad b_4 : x_0^6x_1^3 + x_1^6x_2^3 + x_2^6x_0^3$$

$$b_5 : x_0^6x_2^3 + x_1^6x_0^3 + x_2^6x_1^3 \qquad b_6 : x_0^5x_1^2x_2^2 + x_1^5x_2^2x_0^2 + x_2^5x_0^2x_1^2$$

$$b_7 : x_0^4x_1x_2^4 + x_1^4x_2x_0^4 + x_2^4x_0x_1^4$$

The curve given by the product of the polynomials representing $6x$ intersected with $E$ gives the 18-torsion points on $E$, see Sect. 3.1.

Conclusion:
The 6-torsion points on $E$ are the origin, $E \cap X_2$, $E \cap T_\infty$ and $E \cap Y_{6_i}$, $i = 1..8$.

### 7-torsion points on $E$

| The 7-torsion points on $E$ consist of: |
|---|
| 1 origin |
| 48 primitive 7-torsion points |

1. Compute $7x$:

$$7x = (-x_0(P_7)(\tau P_7)(\tau^2 P_7), \quad -x_1(\sigma P_7)(\tau\sigma P_7)(\tau^2\sigma P_7), \\ -x_2(\sigma^2 P_7)(\tau\sigma^2 P_7)(\tau^2\sigma^2 P_7))$$

The curve given by the product of these polynomials representing $7x$ intersected with $E$ gives the 21-torsion points on $E$, see Sect. 3.1.

2. Common factors of $F_0$ and $F_1 + F_2$ is the curve

$$\begin{aligned} X_7 : \quad & x_0^{12}x_1^2x_2^2 + x_0^9x_1^7 - 2x_0^9x_1^5x_2^2 - x_0^9x_1^4x_2^3 - x_0^9x_1^3x_2^4 - 2x_0^9x_1^2x_2^5 \\ & + x_0^9x_2^7 - x_0^6x_1^9x_2 + 2x_0^6x_1^8x_2^2 - 4x_0^6x_1^7x_2^3 + 5x_0^6x_1^6x_2^4 + 2x_0^6x_1^5x_2^5 \\ & + 5x_0^6x_1^4x_2^6 - 4x_0^6x_1^3x_2^7 + 2x_0^6x_1^2x_2^8 - x_0^6x_1x_2^9 + x_0^3x_1^{12}x_2 - x_0^3x_1^{11}x_2^2 \\ & + x_0^3x_1^{10}x_2^3 - 4x_0^3x_1^9x_2^4 - x_0^3x_1^8x_2^5 + 2x_0^3x_1^7x_2^6 + 2x_0^3x_1^6x_2^7 - x_0^3x_1^5x_2^8 \\ & - 4x_0^3x_1^4x_2^9 + x_0^3x_1^3x_2^{10} - x_0^3x_1^2x_2^{11} + x_0^3x_1x_2^{12} + x_1^{11}x_2^5 - x_1^{10}x_2^6 \\ & + x_1^9x_2^7 - x_1^8x_2^8 + x_1^7x_2^9 - x_1^6x_2^{10} + x_1^5x_2^{11} = 0. \end{aligned}$$

3. $X_7$ is a singular curve and has the following singularties on $T_\infty$:

$$[(1,0,0),4,14,2]\ [(0,1,0),4,14,2]\ [(0,0,1),4,14,2]$$

4. The morphism

$$f : X_{7_{blown-up}} \xrightarrow[48:1]{(x_0x_1x_2, x_0^3+x_1^3+x_2^3)} \mathbf{P^1}$$

has $degR = 144$. $\sum_{q\in X_7}(I_q - r) = (14-2)+(14-2)+(14-2) = 36 = \frac{144}{4}$. So $X_7 \cap E$ are 48 different points.

Conclusion:
The 7-torsion points on $E$ are the origin and $E \cap X_7$.

**8-torsion points on $E$**

| The 8-torsion points on $E$ consist of: |
|---|
| 1 origin |
| 3 primitive 2-torsion points |
| 12 primitive 4-torsion points |
| 48 primitive 8-torsion points |

1. Compute $8x$:

$$\begin{aligned} 8x = (& x_0(P_2)(\tau P_2)(\tau^2 P_2)(P_4)(\tau P_4)(\tau^2 P_4)(P_8)(\tau P_8)(\tau^2 P_8), \\ & x_2(\sigma^2 P_2)(\tau\sigma^2 P_2)(\tau^2\sigma^2 P_2)(\sigma^2 P_4)(\tau\sigma^2 P_4)(\tau^2\sigma^2 P_4) \\ & \quad (\sigma^2 P_8)(\tau\sigma^2 P_8)(\tau^2\sigma^2 P_8), \\ & x_1(\sigma P_2)(\tau\sigma P_2)(\tau^2\sigma P_2)(\sigma P_4)(\tau\sigma P_4)(\tau^2\sigma P_4)(\sigma P_8) \\ & \quad (\tau\sigma P_8)(\tau^2\sigma P_8)) \end{aligned}$$

The curve given by the product of these polynomials representing $8x$ intersected with $E$ gives the 24-torsion points on $E$, see Sect. 3.1.

2. Common factors of $F_0$ and $F_1 + F_2$

$$X_2 : x_2 - x_1 = 0 \ ,$$

$$X_4 : x_0^3x_1 + x_0^3x_2 - x_1^3x_2 - x_2^3x_1 = 0$$

and

$$\begin{aligned} X_8 : \ & x_0^{12}x_1^3x_2 + x_0^{12}x_1x_2^3 - x_0^9x_1^6x_2 - 3x_0^9x_1^4x_2^3 - 3x_0^9x_1^3x_2^4 \\ & -x_0^9x_1x_2^6 - x_0^6x_1^{10} + x_0^6x_1^9x_2 + 6x_0^6x_1^7x_2^3 + 6x_0^6x_1^3x_2^7 \\ & +x_0^6x_1x_2^9 - x_0^6x_2^{10} - x_0^3x_1^{12}x_2 + 3x_0^3x_1^9x_2^4 - 6x_0^3x_1^7x_2^6 \\ & -6x_0^3x_1^6x_2^7 + 3x_0^3x_1^4x_2^9 - x_0^3x_1x_2^{12} + x_1^{10}x_2^6 + x_1^6x_2^{10} = 0. \end{aligned}$$

3. $X_8$ is a singular curve and has the following singularities on $T_\infty$:

$$[(1,0,0),4,20,4]\ [(0,1,0),4,12,2]\ [(0,0,1),4,12,2].$$

4. The morphism

$$f : X_{8_{blown-up}} \xrightarrow[48:1]{(x_0x_1x_2, x_0^3+x_1^3+x_2^3)} \mathbf{P}^1$$

has $degR = 144$. $\sum_{q\in X_8}(I_q - r) = (20-4)+(12-2)+(12-2) = 36 = \frac{144}{4}$. So $X_8 \cap E$ are 48 different points.

Conclusion:
The 8-torsion points on $E$ are the origin, $E \cap X_2$, $E \cap X_4$ and $E \cap X_8$.

### 9-torsion points on $E$

| The 9-torsion points on $E$ consist of: |
|---|
| 1 origo |
| 8 primitive 3-torsjonspunkter |
| 72 primitive 9-torsjonspunkter |

From Sect. 3.1 we know we find the 9-torsion points on $E$ as $T_\infty \cap E$ and $Z(B_i) \cap E$, $(i = 1, \ldots, 8)$, (see p. 138).

The polynomials representing $9x$ are too big to show here, each polynomial has appr. 250 terms.

### 10-torsion points on $E$

| The 10-torsion points on $E$ consist of: |
|---|
| 1 origin |
| 3 primitive 2-torsion points |
| 24 primitive 5-torsion points |
| 72 primitive 10-torsion points |

1. Compute $10x$:

$$\begin{aligned}10x = (&-x_0(P_2)(\tau P_2)(\tau^2 P_2)(P_5)(\tau P_5)(\tau^2 P_5)(P_{10})(\tau P_{10})(\tau^2 P_{10}),\\ &-x_1(\sigma P_2)(\tau\sigma P_2)(\tau^2\sigma P_2)(\sigma P_5)(\tau\sigma P_5)(\tau^2\sigma P_5)\\ &\quad(\sigma P_{10})(\tau\sigma P_{10})(\tau^2\sigma P_{10}),\\ &-x_2(\sigma^2 P_2)(\tau\sigma^2 P_2)(\tau^2\sigma^2 P_2)(\sigma^2 P_5)(\tau\sigma^2 P_5)(\tau^2\sigma^2 P_5)\\ &\quad(\sigma^2 P_{10})(\tau\sigma^2 P_{10})(\tau^2\sigma^2 P_{10}))\end{aligned}$$

   The curve given by the product of these polynomials representing $10x$ intersected with $E$ gives the 30-torsion points on $E$, see Sect. 3.1.
2. Common factors of $F_0$ and $F_1 + F_2$ are

$$X_2 : x_2 - x_1 = 0\ ,$$

$$\begin{aligned}X_5 : {} & x_0^6x_1x_2 + x_0^3x_2^5 - x_2^4x_0^3x_1 - x_0^3x_1^2x_2^3 + x_1^2x_2^6 - x_0^3x_1^3x_2^2\\ & -x_1^3x_2^5 - x_0^3x_1^4x_2 + x_1^4x_2^4 + x_0^3x_1^5 - x_1^5x_2^3 + x_1^6x_2^2 = 0\ ,\end{aligned}$$

   and

$$\begin{aligned}X_{10} : \quad & -x_0^{18}x_1^5x_2 - x_0^{18}x_1^4x_2^2 - x_0^{18}x_1^3x_2^3 - x_0^{18}x_1^2x_2^4 - x_0^{18}x_1x_2^5\\ & +x_0^{15}x_1^9 + 2x_0^{15}x_1^8x_2 + 2x_0^{15}x_1^7x_2^2 + 2x_0^{15}x_1^6x_2^3 + 8x_0^{15}x_1^5x_2^4\\ & +8x_0^{15}x_1^4x_2^5 + 2x_0^{15}x_1^3x_2^6 + 2x_0^{15}x_1^2x_2^7 + 2x_0^{15}x_1x_2^8 + x_0^{15}x_2^9\\ & -x_0^{12}x_1^{11}x_2 - x_0^{12}x_1^{10}x_2^2 - 11x_0^{12}x_1^9x_2^3 - 14x_0^{12}x_1^8x_2^4 - 14x_0^{12}x_1^7x_2^5\\ & +7x_0^{12}x_1^6x_2^6 - 14x_0^{12}x_1^5x_2^7 - 14x_0^{12}x_1^4x_2^8 - 11x_0^{12}x_1^3x_2^9 - x_0^{12}x_1^2x_2^{10}\\ & -x_0^{12}x_1x_2^{11} + x_0^9x_1^{14}x_2 + x_0^9x_1^{13}x_2^2 + 7x_0^9x_1^{12}x_2^3 + 4x_0^9x_1^{11}x_2^4\\ & +4x_0^9x_1^{10}x_2^5 + 3x_0^9x_1^9x_2^6 + 30x_0^9x_1^8x_2^7 + 30x_0^9x_1^7x_2^8 + 3x_0^9x_1^6x_2^9\\ & +4x_0^9x_1^5x_2^{10} + 4x_0^9x_1^4x_2^{11} + 7x_0^9x_1^3x_2^{12} + x_0^9x_1^2x_2^{13} + x_0^9x_1x_2^{14}\\ & -x_0^6x_1^{17}x_2 - x_0^6x_1^{16}x_2^2 - 4x_0^6x_1^{15}x_2^3 + 2x_0^6x_1^{14}x_2^4 + 2x_0^6x_1^{13}x_2^5\\ & +5x_0^6x_1^{12}x_2^6 - 16x_0^6x_1^{11}x_2^7 - 16x_0^6x_1^{10}x_2^8 - 17x_0^6x_1^9x_2^9 - 16x_0^6x_1^8x_2^{10}\\ & -16x_0^6x_1^7x_2^{11} + 5x_0^6x_1^6x_2^{12} + 2x_0^6x_1^5x_2^{13} + 2x_0^6x_1^4x_2^{14} - 4x_0^6x_1^3x_2^{15}\\ & -x_0^6x_1^2x_2^{16} - x_0^6x_1x_2^{17} + x_0^3x_1^{18}x_2^3 + x_0^3x_1^{17}x_2^4 + x_0^3x_1^{16}x_2^5\\ & -2x_0^3x_1^{15}x_2^6 - 2x_0^3x_1^{14}x_2^7 - 2x_0^3x_1^{13}x_2^8 + 4x_0^3x_1^{12}x_2^9 + 14x_0^3x_1^{11}x_2^{10}\\ & +14x_0^3x_1^{10}x_2^{11} + 4x_0^3x_1^9x_2^{12} - 2x_0^3x_1^8x_2^{13} - 2x_0^3x_1^7x_2^{14} - 2x_0^3x_1^6x_2^{15}\\ & +x_0^3x_1^5x_2^{16} + x_0^3x_1^4x_2^{17} + x_0^3x_1^3x_2^{18} - x_1^{14}x_2^{10} - x_1^{13}x_2^{11}\\ & -x_1^{12}x_2^{12} - x_1^{11}x_2^{13} - x_1^{10}x_2^{14} = 0.\end{aligned}$$

3. $X_{10}$ is singular and has the following singularities on $T_\infty$:

$$[(1,0,0),6,18,6]\ [(0,1,0),6,25,3]\ [(0,0,1),6,25,3].$$

4. The morphism

$$f : X_{10_{blown-up}} \xrightarrow[72:1]{(x_0x_1x_2, x_0^3+x_1^3+x_2^3)} \mathbf{P}^1$$

   has $degR = 224$. $\sum_{q\in X_{10}}(I_q - r) = (18-6)+(25-3)+(25-3) = 56 = \frac{224}{4}$. So $X_{10} \cap E$ are 72 different points.

Conclusion:
The 10-torsion points on $E$ are the origin, $E \cap X_2$, $E \cap X_5$ and $E \cap X_{10}$.

### 3.2 A Second Algorithm for Finding the $n$-torsion Points on $E \in \mathcal{H}$

As we saw in the last section it is easy to find the $n$-th division polynomial as long as $3 \mid n$, it is the polynomial $\frac{n}{3}x$. If we want to find the $n$-torsion points when $3 \nmid n$ we can use $nx \cap E$ and divide the points by 3.

Let $k$ be an arbitrary field and let $E$ be an elliptic curve in $\mathcal{H}$.

**Algorithm 2** *An algorithm for finding the n-torsion points on E when* $3 \nmid n$*:*

1. *Compute* $nx = (F_0, F_1, F_2)$*, the* $3n$*-th division polynomial* $P_{3n} = F_0F_1F_2$.
2. *Compute* $E[3n] = \{p \in E \mid 3np = 0\} = E \cap Z(P_{3n})$.
3. *The n-torsion points on E is found by dividing the points in* $E[3n]$ *by* 3*, i.e. find those* $q \in E$ *s.t.* $3q = p$.

Remark; to find those $q \in E$ s.t. $3q = p$ we simply solve the equations $\mathbf{3q} \times \boldsymbol{p} = \mathbf{0}$.

### 3.3 $n$-torsion Points on the Singular Curves in $\mathcal{H}$

The singular curves in $\mathcal{H}$ when $k = K$ are the four triangles $T_\lambda$, $(\lambda \in \{\infty, -3, -3\epsilon, -3\epsilon^2\})$ (see Sect. 2.1 and 2.2).

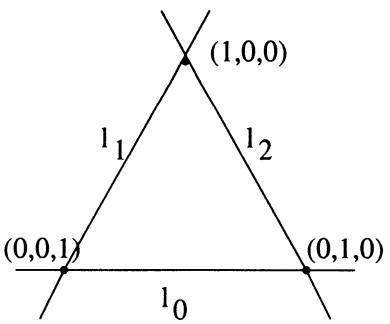



**Fig. 1.** The triangle $T_\infty$.

Let $x, y \in T_\infty \subset \mathbf{P}^2$. The lines $l_i$ are represented by $x_i = 0$, $i = 0, 1, 2$. Let $l_i^*$ be the line $l_i$ minus the two vertices of the triangle that lies on the line.

The rational map

$$\begin{array}{ccc} \mathbf{P}^2 \times \mathbf{P}^2 & \dashrightarrow & \mathbf{P}^2 \\ (x, y) & \mapsto & x + y \end{array}$$

is not defined when both $x$ and $y$ are vertices in the triangle, else:

$$\begin{aligned}
y = (1,0,0) \text{ and } x \in l_0^* &\Rightarrow x + y = (1,0,0)\\
x \in l_1^* &\Rightarrow x + y = (0,1,0)\\
x \in l_2^* &\Rightarrow x + y = (0,0,1)\\
y = (0,1,0) \text{ and } x \in l_0^* &\Rightarrow x + y = (0,1,0)\\
x \in l_1^* &\Rightarrow x + y = (0,0,1)\\
x \in l_2^* &\Rightarrow x + y = (1,0,0)\\
y = (0,0,1) \text{ and } x \in l_0^* &\Rightarrow x + y = (0,0,1)\\
x \in l_1^* &\Rightarrow x + y = (1,0,0)\\
x \in l_2^* &\Rightarrow x + y = (0,1,0)\\
y \in l_0^* \text{ and } x \in l_0^* &\Rightarrow x + y \in l_0^*\\
x \in l_1^* &\Rightarrow x + y \in l_1^*\\
x \in l_2^* &\Rightarrow x + y \in l_2^*\\
y \in l_1^* \text{ and } x \in l_0^* &\Rightarrow x + y \in l_1^*\\
x \in l_1^* &\Rightarrow x + y \in l_2^*\\
x \in l_2^* &\Rightarrow x + y \in l_0^*\\
y \in l_2^* \text{ and } x \in l_0^* &\Rightarrow x + y \in l_2^*\\
x \in l_1^* &\Rightarrow x + y \in l_0^*\\
x \in l_2^* &\Rightarrow x + y \in l_1^*
\end{aligned}$$

Let $T_{\lambda_{ns}}(K)$ denote the nonsingular points of the triangle $T_\lambda \in \mathcal{H}$ over the field $K$. And let $l_0^*$ be the line in $T_\lambda$ s.t. $(0,1,-1) \in l_0^*$, minus the two singular points in $T_\lambda$ that lies on the line.

**Proposition 6.** *The points on $l_0^*$ are a subgroup of $T_{\lambda_{ns}}(K)$ isomorphic to $K^*$.*

*Proof.* Let $x = (0, x_1, x_2), y = (0, y_1, y_2) \in l_0^* \subset T_\infty$. Then

$$M_{\iota(x),y} = \begin{pmatrix} 0 & x_2y_2 & x_1y_1 \\ x_2y_1 & 0 & 0 \\ x_1y_2 & 0 & 0 \end{pmatrix}$$

$M_{\iota(x),y} \cdot z = 0$ implies that $x + y = (0, x_1y_1, -x_2y_2)$. So the following map is a group isomorphism:

$$\begin{array}{ccccc} \psi : & l_0^* & \xrightarrow{\simeq} & \mathbf{C}^* \\ & x = (0,1,\frac{x_2}{x_1}) & \mapsto & -\frac{x_2}{x_1} \end{array}$$

$\psi((0,1,-1) = 1$ and $\psi(x+y) = \psi((0,1,-\frac{x_2y_2}{x_1y_1})) = \frac{x_2y_2}{x_1y_1} = \psi(x)\psi(y)$.

**Corollary 3.** *The $n$-torsion points on the singular curves in $\mathcal{H}$ that lies on the line $l_0^*$ are a subgroup of $K^*$ isomorphic to $U_n$, the multiplicative group of $n$-th roots of unity.*

*Proof.* Let $x = (0, x_1, x_2) \in l_0^* \subset T_\infty$. If $x$ is a $n$-torsion point then $nx = (0, 1, (\frac{x_2}{x_1})^n) \sim (0,1,-1)$, and since $\psi(nx) = 1 \Rightarrow \frac{x_2}{x_1}$ is a $n$-th root of unity in $K^*$.

# A A Proof for the Associativity of the Group Law Defined in Sect. 2.2

Let $x, y, z \in E$ and evaluate $p_1 = (x + y) + z$ and $p_2 = x + (y + z)$ :

$$\begin{aligned}
p_1 = (&-x_0^4y_1^2y_2^2z_1z_2 + 2x_0^2x_1x_2y_0^2y_1y_2z_1z_2 + x_0^2x_1x_2y_1^2y_2^2z_0^2 - x_0x_1^3y_0y_2^3z_0^2\\
&-x_0x_2^3y_0y_1^3z_0^2 - x_1^2x_2^2y_0^4z_1z_2 + x_1^2x_2^2y_0^2y_1y_2z_0^2,\\
&-x_0^3x_1y_1y_2^3z_2^2 + x_0^2x_2^2y_0y_1^2y_2z_2^2 - x_0^2x_2^2y_1^4z_0z_1 + x_0x_1^2x_2y_0^2y_2^2z_2^2\\
&+2x_0x_1^2x_2y_0y_1^2y_2z_0z_1 - x_1^4y_0^2y_2^2z_0z_1 - x_1x_2^3y_0^3y_1z_2^2,\\
&-x_0^3x_2y_1^3y_2z_1^2 + x_0^2x_1^2y_0y_1y_2^2z_1^2 - x_0^2x_1^2y_2^4z_0z_2 + x_0x_1x_2^2y_0^2y_1^2z_1^2\\
&+2x_0x_1x_2^2y_0y_1y_2^2z_0z_2 - x_1^3x_2y_0^3y_2z_1^2 - x_2^4y_0^2y_1^2z_0z_2)
\end{aligned}$$

$$\begin{aligned}
p_2 = (&-x_0^2y_0^2y_1y_2z_1^2z_2^2 + x_0^2y_0y_1^3z_0z_2^3 + x_0^2y_0y_2^3z_0z_1^3 - x_0^2y_1^2y_2^2z_0^2z_1z_2\\
&+x_1x_2y_0^4z_1^2z_2^2 - 2x_1x_2y_0^2y_1y_2z_0^2z_1z_2 + x_1x_2y_1^2y_2^2z_0^4,\\
&x_0x_1y_0^2y_2^2z_1^4 - 2x_0x_1y_0y_1^2y_2z_0z_1^2z_2 + x_0x_1y_1^4z_0^2z_2^2 + x_2^2y_0^3y_1z_1z_2^3\\
&-x_2^2y_0^2y_2^2z_0z_1^2z_2 - x_2^2y_0y_1^2y_2z_0^2z_2^2 + x_2^2y_1y_2^3z_0^3z_1,\\
&x_0x_2y_0^2y_1^2z_2^4 - 2x_0x_2y_0y_1y_2^2z_0z_1z_2^2 + x_0x_2y_2^4z_0^2z_1^2 + x_1^2y_0^3y_2z_1^3z_2\\
&-x_1^2y_0^2y_1^2z_0z_1z_2^2 - x_1^2y_0y_1y_2^2z_0^2z_1^2 + x_1^2y_1^3y_2z_0^3z_2)
\end{aligned}$$

We then factor the crossproduct $\boldsymbol{p_1} \times \boldsymbol{p_2}$ with respect to $detM_{x,y}, detM_{z,y}$ and $detM_{x,z}$. This gives us the following three polynomials:

$$\begin{aligned}
k_1 = ((&-y_1y_2^4 + y_2y_1^4)x_0x_1x_2z_0z_1^4z_2 + y_0^2y_2^3x_1^3z_1^6 - y_0y_1^2y_2^2x_0x_1x_2z_1^6\\
&+y_0y_1^2y_2^2x_0x_1x_2z_2^6 + (-y_1y_2^4 + y_2y_1^4)x_0x_1x_2z_0z_1z_2^4\\
&+(-y_2y_1^4 + y_1y_2^4)x_1^3z_0^2z_1^2z_2^2 - y_0^2y_1^3x_2^3z_1^3z_2^3 - y_0^2y_1^3x_2^3z_2^6\\
&-2z_0z_2^4z_1y_1^2y_2^2y_0x_1^3 + 2z_1^4z_0z_2y_2^2y_1^2y_0x_2^3 + (-y_2y_1^4 + y_1y_2^4)x_2^3z_0^2z_1^2z_2^2\\
&+y_0^2y_2^3x_1^3z_1^3z_2^3)\mathbf{detM_{x,y}}\\
+(&z_1y_2^2z_2z_0y_1^2y_0x_1^6 - z_1y_2^2z_2z_0y_1^2y_0x_2^6 - x_1^3z_1^3y_2^3y_0^2x_2^3 - x_1^6z_1^3y_2^3y_0^2\\
&+x_2^3z_2^3y_1^3y_0^2x_1^3 + x_2^6z_2^3y_1^3y_0^2)\mathbf{detM_{z,y}}\\
+((&-2y_0^2y_1^3y_2^3 - y_0^2y_2^6)x_1^3z_1^3 + (y_0^2y_1^6 + 2y_0^2y_1^3y_2^3)x_2^3z_2^3 + 2y_0y_1^5y_2^2z_0z_1z_2x_1^3\\
&+y_0^2y_1^3y_2^3x_1^3z_2^3 - 2y_0y_1^2y_2^5z_0z_1z_2x_2^3 - y_0^2y_1^3y_2^3x_2^3z_1^3\\
&+(-y_1^7y_2 + y_1y_2^7)x_0x_1x_2z_0z_1z_2 + (y_0y_1^5y_2^2 + y_0y_1^2y_2^5)x_0x_1x_2z_1^3\\
&+(-y_0y_1^5y_2^2 - y_0y_1^2y_2^5)x_0x_1x_2z_2^3)\mathbf{detM_{x,z}}
\end{aligned}$$

$$
\begin{aligned}
k_2 = (&x_0x_1^2y_1^2y_2^3z_0^2z_2^4 - x_0^2x_2y_0y_1^3y_2z_1z_2^5 + x_0^2x_2y_1^2y_2^3z_0z_1^5\\
&+(2y_2^3y_1^2 + y_1^5)x_0^2x_2z_0z_1^2z_2^3 - x_0^2x_2y_0y_1^3y_2z_1^4z_2^2 - x_1x_2^2y_0y_1^3y_2z_0^2z_2^4\\
&+(-2y_2^3y_1^2 - y_1^5)x_1x_2^2z_0^3z_1z_2^2 + (-y_2y_1^3y_0 + y_2^4y_0)x_1x_2^2z_0^2z_1^3z_2\\
&-x_0x_1^2y_0y_2^4z_0z_1^5 + (y_0^3y_1^2 - y_1^5 - y_2^3y_1^2)x_1x_2^2z_1^4z_2^2\\
&+(y_0^3y_1^2 - y_1^5 - y_2^3y_1^2)x_1x_2^2z_1z_2^5 + (y_2y_1^3y_0 - y_2^4y_0)x_0x_1^2z_0z_1^2z_2^3)\mathbf{detM_{x,y}}\\
+((&-y_0^3y_1^2 + y_1^5 + y_2^3y_1^2)x_1x_2^5z_1z_2^2 + x_1x_2^5y_0y_1^3y_2z_0^2z_2\\
&+(-y_0^3y_1^2 + y_1^5 + y_2^3y_1^2)x_1^4x_2^2z_1z_2^2 + x_0x_1^5y_0y_2^4z_0z_1^2 + x_0^2x_2^4y_0y_1^3y_2z_1z_2^2\\
&+x_1^4x_2^2y_0y_1^3y_2z_0^2z_2 - x_0^2x_2^4y_1^2y_2^3z_0z_1^2 - x_0x_1^5y_1^2y_2^3z_0^2z_2)\mathbf{detM_{z,y}}\\
+((&-y_1^5y_2^3 - y_1^2y_2^6)x_0x_1^2z_0^2z_2 + (y_0y_1^6y_2 + 2y_0y_1^3y_2^4)x_1x_2^2z_0^2z_2\\
&+(y_1^8 + 2y_1^5y_2^3 + y_1^2y_2^6)x_1x_2^2z_1z_2^2 + (2y_0y_1^3y_2^4 + y_0y_2^7)x_0x_1^2z_0z_1^2\\
&+(-y_1^5y_2^3 - y_1^2y_2^6)x_0^2x_2z_0z_1^2 - y_1^4y_2^2y_0^2x_1^2x_0z_2^2z_1\\
&-y_0^2y_1y_2^5x_1x_2^2z_0z_1^2)\mathbf{detM_{x,z}}
\end{aligned}
$$

$$
\begin{aligned}
k_3 = (&-x_0x_2^2y_1^3y_2^2z_0^2z_1^4 + x_0^2x_1y_0y_1y_2^3z_1^2z_2^4 - x_0^2x_1y_1^3y_2^2z_0z_2^5\\
&+(-2y_2^2y_1^3 - y_2^5)x_0^2x_1z_0z_1^3z_2^2 + x_0^2x_1y_0y_1y_2^3z_1^5z_2 + x_1^2x_2y_0y_1y_2^3z_0^2z_1^4\\
&+(y_2^5 + 2y_2^2y_1^3)x_1^2x_2z_0^3z_1^2z_2 + (-y_1^4y_0 + y_2^3y_0y_1)x_1^2x_2z_0^2z_1z_2^3\\
&+x_0x_2^2y_0y_1^4z_0z_2^5 + (-y_2^2y_0^3 + y_2^5 + y_2^2y_1^3)x_1^2x_2z_1^5z_2\\
&+(-y_2^2y_0^3 + y_2^5 + y_2^2y_1^3)x_1^2x_2z_1^2z_2^4\\
&+(-y_2^3y_0y_1 + y_1^4y_0)x_0x_2^2z_0z_1^3z_2^2)\mathbf{detM_{x,y}}\\
+(&x_0x_2^5z_1y_2^2z_0^2y_1^3 + x_0^2x_1^4y_1^3y_2^2z_0z_2^2 + (y_2^2y_0^3 - y_2^2y_1^3 - y_2^5)x_1^5x_2z_1^2z_2\\
&-x_0x_2^5y_0y_1^4z_0z_2^2 + (y_2^2y_0^3 - y_2^2y_1^3 - y_2^5)x_1^2x_2^4z_1^2z_2 - x_1^5x_2y_0y_1y_2^3z_0^2z_1\\
&-x_0^2x_1^4y_1z_2z_1^2y_2^3y_0 - x_1^2x_2^4y_0y_1y_2^3z_0^2z_1)\mathbf{detM_{z,y}}\\
+(&y_0^2y_1^5y_2z_0z_2^2x_2x_1^2 + (y_1^6y_2^2 + y_1^3y_2^5)x_0^2x_1z_0z_2^2 + y_0^2y_1^2y_2^4z_1^2z_2x_2^2x_0\\
&+(-y_0y_1^7 - 2y_0y_1^4y_2^3)x_0x_2^2z_0z_2^2 + (y_1^6y_2^2 + y_1^3y_2^5)x_0x_2^2z_0^2z_1\\
&+(-2y_0y_1^4y_2^3 - y_0y_1y_2^6)x_1^2x_2z_0^2z_1\\
&+(-y_1^6y_2^2 - 2y_1^3y_2^5 - y_2^8)x_1^2x_2z_1^2z_2)\mathbf{detM_{x,z}}
\end{aligned}
$$

We see that $\boldsymbol{p_1} \times \boldsymbol{p_2} = \overrightarrow{(k_1, k_2, k_3)} = \mathbf{0}$ as long as $x, y, z \in E$. It then follows that the operation $'+'$ is associative.

## References


1. Aure A., Decker W., Hulek K., Popescu S., Ranestad K. (1993) The Geometry of Bielliptic Surfaces in $\mathbb{P}^4$. International Journal of Mathematics. **6**, 873–902
2. Bix R. (1998) Conics and Cubics, A Concrete Introduction to Algebraic Curves. Springer.
3. Connell I. (1999) Elliptic Curve Handbook. ftp://ftp.math.mcgill.ca /pub/ECH1/

4. Frium H. (1999) Torsjonspunkter på elliptiske kurver. Master's thesis. University of Oslo, Norway.
5. Fulton W. (1974) Algebraic Curves, An Introduction to Algebraic Geometry. Mathematics Lecture Notes Series.
6. Hartshorne R. (1977) Algebraic Geometry. Springer GTM. **52**
7. Hirschfeld J. W. P. (1998) Projective Geometries over Finite Fields. Oxford University Press.
8. Husemöller D. (1987) Elliptic curves. Springer.
9. Menezes A. (1993) Elliptic Curve Public Key Cryptosystems. Kluwer Academic Publishers.
10. Silverman J. H. (1986) The Arithmetic of Elliptic Curves. Springer.
11. Silverman J. H., Tate J. (1992) Rational Points on Elliptic Curves. Springer.

# On Curves with Many Rational Points over Finite Fields

Arnaldo Garcia

IMPA,
Estrada Dona Castorina, 110
22460-320, Rio de Janeiro, RJ, Brazil
e-mail: garcia@impa.br

**Abstract.** We summarize results on maximal curves over $\mathbb{F}_{q^2}$ (i.e., curves attaining the Hasse-Weil upper bound for the number of rational points over finite fields). We discuss the classification problem and the genus spectrum of maximal curves. We present some towers of curves over finite fields attaining the Drinfeld-Vladut bound. Especially interesting is the description of the completely splitting locus (see Formula (20)) of a certain tower of curves, meaning the first description by their coordinates of the supersingular points of the modular curves $X_0(2^n)$, for each $n \in \mathbb{N}$.

## 1 Introduction

The theory of equations over finite fields is a basic topic in Number Theory and Algebraic Geometry. The object of the first investigations in this theory were congruences of the form

$$y^2 \equiv f(x) \quad \text{(modulo a prime number)}, \tag{1}$$

where $f(x)$ is a rational function with integer coefficients. E. Artin associated a zeta-function to Equation (1), in analogy with the one introduced by Dedekind for quadratic number fields, and (assuming Riemann's hypothesis for this zeta-function) he conjectured an upper bound for the number of solutions in the prime field $\mathbb{F}_p$ of congruences such as the ones in (1) above. E. Artin's conjecture was then proved by H. Hasse for polynomials $f(x)$ of degrees 3 and 4 over arbitrary finite fields, and widely generalized by A. Weil (see [30]) as follows: *Let $X$ be a projective geometrically irreducible nonsingular algebraic curve of genus $g = g(X)$, defined over a finite field $\mathbb{F}_\ell$ with $\ell$ elements. Then, its number of rational points $\#X(\mathbb{F}_\ell)$ satisfies*

$$|\#X(\mathbb{F}_\ell) - (\ell + 1)| \leq 2g\sqrt{\ell}. \tag{2}$$

Inequality (2) is equivalent to the validity of Riemann's hypothesis for the zeta-function associated to the curve $X$, and for other proofs of this inequality we refer to [3] and [27].

Curves $X$ that attain the upper bound in (2) are called *maximal*; i.e., a curve $X$ is $\mathbb{F}_{q^2}$-maximal if we have

$$\#X(\mathbb{F}_{q^2}) = q^2 + 1 + 2q \cdot g(X). \tag{3}$$

Y. Ihara noticed that if a curve $X$ is $\mathbb{F}_{q^2}$-maximal then the genus $g(X)$ must be small; more precisely, he showed that its genus satisfies

$$g(X) \leq q(q-1)/2. \tag{4}$$

In order to study the asymptotics of $\mathbb{F}_\ell$-rational points on curves of large genus, Y. Ihara (see [20]) introduced the function

$$A(\ell) = \limsup_{g \to \infty} N_\ell(g)/g, \tag{5}$$

where $N_\ell(g) = \max\{\#X(\mathbb{F}_\ell);\ g(X) = g\}$. To obtain lower bounds on $A(\ell)$ one usually considers towers $\mathcal{F}$ of curves defined over the finite field $\mathbb{F}_\ell$, that is an infinite sequence of curves $X_n$ over $\mathbb{F}_\ell$ with increasing genus, and calculates the *limit* $\lambda(\mathcal{F})$ over $\mathbb{F}_\ell$:

$$\lambda(\mathcal{F}) = \lim_{n \to \infty} \frac{\#X_n(\mathbb{F}_\ell)}{g(X_n)}. \tag{6}$$

We have (for any $\mathbb{F}_\ell$-tower $\mathcal{F}$):

$$A(\ell) \geq \lambda(\mathcal{F}). \tag{7}$$

The best known upper bound for $A(\ell)$ is due to Drinfeld-Vladut (see [7]). It says

$$A(\ell) \leq \sqrt{\ell} - 1, \quad \text{for any } \ell. \tag{8}$$

When the cardinality $\ell$ of the finite field is a square we have equality in (8), and this fact was proved independently by Ihara and by Tsfasman-Vladut-Zink (see [20] and [29]); i.e., we have

$$A(q^2) = q - 1, \quad \text{for any} \quad q. \tag{9}$$

The interest on curves over finite fields was greatly renewed after Goppa's construction of linear codes from such curves (see [13], [28] and [12]). Using Goppa's construction and the equality in (9) above for $q \geq 7$, Tsfasman-Vladut-Zink constructed an infinite sequence of codes of increasing lengths having limit parameters (relative minimum distance and transmission rate) above the so-called Gilbert-Varshamov bound, a result that caused a sensation among specialists in Coding Theory (see [29]).

The purpose of this paper is to survey on results on curves over finite fields in two directions:

1. Maximal Curves (genus and classification).
2. Towers of curves and Drinfeld-Vladut bound.

Throughout this paper we will use the word curve to designate a nonsingular projective geometrically irreducible algebraic curve, defined over a finite field $\mathbb{F}_\ell$ with $\ell$ elements, or its nonsingular projective model. Also, we denote $\overline{\mathbb{F}}_\ell$ an algebraic closure of $\mathbb{F}_\ell$.

## 2 Maximal Curves

A maximal curve $X$ over $\mathbb{F}_{q^2}$ attains the Hasse-Weil upper bound for the number of $\mathbb{F}_{q^2}$-rational points; i.e., we have

$$\#X(\mathbb{F}_{q^2}) = q^2 + 1 + 2q \cdot g(X),$$

where $g(X)$ denotes its genus.

*The genus of a $\mathbb{F}_{q^2}$-maximal curve $X$ satisfies* (see [20]):

$$g(X) \leq q(q-1)/2.$$

The most well-known maximal curve over $\mathbb{F}_{q^2}$ is the so-called *Hermitian curve* (denoted by $H$) which can be given by the affine equation:

$$y^q + y = x^{q+1}. \tag{10}$$

The Hermitian curve $H$ over $\mathbb{F}_{q^2}$ is maximal and has the largest genus possible for a $\mathbb{F}_{q^2}$-maximal curve; i.e., we have

$$g(H) = q(q-1)/2 \quad \text{and} \quad \#H(\mathbb{F}_{q^2}) = 1 + q^3.$$

For a divisor $m$ of $(q+1)$ we denote by $H_m$ the curve over $\mathbb{F}_{q^2}$ defined by the affine equation:

$$y^q + y = x^{\frac{q+1}{m}}. \tag{11}$$

The genus of $H_m$ satisfies

$$2 \cdot g(H_m) = (q-1) \cdot \left( \frac{q+1}{m} - 1 \right).$$

One can check directly that $H_m$ is a $\mathbb{F}_{q^2}$-maximal curve, but this also follows (since $H_m$ is covered by the Hermitian curve $H = H_1$) from the following general result due to J.-P. Serre (see [22]):

*Let $\varepsilon\colon X \twoheadrightarrow Y$ be a surjective map of curves where both curves and the map are defined over $\mathbb{F}_{q^2}$. If $X$ is $\mathbb{F}_{q^2}$-maximal, then $Y$ is also $\mathbb{F}_{q^2}$-maximal.*

Since the genus $g(X)$ of a $\mathbb{F}_{q^2}$-maximal curve $X$ satisfies the upper bound in (4), we have here two natural questions:

1. **Genus Spectrum**
   What are the natural numbers in the interval $(0, q(q-1)/2]$ that are genera of $\mathbb{F}_{q^2}$-maximal curves?
2. **Classification**
   For a fixed genus $g$, what are the $\mathbb{F}_{q^2}$-maximal curves (modulo isomorphisms) that have genus $g$?

## 1. Genus spectrum

Not every integer lying in the interval $(0, q(q-1)/2)$ is the genus of a $\mathbb{F}_{q^2}$-maximal curve. *For example (see [11]) if $X$ is $\mathbb{F}_{q^2}$-maximal and $g(X) < q(q-1)/2$, then $g(X) \leq (q-1)^2/4$.* Note that if $q$ is odd (i.e., the characteristic of the finite field is not two), then the curve $H_2$ in Equation (11) satisfies:

$$g(H_2) = (q-1)^2/4. \tag{12}$$

In order to find lots of entries of the genus spectrum for $\mathbb{F}_{q^2}$-maximal curves (as follows from Serre's result aforementioned), one can determine genera of curves covered by the Hermitian curve $H$ over $\mathbb{F}_{q^2}$. This approach was systematically used in [19] by considering quotient curves $H/G$, where $G$ is a subgroup of automorphisms of $H$. In particular is is shown that for a fixed integer $g \geq 1$, there are $\mathbb{F}_{q^2}$-maximal curves of genus $g$ for infinitely many values of $q$ (see [19], Remark 6.2). *If the characteristic $p$ is odd and considering $p$-subgroups $G$ (writing $q = p^n$), there are $\mathbb{F}_{q^2}$-maximal curves with genus $g$ given by*:

$$g = \frac{1}{2} p^{n-v} \cdot (p^{n-w} - 1), \tag{13}$$

*for each $0 \leq v \leq n$ and for each $0 \leq w \leq (n-1)$.*

In case $p = 2$ it seems a hard problem to determine the pairs $(v, w)$ for which there are $\mathbb{F}_{q^2}$-maximal curves (coming from 2-subgroups $G$ of automorphisms of the Hermitian curve) with genus as in Formula (13) (see also [1]).

*For each divisor $m$ of $(q^2 - q + 1)$, there are $\mathbb{F}_{q^2}$-maximal curves having genus $g$ given by* (see [19], Theorem 5.1):

$$g = \frac{m-1}{2}. \tag{14}$$

The determination of explicit equations for the $\mathbb{F}_{q^2}$-maximal curves leading to the genus formula in (14) above is not so easy (see [5]).

For $k \geq 2$, the following equation gives a $\mathbb{F}_{q^{2k}}$-maximal curve (see [14]):

$$\sum_{j=0}^{k-1} y^{q^j} = \alpha \cdot x^{q^k+1}, \quad \text{with} \quad \alpha^{q^k-1} = -1. \tag{15}$$

The curve given by Equation (15) above has genus $g = q^k(q^k - 1)/2$ and, in particular, its genus appears among those given in Formula (13). For $q = p$, this curve also appears in [[6], Theorem 2.1].

## 2. Classification

The interest on the classification problem for $\mathbb{F}_{q^2}$-maximal curves with a fixed genus was triggered after the following result of Rück-Stichtenoth (see [25]):

*If $X$ is $\mathbb{F}_{q^2}$-maximal and $g(X) = q(q-1)/2$, then $X$ is $\mathbb{F}_{q^2}$-isomorphic to the Hermitian curve $H$ given by Equation* (10).

For $q$ odd, the second largest genus of $\mathbb{F}_{q^2}$-maximal curves is given by $g_2 = (q-1)^2/4$. Here we also have a unicity result (see [10]):

*If $X$ is $\mathbb{F}_{q^2}$-maximal ($q$ odd) and $g(X) = (q-1)^2/4$, then $X$ is $\mathbb{F}_{q^2}$-isomorphic to the curve $H_2$ given by Equation* (11) *with $m = 2$.*

In case of characteristic $p = 2$, the second largest genus of maximal curves over $\mathbb{F}_{q^2}$ is $g_2 = q(q-2)/4$ and the classification problem here has some extra-difficulties (see [2]).

In [[10], Theorem 2.3] it is given a characterization of the $\mathbb{F}_{q^2}$-maximal curves with equations as in (11), but this characterization requires an extra-hypothesis on Weierstrass nongaps at a rational point of the curve.

Write $q = p^t$ and consider the curve $X$ over $\mathbb{F}_{q^2}$ given by the equation

$$\sum_{i=1}^{t} y^{q/p^i} + \alpha \cdot x^{q+1} = 0, \quad \text{with} \quad \alpha^{q-1} = -1. \tag{16}$$

This curve $X$ is $\mathbb{F}_{q^2}$-maximal with $g(X) = q(q-p)/2p$. In case of characteristic $p = 2$, one gets the second largest genus possible. Curves $X$ given by equations as in (16) appear also in [[6], Theorem 2.1] where it is given a classification of the Galois subcoverings of prime degrees of the Hermitian curve.

For $q \equiv 3$ (modulo 4), the following two curves $X_1$ and $X_2$ (both having genus equal to $g = (q-1)(q-3)/8$) are $\mathbb{F}_{q^2}$-maximal and nonisomorphic to each other:

$$X_1 \quad \text{given by} \quad y^q + y = x^{(q+1)/4}, \quad \text{and}$$
$$X_2 \quad \text{given by} \quad x^{(q+1)/2} + y^{(q+1)/2} = 1.$$

*The curve $X_2$ above is the unique maximal curve over $\mathbb{F}_{q^2}$ with genus $g = (q-1)(q-3)/8$ that has a nonsingular plane model over $\mathbb{F}_{q^2}$* (see [4]).

We end up this first part with the following question:

**Question:** Is every $\mathbb{F}_{q^2}$-maximal curve $\mathbb{F}_{q^2}$-covered by the Hermitian curve over $\mathbb{F}_{q^2}$ (i.e., by the curve given by Equation (10))?

A related important result was obtained in [21]. *It is shown that maximal curves over $\mathbb{F}_{q^2}$ lie in nondegenerate Hermitian varieties as curves having degree equal to $(q+1)$.*

## 3 Towers of Curves

In order to study the asymptotic behaviour of the number of $\mathbb{F}_\ell$-rational points on curves of large genus, Y. Ihara introduced the following function:

$$A(\ell) = \limsup_{g\to\infty} N_\ell(g)/g,$$

where $N_\ell(g) = \max\{\#X(\mathbb{F}_\ell);\ g(X) = g\}$.

The study of this function $A(\ell)$ envolves the consideration of infinite sequences of curves defined over the finite field $\mathbb{F}_\ell$ having genera tending to infinity.

**Definition.** A *tower $\mathcal{F}$ over $\mathbb{F}_\ell$* (or a *$\mathbb{F}_\ell$-tower*) is an infinite sequence of curves and surjective and separable maps, both defined over the finite field $\mathbb{F}_\ell$,

$$\cdots \twoheadrightarrow X_{n+1} \twoheadrightarrow X_n \cdots \twoheadrightarrow X_2 \twoheadrightarrow X_1$$

such that $g(X_n) \to \infty$ as $n \to \infty$.

The following limit exists and it is called the *limit of the tower $\mathcal{F}$*:

$$\lambda(\mathcal{F}) = \lim_{n\to\infty} \#X_n(\mathbb{F}_\ell)/g(X_n).$$

Clearly we have $A(\ell) \geq \lambda(\mathcal{F})$, for any $\mathbb{F}_\ell$-tower $\mathcal{F}$.

**Example 1.** Let $\mathcal{W}$ be the tower over $\mathbb{F}_{q^2}$ where (see [16]):

- $X_1$ is the projective line with affine coordinate denoted by $x_1$.
- $X_2$ is the plane projective curve with the equation

$$x_2^q + x_2 = x_1^q/(x_1^{q-1} + 1).$$

- $X_3$ is the curve in $\mathbb{P}^3$ given by the equations

$$x_3^q + x_3 = x_2^q/(x_2^{q-1} + 1) \quad \text{and} \quad x_2^q + x_2 = x_1^q/(x_1^{q-1} + 1).$$

- $X_4$ is the curve in $\mathbb{P}^4$ given by the equations

$$\begin{cases} x_2^q + x_2 = x_1^q/(x_1^{q-1} + 1) \\ x_3^q + x_3 = x_2^q/(x_2^{q-1} + 1) \\ x_4^q + x_4 = x_3^q/(x_3^{q-1} + 1), \end{cases}$$

and so on. The map $X_{n+1} \twoheadrightarrow X_n$ is given by

$$(x_1, x_2, \ldots, x_n, x_{n+1}) \mapsto (x_1, x_2, \ldots, x_n).$$

In this tower $\mathcal{W}$ over $\mathbb{F}_{q^2}$ the ramification points are wildly ramified which makes more subtle the determination of the genus $g(X_n)$, for each $n \in \mathbb{N}$.

One has that

$$\lambda(\mathcal{W}) = q - 1;$$

i.e., the $\mathbb{F}_{q^2}$-tower $\mathcal{W}$ attains the upper bound of Drinfeld-Vladut given in (8). This gives another proof of the Equality (9) (see also [20] and [29]) with the advantages of having the curves in the tower explicitely given by algebraic equations, and also providing an explicit description by coordinates of the $\mathbb{F}_{q^2}$-rational points of the curves in the tower. For the determination of Weierstrass semigroups on this tower we refer to [24].

The tower $\mathcal{W}$ in Example 1 is *recursive*; i.e., it uses the same equation for the construction of all curves $X_n$, $n \in \mathbb{N}$, in the tower. We say that the tower $\mathcal{W}$ is *recursively* given by

$$y^q + y = x^q/(x^{q-1} + 1).$$

**Example 2.** Let $\mathbb{F}_q$ be a nonprime finite field of characteristic $p$ (i.e., $q > p$) and denote $m = (q-1)/(p-1)$. Consider the $\mathbb{F}_q$-tower $\mathcal{G}$ recursively given by the equation (see [18]):

$$y^m + (x+1)^m = 1.$$

More precisely we have

- $X_1$ is the projective line with affine coordinate denoted by $x_1$.
- $X_2$ is the plane projective curve with equation

$$x_2^m + (x_1 + 1)^m = 1.$$

- $X_3$ is the curve in $\mathbb{P}^3$ with equations

$$x_3^m + (x_2 + 1)^m = 1 \quad \text{and} \quad x_2^m + (x_1 + 1)^m = 1,$$

and so on for the curves $X_4, X_5, X_6, \ldots$.

Here one has

$$\lambda(\mathcal{G}) \geq \frac{2}{q-2} > 0. \tag{17}$$

This gives a very simple proof for nonprime finite fields of the following result of Serre ([26]):

$$A(q) > 0, \quad \text{for any} \quad q.$$

Unfortunately the method in [18] does not apply to prime fields $\mathbb{F}_p$ as was pointed out by H.W. Lenstra (see [23]).

The first tower over $\mathbb{F}_{q^2}$ given by explicit algebraic equations that attains the Drinfeld-Vladut bound can be described as follows (see [15] and [9]):

- $X_1$ is the projective line with affine coordinate $x_1$.
- $X_2$ is the plane curve given by

$$z_2^q + z_2 = x_1^{q+1}.$$

- $X_3$ is the curve in $\mathbb{P}^3$ given by

$$z_2^q + z_2 = x_1^{q+1} \text{ and } z_3^q + z_3 = x_2^{q+1}, \text{ with } x_2 = \frac{z_2}{x_1}.$$

- $X_4$ is the curve in $\mathbb{P}^4$ given by

$$\begin{cases} z_2^q + z_2 = x_1^{q+1} \\ z_3^q + z_3 = x_2^{q+1} \\ z_4^q + z_4 = x_3^{q+1}, \text{ with } x_3 = \dfrac{z_3}{x_2}, \end{cases}$$

and so on for the curves $X_5, X_6, X_7, \ldots$. Note that the above description of this tower is not recursive since one makes for each $n \geq 1$, the substitution $x_{n+1} = \dfrac{z_{n+1}}{x_n}$.

Now we review the concepts of ramification point and ramification index. Let $\psi$ be a surjective and separable map between two curves; i.e.,

$$X \overset{\psi}{\twoheadrightarrow} Y.$$

We denote by $d$ the degree of the map $\psi$; i.e., except for finitely many points of the curve $Y$, we have that there are exactly $d$ points on the curve $X$ above each point of $Y$. The finitely many exceptional points of the curve $Y$ having fewer than $d$ preimages in $X$ under the map $\psi$ are called *ramification points*. More precisely, let $P$ be any point of $Y$ and let

$$\psi^{-1}(P) = \{P_1, P_2, \ldots, P_r\} \subseteq X$$

be the preimages of $P$ in $X$ under the map $\psi$. Attached to each $P_j$, $j = 1, 2, \ldots, r$, there is a natural number $e_j \geq 1$ (called *ramification index* of the point $P_j$ over $P$) and we have

$$\sum_{j=1}^{r} e_j = d.$$

Except for finitely many points $P$ in $Y$ one has that $r = d$ (and hence $e_1 = e_2 = \cdots = e_d = 1$). If $e_1 = e_2 \cdots = e_d = 1$ then the point $P$ is called *unramified*. A point $P_j$ is *ramified over the point* $P$ if $e_j \geq 2$. The ramification at $P_j$ is called *wild* if the characteristic $p$ divides the ramification index $e_j$ and

it is called *tame* otherwise. When the map $\psi$ determines a Galois covering one has (for any point $P$ in $Y$):

$$e_1 = e_2 = \cdots = e_r.$$

Hence we have that in Galois coverings of degrees relatively prime to the characteristic all ramifications are tame.

A tower $\mathcal{T}$ over a finite field $\mathbb{F}_\ell$

$$\cdots \twoheadrightarrow X_{n+1} \twoheadrightarrow X_n \cdots \twoheadrightarrow X_2 \twoheadrightarrow X_1$$

is called *tame* if for each map $X_{n+1} \twoheadrightarrow X_n$ all ramifications are tame.

We have two important sets of points on the first curve $X_1$ of the tower. To define these two sets, let us denote by

$$\pi_n \colon X_n \twoheadrightarrow X_1$$

the compositum of the first $(n-1)$ maps in the tower $\mathcal{T}$ over $\mathbb{F}_\ell$.

**1. The ramification locus**

$$S = \{P \in X_1(\overline{\mathbb{F}}_\ell) \mid \text{for some } n \geq 2,\ \exists\, Q \in X_n(\overline{\mathbb{F}}_\ell) \text{ with } \pi_n(Q) = P, \text{ and the point } Q \text{ is ramified over } P\}.$$

**2. The completely splitting locus**

$$T = \{P \in X_1(\mathbb{F}_\ell) \mid \text{for all } n \geq 2, \text{ the point } P \text{ is unramified for the map } \pi_n, \text{ and all its preimages under } \pi_n \text{ are } \mathbb{F}_\ell\text{-rational points of } X_n\}.$$

Tame towers $\mathcal{T}$ over $\mathbb{F}_\ell$ are specially interesting when:
**1.** The ramification locus $S$ is a **finite set**.
**2.** The completely splitting locus $T$ is a **nonempty set**.

This is so because of the following result (see [18]): *For a tame tower $\mathcal{T}$ over $\mathbb{F}_\ell$, its limit satisfies*

$$\lambda(\mathcal{T}) \geq \frac{2t}{2g(X_1) - 2 + s}, \tag{18}$$

*where $t = \#T$ and $s = \#S$.*

For example, one can derive the Inequality (17) in Example 2 from the formula in (18).

**Example 3.** Let $p \geq 3$ be an odd prime number and consider the tower $\mathcal{T}$ over $\mathbb{F}_{p^2}$ defined recursively by the equation (see [17]):

$$y^2 = \frac{x^2+1}{2x}.$$

The ramification locus $S$ of this tower $\mathcal{T}$ is

$$S = \{0, \infty, \pm 1, \pm\alpha\},$$

where $\alpha \in \mathbb{F}_{p^2}$ with $\alpha^2 = -1$. Hence, its cardinality is $s = 6$.

For the determination of the completely splitting locus $T$ of this tower $\mathcal{T}$ we need Deuring's polynomial $H(Z)$ whose roots determine the supersingular elliptic curves in Legendre's form:

$$H(Z) = \sum_{j=0}^{\frac{p-1}{2}} \binom{\frac{p-1}{2}}{j}^2 \cdot Z^j \in \mathbb{F}_p[Z].$$

Here there are two fundamental properties of Deuring's polynomials:

**1. Polynomial identity**

$$H(Z^4) = Z^{p-1} \cdot H\left(\left(\frac{Z^2+1}{2Z}\right)^2\right). \tag{19}$$

**2. Location of the roots**

*All roots of the polynomial $H(Z) \in \mathbb{F}_p[Z]$ are fourth powers in $\mathbb{F}_{p^2}$.*

We then have the following explicit description of the completely splitting locus

$$T = \{\beta \in \mathbb{F}_{p^2};\ H(\beta^4) = 0\}. \tag{20}$$

Since the polynomial $H(Z)$ is separable of degree $(p-1)/2$, it follows from (20) that

$$t = \#T = 2p - 2.$$

Applying Inequality (18) we get

$$\lambda(\mathcal{T}) \geq \frac{2 \cdot (2p-2)}{2 \cdot 0 - 2 + 6} = p - 1.$$

From this it follows that $\lambda(\mathcal{T}) = p - 1$; i.e., the tower $\mathcal{T}$ over $\mathbb{F}_{p^2}$ in Example 3 attains the Drinfeld-Vladut bound.

We conclude with two remarks due to M. Zieve. Firstly that Property 2 (Location of the roots) follows from Property 1 (Polynomial identity) and the fact that all roots of $H(Z)$ lie in $\mathbb{F}_{p^2}$. Secondly that the tower $\mathcal{T}$ in Example 3

corresponds to the modular curves $X_0(2^n)$, for $n \in \mathbb{N}$. This last remark is done by comparison with an explicit modular tower worked out by N. Elkies (see [8] and [17]). *In particular the explicit description of the completely splitting locus $T$ given in* (20) *above represents the first description by their coordinates of the supersingular points of the modular curves $X_0(2^n)$, for each $n \in \mathbb{N}$.*

## References


1. Abdón, M., Quoos, L. (2001) On the genera of subfields of the Hermitian Function Field. Preprint.
2. Abdón, M., Torres, F. (1999) On maximal curves in characteristic two. Manuscripta Math. **99**, 39–53
3. Bombieri, E. (1976) Hilbert's 8th problem: an analogue. Proc. Symposia in Pure Math. **28** (ed. F. Browder), American Math. Society, Providence, 269–274
4. Cossidente, A., Hirschfeld, J., Korchmaros, G., Torres, F. (2000) On plane maximal curves. Compositio Math. **121**, 163–181
5. Cossidente, A., Korchmaros, G., Torres, F. (1999) On curves covered by the Hermitian curve. J. Algebra **216**, 56–76
6. Cossidente, A., Korchmaros, G., Torres, F. (2000) Curves of large genus covered by the Hermitian curve. Comm. Algebra, **28**, 4707–4728
7. Drinfeld, V., Vladut, S. (1983) Number of points of an algebraic curve. Funct. Anal. **17**, 53–54
8. Elkies, N. (1997) Explicit modular towers. In Proc. 35th Annual Allerton Conference on Commun., Control and Computing, Urbana, IL
9. Elkies, N. (2000) Explicit towers of Drinfeld modular curves. To appear in Proc. 3rd European Cong. of Math., Barcelona
10. Fuhrmann, R., Garcia, A., Torres F. (1997) On maximal curves. J. Number Theory **67**, 29–51
11. Fuhrmann, R., Torres, F. (1996) The genus of curves over finite fields with many rational points. Manuscripta. Math. **89**, 103–106
12. Geer van der, G. (2000) Curves over finite fields and codes. To appear in Proc. 3rd European Cong. of Math., Barcelona
13. Goppa, V. (1983) Algebraic - geometric codes. Math. USRR-Izv. **21**, 75–91
14. Garcia, A., Quoos, L. (2001) A construction of curves over finite fields. Acta Arithmetica. **98**, 181–195
15. Garcia, A., Stichtenoth, H. (1995) A tower of Artin-Schreier extensions of function fields attaining the Drinfeld-Vladut bound. Inventiones. Math. **121**, 211–222
16. Garcia, A., Stichtenoth, H. (1996) On the asymptotic behaviour of some towers of function fields over finite fields. J. of Number Theory **61**, No. 2, 248–273
17. Garcia, A., Stichtenoth, H. (2001) On tame towers over finite fields. Preprint
18. Garcia, A., Stichtenoth, H., Thomas, M. (1997) On Towers and Composita of Towers of Function Fields over Finite Fields. Finite Fields and Appl. **3**, 257–274
19. Garcia, A., Stichtenoth, H., Xing, C.P. (2000) On subfields of the Hermitian Function Field. Compositio Math. **120**, 137–170
20. Ihara, Y. (1981) Some remarks on the number of rational points of algebraic curves over finite fields. J. Fac. Sci. Tokyo **28**, 721–724

21. Korchmaros, G., Torres, F. Embedding of a maximal curve in a Hermitian Variety. To appear in Compositio Math.
22. Lachaud, G. (1987) Sommes d'Eisenstein et nombre de points de certaines courbes algébriques sur les corps finis. C.R.Acad. Sci. Paris **305**, 729–732
23. Lenstra, H. W. (2000) On a problem of Garcia, Stichtenoth and Thomas. Preprint
24. Pellikaan, R., Stichtenoth, H., Torres, F. (1998) Weierstrass semigroups in an asymptotically good tower of function fields. Finite Fields and Appl. **4**, 381–392
25. Rück, H. G., Stichtenoth, H. (1994) A Characterization of Hermitian Function Fields over Finite Fields. J. Reine Angew. Math. **457**, 185–188.
26. Serre, J. P. (1983) Sur le nombre des points rationnels d'une courbe algébrique sur un corps fini. C.R. Acad. Sci. Paris **296**, 397–402
27. Stöhr, K. O., Voloch, J. F. (1986) Weierstrass points and curves over finite fields. Proc. London Math. Soc. **52**, 1–19
28. Tsfasman, M. A., Vladut, S. G. (1991) Algebraic Geometric Codes. Kluwer, Dordrecht
29. Tsfasman, M., Vladut, S. G., Zink, T. (1982) Modular curves, Shimura curves and Goppa codes, better than the Varshamov-Gilbert bound. Math. Nachrichten. **109**, 21–28
30. Weil, A. (1971) Courbes algébriques et variétés abéliennes. Hermann, Paris

# VHDL Specification of a FPGA to Divide and Multiply in $GF(2^m)$

Mario Alberto García-Martínez[1] and Guillermo Morales-Luna[2]

[1] Instituto Tecnológico de Orizaba
Departamento de Ingeniería Eléctrica-Electrónica
Oriente 9 no. 852, 94300 Orizaba, Veracruz, Mexico
marioag@prodigy.net.mx

[2] Programa de Ingeniería Molecular, Instituto Mexicano del Petróleo
on leave of absence from Computer Science Section, CINVESTAV-IPN
Av. IPN 2508, 07360 Mexico, D.F.
gmorales@cs.cinvestav.mx

**Abstract.** Some FPGA's are designed to compute division and multiplication on Galois fields. FPGA's are quite cheap programmable logic devices used in digital circuits with the important characteristic of being reprogrammable. Any FPGA can be specified within VHDL which at present is a standard language in the design of digital systems. We describe in VHDL the divider and multiplier basic cells and their whole integration. The structures have scalable systolic architectures. The circuits operate by pipelining; the divider in $GF(2^m)$ requires $5m-1$ clock cycles while the multiplier $3m-1$. The divider proceeds by the Gaussian triangulation algorithm and is uniform with respect to the irreducible polynomial generating the field. The codes, some simulations and performance measurements are provided.

## 1 Introduction

Open networks are extensively applied in bank transactions, e-commerce, e-mail and most communications, and several security threats regarding authentication, data integrity and confidentiality arose. Adequate schemes to guarantee high security levels are increasingly important. Many cryptographic schemes, as well as operations in error correcting codes, switching theory and digital signal processing, require computations over finite fields. High speed and low complexity design for finite fields arithmetic is thus quite useful with respect to wider bandwidths and better security.

In $GF(2^m)$, addition is realized directly as bit-wise exclusive-OR without carries. Multiplication and division are more complex [9] and treated with several algorithms [1]. In general, the parallel architectures are faster than their serial counterparts, and require greater amounts of logical gates when implemented. In addition, different basis are used for the representation of the field elements. Some of these architectures require an additional circuitry for the basis conversion, increasing their hardware complexities. On the other hand, each structure depends on the irreducible polynomial generating the

field, nevertheless it is possible to change the generating polynomial without alteration of the circuitry structures. This is particularly important for cryptographic applications that require frequent change of system keys. Fenn et al. [4] present a similar systolic divider in $GF(2^m)$ with space and time complexities of order $m^2$ and $m$ respectively, but they use a dual basis representation. Lee and Yoo [8] introduce a structure for exponentiation with time complexity of order $3m + 9$ cycles, that uses a canonical basis parallel multiplier. Chin [2] presents a parallel inverter/divider systolic array that uses the canonical basis and has a space complexity of order $m^2(m-1)/2$ cells and a time complexity of order $2m^2 - 3m/2$ cycles. Sunar and Koç [10] introduce a new algorithm for the parallel Mastrovito multiplier with a space complexity of $m^2$ AND gates and $m^2 - 1$ XOR gates. Our main interest is the development of an implementation with FPGA for great values of $m$ optimizing the space resources.

We propose a serial systolic structure for a divider and a multiplier that work using the canonical basis and the algorithm developed in [6], which has space and time complexities of order $m^2$ and $m$ respectively. Moreover this structure is uniform with respect to the irreducible polinomial.

As a first approach, division can be carried out by calculating the multiplicative inverse of the divisor and multiplying by the dividend. This procedure has an unnecessary lengthy run-time. Better methods to divide have been developed by posing an equivalent problem consisting in solving a linear system of equations over $GF(2^m)$. We follow the method introduced in [6] and translate the problem into a system of $m$ linear equations on $GF(2^m)$ ($m$ is the dimension of the finite field, whose order is $2^m$). The algorithm proceeds in two main steps. First, the coefficients matrix of the linear system of equations is built. Second, the system is solved by a triangulation process. Our `Divider` is modular and quite appropriate to handle large values of $m$. In addition, it does not depend on the irreducible polynomial that is used to generate the field. The number of required timing cycles, or clock pulses, is linear with respect to $m$. We present the design of FPGA's (Field Programmable Gate Array) to compute the division and multiplication over Galois fields. FPGA's are quite cheap programmable logic devices used in digital circuits with the important characteristic of being reprogrammable [5,11]. Any FPGA can be specified within VHDL (Hardware Description Language) which at present is a standard language in the design of digital systems [3,7]. We describe the basic cells and their integration into the divider and the multiplier. The structures have systolic architectures and can be expanded easily. The circuit operates by pipelining and requires $5m - 1$ clock cycles to compute division and $3m - 1$ clock cycles for multiplication. The divider uses the Gaussian triangulation algorithm and is uniform with respect to the generating irreducible polynomial. The simulation of the cells were validated. The `Divider` circuit has been recorded as a programmable device.

## 2 The Circuit `Divider`

### 2.1 Algebraic Preliminares

Division in finite fields is reduced to solve a system of linear equations by Hasan and Bhargava in 1992 [6], with smaller calculation time and simpler computing structures. The basic structures and procedures do not depend on the irreducible polynomial generating the field.

Let $a, c$ be two elements over $GF(2^m)$, with $a \neq 0$. The quotient $b = c/a$ can be computed as follows: Let us write $c = ab$, $a = \boldsymbol{a}^T \boldsymbol{A}^{(m)}$, $b = \boldsymbol{b}^T \boldsymbol{A}^{(m)}$, $c = \boldsymbol{c}^T \boldsymbol{A}^{(m)}$, where $\boldsymbol{A}^{(m)} = \left(\alpha^i\right)_{i=0}^{m-1}$ is the matrix whose elements are the powers of a generator $\alpha$ of $GF(2^m)$ and $\boldsymbol{a}, \boldsymbol{b}$ and $\boldsymbol{c}$ are the column-coordinate vectors of $a, b$ and $c$ relative to the basis $\boldsymbol{A}^{(m)}$. Then

$$\boldsymbol{c}^T \boldsymbol{A}^{(m)} = (\boldsymbol{a} * \boldsymbol{b})^T \boldsymbol{A}^{(2m-1)} = (\boldsymbol{a} * \boldsymbol{b})^T \boldsymbol{P}^T \boldsymbol{A}^{(m)}$$

where $\boldsymbol{a} * \boldsymbol{b}$ is the convolution of $\boldsymbol{a}$ with $\boldsymbol{b}$ and $\boldsymbol{P} = (p_{ij})_{i=0,\ldots,m-1}^{j=0,\ldots,2m-2}$ is the coordinate matrix of powers of $\alpha$ in terms of the chosen basis $\boldsymbol{A}^{(m)}$, $\boldsymbol{A}^{(2m-1)} = \boldsymbol{P}^T \boldsymbol{A}^{(m)}$. Hence, $\boldsymbol{c} = \boldsymbol{P}(\boldsymbol{a} * \boldsymbol{b})$, where for each $i = 0, \ldots, m-1$, $c_i = \sum_{k=0}^{2m-2} p_{ik} \sum_{\ell+j=k} a_\ell b_j$. Thus:

$$\boldsymbol{c} = \boldsymbol{Q}\boldsymbol{b} \tag{1}$$

where $\boldsymbol{Q} = (q_{ij})_{i=0,\ldots,m-1}^{j=0,\ldots,m-1}$ has entries $q_{ij} = \sum_{k=0}^{m-1} p_{i,k+j} a_k$. In this way, division can be performed by solving a system of $m$ linear equations over $GF(2^m)$[6].

### 2.2 Serial Structure of the `Divider`

The general design of the serial divider on $GF(2^m)$ is sketched in Fig. 1. It can be used for large values of $m$ and solves indeed Eq. (1).

The input signals are the following: $(g_i)_{i=0}^{m-1}$ is the vector whose entries are the coefficients of the generating polynomal, provided by the user, $(q_i)_{i\geq 0}$ is a sequence of control digits (marking the size of $m$), $(a_i)_{i=0}^{m-1}$ gives the divider element in $GF(2^m)$ and $(c_i)_{i=0}^{m-1}$ gives the dividend element. The output signal gives the sequence of the quotient coefficients $(b_i)_{i=0}^{m-1}$. The first task of the divider is the generation of the coefficients matrix $\boldsymbol{Q}$ in Eq. (1), along $m$ clock cycles. Thereafter, the linear system of equations is solved in a pipeline process requiring, as we will see, $4m-1$ clock cycles. The total time required by the whole divisor is thus $5m-1$ clock cycles. The middle cells $D_i$'s are flip-flops used to synchronize both main structures `Gen-Mat` and `Solution`.

We will describe each component below.

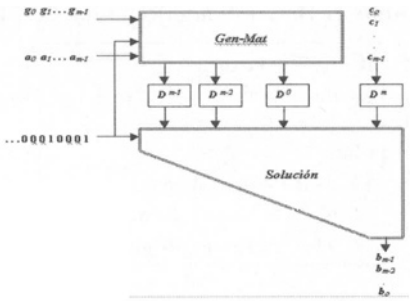
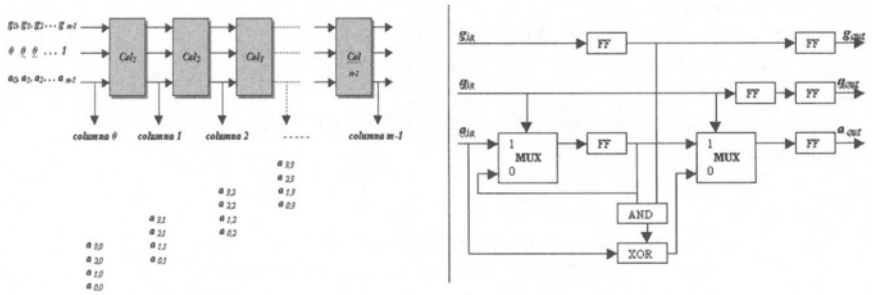



**Fig. 1.** Serial divider on $GF(2^m)$

### 2.3 Array to Compute the Coefficients Matrix

The array `Gen-Mat` that generates the coefficients matrix is shown in Fig. 2-(**a**). The array `Gen-Mat` consists of $m-1$ cells sequentially arranged and




(**a**) (**b**)

**Fig. 2.** (**a**) Array `Gen-Mat`. (**b**) Diagram of `Cel01`

labeled $Cel_1$, $Cel_2$, $Cel_3$, $\cdots$, $Cel_{m-1}$. Each $Cel_i$ is of type `Cel01`, displayed in Fig. 2-(**b**) in which boxes labelled FF represent flip-flops. Indeed, the computation they performed is described in the pseudocode shown at Table 1.

The coordinates of $\boldsymbol{a}$ are introduced in a serial way into $Cel_1$ beginning with the "most significant digit" $a_{m-1}$. Output column 0 is just a copy of $\boldsymbol{a}$. Columns 1 to $m-1$ are the respective outputs of cells $Cel_1$, $Cel_2$, $Cel_3$, $\cdots$, $Cel_{m-1}$. Each cell requires two additional clock cycles to display the first element of each column in its output. The output of cell $Cel_{j-1}$ is introduced into the next processor $Cel_j$.

The VHDL code of circuit `Gen-Mat`, which calculates the coefficients matrix of the linear system (1), is shown in Table 2.

**Table 1.** Description of cell type `Cel01`.

| |
|---|
| **if** $q_{in} = 1$ **then** |
| $\{ \quad a_{out} := r;$ |
| $\quad r := a_{in} \}$ |
| **else** |
| $\{ \quad a_{out} := g_{temp} r + a_{in} \}$ |
| $\quad g_{out} := g_{temp};\ g_{temp} := g_{in};$ |
| $\quad q_{out} := q_{temp};\ q_{temp} := q_{in};$ |

**Table 2.** VHDL description of `Gen-Mat`

```
library IEEE;
use IEEE.std_logic_1164.all;
entity gen_mat4 is
   generic (m: positive:=4);
   port(CLK,CLR,Ginput,Qinput,Ainput:in STD_LOGIC;
      Col_0,Col_1,Col_2,Col_3:out STD_LOGIC);
end gen_mat4;
architecture gen_mat4_arch of gen_mat4 is
component cell01b
   port (CLK,CLR,Gin,Qin,Ain: in STD_LOGIC;
      Gout,Qout,Aout: out STD_LOGIC);
end component;
signal Gtemp,Qtemp,Atemp: STD_LOGIC_VECTOR(0 to m-1);
begin gen_mat:
   for I in 0 to m-1 generate
      cel0: if (I=0) generate
         cell0: cell01b port map (CLK=>CLK,CLR=>CLR,
            Gin=>Ginput,Qin=>Qinput, Ain=>Ainput,
            Gout=>Gtemp(I),Qout=>Qtemp(I), Aout=>Atemp(I));
         end generate cel0;
      cel1: if (I<=m-1 and I>=1) generate
         celda1: cell01b port map (CLK=>CLK, CLR=>CLR,
            Gin=>Gtemp(I-1),Qin=>Qtemp(I-1), Ain=>Atemp(I-1),
            Gout=>Gtemp(I), Qout=>Qtemp(I), Aout=>Atemp(I));
         end generate cel1;
   end generate gen_mat;
   Col_0<=Ainput; Col_1<=Atemp(0); Col_2<=Atemp(1); Col_3<=Atemp(2);
end gen_mat4_arch;
```

### 2.4 Array to Solve the Linear Equations System

The proposed structure that applies the Gauss-Jordan triangulation method to solve the linear equations system (1) is shown in Fig. 3. The signals indi-

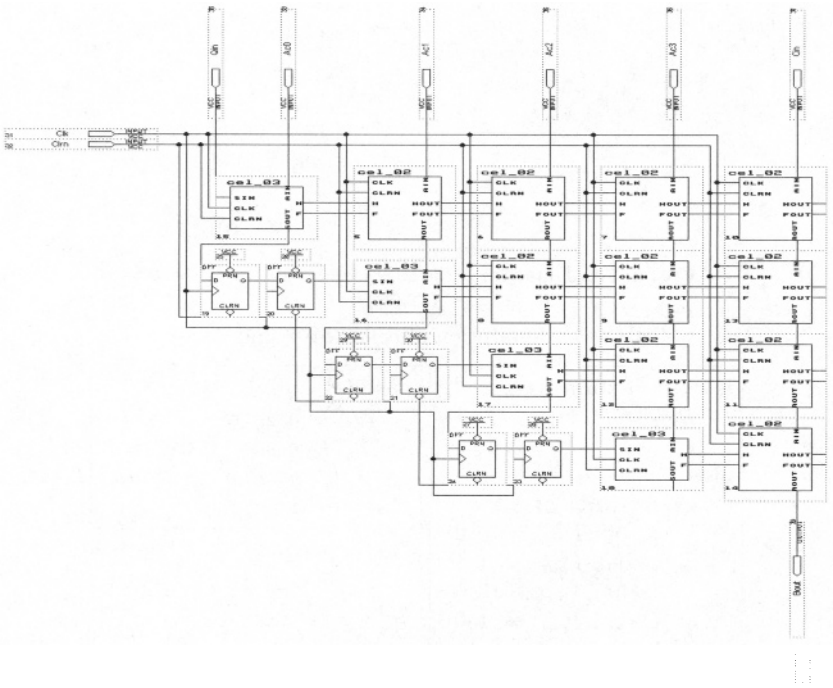


**Fig. 3.** Array to solve the linear equations system

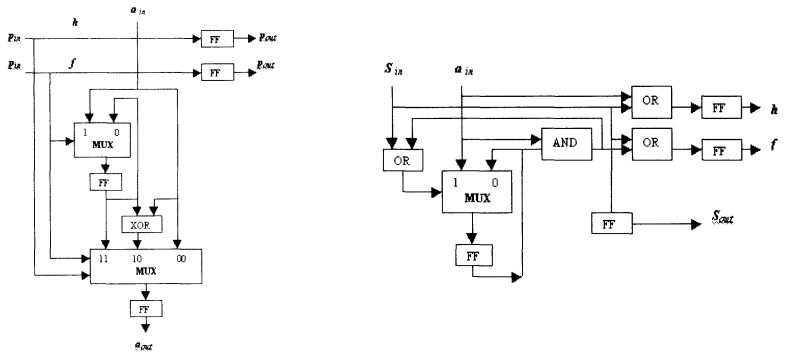
cated in the figure are the following:

`[ ac0 ... ac.[m-1] ]` = Columns of the matrix
`[ b.[m-1] ... b0 ]` = Solution of eq. (1)

`C-in` = Dividend
`s-in` = Control signal
`clk` = Clock

Two types of cells, `Cel02` and `Cel03`, are used here. Corresponding diagrams of `Cel02` and `Cel03` are sketched in Fig. 4. Their operations are described in pseudocodes shown in Table 3, (a) and (b) respectively.

The flip-flops are used for synchronization purposes. The beginning of matrix entries arrival is determined by the control signal. The entries of the solution vector $\boldsymbol{b}$ are output from the $(m-1)$-th cell of type `Cel02`. The $m$-th coordinate appears at cycle $3m-1$. Thereafter, each new coordinate appears at one clock cycle. Hence, in this array encharged to solve the system of equations, the total output of $\boldsymbol{b}$ is obtained at clock cycle $4m-1$. This, together with the $m$ clock cycles used at first array `Gen-Mat` gives $5m-1$ clock cycles for the whole divider. The code of circuit `Solution`, which solves by triangulation the linear system (1), is shown in Tables 4 and 5.




(a) (b)

**Fig. 4.** (a) Diagram of `Cel02`. (b) Diagram of `Cel03`

**Table 3.** Description of cells `Cel02` and `Cel03`.

**if** $s_{in} := 1$ **then**
{ $P_{out} := (1,1)$;
$r := a_{in}$ }
**else**
**case** $(r, a_{in})$ **of**
$(0,0) : P_{out} := (0,0)$;
$(0,1) : P_{out} := (1,1)$;
$(1,0) : P_{out} := (0,0)$;
$(1,1) : P_{out} := (1,0)$;
**endcase** ;
$s_{out} := s_{in}$;

(a) Operation of `Cel02`.

**case** $(P_{in})$ of
$(0,0) : a_{out} := a_{in}$;
$(1,0) : a_{out} := r + a_{in}$;
$(1,1) : a_{out} := r$; $r := a_{in}$;
$P_{out} := P_{in}$;

(b) Operation of `Cel03`

## 3 The Circuit `Multiplier`

### 3.1 Multiplication Procedure

For any $a, b \in GF(2^m)$, the product $c = ab$ can be calculated strightforwardly using Eq. (1) and the matrix $\boldsymbol{Q}$. Thus the following two steps arise naturally:

1. Calculation of the coefficients matrix $\boldsymbol{Q}$.
2. Matrix-vector multiplication to obtain the coordinates of the product $c = ab$.

The first step coincides with that of the divider. For the second step, a trivial recursive relation is used. Namely, from Eq. (1), for $i = 0, \ldots, m-1$:

$$c_i = \sum_{j=0}^{m-1} q_{ij} b_j = \left( \sum_{j=0}^{m-2} q_{ij} b_j \right) + q_{i,m-1} b_{m-1}.$$

**Table 4.** VHDL description of `Solution`

```
library IEEE;
use IEEE.std_logic_1164.all;

entity Solution is generic (m: positive:=4);
   port ( CLK,CLR,Qin,Cin: in STD_LOGIC;
      A0,A1,A2,A3: in STD_LOGIC;
      Bout: out STD_LOGIC);
end Solution;
architecture Solution_arch of Solution is
component cell02b is
   port (CLK,CLR,Hin,Fin,Ain: in STD_LOGIC;
      Hout,Fout,Aout: out STD_LOGIC);
end component cell02b;
component cell03b is
   port (CLK,CLR,Sin,Ain: in STD_LOGIC;
      H,F,Sout: out STD_LOGIC);
end component cell03b;
component ffd is
   port (CLK,CLR,Data_In: in STD_LOGIC;
      Data_Out: out STD_LOGIC);
end component ffd;
signal Htemp,Ftemp:STD_LOGIC_VECTOR(0 to m);
signal Stemp:STD_LOGIC_VECTOR(0 to m);
signal Qtemp,Atemp:STD_LOGIC_VECTOR(0 to m);

begin Sol4:
   for J in 0 to m-1 generate
      row0: if (J=0) generate
      rowa0: for I in 0 to m generate
         cel030: if (I=0) generate
         cell030: cell03b port map (CLK=>CLK,
            CLR=>CLR,Sin=>Qin,Ain=>A0,
            H=>Htemp(I),F=>Ftemp(I),
            Sout=>Stemp(J));
         end generate cel030;
         cel020: if (I<=m and I>=1) generate
         cell020: cell02b port map
            (CLK=>CLK,CLR=>CLR,
            Hin=>Htemp(I-1),
            Fin=>Ftemp(I-1),Ain=>A1,
            Hout=>Htemp(I),Fout=>Ftemp(I),
            Aout=>Atemp(I));
         end generate cel020;
      end generate rowa0;
      end generate row0;          (to be continued ... )
```

**Table 5.** VHDL description of `Solution` (cont')

```
      rowx: if (J<=m-1 and J>=1) generate
      rowax: for I in 0 to m generate
         ff2: if (I=0) generate
         ffd2: ffd port map (CLK=>CLK,CLR=>CLR,
            Data_In=>Stemp(J-1),Data_Out=>Qtemp(J));
         end generate ff2;
         cel03x: if (I=1) generate
         cell03x: cell03b port map (CLK=>CLK,
            CLR=>CLR,Sin=>Qtemp(J),
            Ain=>Atemp(I-1), H=>Htemp(I),F=>Ftemp(I),
            Sout=>Stemp(J));
         end generate cel03x;
         cel02x: if (I<=m and I>1) generate
         celd02x: for K in I to m+1-J generate
         cell02x: cell02b port map
            (CLK=>CLK,CLR=>CLR,
            Hin=>Htemp(I-1),Fin=>Ftemp(I-1),
            Ain=>A1,Hout=>Htemp(I),
            Fout=>Ftemp(I),Aout=>Atemp(I));
         end generate celd02x;
         end generate cel02x;
      end generate rowax;
      end generate rowx;
   end generate Sol4;
end Solucion_arch;
```

Hence, let

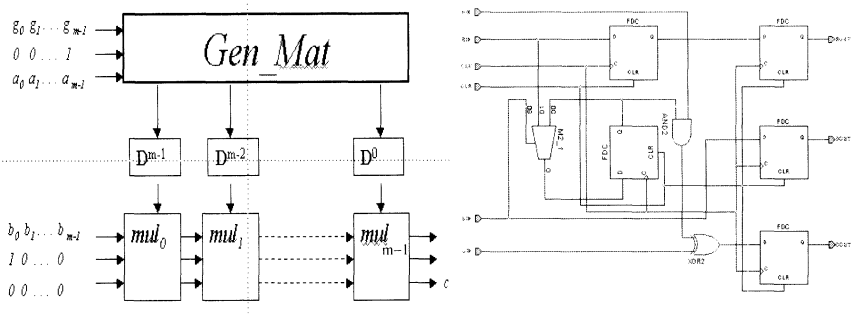
$$\begin{aligned} c_i^{(0)} &= 0 \\ c_i^{(k)} &= c_i^{(k-1)} + q_{i,k-1} b_{k-1} \end{aligned} \tag{2}$$

Obviously, $\forall i : \; c_i = c_i^{(m-1)}$.

### 3.2 Serial Structure `Multiplier`

In Fig. 5-(**a**) we sketch the basic diagram of the structure `Multiplier`. We use `Gen-Mat` for the first step. For the second step, there is a linear array of processors $mul_0$, $mul_1$, ..., $mul_{m-1}$. The operation of cell `mul` is described in the pseudocode in Table 6.

The coordinates of $\boldsymbol{b}$ enter into $mul_0$, with the "most significant bit" coming first. As the coordinates of $\boldsymbol{b}$ pass through the processors, each bit $b_i$ is stored in the internal register of $mul_i$. This processor identifies $b_i$ with the aid of control signal $d$. Once $b_i$ is stored in $mul_i$, $mul_i$ simply performs multiplication and addition operations over $GF(2)$. The coordinates of the product $\boldsymbol{c}$ start emerging from $mul_{m-1}$ at cycle $2m$ at a rate of one coordinate per cycle.




(a) (b)

**Fig. 5.** (**a**) Structure of the multiplier. (**b**) Circuit diagram of cell `mul`

**Table 6.** Description of cell type `mul`.

| |
|---|
| $d_{out} := d_{in}$; |
| $c_{out} := c_{in} + ra_{in}$; |
| $b_{out} := b_{temp}$; |
| $b_{temp} := b_{in}$; |
| **if** $d_{in} := 1$ **then** $r := b_{in}$; |

This gives a computational time of $3m - 1$ cycles. The circuit diagram of cell `mul` is presented in Fig. 5-(**b**). The $D^i$ boxes are delay circuits to synchronize the signal transits from the `Gen-Mat` structure into the multiplier cells.

The code of circuit `Multiplier`, which performs multiplication, is shown in Table 7.

## 4 Some Examples

All tests were made with the field $GF(2^4)$ represented by means of the irreducible polynomial $p(X) = X^4 + X^3 + 1$. Hence, if $\alpha$ is a root of $p(X)$ the cyclic multiplicative group of $GF(2^4)$ is represented as shown in Table 8. $\alpha^4 = \alpha^3 + 1$ is a generator as well. Let us consider $g_{in} = [1\ 0\ 0\ 1]$ and as control signal let $q_{in} = [1\ 0\ 0\ 0]$. As an illustrative example consider as dividend and divisor, respectively:

$$\boldsymbol{a}_{in} = \alpha^6 = \alpha^3 + \alpha^2 + \alpha + 1 = [1\ 1\ 1\ 1]$$
$$\boldsymbol{c}_{in} = \alpha^{14} = \alpha^3 + \alpha^2 = [1\ 1\ 0\ 0]$$

whose quotient evidently is $\boldsymbol{b} = \alpha^8 = [1\ 1\ 1\ 0]$. Indeed, matrices $\boldsymbol{Q}$ and $\boldsymbol{P}$ are:

$$\boldsymbol{P} = \begin{bmatrix} 1\,0\,0\,0\,1\,1\,1\,1\,0\,1\,0\,1\,1\,0\,0 \\ 0\,1\,0\,0\,0\,1\,1\,1\,1\,0\,1\,0\,1\,1\,0 \\ 0\,0\,1\,0\,0\,0\,1\,1\,1\,1\,0\,1\,0\,1\,1 \\ 0\,0\,0\,1\,1\,1\,1\,0\,1\,0\,1\,1\,0\,0\,1 \end{bmatrix} \qquad \boldsymbol{Q} = \begin{bmatrix} 1\,0\,0\,0 \\ 1\,0\,0\,1 \\ 0\,0\,1\,0 \\ 0\,1\,0\,0 \end{bmatrix}$$

**Table 7.** VHDL description of `Solution`

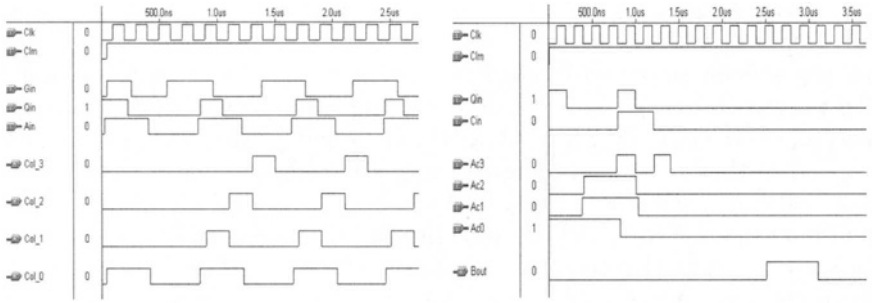
```
library IEEE;
use IEEE.std_logic_1164.all;

entity Sol_multi is generic (m:positive:=4);
   port (CLK,CLR,Binput,Cinput,Dinput: in STD_LOGIC;
      Ainput: in STD_LOGIC_VECTOR(0 to m-1);
      Coutput: out STD_LOGIC);
end Sol_multi;
architecture Sol_multi_arch of Sol_multi is
   component cel_mul
      port (CLK,CLR,Ain,Bin,Cin,Din: in STD_LOGIC;
         Bout,Cout,Dout: out STD_LOGIC);
   end component;
   signal Btemp,Ctemp,Dtemp:STD_LOGIC_VECTOR(0 to m-1);

begin
   Sol: for I in 0 to m-1 generate
      cel0: if (I=0) generate
      cell0:cel_mul port map (CLK=>CLK,CLR=>CLR,
         Ain=>Ainput(I),
         Bin=>Binput, Cin=>Cinput, Din=>Dinput,
         Bout=>Btemp(I),Cout=>Ctemp(I),Dout=>Dtemp(I));
      end generate cel0;
      celx: if (I<=m-1 and I>=1) generate
      celdax: cel_mul port map (CLK=>CLK, CLR=>CLR,
         Ain=>Ainput(I),Bin=>Btemp(I-1),
         Cin=>Ctemp(I-1),Din=>Dtemp(I-1),
         Bout=>Btemp(I),Cout=>Ctemp(I),Dout=>Dtemp(I));
      end generate celx;
   end generate Sol;
   Coutput<=Ctemp(m-1);
end Sol_multi_arch;
```

**Table 8.** Representation of $GF(2^4)^*$ using the irreducible polynomial $p(X)$.

| | | | |
|---|---|---|---|
| $\alpha^0 = 1$ | $\alpha^4 = \alpha^3 + 1$ | $\alpha^8 = \alpha^3 + \alpha^2 + \alpha$ | $\alpha^{12} = \alpha + 1$ |
| $\alpha^1 = \alpha$ | $\alpha^5 = \alpha^3 + \alpha + 1$ | $\alpha^9 = \alpha^2 + 1$ | $\alpha^{13} = \alpha^2 + \alpha$ |
| $\alpha^2 = \alpha^2$ | $\alpha^6 = \alpha^3 + \alpha^2 + \alpha + 1$ | $\alpha^{10} = \alpha^3 + \alpha$ | $\alpha^{14} = \alpha^3 + \alpha^2$ |
| $\alpha^3 = \alpha^3$ | $\alpha^7 = \alpha^2 + \alpha + 1$ | $\alpha^{11} = \alpha^3 + \alpha^2 + 1$ | $\alpha^{15} = 1$ |



(**a**) (**b**)

**Fig. 6.** (**a**) Simulation of `Gen-Mat`. (**b**) Simulation of `Solution`

With this example we obtain the output shown in Fig. 6. It is worth to remark that the first matrix column, marked in graph (a) as *column* 0, initiates in the ascent of pulse 1, the second column, *column* 1, in pulse 2, the third column in pulse 4 and last column in pulse 6. The solution `B_out` = [1 1 1 0] begins to exit after pulse 12. In graph (b) the stages conforming the whole `Divider` are displayed: the `Gen-Mat` circuit, the intermediate stage of flip-flops synchronizing the signals and the solver circuit.

In Fig. 7 we show, using the same factors $a$ and $b$ as above, the behaviour of the multiplier cell `mul`, whose design was shown in Fig. 5-(**b**) and implements the recursion (2).

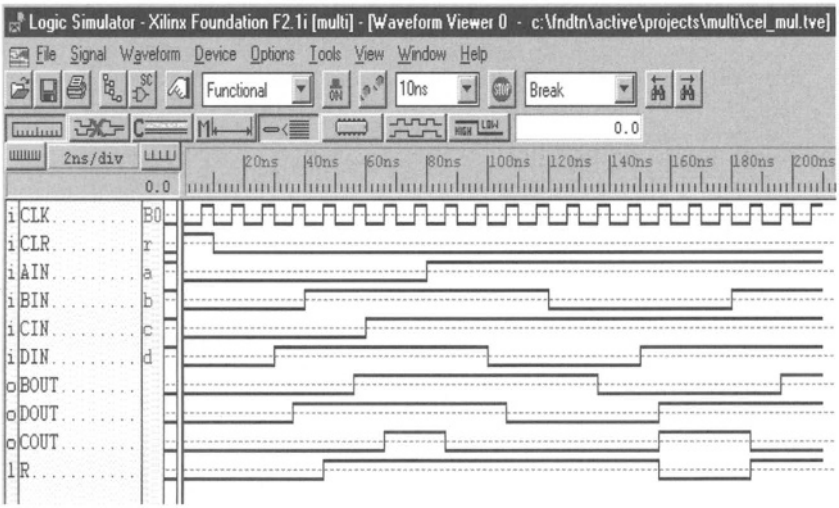



**Fig. 7.** Behaviour of multiplier cell `mul`

## 5 Conclusions

In the present paper we present a serial structure for a `Divider` and a `Multiplier` on $GF(2^m)$. It requires just one control signal and simple and regular interconnections. Regarding its functionning, the clock cycle does not depend on the exponent (coding size, indeed) $m$, the calculations of the division and multiplication require $5m-1$ and $3m-1$, respectively, cycles of clock and both operate in "pipeline" way. In spite that along the current presentation we used an elementary example, we believe that the described circuits are quite adequate for applications of error correction codes for common values of $m$ between 8 and 32. The description with VHDL, for $m = 4, 8, 16$, were validated by simulations using the `Xilinx` [11] package.

A prototype of the divider, for the smaller value $m = 4$, has been built in a PLD, and the construction can be extended up to $m = 16$. For greater values of $m$, up to 256, we plan to construct a proper FPGA. Afterwards, we forsee to use several FPGA's in a cascade composition to treat even greater values of $m$, particularly those suited for cryptographic applications.

**Acknowledgements:** We thank the suggestions of anonymous referees that helped us to improve the final presentation of this paper.

## References

1. Araki K., Fujita I. and Morisue M. "Fast inverter over finite fields based on Euclid's algorithm", *Trans. IEICE*, vol. E-72, pp. 1230-1234, Nov. 1989
2. Chin L.W. "New Systolic Arrays for C + AB2 , Inversion and Division in GF($2^m$)". *IEEE Trans. on Comp.* vol. 49, nr.10, pp. 1120-1125, Oct, 2000
3. Dewey A. M. *Analysis and Design of Digital Systems with VHDL.* PWS Publishing Company, 1997
4. Fenn S. T. J., Benessia M. and Taylor D. , "GF($2^m$) Multiplication and Division Over the Dual Basis", *IEEE Trans. on Comp.*, vol. 45, nr. 3, pp. 319-327, Mar. 1996.
5. García Martínez M. A. "Procesador de División para Campos de Galois en un PLD". Tesis de Maestría, Inst. Tec. de Orizaba, 1999
6. Hasan M. A. and Bhargava V. K. "Bit Serial systolic Divider and Multiplier for Finite Fields GF($2^m$)". *IEEE Trans. Comp.*, vol. 41, nr. 8, pp. 972-979, Aug. 1992
7. Hsu Y.-C., Tsai K. F., Liu J. T. and Lin E. S. *VHDL Modeling for Digital Design Synthesis.* Kluwer Academic Publishers, 1995
8. Lee K.-J. and Yoo K.-Y. "Linear systolic multiplier/squarer for fast exponentiation". *Inf. Proc. Letters*, vol. 76, pp. 105-111, 2000.
9. Peterson W. W. and Weldon E. J. *Error Correcting Codes.* MIT, Cambridge, Massachusetts, 1972
10. Sunar B. and Koç C.K. "Mastrovito Multiplier for All Trinomials". *IEEE Trans. on Comp.* vol. 48, nr. 5, pp. 522-527, May. 1999.
11. *Xilinx. Foundation Series 2.1i Quick Start Guide.* Foundation Series Software from Xilinx, 1999

# Distribution of Irreducible Polynomials over $F_2$

Kenneth H. Hicks[1], Gary L. Mullen[2], and Ikuro Sato[1]

[1] Department of Physics, Ohio University, Athens OH 45701
[2] Department of Mathematics, The Pennsylvania State University, University Park PA 16802, *email:mullen@math.psu.edu*

**Abstract.** Using a polynomial analogue of the wheel sieve, we discuss the distribution of irreducible polynomials over $F_2$. In particular, we provide considerable numerical evidence that in analogue to integer arithmetic progressions, irreducible polynomials over $F_2$ are binomially distributed in the progressions of the wheel sieve. We also present numerical evidence that the irreducibles of fixed degree are binomially distributed by weight. Also briefly discussed is the distribution of self-reciprocal irreducible polynomials. A number of conjectures are raised.

## 1 Introduction

Let $F_q$ denote the finite field of order $q$ where $q$ is a prime number, and let $F_q[x]$ denote the ring of all polynomials over $F_q$ in the variable $x$. It is well known that the ring $Z$ of integers and the polynomial ring $F_q[x]$ share a number of common properties. For example, the ring $Z$ has unique factorization into primes while the ring $F_q[x]$ has unique factorization into irreducible polynomials. Moreover, in each case there are an infinite number of prime elements. In $Z$, this is simply Euclid's Theorem that there are infinitely many primes. In the polynomial setting, this result follows from the fact that for each degree $d \geq 1$, there is an irreducible polynomial of degree $d$ over $F_q$, see [6] page 93.

Dirichlet's Theorem on primes in an arithmetic progression provides a refinement of Euclid's theorem to the effect that if $(a, b) = 1$, then there are infinitely many primes in the progression $an + b$ as $n$ runs through the set of positive integers. In the polynomial ring setting, the analogous result was first proved by Kornblum [5] and states that if $(A(x), B(x)) = 1$, then the progression $A(x)Y + B(x)$ contains infinitely many irreducible polynomials as $Y$ varies through the elements of $F_q[x]$.

While a computer sieve study of the distribution of irreducible polynomials could be conducted for fields of prime or even prime power order, throughout the remainder of this paper we will focus only on the case where $q = 2$. How does one order the polynomials in $F_q[x]$? Corresponding to the polynomial $f(x) = a_n x^n + \cdots + a_1 x + a_0$, we may naturally associate the integer $I_f = a_n 2^n + \cdots + a_1 2 + a_0$. Since each $a_i \in F_2$ and hence may be assumed to be either zero or one, this is of course simply the base 2 representation of

the integer $I_f$. We will often say things like $N_1(x) < N_2(x)$, meaning that subject to the above ordering, the polynomial $N_1(x)$ occurs before $N_2(x)$. While this is an small abuse of the notation, the meaning should be clear from the context.

The wheel sieve for integers was first described by Pritchard [7] as a sublinear algorithm for computer prime number sieve routines. In [3], this technique was used to study the distribution of primes in sets of arithmetic progressions of the form $a + nm_k$ where, if $p_i$ denotes the $i$-th prime, the multiplier $m_k$ is the $k$-th *primorial* number $p_1 \cdot p_2 \ldots p_k$ and $a < m_{k-1}$ is any number relatively prime to $m_k$. The heuristics in [3] show that the primes are distributed binomially among the arithmetic progressions $a + nm_k$, using a binomial probability given by the asymptotic value from Dirichlet's Theorem.

In Section 2 we discuss a polynomial version of the wheel sieve, and in Section 3 we consider the distribution of irreducibles in arithmetic progressions. Section 4 is devoted to a discussion of irreducibles by weight. We close with Section 5, which provides a brief discussion of the distribution of self-reciprical irreducible polynomials.

## 2 The Polynomial Wheel Sieve

For an integer $k \geq 1$, let $M_k(x) = P_1(x) \cdots P_k(x)$ be the product of the first $k$ monic irreducibles in $F_q[x]$. The polynomial $M_k(x)$ corresponds to the $k$-th primorial number $p_1 \cdots p_k$, and will be called the *k-th primorial polynomial.* For each value of $k \geq 1$, the wheel sieve generates a sequence of polynomials, using an interactive process with polynomials from the previous cycle as seeds.

**Definition 1.** For a fixed prime $p_i$, let $W_1 = \{1, 2, \ldots, p_i - 1, x\}$ be the set of initial polynomials. Given $W_k$, let $S_k = \{S \in W_k | \ P_k(x) \not| \ S\}$ be the set after sieving the set $W_k$ by the irreducible $P_k$. Then $W_{k+1} = \{S(x) + N(x)M_k(x) \mid S(x) \in S_k, \ deg(N) < deg(P_k)\}$ and $N(x)$ runs through all polynomials $< P_k$.

Let $\mathbf{W}_k$ be the matrix containing the set $W_k$, with $q^{deg(P_k)}$ columns. The first column is the set $S_{k-1}$, ordered increasingly. And the remaining columns as we move from left to right, contain successive multiples of the primorial polynomial $M_{k-1}(x)$ added to the first polynomial in column 1.

**Example 1:** Let $q = 2$. The first four irreducible polynomials over $F_2$ are $P_1(x) = x$, $P_2(x) = x+1$, $P_3(x) = x^2+x+1$, $P_4(x) = x^3+x+1$, and the first three primorial polynomials are $M_1(x) = x$, $M_2(x) = x^2+x$, $M_3(x) = x^4+x$. Then we have the trivial case

$$W_1 = \{1, x\}, S_1 = \{1\} \ .$$

Continuing we have

$$W_2 = \{1, x+1\}, S_2 = \{1\},$$

and

$$W_3 = \{1, x^2+x+1, x^3+x^2+1, x^3+x+1\}, S_3 = \{1, x^3+x^2+1, x^3+x+1\}$$

or using a more compact notation, where the polynomial $a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ is abbreviated in the form $a_n a_{n-1} \cdots a_0$, we have $S_3 = \{1, 1101, 1011\}$.

For the next case,

$$\mathbf{W_4} = \begin{pmatrix} 1 & 10011 & 100101 & 110111 & 1001001 & 1011011 & 1101101 & 1111111 \\ 1011 & 11001 & 101111 & 111101 & 1000011 & 1010001 & 1100111 & 1110101 \\ 1101 & 11111 & 101001 & 111011 & 1000101 & 1010111 & 1100001 & 1110011 \end{pmatrix}$$

and

$$S_4 = \left\{ \begin{array}{rrrrrrrr} 1 & 10011 & 100101 & 110111 & 1001001 & 1011011 & 1101101 & \\ & 11001 & 101111 & 111101 & 1000011 & 1010001 & 1100111 & 1110101 \\ 1101 & 11111 & 101001 & 111011 & & 1010111 & 1100001 & 1110011 \end{array} \right\}$$

**Remark.** An alternative definition of $S_k$ might be helpful. Since each polynomial in $S_k$ is relatively prime to $M_k(x)$, one could also say $S_k = \{f \in F_q[x] \mid deg(f) < deg(M_k)\ ,\ \gcd(f, M_k) = 1\}$.

As indicated on page 122 of [6], there is a function $\Phi_q$ defined for nonzero polynomials $f$ in $F_q[x]$ which counts the number of polynomials in $F_q[x]$ that are of smaller degree than the degree of $f$ and which are relatively prime to $f$. Lemma 3.69 of [6] provides some of the basic properties of this function, and shows that this function has many of the properties of the Euler function $\phi$ from elementary number theory. The function $\Phi_q$ is multiplicative and if $f \in F_q[x]$ has degree $n \geq 1$, then $\Phi_q(f) = q^n(1-q^{-n_1}) \cdots (1-q^{-n_r})$, where the $n_i$ are the degrees of the distinct monic irreducible polynomials appearing in the canonical factorization of $f$ in $F_q[x]$. This formula can be rewritten to appear to look more like the formula for the usual Euler $\phi$ function. In particular, if $f = P_1^{e_1} \ldots P_r^{e_r}$ where each $P_i$ is irreducible, then

$$\Phi_q(f) = \prod_{i=1}^{r} (q^{n_i e_i} - q^{n_i(e_i-1)}) \quad .$$

**Lemma 1.** *The number of elements in $S_k$ is*

$$\#S_k = \Phi_q(M_k(x)) = \prod_{i=1}^{k} (q^{n_i} - 1)$$

*where $n_i$ is the degree of $P_i(x)$.*

*Proof.* This is a trivial result of the definition of $\Phi_q(f)$.

## 3 Irreducibles in Arithmetic Progressions

The wheel sieve provides a natural framework to study the distribution of irreducible polynomials in sets of arithmetic progressions. Following the notation of Hayes [2], let $H$ be a polynomial over a finite field of $q$ elements and A be a polynomial prime to $H$. If $\pi(r; H, A)$ is the number of irreducibles of degree r which are congruent to $A \pmod{H}$, then a theorem of Artin states

$$\pi(r; H, A) \sim \frac{1}{\Phi_q(H)} \cdot \frac{q^r}{r} \tag{1}$$

with an error term that is $\mathcal{O}(q^{r\nu}/r)$ for some $\nu < 1$, see [2]. We note that the fraction $q^r/r$ is an asymptotic expression for $N_q(r)$, where $N_q(r) = (1/r)\sum_{d|r} \mu(d) q^{r/d}$ is the number of monic irreducibles of degree $r$ over $F_q$, see [6].

If $M_2(n)$ denotes the number of irreducibles over $F_2$ of degree at most $n$, then $M_2(n)$ can be written as the double sum

$$M_2(n) = \sum_{m=1}^{n} \frac{1}{m} \sum_{d|m} \mu(d) 2^{m/d} \ .$$

The asymptotic number of irreducibles in the set $S_k$, after sieving, is given by $M_2(n)$ where $n$ is the largest degree of any polynomial in $S_k$.

Starting with $n = 1$, the first few values of $M_2(n)$ are given by 2, 3, 5, 8, 14, 23, 41, 71, .... A simplified formula or recurrence for $M_2(n)$ would be of interest. A related question is to determine $P_i(x)$, the $i$-th irreducible over $F_2$, subject to the ordering from Section 1.

We are interested in studying heuristics of $\pi(r; H, A)$ and in particular in comparing the error term with the distribution obtained for the polynomial arithmetic progressions of the wheel sieve, where $H = M_k(x)$ and $A$ is taken from the set $S_{k-1}$.

### 3.1 Heuristics of the Distribution of Irreducibles in $S_k$

A computer program was written for $q = 2$ that calculates the elements of the set $S_k$ ordered as in the example for $\mathbf{W_4}$. We chose $q = 2$ because the polynomials can be represented by a string of zeros and ones, as shown for $\mathbf{W_4}$. Each polynomial in $S_k$ was tested for irreducibility using simple bit manipulations such as bit shifts and XOR (exclusive or) operations on the binary string representing the polynomial. The computer output was checked extensively against published tables of irreducible polynomials [6].

The results are given in Table 1 and Figure 1. The numbers given are the distribution of irreducible polynomials found in each "row" of the polynomial arithmetic progressions given in Definition 1. In Example 1, the rows corresponding to the arithmetic progressions are seen clearly for $\mathbf{W_4}$. The

number of irreducibles in each row is easily counted for $k = 4$: in Example 1, $S_4$ has one row has 6 irreducibles and two rows have 7 irreducibles. Similarly, for $k = 5$, there are 11 rows having 6 irreducibles each (see Table 1).

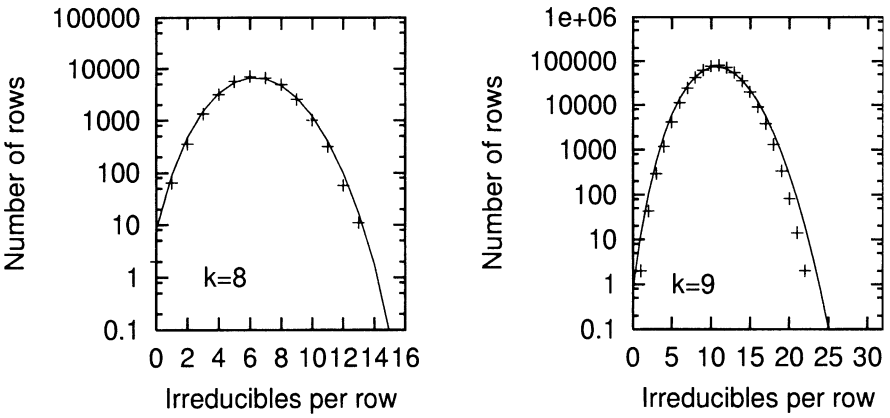



**Fig. 1.** The count of rows with the given number of irreducible polynomials in the matrix $\mathbf{W_k}$ for given $k$. The curve is the prediction from Conjecture 1.

### 3.2 Computational Heuristics Compared to Estimates

The frequency distribution of irreducible polynomials per row in $W_k$ is shown in Figure 1 for $k = 8$ and $k = 9$. When shown in a semi-log plot, this distribution has a parabolic shape which is characteristic of a binomial distribution and so we make

**Definition 2.** A binomial distribution in the parameter $p$ is given by the terms of the expansion $(p + (1 - p))^n$. The mean value of this distribution is $\mu = np$ and the standard deviation is $\sigma = \sqrt{np(1-p)}$.

The solid lines in Figure 1 are calculated using the values of $p$ and $n$ given in the conjecture below. The values of $p$ and $n$ give a mean value $\mu$ for the binomial distribution that, for large $k$, approaches the asymptotic value in (1). The value of $n$ is equal to the number of columns in the matrix representation of $S_k$. Note that the solid lines are calculated and *not* fit to the data. Based upon the data, it is natural to make

*Conjecture 1.* The irreducible polynomials in the progressions given in Definition 1 are distributed so as to asymptotically approach a binomial distribution in the parameter $p = (\Phi_q(M_k(x)))^{-1}(q^r/r)$, where $q = 2$ and $r$ is the degree of $M_k(x)$, and a value of $n = 2^{deg(P_i)} - 1$, where $P_i(x)$ is the $i$-th irreducible polynomial over $F_2$.

**Table 1.** Number of rows with N irreducible polynomials for a given value of $k$.

| | $k$ | | | | |
|---|---|---|---|---|---|
| N | 5 | 6 | 7 | 8 | 9 |
| 0 | 0 | 0 | 0 | 2 | 0 |
| 1 | 0 | 0 | 0 | 64 | 2 |
| 2 | 0 | 0 | 8 | 355 | 43 |
| 3 | 0 | 0 | 10 | 1326 | 294 |
| 4 | 1 | 0 | 69 | 3153 | 1185 |
| 5 | 5 | 1 | 164 | 5648 | 4141 |
| 6 | 11 | 0 | 353 | 7057 | 11166 |
| 7 | 4 | 15 | 522 | 6615 | 24189 |
| 8 | 0 | 32 | 468 | 4936 | 42129 |
| 9 | 0 | 24 | 347 | 2526 | 61697 |
| 10 | 0 | 39 | 193 | 1011 | 76230 |
| 11 | 0 | 32 | 60 | 314 | 80045 |
| 12 | 0 | 4 | 10 | 57 | 71195 |
| 13 | 0 | 0 | 1 | 11 | 54293 |
| 14 | 0 | 0 | 0 | 0 | 35215 |
| 15 | 0 | 0 | 0 | 0 | 19696 |
| 16 | 0 | 0 | 0 | 0 | 9039 |
| 17 | 0 | 0 | 0 | 0 | 3817 |
| 18 | 0 | 0 | 0 | 0 | 1310 |
| 19 | 0 | 0 | 0 | 0 | 341 |
| 20 | 0 | 0 | 0 | 0 | 82 |
| 21 | 0 | 0 | 0 | 0 | 14 |
| 22 | 0 | 0 | 0 | 0 | 2 |

Based on this conjecture, the error term in the distribution of irreducibles is easily computed, based on the standard deviation $\sigma$ of the binomial distribution. The heuristics from the binomial distribution can be compared directly with the error term in (1). The error term given just below (1) has an unknown value of $\nu$ whereas the binomial distribution has all parameters known. For a large number of trials, the binomial distribution may be approximated by a gaussian, with distribution as a function of row $j$,

$$D(j) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(j-\mu)^2/2\sigma^2}$$

where $\mu$ and $\sigma$ are given in Definition 2. When $D(j) = 1$ then the asymptotic estimate in Conjecture 1 is bounded. This occurs when $j-\mu = \pm\sigma\sqrt{\log(2\pi\sigma^2)}$. Now $\mu$ is the location of the peak of the distribution, as given by (1), so the error term is asymptotically $\mathcal{O}(\sigma\sqrt{\log\sigma^2})$. Using the definition of $\sigma$, we see that the error term estimated from Conjecture 1 is $\mathcal{O}(\sqrt{(q^r/r)\log(q^r/r)})$. This estimate for the error term is consistent with the heuristics, as shown in Figure 1, but is dependent on Conjecture 1.

## 4 Distribution of Irreducibles by Weight

Let $W(n,m)$ denote the number of binary irreducibles of degree $n$ and weight $m$, i.e. with $m$ nonzero coefficients. While we are unable to provide a formula, or even a conjecture, for $W(n,m)$, we now provide numerical evidence that the irreducibles of degree $n$ and weight $m$ are binomially distributed. We note that any irreducible must have constant term 1, and have odd weight, otherwise it is divisible by $x$ or $x+1$.

The data for $W(n,m)$ are given in Table 2, for $n$ up to 26, corresponding to the largest degree in cycle $k=9$ of the wheel sieve. The weights appear to be binomially distributed, in this case with binomial probability $p=1/2$. Naively, this is what one might expect from the combinatorics if the weights are randomly chosen. More formally, we make

*Conjecture 2.* The irreducible polynomials of degree $n$ over $F_2$ are binomially distributed by weight.

From the data in Table 2, it seems that for fixed $m$, $W(n,m)$ is an increasing function of $n$, except for $m=3$ and $m=5$. An interesting question to ask is whether $W(n,3)>0$, for "almost all" n when n is large. If one makes a conjecture that the weights follow a binomial distribution, then the weight for the trinomials is easily calculated for monic polynomials over $F_2$ and a binomial probability $p=1/2$ as,

$$2\left(\frac{1}{2}\right)^{n-3}\left(\frac{2^n}{n}\right)=\frac{16}{n}\,, \tag{2}$$

which decreases asymptotically to zero. This would imply that the probability of monic trinomials vanishes as one considers all irreducibles of large degree. Note that (2) does not rule out irreducible trinomials of large degree, but says that the probability of finding one would be vanishingly small for large $n$ if the binomial distribution is an accurate representation of the distribution.

It is difficult to say whether (2) is accurate, because the ends of the weight distribution have small numbers of counts for $W(n,m)$ and thus the statistical errors become significant. For a purely random process with a large number of Bernoulli trials, the distribution follows a Gaussian distribution. Figure 2 shows the data for $W(n,m)$ plotted along with a Gaussian distribution. The amplitude for the gaussian is calculated from $N_2(n)$, the exact number of irreducibles of degree $n$ over $F_2$. The peak of the gaussian is calculated from $(n+3)/2$ for given degree $n$. The standard deviation of the distribution is given by $\sqrt{(n-3)pq}$ where $p=q=0.5$ for an equal probability Bernoulli trial. In other words, there are no free parameters in the Gaussian curve. The agreement between the data for $W(n,m)$ and the Gaussian curve is remarkably good. However, without better heuristics, it is difficult to answer whether $W(n,3)>0$ with finite probability for infinitely many values of $n$.

From [1] we know that $W(2n,3)>0$ for infinitely many $n$. We now raise

**Table 2.** Weight distribution for irreducible polynomials $W(n, m)$ where $n$ is the degree and $m$ is the number of non-zero coefficients.

| | $m$ | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 3 | 5 | 7 | 9 | 11 | 13 | 15 | 17 | 19 | 21 | 23 | 25 |
| 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 2 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 3 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 4 | 10 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 17 | 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 4 | 22 | 28 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 2 | 38 | 44 | 14 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 2 | 46 | 84 | 52 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 4 | 54 | 152 | 110 | 14 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 0 | 66 | 236 | 264 | 60 | 4 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 2 | 73 | 357 | 500 | 214 | 15 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 6 | 98 | 546 | 898 | 546 | 82 | 6 | 0 | 0 | 0 | 0 | 0 |
| 16 | 0 | 94 | 734 | 1587 | 1304 | 337 | 24 | 0 | 0 | 0 | 0 | 0 |
| 17 | 6 | 152 | 1050 | 2674 | 2696 | 1006 | 122 | 4 | 0 | 0 | 0 | 0 |
| 18 | 5 | 124 | 1374 | 4316 | 5406 | 2745 | 531 | 30 | 1 | 0 | 0 | 0 |
| 19 | 0 | 158 | 1774 | 6696 | 10238 | 6766 | 1772 | 190 | 0 | 0 | 0 | 0 |
| 20 | 4 | 199 | 2325 | 9995 | 18405 | 15227 | 5368 | 815 | 39 | 0 | 0 | 0 |
| 21 | 4 | 184 | 2892 | 14988 | 31848 | 32144 | 14698 | 2888 | 212 | 0 | 0 | 0 |
| 22 | 2 | 226 | 3650 | 20993 | 53602 | 64163 | 36877 | 9928 | 1078 | 38 | 0 | 0 |
| 23 | 4 | 296 | 4660 | 29458 | 86626 | 122502 | 86528 | 29748 | 4606 | 286 | 8 | 0 |
| 24 | 0 | 202 | 5191 | 40861 | 136378 | 225569 | 190357 | 81708 | 17063 | 1509 | 32 | 0 |
| 25 | 4 | 406 | 6938 | 55202 | 208988 | 399576 | 399560 | 208542 | 55752 | 6880 | 324 | 4 |
| 26 | 0 | 328 | 8012 | 74404 | 314185 | 685607 | 799042 | 503547 | 166341 | 27390 | 1899 | 40 |

*Conjecture 3.* For fixed odd $m \geq 3$, there are infinitely many values of $n \geq m - 1$ so that $W(n, m) > 0$.

## 5 Distribution of Self-reciprocal Irreducibles

If $f(x)$ is a polynomial of degree $n$, then the *reciprocal* polynomial $f^*(x)$ is defined by $f^*(x) = x^n f(1/x)$, and $f(x)$ is said to be *self-reciprocal* if $f(x) = f^*(x)$. Self-reciprocal irreducibles of degree $> 1$ must have even degree say $2n$, and it is easy to see that if $f(x)$ is irreducible, so is $f^*(x)$. We refer to [4] section 2.7 for a discussion of self-reciprocal irreducibles, including a formula, see page 77, for the number $si(n, 2)$ of self-reciprocal irreducibles of degree $2n$ over $F_2$.

Let $si(n, m, 2)$ be the number of self-reciprocal irreducibles of degree $2n$ and weight $m$ over $F_2$. The distribution of weights for self-reciprocal irreducibles of degree $2n \leq 26$ is given in Table 3.

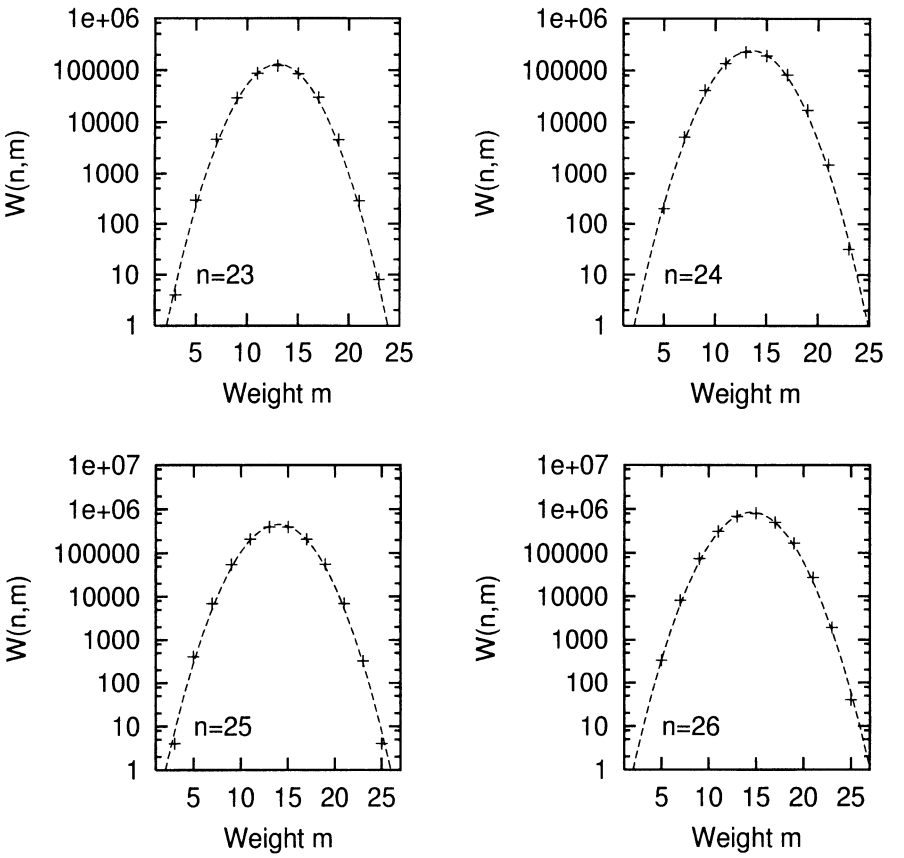



**Fig. 2.** The count of irreducible polynomials of degree $n$ with a given weight $m$. The curve is the prediction from Conjecture 2.

**Table 3.** Distribution of weights for self-reciprocal irreducibles of degree $2n$ with $m$ non-zero coefficients.

| | $m$ | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2n | 3 | 5 | 7 | 9 | 11 | 13 | 15 | 17 | 19 | 21 | 23 | 25 |
| 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 0 | 2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 0 | 0 | 2 | 2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 0 | 1 | 3 | 2 | 2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | 0 | 2 | 2 | 3 | 4 | 5 | 0 | 0 | 0 | 0 | 0 | 0 |
| 18 | 1 | 0 | 2 | 6 | 8 | 7 | 3 | 0 | 1 | 0 | 0 | 0 |
| 20 | 0 | 3 | 1 | 11 | 11 | 11 | 10 | 3 | 1 | 0 | 0 | 0 |
| 22 | 0 | 0 | 4 | 15 | 20 | 19 | 17 | 10 | 8 | 0 | 0 | 0 |
| 24 | 0 | 0 | 7 | 17 | 22 | 37 | 41 | 24 | 15 | 5 | 2 | 0 |
| 26 | 0 | 2 | 2 | 14 | 39 | 77 | 62 | 63 | 35 | 16 | 5 | 0 |

We note from Table 3 for $n \leq 13$ that $si(n, 2)$ is a 0-1 linear combination of the values $si(m, 2)$ for $m < n$. We ask whether this holds for all $n \geq 1$?

It is already known [1] that $si(n, 3, 2) > 0$ for infinitely many values of $n$, and we raise the following:

*Conjecture 4.* For fixed odd $m \geq 3$ there are infinitely many values of $2n > m$ such that $si(n, m, 2) > 0$.

From [1] the self-reciprocal polynomials of degree $2n$ for which $si(n, 3, 2) > 0$ are explicitly given in the form $x^{2n} + x^n + 1$, where $n$ is any non-negative power of 3. It would be of interest to have an analogous explicit form for self-reciprocal irreducibles of degree $2n$ and weight 5, and more generally of degree $2n$ and weight $m$; however this seems to be out of reach at the present.

## References

1. I.F. Blake, S. Gao, R.J. Lambert, *Construction and distribution problems for irreducible trinomials over finite fields*, Applications of Finite Fields, Edited by D. Gollmann, Clarendon Press, Oxford, 1996, 19-32.
2. D.R. Hayes, *The distribution of irreducibles in $GF[q, x]$*, Trans. Amer. Math. Soc. 117(1965), 101-127.
3. K.H. Hicks and I. Sato, *Heuristics of arithmetic progressions in the framework of the wheel sieve*, submitted for publication.
4. D. Jungnickel, Finite Fields: Structure and Arithmetics, Bibliographisches Inst. & F.A. Brockhaus AG, Mannheim, 1993.
5. H. Kornblum, *Über die Primfunktionen in einer arithmetischen Progression*, Math. Z. 5(1919), 100-111.
6. R. Lidl and H. Niederreiter, Finite Fields, Cambridge Univ. Press, 1997.
7. P. Pritchard, *Explaining the wheel sieve*, Acta Informat. 17(1982), 477-485.

# Arithmetic on a Family of Picard Curves

Rolf-Peter Holzapfel and Florin Nicolae

Humboldt-Universität zu Berlin, Institut für Mathematik, Rudower Chaussee 25, D-10099 Berlin, Germany

**Abstract.** The $L$-function of the curve $C_a : Y^3 = X^4 - aX$ over an algebraic number field $k$ which contains $\zeta_9 := \exp(\frac{2\pi i}{9})$ is the inverse of a product of six Hecke $L$-functions with Grössencharakter. The Euler factors at primes of good reduction are determined by means of Jacobi sums associated to certain powers of the 9-th power residue character. The number of points of $C_a$ over a finite field is given in terms of such sums. The jacobian variety of $C_a$ over the field of complex numbers has complex multiplication by the ring $\mathbb{Z}[\zeta_9]$.

Let $k$ be a perfect field of characteristic different from 3. The curves

$$C_a : Y^3 = X^4 - aX, a \in k^*$$

are smooth of genus 3 over $k$, with one point $(0:0:1)$ at infinity. The main result of this paper is that the $L$-function of the curve $C_a$ over an algebraic number field $k$ which contains $\zeta_9 := \exp(\frac{2\pi i}{9})$ is the inverse of a product of six Hecke $L$-functions with Grössencharakter (Theorem 1). As a consequence of this it follows that Hasse's conjecture on the meromorphic continuation and the functional equation of the zeta function is true for the family $C_a$. Since the Jacobians of the curves $C_a$ have complex multiplication, the result on the zeta function fits into the theory of zeta functions of abelian varieties with complex multiplication ([De],[Ta]).

Let $N_1$ denote the number of points of the curve $C_a$ over a finite field $k = \mathbb{F}_q$. If $q \not\equiv 1(\bmod 9)$ then $N_1 = q + 1$. This is proved in propositions 1 and 2. If $q \equiv 1(\bmod 9)$ then

$$N_1 = q + 1 - \mathrm{Tr}_{\mathbb{Q}(\zeta_9)/\mathbb{Q}}(\eta),$$

where

$$\eta := \psi^4(a)\iota(\psi^3, \psi),$$

$\psi$ a character of $k^*$ of order 9, $\iota(\psi^3, \psi)$ the Jacobi sum over $\mathbb{F}_q$ associated to $\psi^3$ and $\psi$. This is proved in proposition 3. Corollaries 1, 2 and proposition 4 give explicit forms of the $L$-polynomial of the curve $C_a$ over $\mathbb{F}_q$ in all cases $q(\bmod 9)$. Proposition 5 gives the arithmetic characterization of the algebraic number $\iota(\psi^3, \psi)$ in the ring $\mathbb{Z}[\zeta_9]$.

Over the field $k = \mathbb{C}$ of complex numbers, all curves $C_a$ are isomorphic to $C_1 : Y^3 = X^4 - X$. The moduli point of $C_1$ is the only orbitally isolated singularity on the modular surface of Picard curves. The endomorphism ring

of the jacobian variety $J(C_1)$ of $C_1$ is the ring $\mathbb{Z}[\zeta_9]$. Up to isomorphism, $C_1$ is the only Picard curve whose jacobian variety has a cyclotomic maximal order as endomorphism ring. This is proved in proposition 7. In proposition 8 is given explicitly a period matrix of $J(C_1)$:

$$\Pi = \begin{pmatrix} -\zeta_9+1 & 0 & -2\zeta_9{}^2-2\zeta_9 & -\zeta_9{}^2-1 & 1 & 2\zeta_9{}^2+\zeta_9 \\ \zeta_9{}^2-1 & 0 & -\zeta_9{}^2+2\zeta_9 & -\zeta_9{}^2+\zeta_9+1 & -1 & \zeta_9{}^2-2\zeta_9 \\ -\zeta_9+1 & 0 & -2\zeta_9{}^2-2\zeta_9 & -\zeta_9{}^2-1 & 1 & 2\zeta_9{}^2+\zeta_9 \end{pmatrix} \cdot \zeta_9{}^3 +$$

$$+ \begin{pmatrix} 2\zeta_9{}^2+\zeta_9+1 & 1 & -\zeta_9+1 & -2\zeta_9{}^2-\zeta_9 & 0 & \zeta_9{}^2+\zeta_9-1 \\ -\zeta_9{}^2+2\zeta_9 & 1 & -2\zeta_9{}^2+2\zeta_9+1 & -\zeta_9+1 & -1 & \zeta_9{}^2-\zeta_9-1 \\ 2\zeta_9{}^2+\zeta_9+1 & 1 & -\zeta_9+1 & -2\zeta_9{}^2-\zeta_9 & 0 & \zeta_9{}^2+\zeta_9-1 \end{pmatrix}.$$

Picard curves of equation type $Y^3 = X^4 - a$ are considered in [Lac].

This research was supported by the Deutsche Forschungsgemeinschaft.

## 1 The Curves $C_a : Y^3 = X^4 - aX$ over $\mathbb{F}_q$

Let $k = \mathbb{F}_q$ be a finite field of characteristic $p \neq 3$ with $q = p^f$ elements, and let $a \in k^*$. The curve

$$C_a : y^3 = x^4 - ax$$

is smooth of genus 3 over $k$. Let $F_a/k$ be the function field of $C_a$, let $\mathbb{P}_{F_a}$ denote the set of places, and let $\mathrm{Div}F_a$ denote the group of divisors of $F_a/k$. The absolute norm $\mathfrak{N}(\mathfrak{P})$ of a place $\mathfrak{P} \in \mathbb{P}_{F_a}$ is the cardinality of its residue class field. It holds $\mathfrak{N}(\mathfrak{P}) = q^{\deg \mathfrak{P}}$, with a natural number $\deg \mathfrak{P} \geq 1$, the degree of $\mathfrak{P}$. The Zeta function of the curve $C_a$ is a meromorphic function in the complex plane, defined for $\Re s > 1$ by

$$\zeta_{C_a}(s) = \prod_{\mathfrak{P} \in \mathbb{P}_{F_a}} \frac{1}{1 - \frac{1}{\mathfrak{N}(\mathfrak{P})^s}} = \sum_{\mathfrak{A} \in \mathrm{Div}F_a,\, \mathfrak{A} \geq 0} \frac{1}{\mathfrak{N}(\mathfrak{A})^s}.$$

Denoting for $n \geq 0$ by $A_n$ the number of positive divisors of degree $n$ it holds

$$\zeta_{C_a}(s) = \sum_{n=0}^{\infty} \frac{A_n}{q^{ns}}.$$

The power series

$$Z_{C_a}(t) := \sum_{n=0}^{\infty} A_n t^n$$

is convergent for $|t| < q^{-1}$ and represents a rational function

$$Z_{C_a}(t) = \frac{L_{C_a}(t)}{(1-t)(1-qt)},$$

where $L_{C_a}(t)$ is a polynomial with coefficients in $\mathbb{Z}$ of the form:

$$L_{C_a}(t) = 1 + a_1 t + a_2 t^2 + a_3 t^3 + q a_2 t^4 + q^2 a_1 t^5 + q^3 t^6.$$

For $r \geq 1$ let $N_r$ be the number of $\mathbb{F}_{q^r}$-rational points of the complete curve $C_a$, and let $S_r := N_r - (q^r + 1)$. It holds

$$a_1 = S_1,$$

$$2a_2 = S_2 + S_1 a_1,$$

$$3a_3 = S_3 + S_2 a_1 + S_1 a_2.$$

The plane curve $C_a$ has only one point at infinity, hence

$$N_1 = N + 1$$

where $N$ is the number of solutions $(x, y)$ in $k$ of the equation

$$y^3 = x^4 - ax.$$

**Proposition 1.** *If $q \equiv 2(\bmod 3)$ then $N_1 = q + 1$.*

P r o o f: If $q \equiv 2(\bmod 3)$ the order $q - 1$ of the cyclic multiplicative group $k^*$ is not divisible by 3, so $k^* = k^{*3}$. This implies that for each $x \in k$ there exists exactly one $y \in k$ with $y^3 = x^4 - ax$. Hence $N = q$. □

**Proposition 2.** *If $q \equiv 4(\bmod 9)$ or $q \equiv 7(\bmod 9)$ then $N_1 = q + 1$.*

P r o o f: If $q \equiv 4(\bmod 9)$ or $q \equiv 7(\bmod 9)$ then the cyclic multiplicative group $k^*$ of order $q - 1$ is equal to the internal direct product of its subgroup of order 3, generated by $\zeta$, and of its subgroup of order $\frac{q-1}{3}$, denoted by $U_{\frac{q-1}{3}}$. Each element $c \in \mathbb{F}_q^*$ can be uniquely written in the form $c = d\zeta^j$ with $d \in U_{\frac{q-1}{3}}$ and $0 \leq j \leq 2$. Let $\chi$ be a character of $k^*$ of order 3. Put $\chi(0) := 0$. The number of solutions in $k$ of the equation $y^3 = x^4 - ax$ is

$$N = q + \sum_{c \in \mathbb{F}_q} \chi(c^4 - ac) + \sum_{c \in \mathbb{F}_q} \chi^2(c^4 - ac) = q + \alpha + \bar{\alpha},$$

where

$$\alpha = \sum_{c \in \mathbb{F}_q} \chi(c^4 - ac) = \sum_{d \in U_{\frac{q-1}{3}}} \sum_{j=0}^{2} \chi(d^4 \zeta^{4j} - ad\zeta^j) =$$

$$= \sum_{d \in U_{\frac{q-1}{3}}} \sum_{j=0}^{2} \chi[\zeta^j (d^4 - ad)] = [\sum_{d \in U_{\frac{q-1}{3}}} \chi(d^4 - ad)] \cdot [\sum_{j=0}^{2} \chi(\zeta^j)] =$$

$$= [\sum_{d \in U_{\frac{q-1}{3}}} \chi(d^4 - ad)] \cdot [\chi(1) + \chi(\zeta) + \chi(\zeta)^2].$$

If $q \equiv 4(\bmod 9)$ or $q \equiv 7(\bmod 9)$ then $\frac{q-1}{3}$ is prime to 3, so $\chi$ is not trivial on the subgroup of $k^*$ of order 3. This implies

$$\chi(1) + \chi(\zeta) + \chi(\zeta)^2 = 0,$$

so $\alpha = 0$ and $N = q$. □

**Corollary 1.** *If* $q \equiv 2(\bmod 9)$ *or* $q \equiv 5(\bmod 9)$ *then*

$$L_{C_a}(t) = 1 + q^3 t^6.$$

P r o o f: If $q \equiv 2(\bmod 9)$ or $q \equiv 5(\bmod 9)$ then $q \equiv 2(\bmod 3)$, $q^2 \equiv 4(\bmod 9)$ or $q^2 \equiv 7(\bmod 9)$, and $q^3 \equiv 2(\bmod 3)$. By Propositions 9 and 10 it holds $N_1 = q + 1$, $N_2 = q^2 + 1$, $N_3 = q^3 + 1$. So $S_i = N_i - (q^i + 1) = 0$ for $i = 1, 2, 3$ and $a_1 = a_2 = a_3 = 0$. Hence $L_{C_a}(t) = 1 + q^3 t^6$. □

For a character $\varphi$ of the multiplicative group $k^*$ let

$$\tau(\varphi) := -\sum_{c \in k^*} \varphi(c) \exp(\frac{2\pi i}{p} \mathrm{Tr}_{k/\mathbb{F}_p} c)$$

be the corresponding Gauss sum ([Da-Ha]). For an element $d \in k^*$ define

$$\tau_d(\varphi) := -\sum_{c \in k^*} \varphi(c) \exp(\frac{2\pi i}{p} \mathrm{Tr}_{k/\mathbb{F}_p} cd).$$

It holds

$$\tau_d(\varphi) = \varphi^{-1}(d)\tau(\varphi). \tag{1}$$

For two characters $\varphi_1$ and $\varphi_2$ of $k^*$ let

$$\iota(\varphi_1, \varphi_2) := -\sum_{c \in k} \varphi_1(c)\varphi_2(1 - c)$$

be the corresponding Jacobi sum. If $\varphi_1 \cdot \varphi_2 \neq 1$ then

$$\iota(\varphi_1, \varphi_2) = \frac{\tau(\varphi_1)\tau(\varphi_2)}{\tau(\varphi_1\varphi_2)}. \tag{2}$$

For each natural number $m \geq 1$ let $\zeta_m := \exp \frac{2\pi i}{m}$ and let $\mu_m := \{\zeta_m^l \mid 0 \leq l \leq m - 1\}$ be the group of complex $m$-th roots of unity.

**Proposition 3.** *If* $q \equiv 1(\bmod 9)$ *then*

$$N_1 = q + 1 - \mathrm{Tr}_{\mathbb{Q}(\zeta_9)/\mathbb{Q}}(\eta),$$

*where*

$$\eta := \psi^4(a)\iota(\psi^3, \psi),$$

$\psi$ *a character of* $k^*$ *of order* 9.

The number of elements of a finite set $X$ is denoted by $|X|$. It holds

**Lemma 1.** *Let $k = \mathbb{F}_q$ be a finite field of characteristic $p \neq 3$, and let $\xi$ be a generator of the cyclic multiplicative group $k^*$. If $B(x) \in k[x]$ is a polynomial with a simple root $x_1 \in k$:*

$$B(x) = (x - x_1)B_1(x), B_1(x) \in k[x], B_1(x_1) \neq 0,$$

*then the number of solutions in $k$ of the equation*

$$y^3 = B(x)$$

*is*

$$N = \frac{1}{3}(|\mathcal{A}_{11}| + |\mathcal{A}_{\xi\xi^2}| + |\mathcal{A}_{\xi^2\xi}|),$$

*where*

$$\mathcal{A}_{11} := \{(t,u) \in k \times k \mid B_1(t^3 + x_1) = u^3\},$$
$$\mathcal{A}_{\xi\xi^2} := \{(t,u) \in k \times k \mid B_1(\xi t^3 + x_1) = \xi^2 u^3\},$$
$$\mathcal{A}_{\xi^2\xi} := \{(t,u) \in k \times k \mid B_1(\xi^2 t^3 + x_1) = \xi u^3\}.$$

P r o o f: I) The case $q \equiv 1(\bmod 3)$. Let $\chi$ be a character of $k^*$ of order 3 such that

$$\chi(\xi) = \omega = e^{\frac{2\pi i}{3}}.$$

Put $\chi(0) := 0$. It holds

$$N = q + \alpha + \bar{\alpha},$$

with

$$\alpha = \sum_{c \in k} \chi(B(c)) = \sum_{c \in k} \chi((c - x_1)B_1(c)) = \sum_{c \in k} \chi(c - x_1)\chi(B_1(c)) =$$

$$= \sum_{i,j=0}^{2} \sum_{c \in A, \chi(c-x_1)=\omega^i, \chi(B_1(c))=\omega^j} \omega^{i+j} =$$

$$= |A_{11}| + |A_{\omega\omega^2}| + |A_{\omega^2\omega}| + \omega(|A_{1\omega}| + |A_{\omega 1}| + |A_{\omega^2\omega^2}|) +$$
$$+\omega^2(|A_{1\omega^2}| + |A_{\omega\omega}| + |A_{\omega^2 1}|),$$

where

$$A := \{c \in k \mid B(c) \neq 0\},$$
$$A_{\omega^i\omega^j} = \{c \in A \mid \chi(c - x_1) = \omega^i, \chi(B_1(c)) = \omega^j\},$$

for $i, j = 0, 1, 2$. It follows that

$$\alpha + \bar{\alpha} = 2(|A_{11}| + |A_{\omega\omega^2}| + |A_{\omega^2\omega}|) + (\omega + \omega^2)(|A_{1\omega}| + |A_{\omega 1}| + |A_{\omega^2\omega^2}|) +$$
$$+(\omega^2 + \omega)(|A_{1\omega^2}| + |A_{\omega\omega}| + |A_{\omega^2 1}|) = 2(|A_{11}| + |A_{\omega\omega^2}| + |A_{\omega^2\omega}|) -$$

$$-(|A_{1\omega}| + |A_{\omega 1}| + |A_{\omega^2\omega^2}|) - (|A_{1\omega^2}| + |A_{\omega\omega}| + |A_{\omega^2 1}|) =$$

$$= 3(|A_{11}| + |A_{\omega\omega^2}| + |A_{\omega^2\omega}|) - \sum_{i,j=0}^{2} |A_{\omega^i\omega^j}| =$$

$$3(|A_{11}| + |A_{\omega\omega^2}| + |A_{\omega^2\omega}|) - |A|, \tag{3}$$

since the sets $A_{\omega^i\omega^j}$, $i, j = 0, 1, 2$, form a partition of the set $A$.
It holds

$$A_{11} = \{c \in A \mid \chi(c - x_1) = 1, \chi(B_1(c)) = 1\} =$$

$$= \{c \in A \mid (\exists)(t, u) \in k^* \times k^* : c - x_1 = t^3, B_1(c) = u^3\}.$$

Let

$$\mathcal{B}_{11} := \{(0, u) \mid u \in k, u^3 = B_1(x_1)\} \cup \{(t, 0) \mid t \in k, B_1(t^3 + x_1) = 0\}.$$

The map

$$g_{11} : \mathcal{A}_{11} \setminus \mathcal{B}_{11} \to A_{11}$$

$$g_{11}(t, u) := t^3 + x_1$$

is precisely 9:1 : For $c \in A_{11}$ and $(t, u) \in g_{11}^{-1}(c)$ it holds:

$$g_{11}^{-1}(c) = \{(\zeta^i t, \zeta^j u) \mid 0 \le i, j \le 2\},$$

where $\zeta$ is an element of $k^*$ of order 3, so $|g_{11}^{-1}(c)| = 9$. Hence

$$|A_{11}| = \frac{1}{9}|\mathcal{A}_{11}| - \frac{1}{9}|\{c \in k \mid c^3 = B_1(x_1)\}| - \frac{1}{9}|\{c \in k \mid B_1(c^3 + x_1) = 0\}|. \tag{4}$$

It holds

$$A_{\omega\omega^2} = \{c \in A \mid \chi(c - x_1) = \omega, \chi(B_1(c)) = \omega^2\} =$$

$$= \{c \in A \mid (\exists)(t, u) \in k^* \times k^* : c - x_1 = \xi t^3, B_1(c) = \xi^2 u^3\}.$$

Let

$$\mathcal{B}_{\xi\xi^2} := \{(0, u) \mid u \in k, \xi^2 u^3 = B_1(x_1)\} \cup \{(t, 0) \mid t \in k, B_1(\xi t^3 + x_1) = 0\}.$$

The map

$$g_{\omega\omega^2} : \mathcal{A}_{\xi\xi^2} \setminus \mathcal{B}_{\xi\xi^2} \to A_{\omega\omega^2}$$

$$g_{\omega\omega^2}(t, u) := \xi t^3 + x_1$$

is also precisely 9:1 : For $c \in A_{\omega\omega^2}$ and $(t, u) \in g_{\omega\omega^2}^{-1}(c)$ it holds:

$$g_{\omega\omega^2}^{-1}(c) = \{(\zeta^i t, \zeta^j u) \mid 0 \le i, j \le 2\},$$

so $|g_{\omega\omega^2}^{-1}(c)| = 9$. Hence

$$|A_{\omega\omega^2}| = \frac{1}{9}|\mathcal{A}_{\xi\xi^2}| - \frac{1}{9}|\{c \in k \mid \xi^2 c^3 = B_1(x_1)\}| -$$

$$-\frac{1}{9}|\{c \in k \mid B_1(\xi c^3 + x_1) = 0\}|. \tag{5}$$

Analogously:

$$|A_{\omega^2\omega}| = \frac{1}{9}|\mathcal{A}_{\xi^2\xi}| - \frac{1}{9}|\{c \in k \mid \xi c^3 = B_1(x_1)\}| -$$

$$-\frac{1}{9}|\{c \in k \mid B_1(\xi^2 c^3 + x_1) = 0\}|. \tag{6}$$

From (3), (4), (5) and (6) it follows that

$$\alpha + \bar{\alpha} = 3(|A_{11}| + |A_{\omega\omega^2}| + |A_{\omega^2\omega}|) - |A| =$$

$$= \frac{1}{3}(|\mathcal{A}_{11}| + |\mathcal{A}_{\xi\xi^2}| + |\mathcal{A}_{\xi^2\xi}|) -$$

$$-\frac{1}{3}(|\{c \in k \mid c^3 = B_1(x_1)\}| + |\{c \in k \mid \xi c^3 = B_1(x_1)\}| +$$

$$+|\{c \in k \mid \xi^2 c^3 = B_1(x_1)\}|) -$$

$$-\frac{1}{3}(|\{c \in k \mid B_1(c^3 + x_1) = 0\}| + |\{c \in k \mid B_1(\xi c^3 + x_1) = 0\}| +$$

$$+|\{c \in k \mid B_1(\xi^2 c^3 + x_1) = 0\}|) - |A| =$$

$$= \frac{1}{3}(|\mathcal{A}_{11}| + |\mathcal{A}_{\xi\xi^2}| + |\mathcal{A}_{\xi^2\xi}|) - 1 - |\{d \in k \mid B_1(d) = 0\}| - |A|.$$

It holds

$$|A| = q - |\{c \in k \mid B(c) = 0\}| = q - 1 - |\{d \in k \mid B_1(d) = 0\}|,$$

hence

$$\alpha + \bar{\alpha} = \frac{1}{3}(|\mathcal{A}_{11}| + |\mathcal{A}_{\xi\xi^2}| + |\mathcal{A}_{\xi^2\xi}|) - q$$

and

$$N = q + \alpha + \bar{\alpha} = \frac{1}{3}(|\mathcal{A}_{11}| + |\mathcal{A}_{\xi\xi^2}| + |\mathcal{A}_{\xi^2\xi}|).$$

II) The case $q \equiv 2 (\bmod 3)$. Each element of $k$ has one and only one third root in $k$. It holds

$$N = q,\ |\mathcal{A}_{11}| = |\mathcal{A}_{\xi\xi^2}| = |\mathcal{A}_{\xi^2\xi}| = q.\square$$

P r o o f of Proposition 3: The polynomial $B(x) = x^4 - ax = x(x^3 - ax)$ has the root $x_1 = 0$ in $k$. Let $B_1(x) := x^3 - a \in k[x]$. With the notations of Lemma 1 it holds:

$$\mathcal{A}_{11} = \{(t,u) \in k \times k \mid B_1(t^3 + x_1) = u^3\} = \{(t,u) \in k \times k \mid -u^3 + t^9 = a\},$$

$$\mathcal{A}_{\xi\xi^2} = \{(t,u) \in k \times k \mid -\xi^2 u^3 + \xi^3 t^9 = a\},$$
$$\mathcal{A}_{\xi^2\xi} = \{(t,u) \in k \times k \mid -\xi u^3 + \xi^6 t^9 = a\}.$$

The equation

$$a_1 u^3 + a_2 t^9 = a_3$$

with $a_1, a_2, a_3 \in k \setminus \{0\}$ has by ([Da-Ha], 6.2 and 6.5)

$$N(a_1,a_2,a_3) =$$

$$= q - \psi^3(-\frac{a_1}{a_2}) - \psi^6(-\frac{a_1}{a_2}) - \sum_{\chi^\mu \neq 1, \psi^\nu \neq 1, \chi^\mu\psi^\nu \neq 1} \frac{\tau_{a_1}(\chi^\mu)\tau_{a_2}(\psi^\nu)}{\tau_{a_3}(\chi^\mu\psi^\nu)} =$$

$$= q - \chi(-\frac{a_1}{a_2}) - \chi^2(-\frac{a_1}{a_2}) - \sum_{1 \leq \mu \leq 2} \sum_{1 \leq \nu \leq 8, 3\mu+\nu \neq 9} \frac{\tau_{a_1}(\psi^{3\mu})\tau_{a_2}(\psi^\nu)}{\tau_{a_3}(\psi^{3\mu+\nu})} =$$

$$= q - \chi(-\frac{a_1}{a_2}) - \chi^2(-\frac{a_1}{a_2}) - \sum_{\nu=1,\nu\neq 6}^{8} \frac{\tau_{a_1}(\psi^3)\tau_{a_2}(\psi^\nu)}{\tau_{a_3}(\psi^{3+\nu})} - \sum_{\nu=1,\nu\neq 3}^{8} \frac{\tau_{a_1}(\psi^6)\tau_{a_2}(\psi^\nu)}{\tau_{a_3}(\psi^{6+\nu})}$$

solutions in $k$. Hence

$$|\mathcal{A}_{11}| = N(-1,1,a) = q - 2 - \sum_{\nu=1,\nu\neq 6}^{8} \frac{\tau_{-1}(\psi^3)\tau_1(\psi^\nu)}{\tau_a(\psi^{3+\nu})} -$$

$$- \sum_{\nu=1,\nu\neq 3}^{8} \frac{\tau_{-1}(\psi^6)\tau_1(\psi^\nu)}{\tau_a(\psi^{6+\nu})},$$

$$|\mathcal{A}_{\xi\xi^2}| = N(-\xi^2, \xi^3, a) =$$

$$= q - \chi(\xi^{-1}) - \chi^2(\xi^{-1}) - \sum_{\nu=1,\nu\neq 6}^{8} \frac{\tau_{-\xi^2}(\psi^3)\tau_{\xi^3}(\psi^\nu)}{\tau_a(\psi^{3+\nu})} -$$

$$- \sum_{\nu=1,\nu\neq 3}^{8} \frac{\tau_{-\xi^2}(\psi^6)\tau_{\xi^3}(\psi^\nu)}{\tau_a(\psi^{6+\nu})} =$$

$$= q + 1 - \sum_{\nu=1,\nu\neq 6}^{8} \frac{\tau_{-\xi^2}(\psi^3)\tau_{\xi^3}(\psi^\nu)}{\tau_a(\psi^{3+\nu})} - \sum_{\nu=1,\nu\neq 3}^{8} \frac{\tau_{-\xi^2}(\psi^6)\tau_{\xi^3}(\psi^\nu)}{\tau_a(\psi^{6+\nu})}$$

and

$$|\mathcal{A}_{\xi^2\xi}| = N(-\xi, \xi^6, a) =$$

$$q - \chi(\xi^{-5}) - \chi^2(\xi^{-5}) - \sum_{\nu=1,\nu\neq 6}^{8} \frac{\tau_{-\xi}(\psi^3)\tau_{\xi^6}(\psi^\nu)}{\tau_a(\psi^{3+\nu})} - \sum_{\nu=1,\nu\neq 3}^{8} \frac{\tau_{-\xi}(\psi^6)\tau_{\xi^6}(\psi^\nu)}{\tau_a(\psi^{6+\nu})} =$$

$$= q + 1 - \sum_{\nu=1,\nu\neq 6}^{8} \frac{\tau_{-\xi}(\psi^3)\tau_{\xi^6}(\psi^\nu)}{\tau_a(\psi^{3+\nu})} - \sum_{\nu=1,\nu\neq 3}^{8} \frac{\tau_{-\xi}(\psi^6)\tau_{\xi^6}(\psi^\nu)}{\tau_a(\psi^{6+\nu})}.$$

It follows that

$$|\mathcal{A}_{11}| + |\mathcal{A}_{\xi\xi^2}| + |\mathcal{A}_{\xi^2\xi}| =$$

$$= 3q - \sum_{\nu=1,\nu\neq 6}^{8} \frac{\tau_{-1}(\psi^3)\tau_1(\psi^\nu) + \tau_{-\xi^2}(\psi^3)\tau_{\xi^3}(\psi^\nu) + \tau_{-\xi}(\psi^3)\tau_{\xi^6}(\psi^\nu)}{\tau_a(\psi^{3+\nu})} -$$

$$- \sum_{\nu=1,\nu\neq 3}^{8} \frac{\tau_{-1}(\psi^6)\tau_1(\psi^\nu) + \tau_{-\xi^2}(\psi^6)\tau_{\xi^3}(\psi^\nu) + \tau_{-\xi}(\psi^6)\tau_{\xi^6}(\psi^\nu)}{\tau_a(\psi^{6+\nu})}. \tag{7}$$

By (1) it holds

$$\tau_{-1}(\psi^3)\tau_1(\psi^\nu) + \tau_{-\xi^2}(\psi^3)\tau_{\xi^3}(\psi^\nu) + \tau_{-\xi}(\psi^3)\tau_{\xi^6}(\psi^\nu) = \psi^{-3}(-1)\tau(\psi^3)\tau(\psi^\nu) +$$

$$+\psi^{-3}(-1)\psi^{-3\nu-6}(\xi)\tau(\psi^3)\tau(\psi^\nu) + \psi^{-3}(-1)\psi^{-6\nu-3}(\xi)\tau(\psi^3)\tau(\psi^\nu) =$$

$$= \tau(\psi^3)\tau(\psi^\nu)(1 + \psi^{-3\nu-6}(\xi) + \psi^{-6\nu-3}(\xi)) =$$

$$= \tau(\psi^3)\tau(\psi^\nu)(1 + \chi^{-\nu-2}(\xi) + \chi^{-2\nu-1}(\xi)) =$$

$$= \tau(\psi^3)\tau(\psi^\nu)(1 + \omega^{-\nu-2} + \omega^{2(-\nu-2)}),$$

so

$$\sum_{\nu=1,\nu\neq 6}^{8} \frac{\tau_{-1}(\psi^3)\tau_1(\psi^\nu) + \tau_{-\xi^2}(\psi^3)\tau_{\xi^3}(\psi^\nu) + \tau_{-\xi}(\psi^3)\tau_{\xi^6}(\psi^\nu)}{\tau_a(\psi^{3+\nu})} =$$

$$= 3\frac{\tau(\psi^3)\tau(\psi)}{\tau_a(\psi^4)} + 3\frac{\tau(\psi^3)\tau(\psi^4)}{\tau_a(\psi^7)} + 3\frac{\tau(\psi^3)\tau(\psi^7)}{\tau_a(\psi)}. \tag{8}$$

Analogously:

$$\tau_{-1}(\psi^6)\tau_1(\psi^\nu) + \tau_{-\xi^2}(\psi^6)\tau_{\xi^3}(\psi^\nu) + \tau_{-\xi}(\psi^6)\tau_{\xi^6}(\psi^\nu) = \psi^{-6}(-1)\tau(\psi^6)\tau(\psi^\nu) +$$

$$+\psi^{-6}(-1)\psi^{-3\nu-12}(\xi)\tau(\psi^6)\tau(\psi^\nu) + \psi^{-6}(-1)\psi^{-6\nu-6}(\xi)\tau(\psi^6)\tau(\psi^\nu) =$$

$$= \tau(\psi^6)\tau(\psi^\nu)(1 + \psi^{-3\nu-12}(\xi) + \psi^{-6\nu-6}(\xi)) =$$

$$= \tau(\psi^6)\tau(\psi^\nu)(1 + \chi^{-\nu-4}(\xi) + \chi^{-2\nu-2}(\xi)) =$$

$$= \tau(\psi^6)\tau(\psi^\nu)(1 + \omega^{-\nu-1} + \omega^{2(-\nu-1)}),$$

so

$$\sum_{\nu=1,\nu\neq 3}^{8} \frac{\tau_{-1}(\psi^6)\tau_1(\psi^\nu) + \tau_{-\xi^2}(\psi^6)\tau_{\xi^3}(\psi^\nu) + \tau_{-\xi}(\psi^6)\tau_{\xi^6}(\psi^\nu)}{\tau_a(\psi^{6+\nu})} =$$

$$= 3\frac{\tau(\psi^6)\tau(\psi^2)}{\tau_a(\psi^8)} + 3\frac{\tau(\psi^6)\tau(\psi^5)}{\tau_a(\psi^2)} + 3\frac{\tau(\psi^6)\tau(\psi^8)}{\tau_a(\psi^5)}. \tag{9}$$

By (7), (8) and (9) it holds

$$|\mathcal{A}_{11}| + |\mathcal{A}_{\xi\xi^2}| + |\mathcal{A}_{\xi^2\xi}| = 3q - 3\frac{\tau(\psi^3)\tau(\psi)}{\tau_a(\psi^4)} - 3\frac{\tau(\psi^3)\tau(\psi^4)}{\tau_a(\psi^7)} - 3\frac{\tau(\psi^3)\tau(\psi^7)}{\tau_a(\psi)} -$$
$$-3\frac{\tau(\psi^6)\tau(\psi^2)}{\tau_a(\psi^8)} - 3\frac{\tau(\psi^6)\tau(\psi^5)}{\tau_a(\psi^2)} - 3\frac{\tau(\psi^6)\tau(\psi^8)}{\tau_a(\psi^5)},$$

by Lemma 1

$$N = q - \frac{\tau(\psi^3)\tau(\psi)}{\tau_a(\psi^4)} - \frac{\tau(\psi^3)\tau(\psi^4)}{\tau_a(\psi^7)} - \frac{\tau(\psi^3)\tau(\psi^7)}{\tau_a(\psi)} -$$
$$-\frac{\tau(\psi^6)\tau(\psi^2)}{\tau_a(\psi^8)} - \frac{\tau(\psi^6)\tau(\psi^5)}{\tau_a(\psi^2)} - \frac{\tau(\psi^6)\tau(\psi^8)}{\tau_a(\psi^5)} =$$
$$= q - \psi^4(a)\iota(\psi^3,\psi) - \psi^7(a)\iota(\psi^3,\psi^4) - \psi(a)\iota(\psi^3,\psi^7) -$$
$$-\psi^8(a)\iota(\psi^6,\psi^2) - \psi^2(a)\iota(\psi^6,\psi^5) - \psi^5(a)\iota(\psi^6,\psi^8),$$

by (1) and (2).
Let $A$ be the automorphism of the field extension $\mathbb{Q}(\zeta_9)/\mathbb{Q}$ defined by $\zeta_9^A := \zeta_9^2$. It holds

$$\eta^A = (\psi^4(a)\iota(\psi^3,\psi))^A = (\psi^4(a))^A(-\sum_{c\in k}\psi^3(c)\psi(1-c))^A =$$
$$= \psi^8(a)(-\sum_{c\in k}\psi^6(c)\psi^2(1-c)) = \psi^8(a)\iota(\psi^6,\psi^2),$$
$$\eta^{A^2} = (\psi^4(a)\iota(\psi^3,\psi))^{A^2} = (\psi^4(a))^{A^2}(-\sum_{c\in k}\psi^3(c)\psi(1-c))^{A^2} =$$
$$= \psi^7(a)(-\sum_{c\in k}\psi^3(c)\psi^4(1-c)) = \psi^7(a)\iota(\psi^3,\psi^4),$$
$$\eta^{A^3} = (\psi^4(a)\iota(\psi^3,\psi))^{A^3} = (\psi^4(a))^{A^3}(-\sum_{c\in k}\psi^3(c)\psi(1-c))^{A^3} =$$
$$= \psi^5(a)(-\sum_{c\in k}\psi^6(c)\psi^8(1-c)) = \psi^5(a)\iota(\psi^6,\psi^8),$$
$$\eta^{A^4} = (\psi^4(a)\iota(\psi^3,\psi))^{A^4} = (\psi^4(a))^{A^4}(-\sum_{c\in k}\psi^3(c)\psi(1-c))^{A^4} =$$
$$= \psi(a)(-\sum_{c\in k}\psi^3(c)\psi^7(1-c)) = \psi(a)\iota(\psi^3,\psi^7),$$
$$\eta^{A^5} = (\psi^4(a)\iota(\psi^3,\psi))^{A^5} = (\psi^4(a))^{A^5}(-\sum_{c\in k}\psi^3(c)\psi(1-c))^{A^5} =$$
$$= \psi^2(a)(-\sum_{c\in k}\psi^6(c)\psi^5(1-c)) = \psi^2(a)\iota(\psi^6,\psi^5),$$

hence

$$N = q - \eta - \eta^A - \eta^{A^2} - \eta^{A^3} - \eta^{A^4} - \eta^{A^5} = q - \mathrm{Tr}_{\mathbb{Q}(\zeta_9)/\mathbb{Q}}(\eta).\square$$

**Corollary 2.** *If* $q \equiv 4(\bmod 9)$ *or* $q \equiv 7(\bmod 9)$ *then*

$$L_{C_a}(t) = 1 - \frac{1}{3}\mathrm{Tr}_{\mathbb{Q}(\zeta_9)/\mathbb{Q}}(\eta)t^3 + q^3t^6,$$

*where* $\eta = \psi^4(a)\iota(\psi^3, \psi)$, $\psi$ *a character of order* 9 *of the multiplicative group of the field* $\mathbb{F}_{q^3}$.
*If* $q \equiv 8(\bmod 9)$ *then*

$$L_{C_a}(t) = 1 - \frac{1}{2}\mathrm{Tr}_{\mathbb{Q}(\zeta_9)/\mathbb{Q}}(\eta)t^2 - q\frac{1}{2}\mathrm{Tr}_{\mathbb{Q}(\zeta_9)/\mathbb{Q}}(\eta)t^4 + q^3t^6,$$

*where* $\eta = \psi^4(a)\iota(\psi^3, \psi)$, $\psi$ *a character of order* 9 *of the multiplicative group of the field* $\mathbb{F}_{q^2}$.

P r o o f: If $q \equiv 4(\bmod 9)$ or $q \equiv 7(\bmod 9)$ then $q^2 \equiv 7(\bmod 9)$ or $q^2 \equiv 4(\bmod 9)$ and $q^3 \equiv 1(\bmod 9)$. By proposition 2 it holds $N_1 = q + 1$ and $N_2 = q^2 + 1$, so the coefficients $a_1$ and $a_2$ of $L_{C_a}(t)$ vanish and the coefficient $a_3$ equals $\frac{1}{3}(N_3 - q^3 - 1)$, which by proposition 3 equals $-\frac{1}{3}\mathrm{Tr}_{\mathbb{Q}(\zeta_9)/\mathbb{Q}}(\eta)$ .
If $q \equiv 8(\bmod 9)$ then $q^2 \equiv 1(\bmod 9)$ and $q^3 \equiv 2(\bmod 9)$. By proposition 1 it holds $N_1 = q + 1$ and $N_3 = q^3 + 1$, so $a_1 = 0$, $a_3 = 0$ and $a_2$ equals $\frac{1}{2}(N_2 - q^2 - 1)$, which by proposition 3 equals $-\frac{1}{2}\mathrm{Tr}_{\mathbb{Q}(\zeta_9)/\mathbb{Q}}(\eta)$. $\square$

*Remark 1.* Corollary 2 explains some computations done in ([CER]).

**Proposition 4.** *If* $q \equiv 1(\bmod 9)$ *then*

$$L_{C_a}(t) = (1 - \eta t)(1 - \eta^A t)(1 - \eta^{A^2} t)(1 - \eta^{A^3} t)(1 - \eta^{A^4} t)(1 - \eta^{A^5} t),$$

*where* $\eta = \psi^4(a)\iota(\psi^3, \psi)$, $\psi$ *a character of order* 9 *of the multiplicative group* $k^*$, $A$ *the automorphism of the field extension* $\mathbb{Q}(\zeta_9)/\mathbb{Q}$ *defined by* $\zeta_9^A := \zeta_9^2$.

P r o o f: The $L$-polynomial of the curve $C_a/k$ can be written in the form $L_{C_a}(t) = \prod_{j=1}^{6}(1 - \alpha_j t)$, where $\alpha_1, \ldots, \alpha_6$ are algebraic integers. For $r \geq 1$ it holds

$$N_r = q^r + 1 - \sum_{j=1}^{6} \alpha_j^r \tag{10}$$

Let $\psi$ be a character of order 9 of the cyclic group $k^*$. The map

$$\psi_r : \mathbb{F}_{q^r}^* \to \mathbb{C}^*, \psi_r(x) := \psi(N_{\mathbb{F}_{q^r}|\mathbb{F}_q}(x))$$

is a character of order 9 of the cyclic group $\mathbb{F}_{q^r}^*$. It holds ([Da-Ha],0.8)

$$\tau_d^{(r)}(\psi_r^l) = \tau_d(\psi^l)^r \tag{11}$$

for $1 \leq l \leq 8$ and $d \in \mathbb{F}_q^*$, where $\tau_d^{(r)}(\psi_r^l)$ denotes the Gauss sum of the character $\psi_r^l$ on $\mathbb{F}_{q^r}$.
By Proposition 4 it holds

$$N_r = q^r + 1 - \frac{\tau^{(r)}(\psi_r^3)\tau^{(r)}(\psi_r)}{\tau_a^{(r)}(\psi_r^4)} - \frac{\tau^{(r)}(\psi_r^3)\tau^{(r)}(\psi_r^4)}{\tau_a^{(r)}(\psi_r^7)} - \frac{\tau^{(r)}(\psi_r^3)\tau^{(r)}(\psi_r^7)}{\tau_a^{(r)}(\psi_r)} -$$

$$- \frac{\tau^{(r)}(\psi_r^6)\tau^{(r)}(\psi_r^2)}{\tau_a^{(r)}(\psi_r^8)} - \frac{\tau^{(r)}(\psi_r^6)\tau^{(r)}(\psi_5)}{\tau_a^{(r)}(\psi_r^2)} - \frac{\tau^{(r)}(\psi_r^6)\tau^{(r)}(\psi_r^8)}{\tau_a^{(r)}(\psi_r^5)},$$

hence by (11)

$$N_r = q^r + 1 - \frac{\tau(\psi^3)^r\tau(\psi)^r}{\tau_a(\psi^4)^r} - \frac{\tau(\psi^3)^r\tau(\psi^4)^r}{\tau_a(\psi^7)^r} - \frac{\tau(\psi^3)^r\tau(\psi^7)^r}{\tau_a(\psi)^r} -$$

$$- \frac{\tau(\psi^6)^r\tau(\psi^2)^r}{\tau_a(\psi^8)^r} - \frac{\tau(\psi^6)^r\tau(\psi^5)^r}{\tau_a(\psi^2)^r} - \frac{\tau(\psi^6)^r\tau(\psi^8)^r}{\tau_a(\psi^5)^r},$$

so one can choose in (10)

$$\alpha_1 = \frac{\tau(\psi^3)\tau(\psi)}{\tau_a(\psi^4)} = \eta, \alpha_2 = \frac{\tau(\psi^3)\tau(\psi^4)}{\tau_a(\psi^7)} = \eta^{A^2}, \alpha_3 = \frac{\tau(\psi^3)\tau(\psi^7)}{\tau_a(\psi)} = \eta^{A^4},$$

$$\alpha_4 = \frac{\tau(\psi^6)\tau(\psi^8)}{\tau_a(\psi^5)} = \eta^{A^3}, \alpha_5 = \frac{\tau(\psi^6)\tau(\psi^5)}{\tau_a(\psi^2)} = \eta^{A^5}, \alpha_6 = \frac{\tau(\psi^6)\tau(\psi^2)}{\tau_a(\psi^8)} = \eta^A. \square$$

Let $m \geq 1$ be a natural number and let $K$ be an algebraic number field with ring of integers $\mathcal{O}_K$ such that $\zeta_m \in \mathcal{O}_K$. Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$ not dividing $m$, and let $x \in \mathcal{O}_K$ not divisible by $\mathfrak{p}$. The number $x^{\frac{N_{K/\mathbb{Q}}(\mathfrak{p})-1}{m}}$ is congruent modulo $\mathfrak{p}$ to one and only one root of unity $\zeta_m^l \in \mu_m$. The map

$$(\mathcal{O}_K/\mathfrak{p}) \setminus \{0\} \to \mu_m,\ x \bmod \mathfrak{p} \mapsto \zeta_m^l$$

is a character of order $m$ of the multiplicative group of the finite field $\mathcal{O}_K/\mathfrak{p}$ called the $m$-th power residue character modulo $\mathfrak{p}$.

**Proposition 5.** *Let $q \equiv 1 (\bmod 9)$ and let $\mathfrak{p}$ be a prime divisor of $p$ in the ring $\mathbb{Z}[\zeta_{q-1}]$. Let $\psi$ be the 9-th power residue character modulo $\mathfrak{p}$ in $\mathbb{Z}[\zeta_{q-1}]$. Identifying the finite field $\mathbb{F}_q$ with the residue class field $\mathbb{Z}[\zeta_{q-1}]/\mathfrak{p}$ it holds:*
*a) The absolute value of the complex number $\iota(\psi^3, \psi)$ is*

$$|\iota(\psi^3, \psi)| = \sqrt{q};$$

*b) The prime ideal decomposition of the principal ideal generated by $\iota(\psi^3, \psi)$ in the ring of integers $\mathbb{Z}[\zeta_9]$ is*

$$\iota(\psi^3, \psi)\mathbb{Z}[\zeta_9] = (\mathfrak{q} \cdot \mathfrak{q}^{A^4} \cdot \mathfrak{q}^{A^5})^{f(\mathfrak{p}|\mathfrak{q})},$$

*where* $\mathfrak{q} := \mathfrak{p} \cap \mathbb{Z}[\zeta_9]$, *A is the automorphism of* $\mathbb{Q}(\zeta_9)/\mathbb{Q}$ *defined by* $\zeta_9^A := \zeta_9^2$ *and* $N_{\mathbb{Q}(\zeta_{q-1})/\mathbb{Q}(\zeta_9)}(\mathfrak{p}) = \mathfrak{q}^{f(\mathfrak{p}|\mathfrak{q})}$.
*c) In the ring* $\mathbb{Z}[\zeta_9]$ *it holds*

$$\iota(\psi^3, \psi) \equiv 1(\bmod\ (\zeta_9 - 1)^4).$$

*The number* $\iota(\psi^3, \psi) \in \mathbb{Z}[\zeta_9]$ *is uniquely determined by the properties a), b) and c).*

P r o o f:
a): Every Jacobi sum in a finite field with $q$ elements has absolute value $\sqrt{q}$.
b): By ([Ha1], p.40, (6.)) it holds

$$\iota(\psi^3, \psi)\mathbb{Z}[\zeta_9] = (\mathfrak{q}^{\sum_J d(-3j,-j)J})^{f(\mathfrak{p}|\mathfrak{q})},$$

where $J$ runs over the set $\{A^k \mid 0 \leq k \leq 5\}$ of automorphisms of $\mathbb{Q}(\zeta_9)$, $j$ mod 9 is defined by

$$\zeta_9^{J^{-1}} = \zeta_9^j$$

and

$$d(-3j, -j) = \frac{r(-3j) + r(-j) - r(-4j)}{9},$$

$r(x)$ the smallest non-negative residue of $x$ mod 9. It holds

$$\zeta_9^{(A^0)^{-1}} = \zeta_9,\ d(-3,-1) = \frac{r(-3) + r(-1) - r(-4)}{9} = 1,$$

$$\zeta_9^{(A^1)^{-1}} = \zeta_9^{A^5} = \zeta_9^5,\ d(-15,-5) = \frac{r(-15) + r(-5) - r(-20)}{9} = 0,$$

$$\zeta_9^{(A^2)^{-1}} = \zeta_9^{A^4} = \zeta_9^7,\ d(-21,-7) = \frac{r(-21) + r(-7) - r(-28)}{9} = 0,$$

$$\zeta_9^{(A^3)^{-1}} = \zeta_9^{A^3} = \zeta_9^8,\ d(-24,-8) = \frac{r(-24) + r(-8) - r(-32)}{9} = 0,$$

$$\zeta_9^{(A^4)^{-1}} = \zeta_9^{A^2} = \zeta_9^4,\ d(-12,-4) = \frac{r(-12) + r(-4) - r(-16)}{9} = 1,$$

$$\zeta_9^{(A^5)^{-1}} = \zeta_9^A = \zeta_9^2,\ d(-6,-2) = \frac{r(-6) + r(-2) - r(-8)}{9} = 1,$$

$$\iota(\psi^3, \psi)\mathbb{Z}[\zeta_9] = (\mathfrak{q}^{1+A^4+A^5})^{f(\mathfrak{p}|\mathfrak{q})} = (\mathfrak{q} \cdot \mathfrak{q}^{A^4} \cdot \mathfrak{q}^{A^5})^{f(\mathfrak{p}|\mathfrak{q})}.$$

c): For $c \in \mathbb{F}_q^*$ it holds

$$\psi(c) \equiv 1 \bmod (\zeta_9 - 1)$$

and

$$\psi^3(c) \equiv 1 \bmod (\zeta_9 - 1)^3.$$

Indeed, if $\psi(c) = \zeta_9^k$, $0 \le k \le 8$, then $\psi(c) - 1 = \zeta_9^k - 1$ is divisible by $\zeta_9 - 1$ in $\mathbb{Z}[\zeta_9]$ and $\psi^3(c) - 1$ is divisible by $\zeta_9^3 - 1$ which is associate with $(\zeta_9 - 1)^3$. Then

$$\iota(\psi^3, \psi) = -\sum_{c \in \mathbb{F}_q} \psi^3(c)\psi(1-c) = -\sum_{c \in \mathbb{F}_q} \psi(c)\psi^3(1-c) =$$

$$= -\sum_{c \neq 1} \psi(c) - \sum_{c \neq 0,1} \psi(c)(\psi^3(1-c) - 1) =$$

$$= 1 - \sum_{c \neq 0,1} \psi(c)(\psi^3(1-c) - 1) \equiv 1 - \sum_{c \neq 0,1} (\psi^3(1-c) - 1) \bmod (\zeta_9 - 1)^4 \equiv$$

$$\equiv 1 - \sum_{c \neq 0,1} \psi^3(1-c) + \sum_{c \neq 0,1} 1 \bmod (\zeta_9 - 1)^4 \equiv$$

$$\equiv 1 + 1 + q - 2 \bmod (\zeta_9 - 1)^4 \equiv q \bmod (\zeta_9 - 1)^4 \equiv 1 \bmod (\zeta_9 - 1)^4.$$

Two numbers in $\mathbb{Z}[\zeta_9]$ with the same absolute value and the same prime ideal decomposition differ by a root of unity. The group of roots of unity in $\mathbb{Z}[\zeta_9]$ is $\mu_{18}$. The only element of $\mu_{18}$ which is $\equiv 1 \bmod (\zeta_9 - 1)^4$ is 1. The properties a), b), c) determine the number $\iota(\psi^3, \psi)$ in $\mathbb{Z}[\zeta_9]$. □

## 2 The Curves $C_a : Y^3 = X^4 - aX$ over an Algebraic Number Field

Let $k$ be an algebraic number field which contains $\zeta_9$. Let $a \in k^*$, and let $\mathfrak{m}_a$ be the product of 3 and of all prime divisors $\mathfrak{p}$ of $k$ which appear in the decomposition of $a$. Let $\mathfrak{p}$ be a prime divisor of $k$ which does not divide $\mathfrak{m}_a$. The curve $C_a$ has good reduction at $\mathfrak{p}$: By reducing modulo $\mathfrak{p}$ the equation $y^3 = x^4 - ax$ one obtains a curve $C_{a(\mathfrak{p})}$ over the residue class field $k(\mathfrak{p})$ at $\mathfrak{p}$ with the equation

$$C_{a(\mathfrak{p})} : y^3 = x^4 - a(\mathfrak{p})x, a(\mathfrak{p}) := a \bmod \mathfrak{p} \in k(\mathfrak{p})^*$$

which is smooth of genus 3 over $k(\mathfrak{p})$. Let $L_{C_{a(\mathfrak{p})}}(t)$ be the $L$-polynomial of $C_{a(\mathfrak{p})}/k(\mathfrak{p})$. By proposition 4 it holds

$$L_{C_a}(t) = \prod_{j=0}^{5} (1 - \eta(\mathfrak{p})^{A^j} t),$$

where $\eta(\mathfrak{p}) := {\psi_{\mathfrak{p}}}^4(a(\mathfrak{p}))\iota({\psi_{\mathfrak{p}}}^3, \psi_{\mathfrak{p}})$, $\psi_{\mathfrak{p}}$ the 9-th power residue character modulo $\mathfrak{p}$, $A$ the automorphism of the field extension $\mathbb{Q}(\zeta_9)/\mathbb{Q}$ defined by $\zeta_9^A := \zeta_9^2$.

The *L-function of* $C_a$ *over* $k$ is defined by

$$L(s, C_a, k) := \prod_{(\mathfrak{p}, \mathfrak{m}_a) = 1} L_{C_{a(\mathfrak{p})}}(N(\mathfrak{p})^{-s}). \tag{12}$$

The product on the right hand side of (12) is absolutely convergent for $\Re s > \frac{3}{2}$ ([Ha1], [We], [De]). It holds

$$L(s, C_a, k) = \prod_{j=0}^{5} L_j(s),$$

where

$$L_j(s) := \prod_{(\mathfrak{p}, \mathfrak{m}_a)=1} (1 - \eta(\mathfrak{p})^{A^j} N(\mathfrak{p})^{-s}), \tag{13}$$

for $j = 0, \ldots, 5$. Extend the function $\eta(\mathfrak{p})$ multiplicatively on the group $\mathrm{Div}_{\mathfrak{m}_a} k$ of divisors of $k$ prime to $\mathfrak{m}_a$ and define

$$\lambda_j : \mathrm{Div}_{\mathfrak{m}_a} k \mapsto \mathbb{C}^*, \lambda_j(\mathfrak{a}) := \frac{\eta(\mathfrak{a})^{A^j}}{\sqrt{N(\mathfrak{a})}},$$

for $j = 0, \ldots, 5$. The functions $\lambda_0, \ldots, \lambda_5$ are *Grössencharaktere* of $k$ ([Ha1], [We]) in the sense of Hecke ([He]). Let $\mathrm{Div}^+_{\mathfrak{m}_a} k$ denote the set of positive divisors in $\mathrm{Div}_{\mathfrak{m}_a} k$. By (13) it holds for $\Re s > \frac{3}{2}$

$$L_j(s)^{-1} = \prod_{(\mathfrak{p}, \mathfrak{m}_a)=1} (1 - \lambda_j(\mathfrak{p}) N(\mathfrak{p})^{-s+\frac{1}{2}})^{-1} =$$

$$= \sum_{\mathfrak{a} \in \mathrm{Div}^+_{\mathfrak{m}_a} k} \frac{\lambda_j(\mathfrak{a})}{N(\mathfrak{a})^{s-\frac{1}{2}}} = L(s - \frac{1}{2}, \lambda_j, k),$$

where

$$L(s, \lambda_j, k) := \sum_{\mathfrak{a} \in \mathrm{Div}^+_{\mathfrak{m}_a} k} \frac{\lambda_j(\mathfrak{a})}{N(\mathfrak{a})^s}, \Re s > 1,$$

is the Hecke $L$-function corresponding to $\lambda_j$, $j = 0, \ldots, 5$. So

**Theorem 1.** *The L-function $L(s, C_a, k)$ of the curve $C_a$ over $k$ equals the product of the inverses of Hecke L-functions $L(s - \frac{1}{2}, \lambda_j, k)$, $j = 0, \ldots, 5$.*

## 3 The Curves $C_a : Y^3 = X^4 - aX$ over $\mathbb{C}$

A complex *Picard curve* is the projective closure of an affine plane curve of equation type $Y^3 = p_4(X)$, where $p_4(X)$ is a polynomial of degree 4. We exclude all polynomials $p_4(X)$ with only one zero. So one avoids unstable curves in order to get a compact algebraic moduli space $\hat{M}$ of (isomorphy classes of semistable) Picard curves, which we choose in a very canonical way. Smooth Picard curves have genus 3. They correspond to a Zariski-open part $M^{\#}$ of $\hat{M}$. Let $K = \mathbb{Q}(\sqrt{-3}) = \mathbb{Q}(\omega)$, $\omega := e^{\frac{2\pi i}{3}}$, be the field of Eisenstein numbers. The cyclic group $\mathbb{Z}/3\mathbb{Z}$ of order 3 acts via $(x, y) \mapsto (x, \omega y)$ on each Picard

curve $C$. If $C$ is smooth, we get $\mathbb{P}^1$ as quotient curve $C/(\mathbb{Z}/3\mathbb{Z})$ with $\mathbb{Z}/3\mathbb{Z}$ as Galois group of $C \,/\, \mathbb{P}^1$. The action of $\mathbb{Z}/3\mathbb{Z}$ induces a $K$-multiplication of type $(2,1)$ on the jacobian variety $J(C)$ of $C$, which means that the diagonalized representation group of $\mathbb{Z}/3\mathbb{Z}$ on the tangent space $T_0J(C)$ of $J(C)$ is generated by $\left(\begin{smallmatrix} \omega & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \bar{\omega} \end{smallmatrix}\right)$. Let

$$\mathbb{B} := \{z = (z_1, z_2) \in \mathbb{C}^2;\ |z|^2 := |z_1|^2 + |z_2|^2 < 1\},$$

be the two-dimensional complex unit ball. The moduli space of abelian threefolds with $K$-multiplication of type $(2,1)$ is the Shimura surface $\mathbb{B}/\Gamma$, $\Gamma = \mathbb{U}((2,1), \mathfrak{O})$, $\mathfrak{O} = \mathfrak{O}_K = \mathbb{Z} + \mathbb{Z}\omega$ the ring of Eisenstein integers. Define the congruence subgroup $\Gamma(\sqrt{-3})$ by the exact group sequence

$$1 \longrightarrow \Gamma(\sqrt{-3}) \longrightarrow \Gamma \longrightarrow \mathbb{U}((2,1), \mathfrak{O}/(1-\omega)\mathfrak{O}) \longrightarrow 1.$$

In ([Ho1], Ch. I, Prop. 3.2.3) it is proved the following

**Theorem 2.** *The Baily-Borel compactification* $\widehat{\mathbb{B}/\Gamma(\sqrt{-3})}$ *coincides with the projective plane* $\mathbb{P}^2$. *The compactifying cusp points are four points* $K_1, K_2, K_3, K_4 \in \mathbb{P}^2$ *in general position. The open part* $\mathbb{P}_2^{\#} \subset \mathbb{P}^2$ *coming from smooth Picard curves is precisely the complement of the six projective lines* $L_{ij} = L_{ji}$ *going through pairs* $K_i, K_j$ *of different cusp points.*

It turns out that

$$M^{\#} = \mathbb{P}_2^{\#}/S_4\ ,\ \hat{M} = \mathbb{P}^2/S_4\ ,\ M = \mathbb{P}_2^*/S_4,$$

where $\mathbb{P}_2^* := \mathbb{P}^2 \setminus \{K_1, K_2, K_3, K_4\}$. Now identify $\mathbb{P}^2$ with

$$\mathbb{P}_0^3 = \{(t_1 : t_2 : t_3 : t_4) \in \mathbb{P}^3;\ t_1 + t_2 + t_3 + t_4 = 0\},$$

and introduce projective coordinates such that

$$K_1 = (-3 : 1 : 1 : 1),\ K_2 = (1 : -3 : 1 : 1),$$
$$K_3 = (1 : 1 : -3 : 1),\ K_4 = (1 : 1 : 1 : -3).$$

Each Picard curve is isomorphic to a *normal form* representative

$$C_{\mathfrak{t}} :\ Y^3 = (X - t_1)(X - t_2)(X - t_3)(X - t_4),\ t_1 + t_2 + t_3 + t_4 = 0.$$

The correspondence

$$C_{\mathfrak{t}} \mapsto \mathfrak{t} = (t_1, t_2, t_3, t_4) \mapsto (t_1 : t_2 : t_3 : t_4) \in \mathbb{P}_2^*$$

restricted to $\mathbb{P}_2^{\#}$ and composed with the $S_4$- quotient map yields the precise parametrisation of isomorphy classes ([Ho1] I, Prop.5.2.3). Especially, all curves of the family

$$C_a\ : Y^3 = X^4 - aX\ , a \in \mathbb{C}^*,$$

are isomorphic over $\mathbb{C}$ to

$$C_1 : Y^3 = X^4 - X,$$

whose moduli point is the image of $(0 : 1 : \omega : \omega^2)$.

The Jacobians of smooth Picard curves are (principally polarized) abelian threefolds. Via period matrices they are represented by points in the generalized Siegel upper half plane

$$\mathbb{H}_3 = \{\Omega \in Mat_3(\mathbb{C});\ {}^t\Omega = \Omega\,,\ Im\,\Omega\ positive\ definite\},$$

uniquely up to $\mathbb{S}p(6,\mathbb{Z})$-equivalence, where

$$\mathbb{S}p(6,\mathbb{Z}) = \{G \in \mathbb{G}l_6(\mathbb{Z});\ {}^tG \cdot \left(\begin{smallmatrix} O & E_3 \\ -E_3 & O \end{smallmatrix}\right) \cdot G = \left(\begin{smallmatrix} O & E_3 \\ -E_3 & O \end{smallmatrix}\right)\}, E_3 := diag(1,1,1),$$

denotes the symplectic group acting on $\mathbb{H}_3$ in the well-known manner. By Torelli's theorem there is a canonical algebraic embedding $M^{\#} \hookrightarrow \mathfrak{A}_3$ into the moduli space $\mathfrak{A}_3 = \mathbb{H}_3/\mathbb{S}p(6,\mathbb{Z})$ of principally polarized abelian threefolds. Restricting to the Zariski-open subspace $\mathfrak{A}_3^{\#} \subset \mathfrak{A}_3$ corresponding to Jacobians of smooth genus 3 curves one gets a closed embedding $M^{\#} \hookrightarrow \mathfrak{A}_3^{\#}$, which determines $M^{\#}$ uniquely, up to isomorphy. The closed algebraic embedding $M^{\#} \hookrightarrow \mathfrak{A}_3^{\#}$ can be *uniformized* in the following sense. In the analytic category there is a commutative *Shimura diagram*

$$\begin{array}{ccccc}
\mathbb{B} & & \hookrightarrow & & \mathbb{H}_3 \\
& \nwarrow & & \nearrow & \\
& \mathbb{B}^{\#} & \hookrightarrow & \mathbb{H}_3^{\#} & \\
\downarrow & \downarrow & & \downarrow & \downarrow \\
& M^{\#} & \hookrightarrow & \mathfrak{A}_3^{\#} & \\
& \swarrow & & \searrow & \\
M & & \to & & \mathfrak{A}_3
\end{array}$$

where $\mathbb{H}_3 \longrightarrow \mathfrak{A}_3$ is the $\mathbb{S}p(6,\mathbb{Z})$-quotient morphism, $\mathbb{H}_3^{\#}$ is the preimage of $\mathfrak{A}_3^{\#}$ in $\mathbb{H}_3$, $\mathbb{B} \hookrightarrow \mathbb{H}_3$ is a closed embedding, $\mathbb{B}^{\#} = \mathbb{B} \cap \mathbb{H}_3^{\#}$, and $\mathbb{B} \longrightarrow M$ is the analytic quotient morphism of the arithmetic group

$$N_{\mathbb{S}p(6,\mathbb{Z})}(\mathbb{B}) := \{G \in \mathbb{S}p(6,\mathbb{Z});\ G(\mathbb{B}) = \mathbb{B}\}$$

acting on $\mathbb{B}$. In ([Ho3]) it is proved that this ball lattice coincides with $\Gamma$.

Identifying for a moment the ball with its image in $\mathbb{H}_3$ we call $\mathbb{B}$ the *period space of Picard curves* and its points are called *Picard period points* (of the family of Picard curves). An element $\gamma \in \Gamma$ is called *elliptic*, iff $\gamma$ has an isolated fixed point $P \in \mathbb{B}$. Let $\Gamma'$ be a subgroup of $\Gamma$. We call the elliptic element $\gamma$ *purely $\Gamma'$-elliptic*, iff all non-trivially on $\mathbb{B}$ acting elements of the stationary group $\Gamma'_P$ are elliptic. The images of purely $\Gamma'$-elliptic points on $\mathbb{B}/\Gamma'$ are isolated (cyclic quotient) singularities. Notice that the fixed point

$P$ is uniquely determined by the elliptic element $\gamma$ because the group of biholomorphic automorphisms of $\mathbb{B}$ coincides with $\mathbb{PU}((2,1),\mathbb{C})$, so $\gamma$ has only one negative eigenline in $V = (\mathbb{C}^3, <.,.>)$ with respect to the hermitian metric $<.,.>$ of signature $(2,1)$ on $\mathbb{C}^3$.
In ([Ho1], Ch. I, 3.4.4) it is proved the following

**Theorem 3.** *(see [Ho1] I, Prop. 3.4.4). The only singularities of $\hat{M}$ are the image points of $S := (0:1:\omega:\omega^2)$ and $N := (1:i:-1:-i)$, along the $S_4$-quotient morphism.* □

This is a simple application of a theorem of Chevalley stating that the singularities of a finite (more generally: locally finite) Galois quotient $X/G$ of a smooth complex manifold X come precisely from points $x \in X$ with isotropy group $G_x$ not generated by reflections at $x$, where reflections at $x$ are defined as elements of $G_x$ acting trivially on a submanifold of $X$ through $x$ of codimension 1. Looking at finite subgroups of $S_4$ and their fixed points on $\mathbb{P}^2$ one finds up to $S_4$-equivalence the points $S$, $N$ as only singular possibilities. The $S_4$-isotropy group of $S$ is generated by the cyclic permutation (234) of order 3. The $S_4$-isotropy group of $N$ is generated by the cyclic permutation (1234) of order 4. The (13)(24)-reflection line on $\mathbb{P}^2$ contains N.

**Proposition 6.** *The set of Picard period points of $C_1$ coincides with the set of purely $\Gamma$-elliptic points on $\mathbb{B}$. It coincides with the $\Gamma$-orbit of*

$$P_{\zeta_9} := (\zeta_9{}^4 - \zeta_9{}^2 : 1 : \zeta_9{}^5 + \zeta_9{}^4 - 1) \in \mathbb{B}.$$

P r o o f: For an arbitrary group $G$ let $G_{tor}$ be the set of elements of finite order of $G$ (torsion elements), and let $G_{k-tor}$ be the subset of elements of precise order $k \in \mathbb{N}_+$. $G$ acts by conjugation on $G_k$ and on $G_{tor}$. It holds

**Lemma 2.** *For $\Gamma = \mathbb{U}((2,1),\mathfrak{O})$ the set $\Gamma_{9-tor}$ is not void. It consists of precisely six $\Gamma$-conjugation classes. They are projected onto two $\mathbb{P}\Gamma$-conjugation classes in $(\mathbb{P}\Gamma)_{3-tor}$.*

P r o o f of Lemma 2: For the first statement we consider the element

$$\varphi_1 := \begin{pmatrix} -\omega^2 & -1 & \omega^2 \\ \omega & 1 & 1 \\ 1 & -1 & \omega^2 - 1 \end{pmatrix}$$

with

$$det\,\varphi_1 = \omega\,,\ \varphi_1^3 = \omega E_3.$$

found by Feustel in [Feu]. It is easy to check that $\varphi_1$ belongs to $\Gamma$. The eigenvalues are $\zeta_9$, $\zeta_9{}^4$, $\zeta_9{}^7$. The powers $\varphi_1^k$, $k = 1,2,4,5,7,8$, yield six different conjugation classes in $\Gamma_{9-tor}$ (compare determinants and eigenvalues) and two conjugation classes in $(\mathbb{P}\Gamma)_{3-tor}$. □

Now let $\varphi$ be an arbitrary element of $\Gamma_{9-tor}$ with eigenvalues $\zeta_9$, $\zeta_9{}^j$, $\zeta_9{}^k$, say. The Galois group of $F := K(\zeta_9)$ over $K$ is generated by $\sigma : \zeta_9 \mapsto$

$\zeta_9{}^4$. The characteristic polynomial $\chi_\varphi(T)$ of $\varphi$ belongs to $K[T]$. Looking at trace and determinant of $\varphi$, which must belong to $K$, it is easy to see that $\varphi$ has three different eigenvalues. They must be conjugated over $K$, hence $\zeta_9{}^j = \zeta_9{}^4 = \sigma(\zeta_9)$, $\zeta_9{}^k = \zeta_9{}^7 = \sigma^2(\zeta_9)$. The eigenvectors $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ of $\zeta_9$, $\sigma(\zeta_9)$, $\sigma^2(\zeta_9)$, respectively, can be choosen in $F^3$. They form an orthogonal basis of $F^3$ endowed with our hermitian $(2,1)$-metric because of different eigenvalues. From $\varphi(\mathfrak{a}) = \zeta_9 \cdot \mathfrak{a}$ it follows that

$$\sigma(\varphi(\mathfrak{a})) = \sigma(\zeta_9)\sigma(\mathfrak{a}) = \zeta_9{}^4\sigma(\mathfrak{a})$$

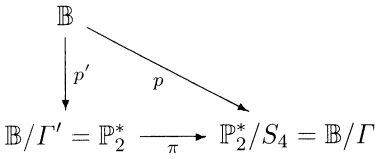
because $\varphi$ belongs to $Mat_3(K)$. Therefore

$$\mathfrak{a},\ \mathfrak{b} = \sigma(\mathfrak{a}),\ \mathfrak{c} = \sigma^2(\mathfrak{a}) \in F^3,$$

satisfying

$$< \mathfrak{a}, \mathfrak{a} > < 0,\ < \mathfrak{b}, \mathfrak{b} > > 0,\ < \mathfrak{c}, \mathfrak{c} > > 0, \tag{14}$$

(without loss of generality) is an orthogonal $\varphi$-eigenbasis of $\mathbb{C}^3$ . The elliptic element $\varphi$ has the unique elliptic fixed point $P = \mathbb{P}\mathfrak{a} \in \mathbb{B}$. We show that $P$ is a purely $\Gamma$-elliptic point. With $\Gamma' := \Gamma(\sqrt{-3})$ we have a commutative diagram of quotient morphisms




In [Ho1] I, Prop. 3.4.4, there are listed on $\mathbb{P}_2^*$ the $p'$- images of all $\Gamma$-elliptic points $Q \in \mathbb{B}$ together with their (abstract) isotropy groups $\Gamma_Q$. Our $P$ cannot be an intersection point of two $\Gamma$-reflection discs because the reflections have eigenvalues only in $K$. Otherwise $P \in \mathbb{B} \subset \mathbb{P}^2$ would be the intersection point of two projective lines (the projectivized orthogonal complements of the one-dimensional eigenspaces) defined over $K$. This leads to $\mathbb{P}\mathfrak{a} = P = \mathbb{P}\mathfrak{a}'$, $\mathfrak{a}' \in K^3$, $\sigma(P) = P$, which contradicts to $\sigma(P) \notin \mathbb{B} = \mathbb{P}V_-$, by (14). There are precisely two $\Gamma$-orbits $\Gamma\tilde{N}$, $\Gamma\tilde{S}$ of $\Gamma$-elliptic points whose isotropy groups are not generated by reflections. The projective isotropy groups $\mathbb{P}\Gamma_{\tilde{N}}$ or $\mathbb{P}\Gamma_{\tilde{S}}$ are cyclic of order 4 or 3, respectively. Since $\mathbb{P}\varphi \in \mathbb{P}\Gamma_P$ is elliptic of order 3 the point $P$ must belong to the second orbit. The image $p(\tilde{S})$ coincides with $p'(S)$, which is an orbitally isolated singularity with respect to $\Gamma$. This means that $\tilde{S}$ is a purely $\Gamma$-elliptic point, hence $\mathbb{P}\Gamma_{\tilde{S}} \cong < \mathbb{P}\varphi >$ of order 3. □

Let $F$ be a number field and $A$ a complex abelian variety of dimension $g$. We say that $A$ has *$F$-multiplication*, if there is a $\mathbb{Q}$-algebra embedding $\iota$ of $F$ into the endomorphism algebra $End^\circ A = \mathbb{Q} \otimes End\, A$ of $A$. If, moreover, the degree $[F : \mathbb{Q}]$ of $F$ is equal to $2g$ and $\iota$ is an isomorphism, then $A$ is called an abelian CM-*variety*. It is well-known in this case that $A$ is simple and $F$ is a CM-field, which is, by definition, a totally imaginary quadratic field extension

of a totally real number field, see [La]. A CM-curve is a (smooth complex) projective curve $C$ whose jacobian variety $J(C)$ is an abelian CM-variety.

**Proposition 7.** *The endomorphism ring $End\, J(C_1)$ is isomorphic to $\mathbb{Z}[\zeta_9]$. Up to isomorphy, $C_1$ is the only Picard CM-curve with a cyclotomic maximal order as endomorphism ring.*

P r o o f: Our special Picard curve $C_1 : Y^3 = X(X^3 - 1)$ has an obvious non-trivial automorphism of 9-th order fixing $\infty = (0 : 0 : 1)$:

$$(x, y) \mapsto (\omega x,\, \zeta_9 y)\ ,\ ({\zeta_9}^3 = \omega).$$

It extends to an automorphism of the Jacobian threefold of $C_1$. With Theorem 6 below we will see that this automorphism generates a subfield in the endomorphism algebra of the Jacobian. Therefore we get embeddings

$$\mathbb{Z}[\zeta_9] \hookrightarrow End\, J(C_1)\ ,\ F = \mathbb{Q}(\zeta_9) \hookrightarrow End^\circ J(C_1). \tag{15}$$

The representing period point $P_{\zeta_9} = \mathbb{P}\mathfrak{a} \in \mathbb{B}$ is purely $\Gamma$-elliptic by Proposition 3, fixed by $\varphi_1$ of nine-th order. Therefore the ring $End_K(\mathfrak{a}, \mathfrak{a}^\perp)$ of $K$-endomorphisms of $V$ with eigenvector $\mathfrak{a}$ and invariant subspace $\mathfrak{a}^\perp$ is bigger than $K$. Such ball points have been called *exceptional* in [Ho2], Corollary 7.10. Moreover, $\mathfrak{a}$ is eigenvector of a simple eigenvalue of $\varphi_1 \in End_K(\mathfrak{a}, \mathfrak{a}^\perp)$. Therefore $P_{\zeta_9}$ is an *isolated exceptional* point in the sense of Definition 7.12 of [Ho2]. The $K$-degree $[K(P_{\zeta_9}) : K]$ of $P_{\zeta_9}$ is equal to 3. Now apply the following theorem to see that $J(C_1)$ is a simple CM-threefold with multiplication field $K(\zeta_9)$.

**Theorem 4.** (see [Ho2], section 7.) *The endomorphism algebra of the jacobian variety $J_\tau \cong J(C_t)$ of a Picard curve with period point $\tau \in \mathbb{B}$ and moduli point $t = (t_1 : t_2 : t_3 : t_4) \in \mathbb{P}_2^*$ is greater than $K$ if and only if $\tau$ is exceptional. $J_\tau$ splits up to isogeny into abelian CM- subvarieties if and only if $\tau$ is an isolated exceptional point. Thereby Jacobians with CM-field $F$ (of degree 3 over $K$) correspond to isolated exceptional points of $K$-degree 3 and $F \cong K(\tau)$. All other isolated exceptional points (of $K$-degree 2 or 1) ly on $K$-discs on $\mathbb{B}$ (defined as non-empty intersections $L \cap \mathbb{B}$, $L$ projective lines on $\mathbb{P}^2$ defined over $K$). Thereby $\tau \in \mathbb{B}(K)$ if and only if $J_\tau$ splits into $E \times E \times E$. The degree 2 case happens if and only if $J_\tau$ splits into $E \times (E'^2)$, where $E$ is an elliptic CM-curve with $K$-multiplication and $E'$ elliptic CM with imaginary quadratic multiplication field $L \neq K$. Moreover, it holds that $K(L) = K(\tau)$ in the latter case.* □

The endomorphism ring of any abelian CM-variety is an order in the corresponding CM-field. Each order of a number field $L$ is contained in the maximal order, the ring $\mathfrak{O}_L$ of integers in $L$. The maximal order of a cyclotomic field $L = \mathbb{Q}(\zeta)$ is equal to $\mathbb{Z}[\zeta]$, $\zeta$ a generating unit root, see e.g. [Neu], I, Prop. 10.2. So the embeddings (15) must be isomorphisms, especially

$$\mathfrak{O}_F = \mathbb{Z}[\zeta_9] \cong End J(C_1) \subseteq End^\circ J(C_1) \cong F.$$

The first part of Proposition 5 is proved.
$F$ is the only cyclotomic field of degree 3 over $K$. Therefore the Jacobian threefolds of CM-Picard curves $C$ with cyclotomic endomorphism algebra $End^{\circ}J(C)$, which must be isomorphic to $F$, have to be isogeneous. There is a bijective correspondence between the ideal classes of $\mathfrak{O}_F$ and the isomorphy classes of principally polarized abelian CM-threefolds A (of same multiplication type) with endomorphism rings $\mathfrak{O}_F$, see e.g. [La], III.2, Cor. 2.7. It is well-known that the class number of $F$ is equal to 1, see e.g. [Ha2], III, end of 29. Therefore, up to isomorphy, there is only one such $A$. Then, by Torelli's theorem, also the isomorphy class of Picard CM-curves with $End\,J(C) \cong \mathfrak{O}_F$ is uniquely determined. This completes the proof of Proposition 5. □

*Remark 2.* The *type* of $F$-multiplication is a lift ($F$-extension) from the type $(2,1)$ of K-multiplication on $J(C_1)$. This lifted type is unique by [La], I.3, Theorem 3.6.

**Proposition 8.** *A period matrix of the Jacobian $J(C_1)$ is:*

$$\Pi = \begin{pmatrix} -\zeta_9+1 & 0 & -2\zeta_9{}^2-2\zeta_9 & -\zeta_9{}^2-1 & 1 & 2\zeta_9{}^2+\zeta_9 \\ \zeta_9{}^2-1 & 0 & -\zeta_9{}^2+2\zeta_9 & -\zeta_9{}^2+\zeta_9+1 & -1 & \zeta_9{}^2-2\zeta_9 \\ -\zeta_9+1 & 0 & -2\zeta_9{}^2-2\zeta_9 & -\zeta_9{}^2-1 & 1 & 2\zeta_9{}^2+\zeta_9 \end{pmatrix} \cdot \omega +$$

$$+ \begin{pmatrix} 2\zeta_9{}^2+\zeta_9+1 & 1 & -\zeta_9+1 & -2\zeta_9{}^2-\zeta_9 & 0 & \zeta_9{}^2+\zeta_9-1 \\ -\zeta_9{}^2+2\zeta_9 & 1 & -2\zeta_9{}^2+2\zeta_9+1 & -\zeta_9+1 & -1 & \zeta_9{}^2-\zeta_9-1 \\ 2\zeta_9{}^2+\zeta_9+1 & 1 & -\zeta_9+1 & -2\zeta_9{}^2-\zeta_9 & 0 & \zeta_9{}^2+\zeta_9-1 \end{pmatrix}.$$

*The set of $\mathbb{H}_3$-(Siegel-)period points of $J(C_1)$ coincides with the $\mathbb{S}p(6,\mathbb{Z})$-orbit of*

$$\begin{pmatrix} \frac{-2rs-1}{3r^2} & \frac{1}{r} & \frac{rs-1}{3r^2} \\ \frac{1}{r} & -1 & 0 \\ \frac{rs-1}{3r^2} & 0 & \frac{-2rs+2}{3r^2} \end{pmatrix} \cdot \omega + \begin{pmatrix} \frac{2rs-2}{3r^2} & \frac{1}{r} & \frac{-rs+1}{3r^2} \\ \frac{1}{r} & -1 & \frac{-1}{r} \\ \frac{-rs+1}{3r^2} & \frac{-1}{r} & \frac{2rs+1}{3r^2} \end{pmatrix}$$

*with*

$$r := -\zeta_9{}^4+\zeta_9{}^3+2\zeta_9{}^2+\zeta_9+1\ ,\ s := -(\zeta_9{}^5+\zeta_9{}^3+2\zeta_9{}^2+\zeta_9).$$

P r o o f: In [Ho3], sections 2.4-2.5, it is described a procedure to receive the period matrices starting from the coordinates of the fixed point $P_{\zeta_9}$. First one has to move the "diagonal ball" $\mathbb{B} \subset \mathbb{P}^2$ by a plane projective linear transformation to the "Picard ball" (Siegel domain) $\mathbb{B}' \subset \mathbb{P}^2$. This is done by the inverse of

$$M := \begin{pmatrix} \omega & 0 & -1 \\ 0 & 1 & 0 \\ -\omega^2 & 0 & -1 \end{pmatrix},$$

(see [Ho3], p. 28) acting on row-vectors from the right. Let $P' := (a:b:c) \in \mathbb{B}'$ be the image point of $P_{\zeta_9} \in \mathbb{B}$. Setting $b=1$ and applying Proposition 3 one gets $a, c \in \mathbb{Z}[\zeta_9]$. From the vector $(a,1,c)$ one gets the period matrices

via orthogonal fillings and $*$-procedure coming from Picard period integrals, all described in [Ho3] around Lemma 2.22. The numbers $r$, $s$ appear in the period matrix $\Pi$ at places $(1,1)$ or $(1,4)$, respectively. □

## References


[CER] *Cherdieu, J.-P., Estrada-Sarlabous, J., Reinaldo-Barreiro, E.*, Efficient Reduction on the Jacobian Variety of Picard Curves, in: Coding Theory, Cryptography and Related Areas, Proceedings of the ICCC-98, J. Buchmann, T. Hohold, H. Stichtenoth, H. Tapia-Recillas (eds.), pp. 13–28, Springer-Verlag, 2000.

[Da-Ha] *Davenport, H., Hasse, H.*, Die Nullstellen der Kongruenzzetafunktionen in gewissen zyklischen Fällen, J.Reine Angew. Math. **172**(1934), 151–182.

[De] *Deuring, M.*, Die Zetafunktion einer algebraischen Kurve vom Geschlechte Eins, Nachr. Akad. Wiss. Göttingen, 1953, 85–94.

[Feu] *Feustel, J.M.*, Kompaktifizierung und Singularitäten des Faktorraumes einer arithmetischen Gruppe, die in der zweidimensionalen Einheitskugel wirkt, Diplomarbeit, Humboldt-Univ. Berlin, 1976 (unpublished)

[Ha1] *Hasse, H.*, Zetafunktion und L-Funktionen zu einem arithmetischen Funktionenkörper vom Fermatschen Typus, Abhandlungen der Deutschen Akademie der Wissenschaften Berlin, Math.-Nat. Kl. 1954, Nr. 4, 5–70

[Ha2] *Hasse, H.*, Zahlentheorie, Akademie Verlag, Berlin, 1963

[He] *Hecke, E.*, Eine neue Art von Zetafunktionen und ihre Beziehungen zur Verteilung der Primzahlen, Math. Zeitschr. **1**(1918), 357-376, **6**(1920), 11–51.

[Ho1] *Holzapfel, R.-P.*, Geometry and Arithmetic around Euler partial differential equations, Dt. Verlag d. Wiss., Berlin/Reidel Publ. Comp., Dordrecht, 1986

[Ho2] *Holzapfel, R.-P.*, Hierarchies of endomorphism algebras of abelian varieties corresponding to Picard modular surfaces, Schriftenreihe Komplexe Mannigfaltigkeiten **190**, Univ. Erlangen, 1994

[Ho3] *Holzapfel, R.-P.*, The ball and some Hilbert problems, Lect. in Math. ETH Zürich, Birkhäuser, Basel-Boston-Berlin, 1995

[Lac] *Lachaud, G.*, Courbes diagonales et courbes de Picard, Prétirage No. 97-30, Institut de Mathematiqués de Luminy, 1997

[La] *Lang, S.*, Complex multiplication, Grundl. Math. Wiss. **255**, Springer, 1983

[Neu] *Neukirch, J.*, Algebraische Zahlentheorie, Springer, Berlin-Heidelberg, 1992

[Ta] *Taniyama, Y.*, L-functions of number fields and zeta functions of abelian varieties, J. Math. Soc. Japan **9**(1957), 330–366.

[We] *Weil, A.*, On Jacobi sums as "Grössencharaktere", Transact. Amer. Math. Soc. **73**(1952), 487–495.

# New Quantum Error-Correcting Codes from Hermitian Self-Orthogonal Codes over GF(4)

Jon-Lark Kim

Department of Mathematics, Statistics, and Computer Science,
322 SEO(M/C 249),
University of Illinois–Chicago,
851 S. Morgan, Chicago, IL 60607-7045, USA

**Abstract.** In order to construct good quantum-error-correcting codes, we construct good Hermitian self-orthogonal linear codes over GF(4). In this paper we construct record-breaking pure quantum-error-correcting codes of length 24 with 2 encoded qubits and minimum weight 7 from Hermitian self-orthogonal codes of length 24 with dimension 11 over GF(4). This shows that length $n = 24$ is the smallest length for any known $[[n, k, d]]$ quantum-error-correcting code with $k \geq 2$ and $d = 7$. We also give a construction method to produce Hermitian self-orthogonal linear codes GF(4) from a shorter length such code.

## 1 Introduction

It was shown [4] in 1995 that there could exist quantum-error-correcting codes (QECC throughout the paper) which would protect quantum information as classical error-correcting codes protect classical information. See [1] for the brief history of QECC. It is also known [1] that the problem of finding QECC is transformed into the problem of finding additive self-orthogonal codes under a certain inner product over the finite field GF(4). These additive self-orthogonal codes include the classical Hermitian self-orthogonal codes over GF(4). So our purpose is to construct good Hermitian self-orthogonal codes in order to construct good QECC using the ideas of [3].

We recall some basic definitions from [1,2]. Let $\mathrm{GF}(4) = \{0, 1, \omega, \overline{\omega}\}$ with the convention that $2 = \omega$ and $3 = \overline{\omega}$ where $\overline{\omega} = \omega^2 = 1 + \omega$. An *additive code* $\mathcal{C}$ *over* GF(4) *of length* $n$ is an additive subgroup of $\mathrm{GF}(4)^n$. As $\mathcal{C}$ is a free GF(2)-module, it has size $2^k$ for some $0 \leq k \leq 2n$. We call $\mathcal{C}$ an $(n, 2^k)$ code. It has a basis, as a GF(2)-module, consisting of $k$ basis vectors; a *generator matrix* of $\mathcal{C}$ will be a $k \times n$ matrix with entries in GF(4) whose rows are a basis of $\mathcal{C}$. The *weight* $\mathrm{wt}(\mathbf{c})$ of $\mathbf{c} \in \mathcal{C}$ is the number of nonzero components of $\mathbf{c}$. The minimum weight $d$ of $\mathcal{C}$ is the smallest weight of any nonzero codeword in $\mathcal{C}$. If $\mathcal{C}$ is an $(n.2^k)$ additive code of minimum weight $d$, $\mathcal{C}$ is called an $(n, 2^k, d)$ code.

To study QECC, we consider a somewhat different inner product from the ordinary inner product. We start with the following trace map. The *trace*

map $\mathrm{Tr} : \mathrm{GF}(4) \to \mathrm{GF}(2)$ is given by

$$\mathrm{Tr}(x) = x + x^2.$$

In particular $\mathrm{Tr}(0) = \mathrm{Tr}(1) = 0$ and $\mathrm{Tr}(\omega) = \mathrm{Tr}(\overline{\omega}) = 1$. The *conjugate* of $x \in \mathrm{GF}(4)$, denoted $\overline{x}$, is the image of $x$ under the Frobenius automorphism; in other words, $\overline{0} = 0$ $\overline{1} = 1$, and $\overline{\overline{\omega}} = \omega$. We now define the *trace inner product* of two vectors $\mathbf{x} = x_1 x_2 \cdots x_n$ and $\mathbf{y} = y_1 y_2 \cdots y_n$ in $\mathrm{GF}(4)^n$ to be

$$\mathbf{x} \star \mathbf{y} = \sum_{i=1}^{n} \mathrm{Tr}(x_i \overline{y_i}) \in \mathrm{GF}(4). \tag{1}$$

If $\mathcal{C}$ is an additive code, its *dual*, denoted $\mathcal{C}^{\perp}$, is the additive code $\{\mathbf{x} \in \mathrm{GF}(4)^n \mid \mathbf{x} \star \mathbf{c} = 0 \text{ for all } \mathbf{c} \in \mathcal{C}\}$. If $\mathcal{C}$ is an $(n, 2^k)$ code, then $\mathcal{C}^{\perp}$ is an $(n, 2^{2n-k})$ code. As usual, $\mathcal{C}$ is *trace self-orthogonal* if $\mathcal{C} \subseteq \mathcal{C}^{\perp}$ and *self-dual* if $\mathcal{C} = \mathcal{C}^{\perp}$. In particular, if $\mathcal{C}$ is trace self-dual, $\mathcal{C}$ is an $(n, 2^n)$ code.

We say that two additive codes $\mathcal{C}_1$ and $\mathcal{C}_2$ are *equivalent* provided there is a map sending the codewords of $\mathcal{C}_1$ onto the codewords of $\mathcal{C}_2$ where the map consists of a permutation of coordinates followed by a scaling of coordinates by elements of $GF(4)$ followed by conjugation of some of the coordinates. Notice that permuting coordinates, scaling coordinates, and conjugating some coordinates of a self-orthogonal (or self-dual) code does not change self-orthogonality (or self-duality) and the weight distribution of the code. The *automorphism group* of $\mathcal{C}$, denoted $\mathrm{Aut}(\mathcal{C})$, consists of all maps which permute coordinates, scale coordinates, and conjugate coordinates that send codewords of $\mathcal{C}$ to codewords of $\mathcal{C}$.

Now we state the relationship between QECC and additive self-orthogonal codes over GF(4).

**Lemma 1 (Theorem 2, [1]).** *Suppose that $\mathcal{C}$ is an additive trace self-orthogonal $(n, 2^{n-k})$ code of $\mathrm{GF}(4)^n$ such that there are no vectors of weight $< d$ in $\mathcal{C}^{\perp} \backslash \mathcal{C}$. Then an additive quantum-error-correcting code with parameters $[[n, k, d]]$ is obtained.*

If there are no nonzero vectors of weight $< d$ in $\mathcal{C}^{\perp}$ in the above lemma, $\mathcal{C}$ is *pure* (or *nondegenerate*); otherwise it is *impure* (or *degenerate*) [1]. A $[[n, k, d]]$ QECC can correct $[(d-1)/2]$ errors, where $k$ is the number of *encoded qubits* (*quantum bits*).

The *Hermitian inner product* is defined as

$$\mathbf{x} \cdot \mathbf{y} = x_1 \overline{y_1} + \cdots + x_n \overline{y_n} \in \mathrm{GF}(4), \tag{2}$$

for two vectors $\mathbf{x} = (x_1, \cdots, x_n)$ and $\mathbf{y} = (y_1, \cdots, y_n)$ in $\mathrm{GF}(4)^n$. A linear code with length $n$, dimension $k$(as a vector space over GF(4)), and minimum weight $d$ is called an $[n, k, d]$ code. The following theorem explains why Hermitian self-orthogonal linear codes are interesting in order to construct QECC.

**Lemma 2 (Theorem 3, [1]).** *A linear code $\mathcal{C}$ is self-orthogonal with respect to (1) if and only if it is self-orthogonal with respect to (2).*

Combining the above two lemmas, we get the following corollary.

**Corollary 1 ([1,5]).** *Let $\mathcal{C}$ be a Hermitian self-orthogonal linear $[n,k]$ code over* GF(4) *such that there are no vectors of weight $< d$ in $\mathcal{C}^{\perp}\backslash\mathcal{C}$, where $\mathcal{C}^{\perp}$ is the Hermitian dual of $\mathcal{C}$. Then there is a quantum-error-correcting $[[n, n-2k, d]]$ code.*

*Proof.* Since the given code $\mathcal{C}$ is linear, it has parameters as an additive code $(n, 2^{2k}) = (n, 2^{n-(n-2k)})$. Thus by Lemma 1 a quantum-error-correcting $[[n, n-2k, d]]$ code is obtained.

## 2 Construction Method

By generalizing the building-up construction [3, Theorem 1] for self-dual codes over GF(4) to self-orthogonal codes, we have the following theorem. We remark that there was an error in [3, Theorem 1] about the definition of $\overline{y_i}$ and so correct it here.

**Theorem 1.** *Let $G_0 = (g_i)$ be a generator matrix(may not be in standard form) of a Hermitian self-orthogonal code $\mathcal{C}_0$ over* GF(4) *of length $n$ with dimension $k$, where $g_i$ are rows of $G_0$ respectively for $1 \leq i \leq k$. Let $\mathbf{x} = (x_1, \cdots, x_n)$ be a vector in* $\mathrm{GF}(4)^n$ *with an odd weight. Suppose that $\overline{y_i} := (x_1, \cdots, x_n) \cdot g_i$ for $1 \leq i \leq k$. Here $\overline{y_i}$ is the conjugate of $y_i$ and $\cdot$ denotes the Hermitian inner product. Then the following matrix*

$$G = \left[\begin{array}{cc|ccccc} 1 & 0 & x_1 & x_2 & \cdots & x_{n-1} & x_n \\ \hline y_1 & \overline{y_1} & & & g_1 & & \\ \vdots & \vdots & & & \vdots & & \\ y_k & \overline{y_k} & & & g_k & & \end{array}\right]$$

*generates a Hermitian self-orthogonal code $\mathcal{C}$ over* GF(4) *of length $n+2$ with dimension $k+1$.*

As an example of the above theorem, let $\mathcal{C}_0$ be a Hermitian self-dual code over GF(4) generated by $\{1010, 0101\}$. If we take $\mathbf{x} = (01\omega\overline{\omega})$, then the code $\mathcal{C}$ is generated by $\{1001\omega\overline{\omega}, \overline{\omega\omega}1010, \overline{\omega\omega}0101\}$ by Theorem 1. This is the unique $[6,3,4]$ Hexacode over GF(4).

As in [3, Theorem 2] we get the converse of the above theorem as follows.

**Theorem 2.** *Any Hermitian self-orthogonal code $\mathcal{C}$ over* GF(4) *of length $n$ and dimension $k > 1$ with minimum weight $d > 2$ is obtained from some Hermitian self-orthogonal code $\mathcal{C}_0$ of length $n-2$ and dimension $k-1$(up to equivalence) by the construction in Theorem 1.*

In the following section, we construct 19 inequivalent linear Hermitian self-orthogonal $[24, 11, 8]$ codes over GF(4) with its dual minimum weight 7. These give record-breaking $[[24, 2, 7]]$ quantum-error-correcting codes.

**Table 1.** Generator matrix of $Q_{22}^1$

$$G(Q_{22}^1) = \begin{bmatrix} 10000001000000133233203 \\ 01000003000202212312 12 \\ 00100001000333030001 20 \\ 00010002000122203320 02 \\ 00001002000210312011 03 \\ 00000102000210012332 10 \\ 00000011000220203121 01 \\ 00000001000001001133 22 \\ 00000000100001111111 \\ 00000000010001011223 3 \end{bmatrix}$$

**Table 2.** New $[[24, 2, 7]]$ quantum-error-correcting codes using $Q_{22}^1$

| codes $C$ | $\mathbf{x} = (x_1, \cdots, x_{11})$ | $A_8$, $B_7$ | codes $C$ | $\mathbf{x} = (x_1, \cdots, x_{11})$ | $A_8$, $B_7$ |
|---|---|---|---|---|---|
| $Q_{24,1}$ | 03001111121 | 117, 171 | $Q_{24,2}$ | 22321301221 | 144, 156 |
| $Q_{24,3}$ | 22131000321 | 141, 186 | $Q_{24,4}$ | 10020033021 | 108, 174 |
| $Q_{24,5}$ | 13013111132 | 99, 156 | $Q_{24,6}$ | 12030021132 | 120, 198 |
| $Q_{24,7}$ | 23310200132 | 105, 132 | $Q_{24,8}$ | 20012000332 | 96, 150 |
| $Q_{24,9}$ | 31100212032 | 105, 165 | $Q_{24,10}$ | 02212022032 | 102, 162 |
| $Q_{24,11}$ | 12010313032 | 126, 183 | $Q_{24,12}$ | 11110021202 | 114, 150 |
| $Q_{24,13}$ | 20223012202 | 96, 159 | $Q_{24,14}$ | 33030202002 | 105, 147 |
| $Q_{24,15}$ | 02311200002 | 108, 147 | $Q_{24,16}$ | 31231302123 | 102, 150 |
| $Q_{24,17}$ | 33321333303 | 102, 159 | $Q_{24,18}$ | 20212031120 | 108, 180 |
| $Q_{24,19}$ | 21121332320 | 90, 144 | | | |

## 3 Existence of $[[24, 2, 7]]$ Quantum-error-correcting Codes

According to [1, Table III], it is known that the highest minimum weight $d$ for $[[24, 2, d]]$ codes is bounded by $6 \leq d \leq 8$. We apply Theorem 1 to a Hermitian self-orthogonal $[22, 10, 8]$ code $Q_{22}^1$ in Table 1 with many possibilities for vectors $\mathbf{x}$ to get 19 inequivalent Hermitian self-orthogonal $[24, 11, 8]$ codes such that their dual codes all have minimum weight $d = 7$. Hence it follows from Corollary 1 that there exist pure $[[24, 2, 7]]$ codes. Moreover length $n = 24$ is the smallest length for any known three error-correcting $[[n, k, 7]]$ codes with $k \geq 2$ according to [1, Table III]. See Table 2 for such codes, where $A_8$(resp, $B_7$) denotes the number of minimum vectors in $\mathcal{C}$(resp, $\mathcal{C}^{\perp}$), justifying the inequivalence of the codes. Here we gave the vectors $\mathbf{x}$ with only first 11 co-

ordinates, the right half being 1's. For example, $\mathbf{x} = (23310200132)$ in $Q_{24,7}$ means $\mathbf{x} = (2331020013211111111111)$. We summarize our result as follows.

**Theorem 3.** *There exist at least 19 inequivalent pure* $[[24, 2, 7]]$ *quantum-error-correcting codes, which are obtained from Hermitian self-orthogonal linear* $[24, 11, 8]$ *codes with its dual minimum weight 7.*

**Acknowledgment.** The author would like to thank Vera Pless for reading the first manuscript and the referee for useful comments.

## References

1. Calderbank, A. R., Rains, E. M., Shor, P. W., Sloane, N. J. A. (1998) Quantum error correction via codes over GF(4). IEEE Trans. Inform. Theory. **44**, 1369–1387
2. Gaborit, P., Huffman, W. C., Kim, J.-L., and Pless, V. (2001) On additive GF(4) codes. DIMACS Workshop on Codes and Association Schemes DIMACS Series in Discrete Math. and Theoret. Computer Science, American Mathematical Society, **56**, 135–149
3. Kim, J.-L. (2001) New self-dual codes over GF(4) with the highest known minimum weights. IEEE Trans. Inform. Theory. **47**, 1575–1580
4. Shor, P. W. (1995) Scheme for reducing decoherence in quantum memory. Phys. Rev. A. **52**, 2493
5. Thangaraj, A., McLaughlin, S. W. (2001) Quantum codes from cyclic codes over $GF(4^m)$. IEEE Trans. Inform. Theory. **47**, 1176–1178

# A Note on the Counter-Example of Patterson–Wiedemann

Philippe Langevin[1] and Jean-Pierre Zanotti[1]

Université de Toulon et du Var,
Groupe de Recherche en Informatique et Mathématiques,
F-83957 La Garde cedex, France.

**Abstract.** Following Patterson and Wiedemann [10], we find new counter-examples for a conjecture of Mykkelveit [9] related to the covering radius of the first order Reed-Muller code. One of them has a remarkable algebraic structure.

## 1 Spectral Magnitude

Let $L$ be an extension of degree $m$ of $\mathbf{F}_2$, the field of order 2. Let $\mu_L$ be the canonical additive character of $L$ that maps $z \in L$ onto $(-1)^{\mathrm{tr}_L(z)}$, where $\mathrm{tr}_L(z)$ is the absolute trace of $z$. Using the fact that the non-degenerate bilinear symmetric form $(x, y) \mapsto \mathrm{tr}_L(xy)$ is non-degenerate, one defines the Fourier coefficient of Boolean function $f\colon L \to \mathbf{F}_2$ at $a \in L$ by :

$$\widehat{f}(a) = \sum_{x \in \mathbf{F}_2^m} (-1)^{f(x)} \mu_L(ax) \tag{1}$$

Note that the Hamming distance between $f$ and the affine function $x \mapsto a.x + b$ ($b \in \mathbf{F}_2$) is $2^{m-1} - \frac{(-1)^b}{2}\widehat{f}(a)$. The *spectral magnitude* of $f$ defined by $\mathrm{Sp}(f) = \sup_{a \in L} |\widehat{f}(a)|$ measures the distance between $f$ and the space of affine functions. The *spectral radius* $R(m) = \inf_f \mathrm{Sp}(f)$ is particularly relevant from the cryptographic point of view. Fourier analysis gives the Parseval's lower bound :

$$\sqrt{2^m} \leq R(m) \tag{2}$$

The functions reaching that bound are the *bent functions* of Rothaus [11], they exist in even dimension only. All along that paper, we will denote $m = 2t$ or $m = 2t + 1$ according whether $m$ is even or odd. It is easy to see that the spectral magnitude of a quadratic form of rank $k$ is $2^{\frac{m+k}{2}}$. In even dimension, non-degenerate quadratic functions are bent, but there are bent functions of degree $d$ for all $d$ between 2 and $t$. In odd dimension, the minimal spectral magnitude of quadratic functions is $2^{t+1}$ whence we obtain the quadratic bound :

$$R(m) \leq 2^{t+1}. \tag{3}$$

We say that $f$ exceeds the quadratic bound if its spectral magnitude is less than $2^{t+1}$. Berlekamp and Welch [2], Mykkelveit [9] and more recently Hou

[6] have proved that if $m \in \{3, 5, 7\}$ then $R(m) = 2^{t+1}$ : such a function does not exist, for these values of $m$. For $m = 9$, 11 and 13 nothing more than the bounds (2) and (3) is known. In their famous note [10], Patterson and Wiedemann construct Boolean functions of spectral magnitude 216 in dimension 15 providing counter-examples to a conjecture of Mykkelveit that claimed $R(m) = 2^{(m+1)/2}$ for odd $m$. They conjecture that $R(m) \sim \sqrt{2^m}$ and explain :

*"We have not succeeded in understanding algebraically the choice of orbits and thus have not succeeded in generalizing our construction to other dimensions although we suspect there is a construction when m is not a prime power"*

In this note, we give a new counter-example that makes a link with the theory of (relative) difference sets.

## 2 Action of Cyclic Groups

Let $G$ be a subgroup of order $d$ of $L^\times$. We say that a Boolean function $f$ is invariant under $G$ if and only if $F(gz) = F(z)$ for all $g \in G$ and $z \in L$. It is equivalent to say that $f$ is an union of cosets of $L^\times$ modulo $G$. We want to study the spectral magnitude of the $G$-invariant Boolean functions in order to obtain functions exceeding the quadratic bound.

Using a slight abuse of notations, we denote by $G$ the characteristic Boolean function of the group $G$ so that

$$F(z) = \sum_{\omega \in \Omega} s(\omega) G(\omega z) = \sum_{\omega \in D} G(\omega z),$$

where $s$ is a numerical Boolean sequence corresponding to a choice of cosets of $G$ in $L^\times$ whose support denoted by $D$ is a set of cardinality say $k$. We denote by $v$ the order of the quotient group $\Omega = L^\times / G$ so that $dv = 2^m - 1$. Later we will make connections with difference set theory that explains our notations : $D$, $v$, $k$. The Fourier coefficient of $F$ at 0 is

$$\widehat{F}(0) = 2^m - 2dk = 2^m - 1 + 1 - 2dk = dv + 1 - 2dk.$$

This equality shows that $d$ must be chosen less than or equal to $2^t + 1$ to obtain bent functions, and less than $2^{t+1}$ to have some chance to exceed the quadratic bound. In particular, if we want $F$ to have a spectral magnitude less than or equal to $R$ we must assume

$$|v - 2k| \leq \frac{R+1}{d}.$$

Let $a$ be a non-zero element of $L$. The Fourier coefficient of $G$ at $a$ depends on the class of $a$ modulo $G$ and is given by

$$\widehat{G}(a) = -\sum_{g \in G} \mu_L(ag) + \sum_{g \notin G} \mu_L(ag) = -2 \sum_{g \in G} \mu_L(ag) = -\frac{2}{v} \sum_{\chi \perp G} \tau_L(\chi) \bar{\chi}(a),$$

where $\tau_L(\chi) = \sum_{z \in L^\times} \chi(z)\mu_L(z)$ is the Gauss sum associated to the multiplicative character $\chi$ and the additive character $\mu_L$. The coefficients of $F$, is equal to the inter-correlation of $s$ by $\widehat{G}$ and is given by :

$$\widehat{F}(a) = \sum_{\omega \in \Omega} s(\omega)\widehat{G}(a\omega) = \sum_{\omega \in D} \widehat{G}(a\omega) = -\frac{2}{v} \sum_{\chi \perp G} \tau_L(\chi) D(\chi) \bar{\chi}(a).$$

In the last equality, we identify $G^\perp$ with the dual of $\Omega$, whence $D(\chi)$ is the (multiplicative) Fourier transform of $D$ at $\chi$. The main question is : what kind of set $D$ could make small the absolute value of the $\widehat{F}(a)$'s ?

*Problem 1.* Find the spectral magnitude of the Legendre construction where $v$ is prime and $D$ is the set of quadratic residues modulo $v$.

## 3 Functions from Sub-fields

Assume that $m$ is even. Let $K$ be the subfield of degree $t$ in $L$. By a theorem of Stickelberger [8], we know that the Gauss sums of order dividing $2^t + 1$ are rational and so equal to $2^t$. It follows that the Fourier transform of $F$ at $a$ is given by :

$$-\frac{1}{2}\widehat{F}(a) = \frac{1}{v} \sum_{\chi \perp G} \tau_L(\chi) s(\chi) \bar{\chi}(a) = \frac{2^t}{v} \sum_{1 \neq \chi \perp G} s(\chi)\bar{\chi}(a) - \frac{k}{v} = 2^t s(a) - k$$

We recover a result of Dillon [4].

**Proposition 1.** *If $D$ is a set of $2^{t-1}$ cosets of the group $G$ in $L^\times$ then the corresponding function is bent.*

*Proof.* Clear.

In [10], Patterson and Wiedemann proceed by analogy looking for Boolean function that are invariant under the multiplicative group of subfields of the field $\mathbf{F}_{2^m}$. The multiplicative group of the field of order $2^{15}$ factors in $7 \times 31 \times 151$ and contains a cyclic group $G$ of order 217 which is the direct product of ${\mathbf{F}_8}^\times$ and ${\mathbf{F}_{32}}^\times$, there are $2^{151}$ $G$-invariant functions but only $2^{12}$ of them are also invariant under the Frobenius automorphism and correspond to the 12 cyclotomic classes modulo 151 represented by the coset leader 0, 1, 3, 5, 7, 11, 15, 17, 23, 35, and 37. Using a computer, one can see that 8 of these functions exceed the quadratic bound, see TAB.1.

We have computed the auto-correlation function of the sequences $s$ corresponding to them, and for example, the auto-correlation of the first row takes only four values : 60 times the value 36, 30 times the value 38, 60 times the value 40 and 1 time 76. Its structure is similar to of that a difference sets but is not a (relative) difference set.

**Table 1.** Spectral distributions of Patterson-Wiedemann functions

| orbits | degree | -216 | -152 | -88 | -24 | 40 | 168 | 232 |
|---|---|---|---|---|---|---|---|---|
| 0 1 7 11 15 17 | 8 | 5 | | 1 | | 1 | 5 | |
| 0 1 3 7 17 35 | 9 | 5 | | 1 | | 1 | 5 | |
| 0 3 5 23 35 37 | 10 | 2 | 4 | 1 | 1 | | 1 | 3 |
| 0 5 11 15 23 37 | 10 | 2 | 4 | 1 | 1 | | 1 | 3 |

## 4 Function from Subgroup

Once again, assume that $m$ is even. We suppose that $G$ is the group of order $d := 2^t + 1$. The function $z \mapsto \mathrm{tr}_L(az^{2^t+1})$ is $G$-invariant, bent if and only if the Kloosterman sum $\mathrm{K}l(a)$ is equal to $-1$, see [5]. Arguing with elliptic curve theory, Lachaud and Wolfmann [7] prove that for any $L$ there are $2^t - 1$ such $a$. Now, we recover a similar result by means of Gauss sums. A classical proposition of Davenport-Hasse [8] asserts that for any character $\chi$ of order $v$, the negative of the Gauss sum $\tau_L(\chi)$ is equal to $\tau_K(\chi)^2$. It follows that

$$-\frac{1}{2}\widehat{F}(a) = \frac{1}{v}\sum_{\chi \perp G} \tau_L(\chi)s(\chi)\bar{\chi}(a) = \frac{1}{v}\sum_{\chi \perp G} \tau_K(\chi)^2 s(\chi)\bar{\chi}(a).$$

Let $c$ be a non-zero element of $K$ and let $E_c = \{z \in K^\times \mid \mathrm{tr}_K(z/c) = 1\}$. It has order $2^{t-1}$ and the multiplicative Fourier transform of $E_c$ at $\chi \neq 1$ is

$$E_c(\chi) = -\frac{1}{2}\sum_{z \in K^\times} [1 - \mu_K(z/c)]\chi(z) = -\frac{1}{2}\tau_K(\chi)\chi(c).$$

Similarly those of the set $D_c = \{z \mid \mathrm{tr}_K(c/z) = 1\}$ is equal to $-\tau_K(\bar{\chi})\chi(c)/2$.

**Proposition 2.** *Let $c \in K^\times$. The set $D_c$ defines a bent function.*

*Proof.* Let $s_c$ be the characteristic function of $E_c$. Let $F$ be the boolean function defined by $D_c$ whose Fourier coefficient at $a$ is :

$$\begin{aligned} -\frac{1}{2}\widehat{F}(a) &= \frac{1}{v}\sum_{\chi \perp G} -\tau_K(\chi)^2 D_c(\chi)\bar{\chi}(a) \\ &= \frac{2^t}{v}\sum_{1 \neq \chi \perp G} -\tau_K(\chi)\chi(c)/2\bar{\chi}(a) + \frac{2^{t-1}}{v} \\ &= -2^t s_c(a) + 2^{t-1}. \end{aligned}$$

*Problem 2.* Prove that the bent functions given by (2) are quadratic.

## 5 A Remarkable Example

As Patterson and Wiedemann have done with the group of order 217, we have looked for boolean functions invariant under the action of the group of order 151. There are $2^{217}$ such functions but only $2^{23}$ are also invariant by the Frobenius automorphism. Four of them (TAB.2) exceed the quadratic bound with spectral magnitude : 248, 234, 232, and 246. It is worse than in the case of the group of order 217 but one of these functions has a remarkable structure. Indeed, the auto-correlation of the function of the second row takes only three values. Its support $D$ is composed of the cyclotomic classes of 1, 7, 9, 11, 25, 37, 49, it defines a $(7.31, 108, 100, 46)$ relative difference set. Moreover, regarding $\mathbf{Z}/151\mathbf{Z}$ as the direct product $\mathbf{Z}/7\mathbf{Z} \times \mathbf{Z}/31\mathbf{Z}$, $D$ is equal to

$$D_1 \times \mathbf{Z}/31\mathbf{Z} \quad \cup \quad \{0\} \times D_2$$

where $D_1$ is the Singer set $\{1, 2, 4\}$ and $D_2$ the Singer set $\{3, 5, 7, \ldots\}$ so that $D$ appears as one of the relative difference sets described by theorem 2.1 of [1].

Concluding remark. Can we construct other functions, related to Singer sets for a greater $m = pq$, where $p$ and $q$ are distinct odd prime ? Unfortunately, the answer is no! Indeed, the condition on $d$ given in section 2 says that

$$\frac{2^{pq} - 1}{(2^p - 1)(2^q - 1)} \leq 2^{(pq+1)/2}$$

implying $p = 3$ and $q = 5$.

**Table 2.** Action of $\mathbf{Z}/151\mathbf{Z}$

| spec. mag. | orbits | degree | correlation |
|---|---|---|---|
| 248 | 3 5 13 19 21 | 12 | 186 [46], 1 [108] |
| | 27 31 33 35 77 | | 10 [99], 10 [100], 10 [101] |
| 234 | 0 3 5 13 19 21 | 15 | 186 [47], 30 [101], 1 [109] |
| | 27 31 33 35 77 | | 186 [47], 30 [101], 1 [109] |
| 232 | 1 7 9 11 15 25 | 10 | 186 [46], 30 [100], 1 [108] |
| | 37 49 93 105 | | 186 [46], 30 [100], 1 [108] |
| 246 | 0 1 7 9 11 15 25 | 15 | 186 [47] , 1 [109] |
| | 37 49 93 105 | | 10 [100], 10 [101], 10 [102] |

## Acknowledgments

The authors wish to thank professors K. T. Arasu and J. Wolfmann for pointing and communicating to us the references [1] and [7].

## References

1. Arasu K. T., Haemers W. H., Jungnickel D., Pott A. Matrix constructions of divisible designs. *Linear of Algebra and its Applications*, 153, 1991.
2. Berlekamp E. R., Welch L. R. Distributions of the cosets of the $(32, 6)$ Reed-Muller code. *IEEE Transactions on Information Theory*, 13(1):203–207, 1972.
3. Constantin J., Courteau B., Wolfmann J. Numerical experiments related to the covering radius of some Reed-Muller codes. In Lecture Notes in Computer Science, editor, *Algebraic Algorithms and Error-Correcting Codes*, volume 229, pages 69–75. 3rd International Conference AAECC, 1986.
4. Dillon J. F. *Elementary Hadamard Difference Sets.* PhD thesis, University of Maryland, 1974.
5. Dillon J. F. Elementary Hadamard difference sets. In *sixth SECCGTC*, 1975.
6. Hou X.-D. Covering radius of the Reed-Muller code $R(1, 7)$ — a simpler proof. *Journal of Combinatorial Theory*, 74, 1996.
7. Lachaud G., Wolfmann J. Sommes de Kloosterman, courbes elliptiques et codes cycliques en caractéristique 2. *Comptes rendus de l'académie des sciences*, 305:881–883, 1987.
8. Lidl R., Niederreiter H. *Finite Fields*, volume 20 of *Encyclopedia of Mathematics and its Applications.* Addison-Wesley, 1983.
9. Mykkelveit J. J. The covering radius of the $(128, 8)$ Reed-Muller codes is 56. *IEEE Transactions on Information Theory*, 26(3):359–362, 1980.
10. Patterson N. J., Wiedemann D. H. The covering radius of the $(1, 15)$ Reed-Muller code is at least 16276. *IEEE Transactions on Information Theory*, 29:354–356, 1983.
11. Rothaus O. S. On bent functions. *Journal of Combinatorial Theory (A)*, 20:300–305, 1976.

# Continued Fractions for Certain Algebraic Power Series over a Finite Field

Alain Lasjaunias

Université de Bordeaux I
CNRS-UMR 5465
351, Cours de la Libération
F33405 Talence Cedex - France
E-mail: lasjauni@math.u-bordeaux.fr

**Abstract.** In this survey we discuss rational approximation properties of certain algebraic power series over a finite field using continued fractions. These algebraic elements are fixed points of the composition of a linear fractional transformation and of the Frobenius homomorphism.

## 1 The Fields of Power Series over $\mathbb{F}_q$

Let $\mathbb{F}_q$ be the field with $q$ elements and let $p$ be its characteristic. We consider the field $\mathbb{F}_q(T)$ of rational functions in the indeterminate $T$, with coefficients in $\mathbb{F}_q$. On this field $\mathbb{F}_q(T)$ we consider the ultrametric absolute value defined by

$$|P/Q| = |T|^{\deg P - \deg Q} \quad \text{and} \quad |0| = 0$$

where $|T| > 1$ is a fixed real number. The field obtained by completion from $\mathbb{F}_q(T)$ for this absolute value will be denoted by $\mathbb{F}(q)$. If $\Theta \in \mathbb{F}(q)$ and $\Theta \neq 0$, we can write it as a power series expansion

$$\Theta = \sum_{k \leq k_0} \theta_k T^k \quad \text{with} \quad k_0 \in \mathbb{Z}, \quad \theta_k \in \mathbb{F}_q \quad \theta_{k_0} \neq 0$$

and the absolute value is extended by $|\Theta| = |T|^{k_0}$. This construction is analogous to the classical construction of the field of real numbers from the ring of integers. The resulting field $\mathbb{F}(q)$ has many similar properties with $\mathbb{R}$ and hence could be called the field of formal numbers over $\mathbb{F}_q$. In the two statements below we consider $\Theta \in \mathbb{F}(q)$ written as $\Theta = P(T) + \sum_{n \geq 0} \theta_n T^{-n}$ where $P(T)$ is the integral (polynomial) part of $\Theta$. The first result is an illustration of the similarity with real numbers.

**Theorem 1.1.** *$\Theta \in \mathbb{F}_q(T)$ if and only if the sequence $(\theta_n)_{n \geq 0}$ is ultimately periodic.*

Since we are concerned with algebraic elements in $\mathbb{F}(q)$ over $\mathbb{F}_q(T)$, it is interesting to mention a deeper result concerning the power series expansion, due to Christol.

**Theorem 1.2.** *$\Theta$ is algebraic over $\mathbb{F}_q(T)$ if and only if the set of subsequences*

$$\{(\theta_{q^i n+r})_{n\geq 0} : i \geq 0 \text{ and } 0 \leq r \leq q^i - 1\}$$

*is finite.*

Clearly the same construction as above can be made from an arbitrary base field $K$ instead of $\mathbb{F}_q$. Then the resulting field is called the field of power series over $K$ and denoted by $K((T^{-1}))$ . Indeed the finiteness of the base field is essential in many results and this makes the field $\mathbb{F}(q)$ particularly interesting. We study here, in the case $K = \mathbb{F}_q$, rational approximation to algebraic power series over $K(T)$. For a study in a larger context and for more references see [L1].

Many classical questions in number theory, which have been studied in the setting of real numbers, can be transposed and studied in fields of power series. The starting point in the study of rational approximation to algebraic real numbers is a famous theorem established by Liouville in 1850. This theorem has been adapted by Mahler [M] in fields of power series with an arbitrary base field.

**Theorem 1.3.** *Let $K$ be a field. Let $\Theta \in K((T^{-1}))$ be an algebraic element over $K(T)$, of degree $n > 1$. Then there is positive real number $C$ such that*

$$|\Theta - P/Q| \geq C|Q|^{-n}$$

*for all $P, Q \in K[T]$, with $Q \neq 0$.*

In the case of real numbers, we know that Liouville's theorem was the first step in the study of rational approximation to algebraic numbers. A deeper result was obtained with Roth's theorem established in 1955. This last theorem can be transposed in fields of power series if and only if the base field has characteristic zero. In this case the exponent $n$ in the right hand side of the inequality in the above theorem can be replaced by $2 + \epsilon$ for all $\epsilon > 0$ with the constant $C$ depending upon $\epsilon$. But this is not so in the case of the field $\mathbb{F}(q)$ and consequently the study of rational approximation to algebraic elements becomes more complex.

## 2 The Continued Fraction Algorithm

As in the classical context of the real numbers, we have a continued fraction algorithm in $\mathbb{F}(q)$. For a general reference on this subject see [S]. If $\Theta \in \mathbb{F}(q)$

we can write

$$\Theta = a_0 + 1/(a_1 + 1/(a_2 + ... = [a_0, a_1, a_2, ....] \quad \text{where} \quad a_i \in \mathbb{F}_q[T].$$

The $a_i$'s are called the *partial quotients* and we have $\deg a_i > 0$ for $i > 0$. This continued fraction expansion is finite if and only if $\Theta \in \mathbb{F}_q(T)$. As in the classical theory we define recursively the two sequences of polynomials $(x_n)_{n\geq 0}$ and $(y_n)_{n\geq 0}$ by

$$x_n = a_n x_{n-1} + x_{n-2} \quad \text{and} \quad y_n = a_n y_{n-1} + y_{n-2},$$

with the initial conditions $x_0 = a_0$, $x_1 = a_0 a_1 + 1$, $y_0 = 1$ and $y_1 = a_1$. We have $x_{n+1}y_n - y_{n+1}x_n = (-1)^n$, whence $x_n$ and $y_n$ are coprime polynomials. The rational $x_n/y_n$ is called a *convergent* to $\Theta$ and we have $x_n/y_n = [a_0, a_1, a_2, ..., a_n]$. Because of the ultrametric absolute value we have

$$|\Theta - x_n/y_n| = |x_{n+1}/y_{n+1} - x_n/y_n| = |y_n y_{n+1}|^{-1} = |a_{n+1}|^{-1}|y_n|^{-2}. \quad (1)$$

We mention an important result which is an analogue of Lagrange's theorem (see [S]).

**Theorem 2.1.** *Let $\Theta \in \mathbb{F}(q)$ be irrational. Then the sequence of partial quotients in the continued fraction expansion of $\Theta$ is ultimately periodic if and only if $\Theta$ is quadratic over $\mathbb{F}_q(T)$.*

## 3 The Approximation Exponent

Let $\Theta \in \mathbb{F}(q)$ be an irrational element. We define the *approximation exponent* of $\Theta$ by

$$\nu(\Theta) = \limsup_{|Q|\to\infty}(-\frac{\log|\alpha - P/Q|}{\log|Q|})$$

where $P$ and $Q$ run over polynomials in $\mathbb{F}_q[T]$ with $Q \neq 0$. Let us consider the continued fraction expansion $\Theta = [a_0, a_1, \ldots, a_n, \ldots]$. Since the convergents are the best rational approximations to $\Theta$, it is clear, using (1), that the approximation exponent can also be defined directly by

$$\nu(\Theta) = 2 + \limsup_k(\deg a_{k+1}/\deg y_k).$$

Observe that $\deg y_k = \sum_{1\leq i\leq k} \deg a_i$ and therefore $\nu(\Theta)$ is directly connected to the growth of the sequence $(\deg a_i)_{i\geq 1}$. In particular if the sequence $(\deg a_i)_{i\geq 1}$ is bounded then $\nu(\Theta) = 2$. Clearly we may have $\nu(\Theta) = 2$ without this assumption.

Because of Mahler's theorem, for all $\Theta \in \mathbb{F}(q)$ algebraic over $\mathbb{F}_q(T)$ and of degree $n > 1$, we have

$$\nu(\Theta) \in [2, n].$$

We give now two classical examples, in some sense dual of each other, of algebraic elements for which the approximation exponent is maximal. The second was first introduced in [M]. Here $r = p^t$ where $t \geq 0$ is an integer.

*Example 1* : We define $\Theta \in \mathbb{F}(p)$ by

$$\Theta = [0, T, T^r, \ldots, T^{r^k}, \ldots].$$

Because of the Frobenius homomorphism, we have $\Theta = 1/(T + \Theta^r)$. It is easy to see that $\nu(\Theta) = r + 1$.

*Example 2* : Here we assume that $t \neq 0$ and we define $\Theta \in \mathbb{F}(p)$ by

$$\Theta = \sum_{k \geq 0} T^{-r^k}.$$

In that case we have $\Theta = 1/T + \Theta^r$ and $\nu(\Theta) = r$ (see [M]). It is interesting to observe that the continued fraction for this algebraic element can be given explicitly.

**Theorem 3.1.** *Assume that $r > 2$. We define recursively on $n \geq 1$ a finite sequence $\Omega_n$ of elements in $\mathbb{F}_p[T]$ such that*

$$\Omega_1 = T \quad \text{and} \quad \Omega_n = \Omega_{n-1}, -T^{(r-2)r^{n-2}}, -\widetilde{\Omega}_{n-1}$$

*If $\Omega_n = a_1, a_2, \ldots, a_m$ then $\widetilde{\Omega}_n = a_m, \ldots, a_2, a_1$ and $-\Omega_n = -a_1, -a_2, \ldots, -a_m$. Further $\Omega_\infty$ denotes the infinite sequence beginning by $\Omega_n$ for all $n \geq 1$. Then the continued fraction expansion for $\Theta$ is $[0, \Omega_\infty]$.*

We recall that no explicit continued fraction expansion is known for a non-quadratic algebraic real number.

## 4 Algebraic Elements of Class

Let $r = p^t$ where $t \geq 0$ is an integer. We denote by $\mathcal{H}(r, q)$ the subset of irrational elements in $\mathbb{F}(q)$ such that there exist $A, B, C, D \in \mathbb{F}_q[T]$ with

$$\Theta = \frac{A\Theta^r + B}{C\Theta^r + D} \tag{2}$$

We put $\mathcal{H}(q) = \bigcup_{r=p^t, t \geq 0} \mathcal{H}(r, q)$. The elements of $\mathcal{H}(q)$ are called algebraic elements of class I. Observe that the elements which are quadratic and also cubic over $\mathbb{F}_q(T)$ are algebraic of class I (indeed the four elements $1, \Theta, \Theta^p$ and $\Theta^{p+1}$ are linked over $\mathbb{F}_q(T)$).

These algebraic elements were introduced by Baum and Sweet [BS1] when the base field was $\mathbb{F}_2$. Later they were considered by Mills and Robbins [MR]

in a general context. We give below a theorem which gives the main known rational approximation properties of these elements.

**Theorem 4.1.** *We have the following properties :*
(i) *If* $\Theta \in \mathcal{H}(q)$ *then* $\nu(\Theta) \in \mathbb{Q}$ *and* $\liminf_{|Q|\to\infty} |Q|^{\nu(\Theta)}|\Theta - P/Q| \neq 0, \infty$.
(ii) *If* $\Theta \in \mathcal{H}(r,q)$ *and if in equation (2) we have* $\deg(AD-BC) < r-1$ *then* $\nu(\Theta) > 2$.
(iii) *If* $\Theta \in \mathbb{F}(q)$ *is algebraic of degree* $n > 1$ *and* $\Theta \notin \mathcal{H}(q)$ *then* $\nu(\Theta) \leq [n/2] + 1$.

For (i) see [dM] and [V], for (ii) see [L3] and for (iii) see [LdM1] and [LdM2]. Observe that the last property implies that the algebraic elements which are best approximable by rational elements must belong to $\mathcal{H}(q)$ (as both examples above). A consequence of the first property in this theorem is that there exists a natural partition of the set $\mathcal{H}(q)$ into two subsets $\mathcal{H}_1(q)$ and $\mathcal{H}_2(q)$.

**Corollary 4.2.** *Let* $\Theta = [a_0, a_1, a_2, \dots] \in \mathcal{H}(q)$. *Then we have either*
- $\Theta \in \mathcal{H}_1(q)$ *: there is a real number* $\mu > 0$ *such that*

$$\limsup_k (\deg a_{k+1} / \sum_{1 \leq i \leq k} \deg a_i) = \mu \qquad (i.e. \quad \nu(\Theta) > 2)$$

*or*
- $\Theta \in \mathcal{H}_2(q)$ *: there is an integer* $B$ *such that*

$$\deg a_i \leq B \quad for \quad i \geq 0.$$

The two examples given above belong to $\mathcal{H}_1(q)$ when they are not quadratic. Clearly $\mathcal{H}_2(q)$ contains the quadratic formal numbers. The existence of non-quadratic elements in $\mathcal{H}_2(q)$ was first observed by Baum and Sweet [BS1] and latter by Mills and Robbins [MR]. It is interesting to remark that evident computer calculation shows that $\mathcal{H}_1(q)$ is a much larger set than $\mathcal{H}_2(q)$. It is also important to observe that both subsets $\mathcal{H}_1(q)$ and $\mathcal{H}_2(q)$ are stable under three transformations : 1) the Frobenius homomorphism, 2) a linear fractional transformation with polynomial coefficients, 3) the change of $T$ into a polynomial of $T$. Moreover these three transformations preserve the degree of an algebraic element over $\mathbb{F}_q(T)$.

## 5 A Particular Subclass in $\mathcal{H}(q)$

If we look for an analogue of the subset $\mathcal{H}(q)$ in the setting of the real numbers, we should consider the subset of quadratic real numbers. Indeed these numbers are fixed points of a linear fractional transformation with integer

coefficients. This implies the peculiar pattern of their continued fraction expansions. Unluckily the possibility of describing explicitly the continued fraction expansion for all the elements of $\mathcal{H}(q)$ seems yet out of reach (see [MR]). Nevertheless this description is possible for a particular subclass.

We will say that an element in $\mathcal{H}(q)$ is of class IA if $AD - BC \in \mathbb{F}_q^*$ in equation (2). Example 1 given above belongs to this subclass. Observe that, according to the second property of Theorem 4.1, if $\Theta$ is of class IA and $r \neq 1$ then $\Theta \in \mathcal{H}_1(q)$. Such algebraic elements have been studied by Schmidt [S] and also by Thakur [T] who proved independently the following theorem.

**Theorem 5.1.** *$\Theta \in \mathbb{F}(q)$ is algebraic of class IA if and only if there exist $k \geq -1$, $a_j, c_i \in \mathbb{F}_q[T]$ with $0 \leq j \leq k$ and $i \geq 1$, $t \in \mathbb{N}^*$ and $\epsilon \in \mathbb{F}_q^*$ such that*

$$\Theta = [a_0, a_1, \ldots, a_k, c_1, c_2, \ldots, c_n, \ldots]$$

*where for $l \geq 1$ we have*

$$c_{l+t} = \begin{cases} \epsilon c_l^r & \text{if } l \text{ is odd,} \\ \epsilon^{-1} c_l^r & \text{if } l \text{ is even.} \end{cases}$$

Observe that for $r = 1$ the corresponding element is quadratic and the expansion becomes ultimately periodic. The fact that the continued fraction expansion can be obtained explicitly for algebraic numbers of class IA implies the following result.

**Corollary 5.2.** *Let $\mu$ be a rational real number with $\mu \geq 2$ then there is an element $\Theta$ in $\mathcal{H}(q)$ such that $\nu(\Theta) = \mu$.*

## 6 A Particular Subset of $\mathcal{H}_2(q)$

As we noted above non-quadratic elements in $\mathcal{H}_2(q)$ appear to be exceptional. The first examples were given in [BS1], [BS2] and [MR]. In [L2] and later in a joint work with J.-J. Ruch, [LR1] and [LR2], we have searched for these elements with all partial quotients of degree one. For the theorem below, we need to introduce a new notation. If $(a_n)_{n\geq 1}$ is a sequence of polynomials in $\mathbb{F}_q[T]$ then, for $i \geq 1$ and $k \geq 0$, we define the polynomial $x_{i,k}$ as the numerator of the finite continued fraction $[0, a_{k+1}, a_{k+2}, \ldots, a_{k+i}]$.

**Theorem 6.1.** *Let $r = p^t$ with $t \geq 0$ and $l \geq r$ be two integers. Let $\Theta$ be an irrational element in $\mathbb{F}(q)$. Assume that $\Theta = [0, a_1, a_2, \ldots]$ with $\deg a_i = 1$ for $i \geq 1$. Then there exists $\epsilon \in \mathbb{F}_q^*$ such that*

$$(E) \quad \Theta = \frac{\epsilon x_l + x_{l-r}\Theta^r}{\epsilon y_l + y_{l-r}\Theta^r}$$

*if and only if there is a sequence* $(\epsilon_n)_{n\geq 0}$ *of elements in* $\mathbb{F}_q^*$, *with* $\epsilon_0 = 1$ *and* $\epsilon_1 = \epsilon$, *such that for* $n \geq 1$ *we have*

$$(S) \quad \begin{cases} \epsilon_{n-1} x_{2r,(n-2)r+l} = \epsilon_n a_n^r x_{r,(n-1)r+l} \\ \epsilon_{n+1} x_{r,(n-1)r+l} = \epsilon_{n-1} x_{r,(n-2)r+l} \end{cases}$$

In the trivial case $r = 1$, we meet with a particular case of Theorem 5.1. Indeed $\Theta$ is then quadratic and $(S)$ simply becomes $a_{l+k} = a_k \epsilon^{(-1)^{k+1}}$ for $k \geq 1$. In the general case the existence of a sequence $(a_n)_{n\geq 1}$ solution of $(S)$ will depend upon the choice of $\epsilon$ and of the first $l$ partial quotients. It is remarkable that for all $q$, $r$ and $l \geq r$ there is a trivial solution of $(S)$ given by $a_i = T$ for $i \geq 1$ and $\epsilon_i = 1$ for $i \geq 0$. The corresponding algebraic element is the one given as Example 1 above with $r = 1$ and is thus quadratic. We have given ( in [L2] for $q = 3$ and in [LR1] for general $q$) families of examples of sequences $(a_n)_{n\geq 1}$ satisfying $(S)$. As an illustration in odd characteristic, we give the following corollary [LR1]. Here, if $S$ is a finite sequence and $k \geq 0$ an integer, then $S^{[k]}$ denotes the empty sequence if $k = 0$ or else the sequence $S$ repeated $k$ times. Furt! her if $S_1$ and $S_2$ are two finite sequences then $S_1 \oplus S_2$ denotes the sequence obtained by concatenation.

**Corollary 6.2.** *Let* $q = p^s$ *with* $p \neq 2$ *and* $s \geq 1$. *Let* $\alpha, \beta \in \mathbb{F}_q^*$ *with* $\alpha + \beta = 2$. *Let* $k \geq 0$ *be an integer. Let* $\Theta \in \mathbb{F}(q)$ *be defined by the following continued fraction expansion*

$$\Theta = [0, T^{[k]}, \oplus_{i\geq 1}(T, (\alpha T, \beta T)^{[(q^i-1)/2]})^{[k+1]}\ ].$$

*Then* $\Theta$ *satisfies the algebraic equation*

$$y_k X^{q+1} - x_k X^q + (\alpha\beta)^{(q-1)/2} y_{q+k} X - (\alpha\beta)^{(q-1)/2} x_{q+k} = 0.$$

Clearly the complexity of the system $(S)$ in theorem 6.1 is growing with $r$. In the case of even characteristic, we can choose $r = 2$. By studying the simplest case where the partial quotients are all linear in $T$, we can prove the following corollary [LR2].

**Corollary 6.3.** *Let* $q = 2^s$ *with* $s \geq 1$ *and* $l \geq 2$ *be integers. Let* $\lambda_1, \lambda_2, \ldots, \lambda_l$ *and* $\epsilon$ *be given in* $\mathbb{F}_q^*$. *We consider the sequence* $(\lambda_i)_{i\geq 1}$ *defined recursively for* $n \geq 1$ *by*

$$\begin{cases} \lambda_{l+2n-1} = \lambda_n^2 \lambda_l^{-1} \epsilon^{(-1)^{n+1}} \\ \lambda_{l+2n} = \lambda_l. \end{cases}$$

*Let* $\Theta$ *be the irrational element in* $\mathbb{F}(q)$ *defined by the continued fraction expansion*

$$\Theta = [0, \lambda_1 T, \lambda_2 T, \ldots, \lambda_n T, \ldots].$$

*Then $\Theta$ satisfies the algebraic equation*

$$y_{l-2}\Theta^3 + x_{l-2}\Theta^2 + \epsilon y_l \Theta + \epsilon x_l = 0.$$

We conclude by making a last observation. Let us denote by $\mathcal{F}(q)$ the subset of $\mathcal{H}(q)$ containing all the elements satisfying an equation of type $(E)$ as defined in Theorem 6.1. From the subset $\mathcal{F}(q)$, using the three transformations mentioned at the end of section 4, we obtain a wider set of badly approximable algebraic elements (that is to say with bounded partial quotients). Does this set cover $\mathcal{H}_2(q)$? The answer is no if $q = 2$. In that case Baum and Sweet have described all the power series with partial quotients of degree one (see [BS2]). There are among them algebraic elements which are not of class I (see [L1] p. 225). On the other hand Baum and Sweet have given the example of a cubic element with bounded partial quotients (see [BS1] and [L3]). The case of characteristic 2 might be specific since then the existence of badly approximable elements comes from arguments of differential algebra (see [LdM2] p. 5). Consequently it is natural to a! sk : if the characteristic is different from 2, are there badly approximable algebraic elements which are not of class I ? This last question forces us to think of an open problem in number theory : are there badly approximable algebraic real numbers which are not quadratic ?

## References


[BS1] Baum L. and Sweet M., Continued fractions of algebraic power series in characteristic 2, Annals of Mathematics, 103(1976), 593–610.

[BS2] Baum L. and Sweet M., Badly approximable power series in characteristic 2, Annals of Mathematics, 105(1977), 573–580.

[L1] Lasjaunias A., A survey of diophantine approximation in fields of power series, Monatshefte für Mathematik, 130(2000), 211–229.

[L2] Lasjaunias A., Quartic power series in $\mathbb{F}_3((T^{-1}))$ with bounded partial quotients, Acta Arithmetica, XCV.1(2000), 49–59.

[L3] Lasjaunias A., Continued fractions for algebraic power series over a finite field, Finite Fields and their Applications, 5(1999), 46–56.

[LdM1] Lasjaunias A. and de Mathan B., Thue's Theorem in positive characteristic, Journal für die reine und angewandte Mathematik, 473(1996), 195–206.

[LdM2] Lasjaunias A. and de Mathan B., Differential equations and diophantine approximation in positive characteristic, Monatshefte für Mathematik, 128(1999), 1–6.

[LR1] Lasjaunias A. and Ruch J-J., Algebraic and badly approximable power series over a finite field, Finite Fields and their Applications, 8(2002), 91–107.

[LR2] Lasjaunias A. and Ruch J-J., Flat power series over a finite field, Journal of Number Theory, to appear.

[M] Mahler K., On a theorem of Liouville in fields of positive characteristic, Canadian Journal of Mathematics, 1(1949), 397–400.

[dM] de Mathan B. Approximation exponents for algebraic functions, Acta Arithmetica, LX.4(1992), 359–370.

[MR] Mills W. and Robbins D., Continued fractions for certain algebraic power series, Journal of Number Theory, 23(1986), 388–404.

[S] Schmidt W., On Continued fractions and diophantine approximation in power series fields, Acta Arithmetica, XCV.2(2000), 139–165.

[T] Thakur D., Diophantine approximation exponents and continued fractions for algebraic power series, Journal of Number Theory, 79(1999), 284–291.

[V] Voloch J-F., Diophantine approximation in positive characteristic, Periodica Mathematica Hungarica, 19.3(1988), 217-225.

# Linear Complexity and Polynomial Degree of a Function Over a Finite Field

Wilfried Meidl and Arne Winterhof

Institute of Discrete Mathematics, Austrian Academy of Sciences,
Sonnenfelsgasse 19/2, 1010 Vienna, Austria,
E-Mail: {wilfried.meidl, arne.winterhof}@oeaw.ac.at

**Abstract.** We compare the complexities of the polynomial representation and the periodic sequence representation of a function over a finite field in the complexity measures degree and linear complexity. We prove a sharp inequality describing the relation between degree and linear complexity. These investigations are motivated by results on some cryptographic functions. In particular, as an application of the above mentioned inequality we prove new lower bounds on the linear complexity of sequences related to the Diffie-Hellman mapping.

## 1 Introduction

One way functions are important topics in cryptography. For example the discrete logarithm is an attractive candidate for the inverse of a one way function. Various cryptographic protocols as the Diffie-Hellman key exchange depend on the intractability of the discrete logarithm (see e. g. [10, Chapter 3]). Unfortunately, there exists no exact definition for intractability of a function and we have to compensate this lack with several complexity measures. In the present paper we consider functions over finite fields and their representations as polynomials and as periodic sequences and compare the complexity measures degree and linear complexity.

Let $q$ be a prime power and fix an ordering $F_q = \{\xi_0, \ldots, \xi_{q-1}\}$ of the elements of the finite field $F_q$. A $q$-periodic sequence $(\sigma_n)$ of elements of $F_q$ can be represented by a uniquely determined polynomial $f(X) \in F_q[X]$ of degree at most $q-1$. Conversely, every polynomial $f(X) \in F_q[X]$ defines a unique $q$-periodic sequence over $F_q$. In other words, we have

$$\sigma_n = f(\xi_n) \in F_q \quad \text{for } 0 \le n < q \quad \text{and} \quad \sigma_{n+q} = \sigma_n \quad \text{for } n \ge 0. \tag{1}$$

Since $\xi^q = \xi$ for all $\xi$ in $F_q$ we may restrict ourselves to the case that the degree of $f(X)$ is at most $q-1$ in the sequel.

The *linear complexity* $L(\sigma_n)$ of the sequence $(\sigma_n)$ is the shortest positive integer $L$ such that there are constants $\gamma_1, \ldots, \gamma_L \in F_q$ satisfying

$$-\sigma_n = \gamma_1\sigma_{n-1} + \gamma_2\sigma_{n-2} + \ldots + \gamma_L\sigma_{n-L} \qquad \text{for all } n \ge L.$$

If $q = p$ is a prime, $F_p = \{0, 1, \ldots, p-1\}$, and $\deg(f) < p$ then it can easily be shown (see e. g. [2, Theorem 8]) that

$$L(\sigma_n) = \deg(f) + 1. \tag{2}$$

For arbitrary finite fields $F_q$ of order $q = p^r$ with an integer $r > 1$ the situation is different. For example in $F_4 = F_2(\rho) = \{0, 1, \rho, \rho + 1\}$ where $\rho$ is a zero of the polynomial $g(X) = X^2 + X + 1 \in F_2[X]$ the sequence $(\sigma_n)$ defined by the polynomial $f(X) = \rho X + X^2 \in F_4[X]$ satisfies the recurrence relation $\sigma_n = \sigma_{n-2}$ for $n \geq 2$ and we have $L(\sigma_n) = \deg(f) = 2$. On the other hand the sequence $(\sigma_n)$ defined by the polynomial $f(X) = X$ does not satisfy any recurrence relation of order at most 2 and we have $L(\sigma_n) = \deg(f) + 2 = 3$. (The sequence $(\sigma_n)$ satisfies the recurrence relation $\sigma_n = \sigma_{n-1} + \sigma_{n-2} + \sigma_{n-3}$ for $n \geq 3$.) In the present paper we investigate how far (2) holds true for polynomials $f(X)$ and sequences $(\sigma_n)$ over arbitrary finite fields defined by (1) where for $0 \leq n < q$ the element $\xi_n \in F_q$ is defined by

$$\xi_n = n_1\beta_1 + n_2\beta_2 + \ldots + n_r\beta_r \tag{3}$$

if

$$n = n_1 + n_2 p + \ldots + n_r p^{r-1} \quad \text{with } 0 \leq n_k < p \text{ for } 1 \leq k \leq r$$

for a fixed basis $\{\beta_1, \ldots, \beta_r\}$ of $F_q$ over $F_p$. (Note that $L(\sigma_n)$ depends on the ordering of $F_q$.)

After some preliminaries in Section 2 we prove the following extension of (2) in Section 3.

**Theorem 1.** *Let $f(X) \in F_q[X]$ be a polynomial of degree at most $q-1$ and $(\sigma_n)$ be the sequence defined by* (1) *and* (3). *Then we have*

$$(\deg(f) + 1 + p - q)\frac{q}{p} \leq L(\sigma_n) \leq (\deg(f) + 1)\frac{p}{q} + q - p$$

*or equivalently*

$$(L(\sigma_n) + p - q)\frac{q}{p} - 1 \leq \deg(f) \leq L(\sigma_n)\frac{p}{q} + q - p - 1.$$

Moreover, we show that the lower bound on $L(\sigma_n)$ is attained if $L(\sigma_n) = q - sq/p$ with $0 \leq s \leq 1$ and the upper bound on $L(\sigma_n)$ is attained if $\deg(f) = q - 1 - sq/p$ with $0 \leq s < p$.

These investigations are motivated by results on some cryptographic functions: the Diffie-Hellman mapping and the discrete logarithm [8,9]. In particular, we combine Theorem 1 and results of Shparlinski [12, Chapter 8] to obtain new bounds on the linear complexity of sequences related to the Diffie-Hellman problem in Section 4.

## 2 Preliminaries

We define the polynomial $S^q(X) \in F_q[X]$ by

$$S^q(X) = \sum_{n=0}^{q-1} \sigma_n X^n.$$

By [3, Lemma 8.2.1] the linear complexity $L(\sigma_n)$ is given by

$$L(\sigma_n) = q - \deg(\gcd(X^q - 1, S^q(X))) = q - v, \tag{4}$$

where $v$ denotes the multiplicity of 1 as zero of $S^q(X)$. If $S^q(1) \neq 0$, then $v$ is defined to be 0. In particular we have $L(\sigma_n) = q$ if and only if $S^q(1) \neq 0$. Put

$$f(X) = \sum_{j=0}^{q-1} \alpha_j X^j.$$

Then we have

$$S^q(1) = \sum_{\xi \in F_q} f(\xi) = \sum_{j=0}^{q-1} \alpha_j \sum_{\xi \in F_q} \xi^j = -\alpha_{q-1}$$

and thus

$$L(\sigma_n) = q \quad \text{if and only if} \quad \deg(f) = q - 1. \tag{5}$$

To estimate the multiplicity $v$ we evaluate the *Hasse-Teichmüller derivatives* (see [5], [13])

$$S^q(X)^{(t)} = \sum_{n=0}^{q-1} \binom{n}{t} \sigma_n X^{n-t}$$

in 1. For the *dual basis* $\{\delta_1, \ldots, \delta_r\}$ of $\{\beta_1, \ldots, \beta_r\}$ (see [6, Definition 2.30]), i. e.

$$\mathrm{Tr}\,(\delta_k \beta_j) = \begin{cases} 1 \text{ if } k = j, \\ 0 \text{ otherwise,} \end{cases}$$

we have for $0 \leq n < q$,

$$n = \sum_{k=1}^{r} \mathrm{Tr}\,(\delta_k \xi_n)\, p^{k-1}, \tag{6}$$

where $\mathrm{Tr}(X) = \sum_{k=1}^{r} X^{p^{k-1}}$ is the (absolute) trace of $F_q$ with the convention $0 \leq \mathrm{Tr}(\xi) < p$ for $\xi \in F_q$. By (6) and Lucas' congruence (see e. g. [4] or [7]) for $t = t_1 + \ldots + t_r p^{r-1}$, $0 \leq t_i < p$ we have

$$\binom{n}{t} \equiv \binom{\mathrm{Tr}\,(\delta_1 \xi_n)}{t_1} \cdots \binom{\mathrm{Tr}\,(\delta_r \xi_n)}{t_r} \bmod p.$$

Thus

$$S^q(1)^{(t)} = \sum_{\xi \in F_q} \binom{\mathrm{Tr}\,(\delta_1 \xi)}{t_1} \cdots \binom{\mathrm{Tr}\,(\delta_r \xi)}{t_r} f(\xi). \tag{7}$$

## 3 Bounds on the Linear Complexity

In this section we prove Theorem 1 as an easy consequence of the following Lemmas and show that the inequalities of Theorem 1 are best possible if we don't restrict the range of $\deg(f)$ or of $L(\sigma_n)$.

**Lemma 1.** *Let* $f(X) = \sum_{j=0}^{q-2} \alpha_j X^j \in F_q[X]$. *If* $L(\sigma_n) = q-s$ *with* $0 \le s < p$ *then some coefficients* $\alpha_{q-1-p^{m_1}-\ldots-p^{m_s}}$ *of* $f(X)$ *with* $0 \le m_1, \ldots, m_s < r$ *are nonzero.*

*Proof.* For $0 \le t < s$ we have $S^q(1)^{(t)} = 0$ and $S^q(1)^{(s)} \neq 0$ by (4). Since the polynomials $p_0(X) = 1$ and

$$p_t(X) = \frac{1}{t!} X(X-1)\cdots(X-t+1) \in F_p[X], \quad 1 \le t \le s,$$

form a basis of the linear space of polynomials of degree at most $s$ we can write $X^s/s!$ as linear combination of the form

$$\frac{X^s}{s!} = \sum_{t=0}^{s} c_t p_t(X) \quad \text{with } c_s = 1.$$

Hence by (7),

$$\begin{aligned}
S^q(1)^{(s)} &= \sum_{\xi \in F_q} \binom{\mathrm{Tr}(\delta_1 \xi)}{s} f(\xi) = \sum_{\xi \in F_q} p_s(\mathrm{Tr}(\delta_1 \xi)) f(\xi) \\
&= \sum_{\xi \in F_q} \left( \frac{\mathrm{Tr}(\delta_1 \xi)^s}{s!} - \sum_{t=0}^{s-1} c_t p_t(\mathrm{Tr}(\delta_1 \xi)) \right) f(\xi) \\
&= \sum_{\xi \in F_q} \frac{\mathrm{Tr}(\delta_1 \xi)^s}{s!} f(\xi) - \sum_{t=0}^{s-1} c_t S^q(1)^{(t)} = \sum_{\xi \in F_q} \frac{\mathrm{Tr}(\delta_1 \xi)^s}{s!} f(\xi) \\
&= \frac{1}{s!} \sum_{j=0}^{q-2} \alpha_j \sum_{\xi \in F_q} \mathrm{Tr}(\delta_1 \xi)^s \xi^j = \frac{1}{s!} \sum_{j=0}^{q-2} \alpha_j \sum_{\xi \in F_q} \left( \sum_{m=0}^{r-1} (\delta_1 \xi)^{p^m} \right)^s \xi^j \\
&= \frac{1}{s!} \sum_{m_1, \ldots, m_s = 0}^{r-1} \delta_1^{p^{m_1} + \ldots + p^{m_s}} \sum_{j=0}^{q-2} \alpha_j \sum_{\xi \in F_q} \xi^{p^{m_1} + \ldots + p^{m_s} + j} \\
&= -\frac{1}{s!} \sum_{m_1, \ldots, m_s = 0}^{r-1} \delta_1^{p^{m_1} + \ldots + p^{m_s}} \alpha_{q-1-p^{m_1}-\ldots-p^{m_s}} \neq 0, \qquad (8)
\end{aligned}$$

which proves the assertion of the lemma. □

**Lemma 2.** *Let* $0 \le s < p$ *and* $f(X) = \sum_{j=0}^{q-2} \alpha_j X^j \in F_q[X]$ *with*

$$\alpha_{q-1-p^{m_1}-\ldots-p^{m_s}} \neq 0 \quad \textit{for some } 0 \le m_i < r,\ 1 \le i \le s.$$

*Then*

$$L(\sigma_n) \geq q - sq/p.$$

*Proof.* We assume $L(\sigma_n) < q - sq/p$ and thus $S^q(1)^{(t)} = 0$ for $0 \leq t \leq sq/p$. Similarly as in the proof of the previous Lemma we can write for $0 \leq t \leq sq/p$,

$$S^q(1)^{(t)} = \sum_{\xi \in F_q} \frac{\mathrm{Tr}(\delta_1 \xi)^{t_1}}{t_1!} \cdots \frac{\mathrm{Tr}(\delta_r \xi)^{t_r}}{t_r!} f(\xi) = 0, \tag{9}$$

where $t = t_1 + \ldots + t_r p^{r-1}$ with $0 \leq t_i < p$ for $1 \leq i \leq r$. By the linearity of the trace for all $\alpha = \sum_{k=1}^{r} a_k \delta_k \in F_q$ with $a_k \in F_p$ we get

$$\begin{aligned}
&\sum_{\xi \in F_q} \mathrm{Tr}(\alpha\xi)^s f(\xi) = \sum_{\xi \in F_q} \left( \sum_{k=1}^{r} a_k \mathrm{Tr}(\delta_k \xi) \right)^s f(\xi) \\
&= \sum_{\xi \in F_q} \sum_{k_1, \ldots, k_s = 1}^{r} a_{k_1} \cdots a_{k_s} \mathrm{Tr}(\delta_{k_1} \xi) \cdots \mathrm{Tr}(\delta_{k_s} \xi) f(\xi) \\
&= \sum_{k_1, \ldots, k_s = 1}^{r} a_{k_1} \cdots a_{k_s} \sum_{\xi \in F_q} \mathrm{Tr}(\delta_{k_1} \xi) \cdots \mathrm{Tr}(\delta_{k_s} \xi) f(\xi).
\end{aligned} \tag{10}$$

By (9) the polynomial

$$H_s(X) := \sum_{\xi \in F_q} \mathrm{Tr}(\xi X)^s f(\xi)$$

has $q$ zeros. Since $\deg(H_s(X)) \leq sq/p < q$ we have $H_s(X) \equiv 0$. On the other hand analogously to the proof of the previous lemma we get

$$\begin{aligned}
H_s(X) &= \sum_{j=0}^{q-2} \alpha_j \sum_{\xi \in F_q} \mathrm{Tr}(\xi X)^s \xi^j \\
&= \sum_{j=0}^{q-2} \alpha_j \sum_{\xi \in F_q} \left( \sum_{m=0}^{r-1} (\xi X)^{p^m} \right)^s \xi^j \\
&= \sum_{m_1, \ldots, m_s = 0}^{r-1} \sum_{j=0}^{q-2} \alpha_j \sum_{\xi \in F_q} \xi^{p^{m_1} + \ldots + p^{m_s} + j} X^{p^{m_1} + \ldots + p^{m_s}} \\
&= - \sum_{m_1, \ldots, m_s = 0}^{r-1} \alpha_{q-1-p^{m_1} - \ldots - p^{m_s}} X^{p^{m_1} + \ldots + p^{m_s}} \\
&= - \sum_{j=0}^{q-1} k_{q-1-j} \alpha_j X^j \equiv 0
\end{aligned} \tag{11}$$

with

$$k_j = \begin{cases} 0 & \text{if } j_1 + \ldots + j_r \neq s, \\ \binom{s}{j_1}\binom{s-j_1}{j_2}\binom{s-j_1-j_2}{j_3}\cdots\binom{s-j_1-\ldots-j_{r-1}}{j_r} & \text{if } j_1 + \ldots + j_r = s, \end{cases}$$

where $j = j_1 + \ldots + j_r p^{r-1}$ with $0 \le j_i < p$ for $1 \le i \le r$. Since $k_j \neq 0$ if and only if $j_1 + \ldots + j_r = s$ we get $\alpha_{q-1-p^{m_1}-\ldots-p^{m_s}} = 0$ for all $0 \le m_1, \ldots, m_s < r$, which contradicts the assumption of the Lemma. □

*Proof of Theorem 1.* The upper bound on $L(\sigma_n)$ is trivial if $L(\sigma_n) \le q - p$ and we may suppose that

$$L(\sigma_n) = q - s \quad \text{with } 0 \le s < p.$$

By Lemma 1 we have

$$\deg(f) \ge q - 1 - s\frac{q}{p}$$

and thus

$$L(\sigma_n) \le (\deg(f) + 1)\frac{p}{q} + q - p.$$

The lower bound on $L(\sigma_n)$ is trivial if $\deg(f) \le q - 1 - p$ and we may suppose

$$\deg(f) = q - 1 - s \quad \text{with } 0 \le s < p.$$

By Lemma 2 we have

$$L(\sigma_n) \ge q - s\frac{q}{p} = (\deg(f) + 1 + p - q)\frac{q}{p}.$$

The second inequality of Theorem 1 is obviously equivalent to the first one. □

The first inequality of Theorem 1 is only nontrivial if $\deg(f)$ is at least $q - p$ but Lemma 2 yields nontrivial lower bounds on $L(\sigma_n)$ for many other degrees. For example we have the following result.

**Corollary 1.** *If*

$$\deg(f) \ge q - 2p + 1$$

*then we have*

$$L(\sigma_n) \ge \frac{q}{p}.$$

Corollary 1 is best possible in the sense that $L(\sigma_n)$ can fall below the benchmark $q/p$ if $\deg(f) \le q - 2p$.

*Example.* Consider $F_9 = F_3(\alpha)$ with $\alpha^2 + 1 = 0$ and the basis $\{\beta_1, \beta_2\} = \{1, \alpha\}$. The sequence $(\sigma_n)$ defined by the polynomial $f(X) = X^3 + X$ satisfies $\sigma_n = -\sigma_{n-1} - \sigma_{n-2}$, $n \ge 2$, and we have $L(\sigma_n) = 2$.

Lemma 2 yields also a relation between $L(\sigma_n)$ and the number of nonzero coefficients of $f(X)$.

**Corollary 2.** *Let $s$ be an integer with $0 \leq s < p$. If $L(\sigma_n) < q - sq/p$ then the weight of $f(X)$ is at most $q - \binom{r+s}{s}$.*

*Proof.* ¿From (5) and Lemma 2 we know that $L(\sigma_n) < q - sq/p$ implies $\alpha_{q-1} = 0$ and $\alpha_{q-1-p^{m_1}-p^{m_2}-\ldots-p^{m_t}} = 0$ for all $0 \leq m_i < r$, $1 \leq i \leq t$, for $1 \leq t \leq s$. Hence the weight of $f(X)$ is at most $q - N$, where $N$ is the number of integers $i$ of the form $i = q - 1 - j$ with $j_1 + \ldots + j_r \leq s$, where $j = j_1 + \ldots + j_r p^{r-1}$ with $0 \leq j_i < p$ for $1 \leq i \leq r$. It can easily be verified that $N = \binom{r+s}{s}$. □

Now we show that the lower bound on $L(\sigma_n)$ in Theorem 1 is sharp.

**Corollary 3.** *If $L(\sigma_n) = q - sq/p$ with $0 \leq s \leq 1$ then we have*

$$L(\sigma_n) = (\deg(f) + 1 + p - q)\frac{q}{p}.$$

*Proof.* For $s = 0$ the result is equivalent to (5). For $s = 1$ Lemma 2 yields $\deg(f) \leq q - 2$. Since (9) is valid for $0 \leq t < q/p$ from (10) and (11) we know that

$$H_1(X) = -\sum_{m=0}^{r-1} \alpha_{q-1-p^m} X^{p^m}$$

has $q/p$ distinct zeros, namely all elements $\alpha = \sum_{k=1}^{r} a_k \delta_k$ with $a_r = 0$. Since $\deg(H_1) \leq q/p$ all these zeros have multiplicity 1. Hence the first derivative of $H_1(X)$ is not the zero polynomial, i. e.

$$H_1(X)' = -\sum_{m=0}^{r-1} \alpha_{q-1-p^m} p^m X^{p^m-1} = -\alpha_{q-2} \neq 0$$

and thus $\deg(f) \geq q - 2$. Now we have

$$\deg(f) = q - 2 = L(\sigma_n)\frac{p}{q} + q - p - 1$$

which is equivalent to the assertion. □

Now we prove that the upper bound on $L(\sigma_n)$ in Theorem 1 is sharp, as well.

**Corollary 4.** *If $\deg(f) = q - 1 - sq/p$ with $0 \leq s < p$ then we have*

$$L(\sigma_n) = (\deg(f) + 1)\frac{p}{q} + q - p.$$

*Proof.* For $s = 0$ the result follows by (5). For $s \geq 1$ the assumption $L(\sigma_n) = q - s'$ with $0 \leq s' < s$ would imply $\deg(f) \geq q - 1 - s'q/p > q - 1 - sq/p$ by Lemma 1. Hence we have $L(\sigma_n) \leq q - s$. By (8) we have

$$S^q(1)^{(s)} = -\frac{1}{s!}\delta_1^{sq/p}\alpha_{q-1-sq/p} \neq 0$$

and thus $L(\sigma_n) \geq q - s$. Hence,

$$L(\sigma_n) = q - s = (\deg(f) + 1)\frac{q}{p} + q - p.$$

□

*Remark.* For a positive integer $m$ the method in the proof of Theorem 1 can be applied to $q^m$-periodic sequences, i. e.

$$\sigma_{n_1+n_2q+\ldots+n_mq^{m-1}} = f(\xi_{n_1}, \xi_{n_2}, \ldots, \xi_{n_m}), \quad 0 \leq n_i < q,\ 1 \leq i \leq m,$$

where $f(X_1, \ldots, X_m) \in F_q[X_1, \ldots, X_m]$ (cf. [2, Theorem 8]). The method yields also results on the linear complexity of $q(q-1)$ periodic sequences $(\sigma_n)$ of the form $\sigma_n = \sum_i f_i(\xi_n)\lambda_i^n$, $n \geq 0$, where $f_i(X) \in F_q[X]$ and $\lambda_i \in F_q$.

## 4 Applications

The Diffie-Hellman problem in $F_q$ is the following: Let $\gamma$ be a primitive element of $F_q$ and $\gamma^i, \gamma^j$ be nonzero elements of $F_q$. Find $\gamma^{ij}$ without knowing $i$ and $j$.

Since

$$\gamma^{2ij} = \gamma^{(i+j)^2}\gamma^{-i^2}\gamma^{-j^2}$$

and square roots in finite fields can efficiently be calculated (see e. g. [1, Chapter 7]) in order to investigate the Diffie-Hellman problem we may consider the polynomial $h(X) \in F_q[X]$ of degree at most $q-2$ defined by

$$h\left(\gamma^i\right) = \gamma^{i^2} \qquad \text{for } 0 \leq i \leq q-2.$$

**Corollary 5.** *Let $S \subseteq \{1, \ldots, q-1\}$ be of cardinality $|S| = q - 1 - s$ and $(\sigma_n)$ be a sequence satisfying*

$$\sigma_n = h(\xi_n) \quad \textit{for } n \in S.$$

*Then we have*

$$L(\sigma_n) \geq q - (2s+3)\frac{q}{p}.$$

*If $s \leq p-3$ then we have*

$$L(\sigma_n) \geq \frac{q}{p}.$$

*Proof.* By [12, Theorem 8.1] a polynomial $f(X) \in F_q[X]$ satisfying

$$f\left(\xi_n\right) = h(\xi_n) \qquad \text{for } n \in S,$$

has degree at least $q - 2s - 4$. The first result follows from Theorem 1 and the second from Corollary 1. □

*Remark.* For $s = 0$ we have

$$\deg(f) = \begin{cases} q-3 & \text{if } q \equiv 1 \bmod 4, \\ q-2 & \text{otherwise,} \end{cases}$$

by [8, Theorem 1] and thus

$$L(\sigma_n) \geq \begin{cases} q - 2q/p & \text{if } q \equiv 1 \bmod 4, \\ q - q/p & \text{otherwise.} \end{cases}$$

The *discrete logarithm* (or *index*) of a nonzero element $\xi \in F_q$ to the base $\gamma$, denoted $\mathrm{ind}_\gamma(\xi)$, is the unique integer $l$ with $0 \leq l \leq q-2$ such that $\xi = \gamma^l$. Obviously, the Diffie-Hellman problem depends on the intractability of the discrete logarithm, which can be identified with the polynomial $P_\gamma(X) \in F_q[X]$ defined by

$$P_\gamma(\gamma^n) = \xi_n \quad \text{for } 0 \leq n \leq q-2 \quad \text{and} \quad P_\gamma(0) = \xi_{q-1}.$$

(The coefficients of $P_\gamma(X)$ in the special case of a polynomial basis were determined by Mullen and White [11].) For $S \subset \{1, \ldots, q-1\}$ with $|S| = q-1-s$ a polynomial $f(X) \in F_q[X]$ with

$$f(\xi_n) = P_\gamma(\xi_n) \quad \text{for } n \in S$$

must satisfy

$$\deg(f) \geq q - \frac{q}{p} - 1 - 2s$$

by [14, Theorem 1]. Hence, for $q = p^2$ the linear complexity of a $q$-periodic sequence $(\sigma_n)$ satisfying

$$\sigma_n = P_\gamma(\xi_n) \quad \text{for } n \in S$$

is at least $p$ if $s \leq (p-3)/2$ by Corollary 1.

*Remark.* For $s = 0$ and arbitrary $q$ a direct application of the method in the proof of Theorem 1 yields $L(\sigma_n) \geq q - q/p$ (see [9]).

## References

1. E. Bach and J. O. Shallit, *Algorithmic number theory 1: Efficient algorithms.* Cambridge: MIT Press, 1996.
2. S. R. Blackburn, T. Etzion, and K. G. Paterson, Permutation polynomials, de Bruijn sequences, and linear complexity, *J. Combin. Theory Ser. A* **76** (1996), 55–82.
3. T. W. Cusick, C. Ding, and A. Renvall, *Stream ciphers and number theory.* Amsterdam: North-Holland, 1998.

4. A. Granville, Arithmetic properties of binomial coefficients. I. Binomial coefficients modulo prime powers, in *Organic mathematics*, Burnaby, BC, 1995, 253–276, *CMS Conf. Proc.* **20**, Amer. Math. Soc., Providence: RI, 1997.
5. H. Hasse, Theorie der höheren Differentiale in einem algebraischen Funktionenkörper mit vollkommenem Konstantenkörper bei beliebiger Charakteristik, *J. Reine Angew. Math.* **175** (1936), 50–54.
6. R. Lidl and H. Niederreiter, *Finite Fields*. Reading, MA: Addison-Wesley, 1983.
7. R. J. McIntosh, A generalization of a congruential property of Lucas, *Amer. Math. Monthly* **99** (1992), 231–238.
8. W. Meidl and A. Winterhof, A polynomial representation of the Diffie-Hellman mapping, Preprint.
9. W. Meidl and A. Winterhof, Lower bounds on the linear complexity of the discrete logarithm in finite fields, *IEEE Trans. Inf. Th.* **47** (2001), 2807–2811.
10. A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, *Handbook of applied cryptography*. Boca Raton: CRC Press, 1997.
11. G. L. Mullen and D. White, A polynomial representation for logarithms in $GF(q)$, *Acta Arith.* **47** (1986), 255–261.
12. I. E. Shparlinski, *Number theoretic methods in cryptography*. Basel: Birkhäuser, 1999.
13. O. Teichmüller, Differentialrechnung bei Charakteristik $p$, *J. Reine Angew. Math.* **175** (1936), 89–99.
14. A. Winterhof, Polynomial interpolation of the discrete logarithm, *Des. Codes Crytogr.* **25** (2002), 63–72.

# Primitive Roots in Cubic Extensions of Finite Fields

Donald Mills[1] and Gavin McNay[2]

[1] Department of Mathematical Sciences
U.S. Military Academy, West Point, NY 10996
ad3943@usma.edu
[2] 20 Siddons Road, Tottenham
London N17, England, UK
gavin.mcnay@nomura.co.uk

**Abstract.** N. Katz has shown that the absolute value of sums of the form $\sum_{b \in \mathbf{F}_q} \chi(\theta + b)$, $\mathbf{F}_q$ the finite field of $q$ elements, $\chi$ a nontrivial multiplicative character of $\mathbf{F}_{q^n}$, and $\theta$ a $\mathbf{F}_q$-generator of $\mathbf{F}_{q^n}$, is bounded from above by $(n-1)\sqrt{q}$. We use this result in conjunction with a sieve due to S. Cohen to show the following for $n = 3$: For any prime power $q$ and any $\mathbf{F}_q$-generator $\theta$ of $\mathbf{F}_{q^3}$, there exists a primitive element of the form $a\theta + b \in \mathbf{F}_{q^3}$ for some $a, b \in \mathbf{F}_q$, $a \neq 0$. We discuss an application of these primitive sums in their use as pseudorandom vector generators, and conclude by discussing the harder problem of guaranteeing the existence of such roots when $a$ is forced to be 1.

## 1 Introduction

We are interested in the following problem: Given a prime power $q$ and positive integer $n$, prove (or disprove) that for any element $\theta$ such that $\mathbf{F}_q(\theta) = \mathbf{F}_{q^n}$ there exist elements $a$, $b \in \mathbf{F}_q$ such that $a\theta + b$ is a primitive root of $\mathbf{F}_{q^n}$. A counting argument shows that the problem is trivial for $n = 2$; Cohen [5] addressed a non-trivial version of this problem (comparable to $a = 1$) for $n = 2$ with the following result.

**Theorem 1.** *Let $\alpha$, $\theta$ belong to $\mathbf{F}_{q^2}$ with $\alpha \neq 0$ and $\theta \notin \mathbf{F}_q$. Then there exists a primitive root of $\mathbf{F}_{q^2}$ of the form $\alpha(\theta + b)$ for some $b \in \mathbf{F}_q$.*

Other results of Cohen [6–8] show as well that such primitive sums do exist although Mullen [16], when commenting about these, limits himself to the case $a = \pm 1, b = 1$ and $\theta$ primitive, where $\mathbf{F}_q(\theta) = \mathbf{F}_{q^n}$. Here the problem is made dependent on $\theta$ but no mention of how to find such an element is given. We submit that it is better to fix $\theta$ and consider the problem as depending on $a, b$. Indeed in the vein of Theorem 1, we prove the following for cubic extensions.

**Theorem 2.** *For any prime power $q$ and any element $\theta$ such that $\mathbf{F}_q(\theta) = \mathbf{F}_{q^3}$, there exist elements $a$, $b \in \mathbf{F}_q$ such that $a\theta + b$ is a primitive root of $\mathbf{F}_{q^3}$.*

Following the proof of Theorem 2, we give an application of these primitive sums, specifically we demonstrate their use as pseudorandom vector generators. Niki [18,19] showed how primitive roots of the form $a\theta + b$, for some $\mathbf{F}_p$-generator $\theta \in \mathbf{F}_{p^n}$ and $a, b \in \mathbf{F}_p$, can be used to construct a pseudorandom vector generator. (It should be noted, however, that Niki does not show the existence of such primitive sums. Niederreiter [17] discusses the statistical properties of this generator and its generalization.) After showing the existence of these primitive sums, we familiarize the reader with Niki's generator and present our special case of it, namely for cubic extensions.

The more difficult problem of determining whether one can always find a primitive root of the form $\theta + b$, $\theta$ an $\mathbf{F}_q$-generator of $\mathbf{F}_{q^n}$ and $b \in \mathbf{F}_q$, is then discussed. Specifically, we show how Katz's bound and the Cohen sieve can be used to give an explicit bound that assures the existence of primitive elements of the desired form for $q \geq q_0$, where $q_0$ is specified, and on the basis of our sieving and other computational work we are able to give a conjecture that parallels Theorem 2.

We note that the above two questions about $a\theta + b$ and $\theta + b$ being primitive are equivalent to: Given an irreducible cubic $f(x)$, can we produce a primitive polynomial of the form $f(ax + b)$ or $f(x + b)$? This equivalence will be exploited as necessary throughout the paper.

The machinery used in our proofs consists of a character sum analysis and an application of a sieving method. We use a bound on sums of the form

$$\sum_{b \in \mathbf{F}_q} \chi(\theta + b)$$

where $\chi$ is a non-trivial multiplicative character of $\mathbf{F}_{q^n}$ and here $\theta$ is an $\mathbf{F}_q$-generator of $\mathbf{F}_{q^n}$. N. Katz [14] has bounded these sums in a more general setting, and as such his proof uses sophisticated properties of the $l$-adic cohomology of schemes over $\mathbf{F}_q$, of the kind employed by Deligne [10,11] in his proof of Weil's conjectures for general varieties defined over $\mathbf{F}_q$. Specifically, Katz shows that

$$\left| \sum_{b \in \mathbf{F}_q} \chi(\theta + b) \right| \leq (n-1)\sqrt{q}. \tag{1}$$

This bound is, like the famous Weil bound, the best possible bound provided there is no *a priori* knowledge of the sum. We will use this bound in conjunction with a sieving method due to S. Cohen to prove Theorem 2.

The history of the problem of finding primitive sums is a long one, dating back to Davenport's work in the late 1930's. Specifically, in the general case of a degree $n$ extension, Davenport [9] discussed the asymptotics of the $\theta + b$ problem, and then Carlitz [2] generalized some of Davenport's ideas. Shparlinski and Perel'muter [20] used ideas similar to ours to derive the asymptotics and also to get an upper bound of $\sqrt{q}$ on the size of the least such $b$. So far as

the applications are concerned, Chung used (1) to prove results concerning the diameter of a particular class of graphs.

We give the pertinent definitions and notation in Section 2, prove Theorem 2 in Section 3, present the pseudo-random vector generation application in Section 4, and discuss the $\theta + b$ problem in Section 5.

## 2 Definitions and Notation

Throughout $\mathbf{F}_q$ will denote the finite field of $q$ elements and $\theta$ a (fixed) $\mathbf{F}_q$-generator of $\mathbf{F}_{q^n}$, i.e. $\mathbf{F}_q(\theta) = \mathbf{F}_{q^n}$. Here $n \geq 3$ since, as noted, Cohen has resolved the $n = 2$ case.

By saying that an element $x$ is *no kind of dth power*, we mean that if $x = y^k$ with $k \mid d$ then we must have $k = 1$. With this definition in hand, for $e \mid q^n - 1$ we define $N_\theta(e)$ to be the number of elements of the form $a\theta + b$, $a$, $b \in \mathbf{F}_q$ with $a \neq 0$ that are no kind of $e$th power. Thus, $N_\theta(q^n - 1)$ denotes the number of primitive roots of the form $a\theta + b$. Likewise one can define $N'_\theta(e)$ to be the number of elements of the form $\theta + b$, $b \in \mathbf{F}_q$, that are no kind of $e$th power.

Let $Z$ denote the ring of integers. Let $\mu$, $\phi$ denote the Möbius and Euler phi functions, respectively, and let $\sum_{\psi_d}$ denote the sum over all $\phi(d)$ multiplicative characters of $\mathbf{F}_{q^n}$ with exact order $d$. Further let $\omega(m)$ denote the number of distinct prime factors of $m$. We let $(a, b)$ denote the greatest common divisor of $a$ and $b$ and $\{a, b\}$ their least common multiple.

Using a well-known expression due to Vinogradov (see for example Section 7.5 in [13]), we can obtain the following characteristic function for $x$ being no kind of $e$th power, $e \mid q^n - 1$, namely

$$\frac{\phi(e)}{e} \sum_{d \mid e} \frac{\mu(d)}{\phi(d)} \sum_{\psi_d} \psi_d(x). \tag{2}$$

For a multiplicative character $\psi$ of $\mathbf{F}_{q^n}$ we let

$$S_\theta(\psi) = \sum_{a,b \in \mathbf{F}_q, a \neq 0} \psi(a\theta + b), \tag{3}$$

$$S'_\theta(\psi) = \sum_{b \in \mathbf{F}_q} \psi(\theta + b). \tag{4}$$

With $\psi$ as above we have

$$\begin{aligned} S_\theta(\psi) &= \sum_{a \in \mathbf{F}_q^*} \psi(a) \sum_{b \in \mathbf{F}_q} \psi(\theta + ba^{-1}) \\ &= \sum_{a \in \mathbf{F}_q^*} \psi(a) S'_\theta(\psi), \end{aligned} \tag{5}$$

which is equal to $(q-1)S'_\theta(\psi)$ when the order of $\psi$ divides $\frac{q^n-1}{q-1}$ and is equal to zero otherwise (Section 5.1 of [15]). From this we conclude that in order to gain the correct asymptotic result for $N_\theta(q^n-1)$ we need only bound $S'_\theta(\psi) = S(\psi)$. However, we already have the bound in hand, thanks to Katz, as (1) shows that $|S(\psi)| \leq (n-1)\sqrt{q}$.

## 3 The Main Result

We now prove Theorem 2. Using Vinogradov's result, Katz's bound, and the fact that the number of square-free divisors of an integer $m$ is $2^{\omega(m)}$, we have

$$\begin{aligned} N_\theta(e) &= \frac{\phi(e)}{e} \sum_{d|e} \frac{\mu(d)}{\phi(d)} \sum_{\psi_d} S_\theta(\psi_d) \\ &\geq (q-1)\frac{\phi(e)}{e}\left\{q - ((n-1)\sqrt{q})(2^{\omega((e,Q))}-1)\right\} \end{aligned} \tag{6}$$

where $Q = (q^n-1)/(q-1)$. Thus only divisors of $q^n-1$ that are also divisors of $Q$ contribute to the bound, and we will use this to our advantage. For $n=3$ we have $Q = q^2+q+1$; let $p>3$ be a prime, then $x^2+x+1 \equiv 0$ is solvable in $\mathbf{F}_p$ if and only if $-3$ is a square in $\mathbf{F}_p$, i.e. when $\left(\frac{-3}{p}\right) = 1$ or equivalently when $\left(\frac{p}{3}\right) = 1$, that is precisely when $p \equiv 1 \pmod 6$. Thus the only prime factors of $q^2+q+1$ are 3 or primes $p \equiv 1 \pmod 6$. Observe from (6) that, for $n=3$, $N>0$ whenever $q > 2^{2(\omega(Q)+1)}$, and from this it is easily determined that $N>0$ for $\omega(Q) \geq 10$ or $q > 8736506$. We aim to greatly reduce the amount of work necessary to prove Theorem 2, thus in order to use (6) to our greatest advantage, we require the following sieving technique, due to Cohen [5]. The proof below mirrors that of Proposition 6.1 of [3] with $r=2$, $m_1 = e$, $m_2 = f$, $m = \{m_1, m_2\}$, and $m_0 = (e,f)$.

**Lemma 1.** *Let $e$, $f \mid q^n-1$ for positive integer $n$. Then if $N_a$ denotes the number of elements of $\mathbf{F}_{q^n}$ that are no kind of $a$th power, we have*

$$N_{\{e,f\}} \geq N_e + N_f - N_{(e,f)}. \tag{7}$$

*Proof.* Let $U_a$ denote the set of elements that are not any kind of $a$th power. As $U_e \cap U_f = U_{lcm(e,f)}$ and $U_e \cup U_f \subseteq U_{gcd(e,f)}$, (7) follows by examination of cardinalities.

Since only divisors of $Q$ contribute to the bound in (6), it suffices to use (7) with $\{e,f\} = Q$ to prove Theorem 2. If $ef = q^2+q+1$ and $(e,f)=1$ then from the definition of $S_\theta(\psi)$ as well as (6) and (7) we have $N>0$ when

$$\frac{(q-1)\phi(e)}{e}(q-(2\sqrt{q})(2^{\omega(e)}-1)) + \frac{(q-1)\phi(f)}{f}(q-(2\sqrt{q})(2^{\omega(f)}-1)) - q(q-1) > 0.$$

This can be rewritten to give

$$q > 4\left(\frac{\frac{\phi(e)}{e}(2^{\omega(e)}-1)+\frac{\phi(f)}{f}(2^{\omega(f)}-1)}{\frac{\phi(e)}{e}+\frac{\phi(f)}{f}-1}\right)^2, \tag{8}$$

and so it is our goal to, for all values of $\omega(q^2+q+1)$ greater than some value $\omega_0$, satisfy (8).

To apply this we consider $q^2+q+1 = q_1 \dots q_r$, where $q_i$ are members of $\{3, p \ : \ p \equiv 1 \pmod 6\}$. (We let $p_i$ be the $i$th member of this set.) We then take $ef = q_1 \dots q_r$ with $(e, f) = 1$; usually, $e = q_1 \dots q_k$ and $f = q_{k+1} \dots q_r$ where $k = \left[\frac{r+1}{2}\right]$, $[a]$ denoting the least integer less than or equal to $a$, provides the best results. That is, such a factorization is most likely to minimize the right-hand side of (8). Clearly the worst possible case is when $q_i = p_i$ for $1 \leq i \leq r$, that is, if we can satisfy (8) for a given value of $r = \omega(q^2+q+1)$ using only the first $r$ primes $p$ of the form either $p = 3$ or $p \equiv 1 \mod 6$ then any $q$ with $\omega(Q) = r$ will satisfy (8) as well. Using these values we can get a bound $q$ for when $N > 0$ for a fixed value of $r = \omega(q^2+q+1)$, with results given in the table below. The first column gives the value of $r$. The second column gives the minimum $q$ such that $Q - A_r > 0$ where $A_r$ is the product of the first $r$ primes of the form $p = 3$ or $p \equiv 1 \mod 6$, and the third column gives the minimum value of the right-hand side of (8) when using the technique described above. By virtue of the manner in which we build the table, it is clear that if there is an $r = r_0$ for which the second column value exceeds the third column value, then the second column value will be greater than the third column value for all $r \geq r_0$.

| $\omega(q^2+q+1)$ | $q \geq$ | $N > 0$ when $q >$ |
|---|---|---|
| 1 | 1 | 4 |
| 2 | 4 | 33.9 |
| 3 | 16 | 113.8 |
| 4 | 71.5 | 378.5 |
| 5 | 400.5 | 841.2 |
| 6 | 2438.7 | 2244.2 |

The prime power values for which Theorem 2 may not hold, then, are precisely those $q$ for which $\omega(Q) < 6$ and which lie between the second and third column values. The list of such $q$ is enumerated below.

1. $\omega(Q) = 1$: $q = 2, 3$
2. $\omega(Q) = 2$: $q = 4, 7, 9, 11, 13, 19, 23, 29, 31$
3. $\omega(Q) = 3$: $q = 16, 25, 37, 49, 61, 64, 67, 79, 81, 107, 109$
4. $\omega(Q) = 4$: $q = 121, 163, 211, 256, 277, 289, 373$
5. $\omega(Q) = 5$: no possible exceptions

Fortunately, these values are small enough to be directly checked via computer, and we do this using the number theory package pari. The primes are

easy enough to check, as one considers all irreducible cubics $x^3+c_1x^2+c_2x+c_3 \in \mathbf{F}_p[x]$ and asks whether there exist $a$, $b \in \mathbf{F}_p$ with $a \neq 0$ such that $f(ax+b)$ is a primitive polynomial. (Of course, one may also take $\theta \in \mathbf{F}_{p^3}$ with $\theta^3+c_1\theta^2+c_2\theta+c_3=0$ and ask whether $a\theta+b$ is primitive for some $a$, $b \in \mathbf{F}_p$ with $a \neq 0$.) For $q=p^e$, $e>1$, we build $\mathbf{F}_{q^3}$ by using a primitive element $u$ of degree $3e$ over $\mathbf{F}_p$ and, for each $i$ between 1 and $p^{3e}-2$ with $(i,q^3-1)>1$ and $\mathrm{ord}(u^i)=\frac{q^3-1}{(i,q^3-1)}$ not a divisor of $p^j-1$ for $j$ a proper divisor of $3e$ (that is, irreducibles of degree 3 over $\mathbf{F}_q$ that are not primitive), ask whether there exist $a$, $b \in \mathbf{F}_q$ with $a \neq 0$ such that $au^i+b$ is a primitive element of degree 3 over $\mathbf{F}_q$. For each possible exception listed above, we find using the methods just described that none of the above-listed $q$ qualifies as a genuine exception. That is, for each $q$ listed above and for all $\theta \in \mathbf{F}_{q^3}$ with $\mathbf{F}_q(\theta)=\mathbf{F}_{q^3}$ there exist $a$, $b \in \mathbf{F}_q$ with $a \neq 0$ such that $a\theta+b$ is a primitive element of degree 3 over $\mathbf{F}_q$. We have proved Theorem 2.

## 4 An Application to Pseudo-Random Vector Generation

Before considering the primitive sum problem for cubic extensions with $a=1$, we turn aside to mention a special case of Niki's pseudorandom vector generator, namely when $n=3$.

Throughout, we let $f \in \mathbf{F}_p[x]$ be an irreducible cubic with root $\alpha \in \mathbf{F}_{p^3} \setminus \mathbf{F}_p$. Niki's result assumes the existence of a primitive root $\beta$ of the form $\beta=a\alpha+b$, with $a,b \in \mathbf{F}_p$. This assumption is made for speed of calculation. Of course we have proved the existence of such a $\beta$. We discuss a special case that will give even greater speed since the reduction modulo $f$ will be as fast as possible. We give the matrix equation for the generator's operation and mention when it will work.

Clearly, since $\mathbf{F}_{p^3}$ is a vector space over $\mathbf{F}_p$ of dimension 3 and $\alpha$ is a generator, all elements can be written in the form $a_2\alpha^2+a_1\alpha+a_0$ for some appropriate choice of $a_i \in \mathbf{F}_p$.

We take the 'seed' $\alpha_0 \in \mathbf{F}_{p^3}$ and then define recursively

$$\alpha_n = \beta^n\alpha_0$$
$$\alpha_n = \nu_n^{(2)}\alpha^2+\nu_n^{(1)}\alpha+\nu_n^{(0)} \quad (\nu_n^{(j)} \in \mathbf{F}_p).$$

This gives rise to a pseudorandom vector on $[0,1]^3$ by thinking of the field elements as integers and looking at

$$\frac{1}{p}\left(\nu_n^{(2)},\nu_n^{(1)},\nu_n^{(0)}\right). \tag{9}$$

That this generator is a type of linear shift register was noted in [17]. We have shown that *any* generator of $\mathbf{F}_{q^3}$ can be used for $\alpha$ in the random number generator. So in particular, when $x^3-2$ is irreducible over $\mathbf{F}_q$ we can

pick $\alpha = \sqrt[3]{2}$. In fact $x^3 - 2$ is irreducible over $\mathbf{F}_p$ when $p \equiv 1 \pmod 6$ and $p \neq A^2 + 27B^2$ for some $A, B \in Z$, since 2 is a cube when $p = A^2 + 27B^2$ (page 119, Proposition 9.6.2 of [12]). Here we have that

$$\alpha_n = \nu_n^{(2)}\alpha^2 + \nu_n^{(1)}\alpha + \nu_n^{(0)} \quad n \geq 0$$

so that

$$\begin{aligned}\alpha_{n+1} &= \beta\alpha_n \\ &= (b\nu_n^{(2)} + a\nu_n^{(1)})\alpha^2 + (b\nu_n^{(1)} + a\nu_n^{(0)})\alpha + (b\nu_n^{(0)} + 2a\nu_n^{(2)}).\end{aligned}$$

That is,

$$\left(\nu^{(0)}, \nu^{(1)}, \nu^{(2)}\right) \mapsto \left((2a\nu^{(2)} + b\nu^{(0)}), (a\nu^{(0)} + b\nu^{(1)}), (a\nu^{(1)} + b\nu^{(2)})\right). \quad (10)$$

The equivalent statement in matrix form is

$$\begin{pmatrix} \widehat{\nu^{(0)}} \\ \nu^{(1)} \\ \nu^{(2)} \end{pmatrix} = \begin{pmatrix} b & 0 & 2a \\ a & b & 0 \\ 0 & a & b \end{pmatrix} \begin{pmatrix} \nu^{(0)} \\ \nu^{(1)} \\ \nu^{(2)} \end{pmatrix} \quad (11)$$

This has applications as a random number generator and also as a simulation for a one-time pad. The output can be XOR'd with a message producing a ciphertext. However, since only a 'small' portion of the message is required to crack a shift register sequence, a higher degree extension would be better.

We remark that this vector generator can be used to generate random numbers in two simple ways: Either take each component separately and use these for random integers in the integers modulo $p$, written $Z_p$, or take $c_1 + c_2 p + c_3 p^2$ as a stream of random integers in $Z_{p^3}$. These sequences of random numbers will be well-distributed provided the initial sequence is also well-distributed.

## 5 The $\theta + b$ Problem When $n = 3$

The primitive sum problem with $a$ fixed (here $a = 1$) is a considerably more difficult problem in general, not only by its nature but in view of the methods we used to resolve the $a\theta + b$ problem for cubic extensions. To contrast the two problems for the case $n = 3$, note first of all that the appropriate inequality for $N'_\theta(e)$ in general is

$$\begin{aligned}N'_\theta(e) &= \frac{\phi(e)}{e} \sum_{d|e} \frac{\mu(d)}{\phi(d)} \sum_{\psi_d} S'_\theta(\psi_d) \\ &\geq \frac{\phi(e)}{e} \left\{ q - ((n-1)\sqrt{q})(2^{\omega(e)} - 1) \right\}, \end{aligned} \quad (12)$$

where $e \mid q^n - 1$. So, as opposed to inequality (6), we can no longer restrict ourselves to divisors of $Q$. For us, this means we must consider all prime

divisors of $q^3 - 1$, which have no special form. Thus our sieving work will not prove nearly as fruitful as for the $a\theta + b$ problem.

It is an easy task to confirm that we may still use (8) for our sieve inequality, with $ef = q^3 - 1$ now. Again, we will take $e$ and $f$ to be coprime. By proceeding as in Section 3, with (12) being used to gain the third-column entries for $\omega(q^3 - 1) \leq 6$ (for these give better results than (8)), we build the following table.

| $\omega(q^3-1)$ | $q \geq$ | $N > 0$ when $q >$ |
|---|---|---|
| 1 | 1.44 | 4 |
| 2 | 1.91 | 36 |
| 3 | 3.14 | 196 |
| 4 | 5.95 | 900 |
| 5 | 13.22 | 3844 |
| 6 | 31.08 | 15876 |
| 7 | 79.92 | 55851.15 |
| 8 | 213.26 | 125118.49 |
| 9 | 606.50 | 349943.02 |
| 10 | 1863.35 | 962317.64 |
| 11 | 5853.49 | 2371021.36 |
| 12 | 19505.14 | 6123335.26 |
| 13 | 67257.95 | 14707545.65 |
| 14 | 235631.39 | 40213302.64 |
| 15 | 850352.69 | 109323610.97 |
| 16 | 3194167.71 | 248825773.15 |
| 17 | 12434883.46 | 687106165.47 |
| 18 | 48949883.70 | 1573702162.76 |

| $\omega(q^3-1)$ | $q \geq$ | $N > 0$ when $q >$ |
|---|---|---|
| 19 | 198812307.15 | 4187541882.88 |
| 20 | 823245530.25 | 10091738285.57 |
| 21 | 3440622312.80 | 24465331612.66 |
| 22 | 14763161313.90 | 65633354775.93 |
| 23 | 64397952985.33 | 144076300790.86 |
| 24 | 287520444756.94 | 386684814620.55 |
| 25 | 1321070444052.47 | 860102588235.88 |

Thus we are only assured of $N = N'_\theta(q^3 - 1) > 0$ when $\omega(q^3 - 1) \geq 25$. Arguing as before, we conclude that the values of $q$ for which there may not exist an element $\theta \in \mathbf{F}_{q^3}$ of degree 3 over $\mathbf{F}_q$ such that $\theta + b$ is primitive for some $b \in \mathbf{F}_q$ are those values of $q$ with $\omega(q^3 - 1) \leq 24$ and with $S \leq q \leq T$ where $S$ and $T$ are the appropriate second and third-column values respectively. Not surprisingly, there are a great many more possible exceptions; indeed there are a total of 3836 possible exceptions for $1 \leq \omega(q^3 - 1) \leq 17$ (we were not able to check the values of $\omega$ between 18

and 24 due to computational constraints). Because of the great number of possible exceptions, in this paper we list only for each value of $\omega(q^3-1) \leq 17$ the number of possible exceptions as well as the smallest and largest possible exceptions for each value of $\omega$. A complete listing of the possible exceptions for $\omega(q^3-1) \leq 17$ can be accessed at the first author's home page at http://www.dean.usma.edu/math/people/millsd/default.htm.

1. $\omega(q^3-1) = 1$: 1 ($q = 2$)
2. $\omega = 2$: 5 (from $q = 3$ to $q = 17$)
3. $\omega = 3$: 15 (from $q = 7$ to $q = 173$)
4. $\omega = 4$: 64 (from $q = 11$ to $q = 887$)
5. $\omega = 5$: 195 (from $q = 61$ to $q = 3833$)
6. $\omega = 6$: 536 (from $q = 121$ to $q = 15803$)
7. $\omega = 7$: 972 (from $q = 211$ to $q = 55837$)
8. $\omega = 8$: 816 (from $q = 2671$ to $q = 125107$)
9. $\omega = 9$: 661 (from $q = 9811$ to $q = 349801$)
10. $\omega = 10$: 374 (from $q = 37951$ to $q = 959083$)
11. $\omega = 11$: 137 (from $q = 106031$ to $q = 2356831$)
12. $\omega = 12$: 45 (from $q = 955711$ to $q = 6059719$)
13. $\omega = 13$: 15 (from $q = 1462171$ to $q = 14535511$)
14. $\omega = 14$: no possible exceptions
15. $\omega = 15$: no possible exceptions
16. $\omega = 16$: no possible exceptions
17. $\omega = 17$: no possible exceptions

The lack of possible exceptions from $\omega(q^3-1) = 14$ to $\omega(q^3-1) = 17$ leads one to conjecture that there are also no possible exceptions for each of $\omega(q^3-1) = 18, ..., 24$.

To eliminate as many of these values of $q$ as possible, we first use (8) on each of the 3836 possible exceptions, and in so doing we are able to eliminate 3407 values of $q$, leaving us with the following 429 possible exceptions (see http://www.dean.usma.edu/math/people/millsd/primroots.htm for a complete listing of those values of $q$ which satisfy the sieve inequality, with values of $e$ and $f$ such that $(e, f) = 1$ and $ef = q^3 - 1$ given for each one):

- $\omega = 1$ (1 possible exception): $q = 2$
- $\omega = 2$ (5): $q = 3, 4, 5, 8, 17$
- $\omega = 3$ (12): $q = 7, 9, 13, 19, 27, 32, 41, 59, 73, 89, 97, 101$
- $\omega = 4$ (35): $q = 11, 16, 23, 25, 29, 31, 37, 43, 47, 49, 53, 64, 71, 83, 103, 109, 113, 125, 127, 131, 157, 179, 193, 197, 199, 223, 227, 233, 241, 243, 251, 271, 313, 317, 409$
- $\omega = 5$ (74): $q = 61, 67, 79, 81, 107, 137, 139, 149, 151, 163, 169, 181, 229, 239, 263, 269, 281, 283, 289, 307, 311, 337, 343, 347, 349, 359, 361, 367, 379, 397, 419, 421, 439, 443, 457, 461, 491, 499, 521, 523, 541, 599, 601, 607, 613, 617, 619, 643, 647, 653, 659, 661, 683, 709, 733, 739, 751, 757, 787, 829, 853, 859, 881, 907, 911, 1009, 1021, 1051, 1061, 1093, 1117, 1153, 1201, 1213$

- $\omega = 6$ (97): $q$ = 121, 191, 256, 277, 331, 373, 431, 463, 529, 547, 625, 631, 691, 729, 809, 811, 821, 823, 877, 947, 961, 967, 971, 997, 1024, 1031, 1033, 1069, 1087, 1103, 1123, 1129, 1171, 1283, 1291, 1321, 1429, 1453, 1471, 1493, 1531, 1543, 1571, 1607, 1621, 1663, 1681, 1693, 1699, 1721, 1741, 1759, 1789, 1823, 1873, 1901, 1913, 1933, 1951, 1979, 1987, 1999, 2003, 2053, 2081, 2083, 2113, 2131, 2143, 2221, 2237, 2281, 2293, 2341, 2371, 2381, 2389, 2503, 2521, 2683, 2707, 2713, 2791, 2809, 2861, 2887, 3001, 3061, 3181, 3271, 3301, 3313, 3331, 3469, 3511, 3697, 3851
- $\omega = 7$ (95): $q$ = 211, 571, 841, 919, 991, 1231, 1303, 1327, 1369, 1381, 1451, 1511, 1597, 1831, 1849, 1871, 2011, 2209, 2311, 2347, 2401, 2473, 2531, 2551, 2557, 2591, 2731, 2851, 2857, 2971, 3229, 3389, 3481, 3539, 3541, 3691, 3877, 3911, 3917, 3931, 4027, 4271, 4423, 4447, 4489, 4523, 4621, 4643, 4651, 4657, 4663, 4691, 4799, 4831, 4903, 4909, 5021, 5101, 5171, 5209, 5419, 5441, 5479, 5521, 5581, 5659, 5743, 5749, 6007, 6067, 6073, 6133, 6163, 6217, 6241, 6271, 6421, 6427, 6451, 6561, 6661, 6691, 6733, 6841, 6889, 6961, 7039, 7177, 7333, 7393, 7459, 7621, 7921, 8101, 8581
- $\omega = 8$ (68): $q$ = 2671, 3571, 3721, 4096, 4111, 4561, 4951, 5791, 5821, 6091, 6581, 6763, 6871, 6997, 7151, 7639, 7879, 8647, 8779, 8941, 8971, 9109, 9181, 9241, 9283, 10111, 10201, 10609, 11027, 11311, 11449, 11677, 11743, 11881, 12301, 12541, 12547, 12973, 12979, 13003, 13421, 13729, 14389, 14821, 14947, 15031, 15271, 15401, 15541, 15661, 16831, 16921, 16993, 17161, 17491, 17761, 17851, 19321, 19891, 20341, 21121, 21211, 21319, 22051, 22441, 23011, 25741, 28051
- $\omega = 9$ (29): $q$ = 9811, 10627, 14431, 14641, 15439, 19141, 20101, 22621, 23431, 24091, 26107, 27967, 28711, 32041, 32761, 33931, 34981, 35863, 36871, 38011, 40231, 40471, 42331, 42571, 43891, 46861, 51151, 51613, 56611
- $\omega = 10$ (10): $q$ = 37951, 44521, 44851, 88741, 97171, 97861, 111211, 132661, 140071, 162691
- $\omega = 11$ (3): $q$ = 106031, 139129, 220411
- $\omega = 12$: No possible exceptions.
- $\omega = 13$: No possible exceptions.

Using high-performance computers provided by the U.S. Army Research Lab, the first author was able to determine that for each $q < 500$ listed above (there are 100 such values) and for each $\theta \in \mathbf{F}_{q^3}$ of degree 3 over $\mathbf{F}_q$ there exists $b \in \mathbf{F}_q$ such that $\theta + b$ is primitive, with the exceptions $q$ = 3, 7, 9, 13, and 37. For each genuine exception $q$, we list an exceptional polynomial below.

- $q = 3$: $x^3 + 2x + 2$
- $q = 7$: $x^3 + 2$
- $q = 9$: $x^3 + u(u^3 + u^2 + 2u + 1)x + (u^4 + u^3 + 2u^2 + u + 2)$ where $u^6 = 2u + 1$ and $u$ generates $\mathbf{F}_{9^3}$

- $q = 13$: $x^3 + 4$
- $q = 37$: $x^3 + 2$

Since the likelihood of a given $q$ value being a genuine exception decreases as $q$ increases, we feel safe in conjecturing the following.

*Conjecture 1.* For any prime power $q \neq 3, 7, 9, 13, 37$ and any element $\theta$ such that $\mathbf{F}_q(\theta) = \mathbf{F}_{q^3}$, there exists $b \in \mathbf{F}_q$ such that $\theta + b$ is a primitive root of $\mathbf{F}_{q^3}$.

## 6 Acknowledgements

Both authors express their thanks to Stephen Cohen of the University of Glasgow for his advice and encouragement during the completion of this paper. The first author would also like to thank the people with whom he discussed this problem at the $\mathbf{F}_{q^6}$-conference, particularly Michael Tsfasman of the Independent University of Moscow and Igor Shparlinski of Macquarie University. As a National Research Council postdoctoral fellow, he also wishes to thank both the U.S. Army Research Laboratory and the U.S. Military Academy for the use of their facilities and equipment in writing this paper.

## References

1. Bombieri, E. *Counting points on curves over finite fields (d'après S. A. Stepanov)*, Sèm. Bourbaki (1972/3), Exp. **430**.
2. Carlitz, L. *Distribution of primitive roots in a finite field*, Quart. J. Math. (2), **4** (1953), 4–10.
3. Chou, W.-S.; Cohen, S.D. *Primitive elements with zero traces*, Finite Fields Appl. **7** (2001), 125–141.
4. Chung, F.R.K. *Diameters and eigenvalues*, J. Amer. Math. Soc. **2** (1989), 187-196.
5. Cohen, S.D. *Primitive roots in the quadratic extension of a finite field*, J. London Math. Soc. (2), **27** (1983), 221-228.
6. Cohen, S.D. *Consecutive primitive roots in a finite field*, Proc. Amer. Math. Soc. (2), **93** (1985), 189–197.
7. Cohen, S.D. *Consecutive primitive roots in a finite field II*, Proc. Amer. Math. Soc. (2), **94** (1985), 605–611.
8. Cohen, S.D.; Mullen, G.L. *Primitive elements in finite fields and Costas arrays*, App. Alg. Engr. Comm. Comp. **2** (1992), 297–299.
9. Davenport, H. *On primitive roots in finite fields*, Quart. J. Math. Oxford **8** (1937), 308–312.
10. Deligne, P. *La conjecture de Weil (I)*, Publ. Math. IHES **43** (1974), 273–307.
11. Deligne, P. *La conjecture de Weil (II)*, Publ. Math. IHES **52** (1981), 313–428.
12. Ireland, K.; Rosen, M. *A classical introduction to modern number theory*, Springer-Verlag, New York, 1982.
13. Jungnickel, D. *Finite fields: structure and arithmetic*, Wissenschaftsverlag, Mannheim, 1993.

14. Katz, N.M. *An estimate for character sums*, J. Amer. Math. Soc. **2** (1989), 197–200.
15. Lidl, R.; Niederreiter, H. *Finite fields*, Encyclopedia of Mathematics and Its Applications **20**, Cambridge Univ. Press, Cambridge, 1983.
16. Mullen, G.L. *A note on a finite field pseudorandom vector generator of Niki*, Math. Japonica **38** (1993), 59–60.
17. Niederreiter, H. *A pseudorandom vector generator based on finite field arithmetic*, Math. Japonica **31** (1986), 759–774.
18. Niki, N. *Generation of n-space uniform pseudorandom numbers based on computations in a finite field* (Japanese), Proc. Symp. Applications of Number Theory on Numerical Analysis, Kyoto, 1984; Lecture Notes No. 537, 100–111, Research Inst. Math. Sciences, Kyoto, 1984.
19. Niki, N. *Finite field arithmetics and multidimensional uniform pseudorandom numbers* (Japanese), Proc. Inst. Statist. Math. **32** (1984), 231–239.
20. Perel'muter, G.I.; Shparlinski, I.E. *On the distribution of primitive roots in finite fields*, Uspechi Matem. Nauk. **45** (1990), 185–186.

# On Polynomial Families in $n$ Indeterminates over Finite Prime Fields Coming from Planar Functions

Nobuo Nakagawa[1]

Kinki University, Department of Mathematics, Higashi-Osaka,
Osaka 577-8502, Japan

**Abstract.** It is shown that there is a relation between planar functions of elementary abelian groups and bent polynomials. Moreover we prove several results concerning them.

## 1 Introduction

Bent polynomials are a generalization of bent functions to any characteristic $p (p \neq 0)$. The aim of this article is to show that there is a relation between planar functions of elementary abelian groups and bent polynomials and prove several results concerning them.

In section 2 I mention that the existence of a planar function of degree $n$, the existence of a regular affine plane of order $n$ satisfying three conditions and the existence of a split $(n, n, n, 1)$- relative difference set are equivalent.

After that examples are given and I will mention the main known results about planar functions. Moreover I mention some results concerning planar functions on $GF(p^n)$ of monomial type.

In section 3 we will look a property of a Gauss sum with respect to planar functions and the definition of bent polynomials coming from planar functions of elementary abelian groups.

In section 4 a relation between planar functions and bent polynomials is shown and results concerning m-forms in two indeterminates are proved.

## 2 Equivalent Conditions of Planar Functions and Known Results

**Definition 1.**
Let $G$ and $H$ be groups of order $n$. For a mapping

$$f \ : \ G \longrightarrow H \ : x \ \longmapsto f(x)$$

and $u \in G$ (fixed), the mapping $f_u$ is defined as

$$f_u \ : \ G \longrightarrow H \ : x \ \longmapsto f(ux)f(x)^{-1}.$$

Then $f$ is called a planar function of degree $n$ from $G$ into $H$ if and only if $f_u$ is bijective for $\forall u \in G\ (u \neq 1)$.

**Theorem 1.** *Let $G$ and $H$ be finite groups of order $n$.*
*The following three statements are equivalent.*
(1) *There exists a planar function from $G$ into $H$.*
(2) *There existes a $(n, n, n, 1)$-relative diffrence set in $G \times H$ relative to $H$.*
(3) *There existes an affine plane $\mathbf{A}$ of order $n$ satisfying the following conditions.*
(a1) *$G \times H$ acts on the points of $\mathbf{A}$ regularly.*
(a2) *each element of $H$ acts as an elation on $\mathbf{A}$ with axis $\ell_\infty$ and center $(\infty)$.*
(a3) *$G$ is transitive on $(\ell_\infty) \setminus (\infty)$ and $((\infty)) \setminus \{\ell_\infty\}$.*

This theorem is well known. In fact under assumption (1),
$D = \{(x, f(x)) \,|\, x \in G\}$ is a relative difference set in $G \times H$ relative to $H$ where $f$ is a planar function from $G$ into $H$.
Moreover if we take $\mathbf{P} = G \times H$ as the set of points and $\mathbf{L} = \{Hx \,|\, x \in G\} \cup \{Dv \,|\, v \in G \times H\}$ as the set of lines, we can prove $(\mathbf{P}, \mathbf{L}, \in)$ is an affine plane satisfying the condition (a1),(a2),(a3) easily. Conversely (1) comes from (2) or (3) immediately.

*Example 1.*

$$f : GF(q) \longrightarrow GF(q)\ : x \longmapsto x^2$$

where $GF(q)$ is the additive group of the Galois field of q elements for an odd prime power $q$ . (An affine plane corresponding to this function is Desarguesian.)

The following example was given by R.S.Coulter and R.W.Mattews([2]).

*Example 2.*

$$f :\ GF(3^e) \longrightarrow GF(3^e)\ : x \longmapsto x^{(3^\alpha+1)/2}$$

where g.c.d.$(\alpha, 2e) = 1$.(Affine planes corresponding to these functions are not translation planes if $1 < \alpha < 2e$.)

*Example 3.*

$$f : GF(3^4) \longrightarrow GF(3^4)\ : x \longmapsto a(x^6 + x^{30} + x^{54}) - x^{10} - x^{18}$$

where $a^2 = -1$. (An affine plane corresponding to this function is a semifield plane(not Desarguesian.))

Known results with respect to planar functions.

**Theorem 2.** *(Ganley [5]) Suppose that there exists a planar function of degree $n$. Then $n$ is odd.*

**Theorem 3.** *(Gluck,Hiramine,Ronyai and Szonyi [6],[7],[17])*
*Suppose that there exists a planar function $f$ of degree $p$ for an odd prime $p$. Then $f$ is a quadratic polynomial and an affine plane corresponding to $f$ is Desarguesian.*

**Theorem 4.** *There is not a planar function from $\mathbf{Z}_n$ into $\mathbf{Z}_n$ for the following $n$.*
(1)*(Fung,Siu and Ma [4])*
*$n$ is divisible by $p^2$ for a prime $p$.*
(2)*(Hiramine [9])*
*$n = 3p$ for any prime $p$.*
(3)*(Ma [13])*
*$n = pq$ for all distinct primes $p, q$.*
(4)*(Leung,Ma,Tan [11])*
*$n = 3pq$ for distinct primes $p, q$ larger than 3.*

**Theorem 5.** *(Nakagawa [15],[16]) Suppose that $G$ and $H$ are finite abelian groups of order $p^n$ for an odd prime $p$, and there exists a planar function from $G$ into $H$. Then*

$$exp(H) \leq \begin{cases} p^{(n+1)/2} & \text{(n : odd)} \\ p^{n/2} & \text{(n : even)} \end{cases}$$

*Moreover $G$ is not cyclic if $n \geq 2$.*

**Theorem 6.** *(Blokhuis, Jungnickel and Schmidt [1]) If there exists a planar function of degree $n$ between abelian groups, then $n$ is a prime power.*

This elegant result by Blokhuis, Jungnickel and Schmidt is a corollary of the following theorem in the context of the prime power conjecture.

**Theorem 7.** *( Blokhuis, Jungnickel and Schmidt [1]) Let $G$ be an abelian collineation group of order $n^2$ of a projective plane of order $n$. Then $n$ is a prime power, say $n = p^n$. If $p > 2$, then the $p$-rank of $G$ is at least $n + 1$.*

Now I will pose a problem.

*Problem 1.* Determine all monomial polynomials over $GF(p^n)$ which are planar functions from $GF(p^n)$ to $GF(p^n)$ as additive groups.

If $m = 2p^i$ $(i = 0, 1, \cdots)$, then obviously $f(x) = x^m$ is a planar function. The following two lemmas are used to obtain several results concerning the problem above.

**Lemma 1.** *(cf.[12]) Set* $A = \{a_1, a_2, \cdots, a_{p^n}\}$ *where* $a_1, a_2, \cdots, a_{p^n}$ *are elements of* $GF(p^n)$. *Then* $A = GF(p^n)$ *if and only if*

$$\sum_{i=1}^{p^n} a_i^k = \begin{cases} -1 & \text{if } (p^n - 1)|k \\ 0 & \text{otherwise.} \end{cases}$$

We define Dickson polynomials $g_k(x)$ on $GF(p^n)$ inductively as the following.

$$g_0(x) = 2, g_1(x) = x, g_2(x) = x^2 - 2 \text{ and } g_k(x) = xg_{k-1}(x) - g_{k-2}(x).$$

**Lemma 2.** *(cf.[12])* (a) $g_k(y + y^{-1}) = y^k + y^{-k}$ *for all positive integer* $k$.
(b) $g_k(x)$ *is a permutation polynomial if and only if* g.c.d.$(k, p^{2n} - 1) = 1$.

We have the following theorem.

**Theorem 8.** (a) *A monomial polynomial* $f(x) = x^m$ *is a planar function from* $GF(p^n)$ *into* $GF(p^n)$ *if and only if the polynomial* $(x+1)^m - (x-1)^m$ *is a permutation polynomial on* $GF(p^n)$ *where* $p$ *is an odd prime.*
(b) *Suppose that one of the following conditions is satisfied.*
(1) g.c.d.$(m, p^n - 1) \neq 2$.
(2) $p^n - 1$ *is divisible by* $m - 1$, $m \neq 2$ *and* $m$ *is not divisible by* $p$.
(3) $p \geq 5$ *and* $m = (p^a + 1)/2$ $(a = 0, 1, 2, \cdots)$.
*Then* $f(x) = x^m$ *is not a planar function on* $GF(p^n)$ $(n = 1, 2, \cdots)$.

*Proof* (a) Suppose that $f(x) = x^m$ is a planar function. Then $f_u(x) = (u + x)^m - x^m$ is bijective for any $u \neq 0$. Therefore $(1/u^m)f_u(x) = (1 + x/u)^m - (x/u)^m$ is bijective. Set $y = x/u$. Then the polynomial $(y + 1)^m - y^m$ is bijective. Set $z = y - a$ where $a = (p - 1)/2$. Then $(z - a)^m - (z + a)^m$ is bijective because $2a + 1 = 0$. Moreover we have $(x + 1)^m - (x - 1)^m$ is bijective by putting $x = z/a$ again after we multiply the above polynomial by $(-1)^m/a^m$. The inverse is obtained by following the reverse of this argument.

(b)(1) Suppose that an equation $(x + 1)^m - (x - 1)^m = 0$ over $GF(p^n)$ has a unique solution. Then the equation $((x+1)/(x-1))^m = 1$ has a unique solution. On the other hand an equation $y^m = 1$ has exactly $d$ solutions where $d =$g.c.d.$(m, p^n - 1)$. However $y = (x + 1)/(x - 1)$ iff $x = (y + 1)/(y - 1)$ under $x \neq 1, y \neq 1$. Thus excluding 1, $y^m = 1$ has exactly $d - 1$ solutions. Therefore from our assumption we have $d - 1 = 1$, or equivalently $d = 2$. Hence if g.c.d.$(m, p^n - 1) \neq 2$, then $(x+1)^m - (x-1)^m$ is not a permutation polynomial, or equivalently $f(x) = x^m$ is not a planar function from (a).

(b)(2) Suppose that $p^n - 1$ is divisible by $m - 1$, $m \neq 2$ and $m$ is not divisible by $p$. Then there is a positive integer $h$ such that $(m-1)h = p^n - 1$. Note that $h < p^n - 1$ because $m \neq 2$. We calculate

$$\sum_{x \in GF(p^n)} ((x + 1)^m - (x - 1)^m)^h.$$

This polynomial equals

$$\sum_{x\in GF(p^n)} \{2 \sum_{0\le r\le m(r:\text{odd})} \binom{m}{r} x^{m-r}\}^h$$

$$= 2^h \{\sum_{\ell=h}^{hm} \sum_{r_1+\cdots+r_h=\ell(r_i:\text{odd})} \binom{m}{r_1}\cdots\binom{m}{r_h} (\sum_{x\in GF(p^n)} x^{hm-\ell})\}.$$

However $hm-\ell \le hm-h = p^n-1$ and the equality holds iff $\ell = h$ and $r_1 = \cdots = r_h = 1$. If $h < \ell$, then $\sum_{x\in GF(p^n)} x^{hm-\ell} = 0$ from Lemma 1 and $\sum_{x\in GF(p^n)} x^{hm-h} = -1$ from Lemma 1 again. Thus we have

$$\sum_{x\in GF(p^n)} ((x+1)^m - (x-1)^m)^h = -2^h m^h \neq 0$$

from our assumption $m$ is not divisible by $p$. Hence $(x+1)^m - (x-1)^m$ is not bijective on $GF(p^n)$ from Lemma 1. Therefore $f(x) = x^m$ is not a planar function from (a).

(b)(3) Suppose that $p \ge 5$ and $m = (p^a+1)/2$. Obviously $(x+1)^m - (x-1)^m$ is bijective if and only if $(2x+2)^m - (2x-2)^m$ is bijective. Moreover there is an element $y \in GF(p^{2n})$ such that $2x = y + y^{-1}$. Then we have $(2x+2)^m - (2x-2)^m = (y+y^{-1}+2)^m - (y+y^{-1}-2)^m = y^{-m}((y^2+1+2y)^m - (y^2+1-2y)^m)$. This polynomial equals $y^{-m}((y+1)^{p^a+1} - (y-1)^{p^a+1}) = 2(y^{(p^a-1)/2} + y^{-((p^a-1)/2)})$ because $m = (p^a+1)/2$. Thus we have $(2x+2)^m - (2x-2)^m = 2g_{(p^a-1)/2}(y+y^{-1}) = 2g_{(p^a-1)/2}(2x)$ from (a) of Lemma 2. Hence $(x+1)^m - (x-1)^m$ is bijective if and only if g.c.d.$((p^a-1)/2, p^{2n}-1) = 1$ from (b) of Lemma 2. However if $p \ge 5$, it is clear that g.c.d.$((p^a-1)/2, p^{2n}-1) \neq 1$ for any positive integer $a$. Therefore $(x+1)^m - (x-1)^m$ is not bijective, or equivalently $f(x) = x^m$ is not a planar function. □

## 3 Gauss Sum of Planar Functions and Bent Polynomials

Let $f$ be a planar function from $G$ into $H$ where $G$ and $H$ are abelian groups. We define a Gauss sum of $f$ with respect to $\chi \in \hat{G}$ and $\rho \in \hat{H}$ where $\hat{G}$ and $\hat{H}$ are the character groups of $G$ and $H$ respectively.

**Definition 2.**

$$z_{\chi,\rho} = \sum_{x\in G} \chi(x)\rho(f(x))$$

Then we have the following theorem.

**Theorem 9.** *Let $f$ be a function from $G$ into $H$ of degree $n$. Then $f$ is a planar function if and only if $z_{\chi,\rho}\overline{z_{\chi,\rho}} = n$ for any $\chi \in \hat{G}$ and any nontrivial $\rho \in \hat{H}$.*

*Proof* Suppose that $f$ is planar. Then

$$z_{\chi,\rho}\overline{z_{\chi,\rho}} = (\sum_{x\in G}\chi(x)\rho(f(x))\overline{(\sum_{y\in G}\chi(y)\rho(f(y))}$$

$$= \sum_{x,y\in G}\chi(xy^{-1})\rho(f(x)f(y)^{-1}) = \sum_{u\in G}\chi(u)(\sum_{y\in G}\rho(f(uy)f(y)^{-1})$$

$$= n + \sum_{(1\neq)u\in G}\chi(u)(\sum_{y\in G}\rho(f_u(y)) = n$$

The proof of the inverse is also shown easily. We note the following two points. Firstly if $\sum_{(1\neq)u\in G} a_u\chi(u) = 0$ for any $\chi \in \hat{G}$, then $a_u = 0$ for all $u \in G(u \neq 1)$.
Secondly suppose that $h_1, \cdots, h_n$ are elements of $H$. If $\sum_{i=1}^{i=n}\rho(h_i) = 0$ for any nontrivial $\rho \in \hat{H}$, then $\{h_1, \cdots, h_n\} = H$. □

A linear polynomial at $\mathbf{a}$ in $n$ indeterminates over $\mathbf{Z}_p$ is defined as

$$L_{\mathbf{a}}(\mathbf{x}) = a_1x_1 + \cdots + a_nx_n$$

where $\mathbf{a} = (a_1, \cdots, a_n)$ and $\mathbf{x} = (x_1, \cdots, x_n)$.
We put the number of solutions of $g$ at $k \in \mathbf{Z}_p$ as

$$c_k(g) = |\{(x_1, \cdots, x_n) \in \mathbf{Z}_p^n \mid g(x_1, \cdots, x_n) = k\}|$$

for a polynomial $g(x_1, \cdots, x_n)$ in $n$ determinates over $\mathbf{Z}_p$.
Then the bent polynomials $\mathcal{F}_p(n)$ are defined as follows.

$$\mathcal{F}_p(n) := \{f(x_1, \cdots, x_n) \mid f \text{ satisfies the condition } (H_1) \text{ or } (H_2)$$
$$\text{below for any vector } \mathbf{a} \in \mathbf{Z}_p^n\}$$

Suppose that $n$ is even.

$$c_k(L_{\mathbf{a}} + f) = \begin{cases} p^{n-1} \pm p^{n/2} \mp p^{(n-2)/2} & (k = k_0) \\ p^{n-1} \mp p^{(n-2)/2} & (k \neq k_0) \end{cases} \qquad (H_1)$$

where $k_0 = k_0(\mathbf{a})$ is a fixed suitable element of $\mathbf{Z}_p$.

Suppose that $n$ is odd.

$$c_k(L_{\mathbf{a}} + f) = \begin{cases} p^{n-1} & (k = k_0) \\ p^{n-1} + p^{(n-1)/2} & (k \in A) \\ p^{n-1} - p^{(n-1)/2} & (k \in B) \end{cases} \qquad (H_2)$$

where $k_0 = k_0(\mathbf{a})$ is a fixed suitable element of $\mathbf{Z}_p$ and $\mathbf{Z}_p = \{k_0\} \cup A \cup B$ such that $|A| = |B| = (p-1)/2$.
(Remark: Bent polynomials are already defined by Kumar, Scholtz and Welch. (cf. [10]))

*Example 4.* If a polynomial $g$ in $n$ indeterminates over $\mathbf{Z}_p$ is a nondegenerate quadratic form, then $g \in \mathcal{F}_p(n)$([12], pp. 282–283).

## 4 A Relation Between Planar Functions of Elementary Abelian Groups and Bent Polynomials

The following theorem shows a relation betweem planar functions and bent polynomials. Let $G$ and $H$ be elementary abelian groups. Hence we may assume $G \cong \mathbf{Z}_p^n \cong H$.

**Theorem 10.** *A function* $f(\mathbf{x}) = (f_1(\mathbf{x}), \cdots, f_n(\mathbf{x}))$ *from* $G$ *into* $H$ *is a planar function if and only if*

$$s_1 f_1 + \cdots + s_n f_n$$

*is a bent polynomial for each* $(s_1, \cdots, s_n) \in \mathbf{Z}_p^n$ *such that* $(s_1, \cdots, s_n) \neq (0, \cdots, 0)$.

*Proof* Suppose that $f$ of the theorem above is a planar function. Let $\omega$ be a primitive p-th root of unity. Moreover suppose that $n$ is even. Then from Theorem 9 we have $z_{\chi,\rho} = \pm\omega^t p^{n/2}$ for some $t \in \mathbf{Z}$ and any $\chi \in \hat{G}$ and any non trivial $\rho \in \hat{H}$. Since $G$ is an elementary abelian $p$-group, any character of $G$ corresponds to a unique $n$-sequence $(a_1, \cdots, a_n)$ of $\mathbf{Z}_p$ such that

$$\chi(x) = \omega^{a_1 x_1 + \cdots + a_n x_n}$$

for all $x = (x_1, \cdots, x_n) \in G$. Then we may write $\chi = \chi_{(a_1,\cdots,a_n)}$. Similarly $\rho = \rho_{(s_1,\cdots,s_n)}$ for a unique $(s_1, \cdots, s_n) \in \mathbf{Z}_p^n$. Now from the equation above and the definition of $z_{\chi,\rho}$, we obtain

$$\sum_{x=(x_1,\cdots,x_n)\in G} \omega^{a_1 x_1 + \cdots + a_n x_n + s_1 f_1(x) + \cdots + s_n f_n(x)} = \pm\omega^t p^{n/2}$$

for any $(a_1, \cdots, a_n)$ and any $(s_1, \cdots, s_n)$ such that $(s_1, \cdots, s_n) \neq (0, \cdots, 0)$.

Here we fix $(s_1, \cdots, s_n)$, and set as $g(x) = s_1 f_1(x) + \cdots + s_n f_n(x)$. Then we have

$$\sum_{x\in G} \omega^{L_a(x)+g(x)-t} = \pm p^{n/2}$$

for all $\mathbf{a} = (a_1, \cdots, a_n)$. We now denote the cardinality $|\{\ x \in G \mid L_{\mathbf{a}}(x) + g(x) - t = k\ \}|$ by $c_k$ for any $k \in \mathbf{Z}_p$. Then from the equation above it follows that

$$c_0 + c_1\omega + \cdots + c_{p-1}\omega^{p-1} = \pm p^{n/2}.$$

Since $c_0 + c_1 + \cdots + c_{p-1} = p^n$, $1 + \omega + \cdots + \omega^{p-1} = 0$ and $\{1, \omega, \cdots, \omega^{p-2}\}$ are linearly independent over $\mathbf{Q}$ we obtain the following equations from simple calculations.

$$c_0 = p^{n-1} \pm p^{n/2} \mp p^{(n-2)/2},\ c_1 = c_2 = \cdots = c_{p-1} = p^{n-1} \mp p^{(n-2)/2}$$

Thus we have

$$c_k(L_{\mathbf{a}}+g)=\begin{cases} p^{n-1}\pm p^{n/2}\mp p^{(n-2)/2} & (k=t)\\ p^{n-1}\mp p^{(n-2)/2} & (\text{otherwise})\end{cases}$$

Therefore $g(x)=s_1f_1(x)+\cdots+s_nf_n(x)$ is a bent polynomial for any $(s_1,\cdots,s_n)\neq\mathbf{0}$. We can also prove our assertion in the case that $n$ is odd similarly though we use the equation

$$\sum_{x\in G}\omega^{L_a(x)+g(x)-t}=\pm p^{(n-1)/2}\tau$$

where $\tau=\sum_{i=0}^{i=p-1}\lambda(i)\omega^i$ and $\lambda$ is the quadratic character of $\mathbf{Z}_p^{\times}$ with $\lambda(0)=0$.

Follow the inverse of the argument above. Then it can be verified that $f(x)=(f_1(x),\cdots,f_n(x))$ is planar if $s_1f_1(x)+\cdots+s_nf_n(x)$ is a bent polynomial for any $(s_1,\cdots,s_n)\neq\mathbf{0}$. □

*Example 5.* Let $V:=GF(5^4)$ be a 4 dimensional vector space over $GF(5)$ with the basis $\{1,\omega,\omega^2,\omega^3\}$, where $\omega^4=2$. Let $f$ be a planar function on $V$ defined as $f(x)=x^2$. Then we have

$$f_1(x_1,x_2,x_3,x_4)=x_1^2+2x_3^2+4x_2x_4,\quad f_2(x_1,x_2,x_3,x_4)=2x_1x_2+4x_3x_4,$$

$$f_3(x_1,x_2,x_3,x_4)=x_2^2+2x_4^2+2x_1x_3,\quad f_4(x_1,x_2,x_3,x_4)=2x_1x_4+2x_2x_3.$$

Therefore $s_1f_1+s_2f_2+s_3f_3+s_4f_4$ is bent polynomials for all $(s_1,s_2,s_3,s_4)\neq\mathbf{0}$ from Theorem 10.

Now I will pose a problem.

*Problem 2.* Find as many bent polynomials as possible except non degenerate quadratic form over $GF(p)$. Under what conditions are $m$-forms in $n$ determinates over the prime field bent polynomials?

The following theorem is partial solutions of the problem above.

**Theorem 11.** *Let $f(x,y):=a_0x^m+a_1x^{m-1}y+\cdots+a_my^m$ be a $m$-form in two indeterminates on $GF(p)$ where $p$ is an odd prime.*
(a) *We set the polynomial related to $f(x,y)$ as $\varphi_f(t)=a_0+a_1t+\cdots+a_mt^m$ where $t$ is an indeterminate.*
*Suppose that the number of solutions of the equation $\varphi_f(t)=0$ is neither 1 nor 2. Moreover suppose* g.c.d.$(m,p-1)\neq 2$. *Then $f(x,y)$ is not a bent polynomial.*
(b) *If $m=3$, then $f(x,y)$ is not a bent polynomial.*

*Proof* (a) We denote the number of solutions of the equation $\varphi_f(t) = 0$ by $N$. Suppose that $N = 0$ or $N \geq 3$ and g.c.d.$(m, p-1) \neq 2$. Moreover assume that $f(x, y)$ is a bent polynomial. Then we have the following assumption ($\natural$).

$$(\natural)\ N(ix+jy+f(x,y) = k) \in \{1, p-1, p+1, 2p-1\} \quad \text{for} \quad \forall i, \forall j, \forall k \in GF(p)$$

where $N(ix + jy + f(x, y) = k)$ is the number of solutions of a equation $ix+jy+f(x, y) = k$. We will obtain a contradiction from these assumptions.

(1) We may assume that $a_m \neq 0$.

Suppose that $a_m = 0$. We examine $N(f(x, y) = 0)$. If $x = 0$, then $(0, y)$ is a solution of $f(x, y) = 0$ for any $y \in GF(p)$. If $x \neq 0$, then $x^m \varphi_f(t) = 0$ where $t = y/x$. Therefore $\varphi_f(t) = 0$. Hence if $N = 0$, then $N(f(x, y) = 0)$ is $p$, this contradicts to our assumption ($\natural$). If $N \geq 3$, then there are at least three distinct elements $t_1, t_2, t_3$ such that $\varphi_f(t_i) = 0$ for $i = 1, 2, 3$. Then $(x, t_i x)$ is a solution for each $t_i (i = 1, 2, 3)$ and any $x \in GF(p)^\times$. Therefore $N(f(x, y) = 0)$ is larger than $2p - 1$, also a contradiction.

(2) We may assume $N = 0$ from the same arguments as the proof of (1).

(3) g.c.d.$(m, p - 1) = 1$.

We may assume $a_m = 1$ from (1). We have $N(f(x, y) = 0)$ is 1 from (1) and (2). Therefore $N(f(x, y) = k)$ is $p + 1$ for any $k \in GF(p)^\times$ from our assumption that $f(x, y)$ is a bent polynomial. Set $m_0 =$ g.c.d.$(m, p - 1)$ We examine $N(f(x. y) = 1)$. If $x = 0$, then $y^m = 1$. There are exactly $m_0$ elements such that $y^m = 1$. Therefore there are exactly $m_0$ solutions of $f(x, y) = 1$ satisfying $x = 0$. If $x \neq 0$, then $x^m \varphi_f(t) = 1$ where $t = y/x$. Therefore $x^m = 1/\varphi_f(t)$. Set

$$\ell = |\{t \in GF(p) \mid \varphi_f(t) \in (GF(p)^\times)^{m_0}\}|.$$

Then there are exactly $m_0$ solutions of the equation $x^m = 1/\varphi_f(t)$ for each $t$ such that $\varphi_f(t) \in (GF(p)^\times)^{m_0}$. Therefore $N(f(x, y) = 1) = m_0 + m_0\ell$. Hence $m_0 + m_0\ell = p + 1$. Thus we have $m_0 | (p + 1)$. On the other hand $m_0 | (p - 1)$. Therefore $m_0 = 1$ or 2. However $m_0 \neq 2$ from our assumption. Thus $m_0 =$ g.c.d.$(m, p - 1) = 1$.

Put $m_1 =$ g.c.d.$(m - 1, p - 1)$. We note that $m_1$ is even from (3). Set

$$GF(p)^\times = (GF(p)^\times)^{m_1} v_1 \cup (GF(p)^\times)^{m_1} v_2 \cup \cdots \cup (GF(p)^\times)^{m_1} v_{m_1}$$

where $v_1 = 1$. We examine $N(-v_j y + f(x, y) = 0)$ for each $j$. If $x = 0$, then $-v_j y + y^m = 0$, we have $y = 0$ or $y \neq 0$ and $y^{m-1} = v_j$. Therefore $-v_j y + f(0, y) = 0$ has exactly $m_1 + 1$ solutions if $j = 1$, and $(0, 0)$ is only one solution of $-v_j y + f(0, y) = 0$ with respect to $y$, if $2 \leq j$. If $x \neq 0$, then we have $x^{m-1} = v_j t/\varphi_f(t)$ from $-v_j y + f(x, y) = -v_j x t + x^m \varphi_f(t) = 0$ where $t = y/x$. Set

$$n_j = |\{\, t \in GF(p)^\times \mid \varphi_f(t) \in (GF(p)^\times)^{m_1} v_j t \,\}|$$

for each $1 \leq j \leq m_1$. Here we consider two cases.

Case (i): $n_1 \neq 0$. In this case we may assume that $n_1 \neq 0, \cdots, n_s \neq 0$ and $n_{s+1} = \cdots = n_{m_1} = 0$ for some positive integer $s$. Then we have

$$N(-v_j y + f(x,y) = 0) = \begin{cases} 1 + m_1 + m_1 n_1 & \text{if} \quad j = 1 \\ 1 + m_1 n_j & \text{if} \quad 2 \leq j \leq s \end{cases}$$

We note that $1 + m_1 + m_1 n_1$ and $1 + m_1 n_j$ is an odd integer larger than 1. Therefore from ($\natural$) $1 + m_1 + m_1 n_1 = 2p - 1$ and $1 + m_1 n_j = 2p - 1$ each $2 \leq j \leq s$. Thus by taking the sum of these we have $m_1 p = s(2p-2)$ since $n_1 + n_2 + \cdots + n_s = p - 1$. Then $p|s$. However $s \leq m_1$ and $m_1|(p-1)$. Hence $s < p$. This contradicts $p|s$..

Case (ii): $n_1 = 0$. In this case, $N(-v_1 y + f(x,y) = 0) = 1 + m_1$. As the same arguments above we have $1 + m_1 = 2p - 1$. However $m_1|(p-1)$. Therefore $2p - 1 \leq p$, a contradiction. This complete the proof of (a).

(b) Suppose that $m = 3$ and $f(x,y)$ is a bent function. Since g.c.d.$(3, p-1) \neq 2$, we may assume that $N = 2$ or $N = 1$ from (a).

(1) We may assume $N \neq 2$.

Suppose that $N = 2$. Let $t_1$ and $t_2$ be distinct roots of $\varphi_f(t) = 0$. If $a_3 = 0$, then $(0, y)$ is a root of $f(x,y) = 0$ for each $y \in GF(p)$. Moreover $(x, xt_1)$ and $(x, xt_2)$ are roots of $f(x,y) = 0$ for each $x \in GF(p)^\times$. Therefore $N(f(x,y) = 0)$ is larger than $2p-1$, and this contradicts the assumption ($\natural$). Thus we have $a_3 \neq 0$. Now let $\{w_1, w_2\}$ be a representative of the cosets of $GF(p)^\times$ by the subgroup $(GF(p)^\times)^2$. We examine $N(-w_i x + f(x,y) = 0)$ for $i = 1, 2$. If $x = 0$, then $a_3 y^3 = 0$. Therefore $(0,0)$ is a unique solution of $-w_i x + f(x,y) = 0$ satisfying $x = 0$. If $x \neq 0$, then $x^3 \varphi_f(t) = w_i x$ where $t = x/y$. Hence $x^2 = w_i/\varphi_f(t)$ if $\varphi_f(t) \neq 0$. Put

$$m_i = |\{\, t \in GF(p) \setminus \{t_1, t_2\} \mid w_i/\varphi_f(t) \in (GF(p)^\times)^2 \,\}|$$

for $i = 1, 2$. Then we have $m_1 + m_2 = p - 2$ and $N(-w_i x + f(x,y) = 0)$ is $1 + 2m_i$. Since $1 + 2m_i$ is odd, $1 + 2m_i = 1$ or $2p - 1$ from the assumption ($\natural$). Thus $(m_1, m_2) = (0,0), (0, p-1), (p-1, 0)$ or $(p-1, p-1)$. However none of these satisfy $m_1 + m_2 = p - 2$, a contradiction. Thus we may assume that $N = 1$ in the rest.

(2) We may assume $a_0 \neq 0$.

Suppose that $a_0 = a_3 = 0$. Take a nonsingular linear transformation $x = aX + bY, y = cX + dY$, and choose $a, b, c, d$ for which the coefficient of $X^3$ is non-zero.

(3) We may assume that $a_1 = 0$.

If $a_1 \neq 0$, then by the linear transformation $x = X + Y, y = -3a_0/a_1 Y$ we have the coefficient of $X^2 Y$ is zero.

We way assume $\varphi_f(1) = 0$ by giving a linear transformation such that $x = X, y = aY$ for a suitable $a \in GF(p)^\times$ if necessary.

From (2),(3) we have $\varphi_f(t) = 1 + a_2 t^2 + a_3 t^3$.

(4) We may assume $a_3 \neq 0$.

If $a_3 = 0$, $\varphi_f(1) = \varphi_f(-1) = 0$. This contradicts (1).

(5) We may assume that $\varphi_f(i) = 0$ for some $i > 1$.
If $\varphi_f(i) \neq 0$ for all $i \in GF(p)^\times \setminus \{1\}$, we have $f(x,x) = 0$ for all $x \in GF(p)$ and $f(x,y) \neq 0$ for any other $(x,y) \in GF(p) \times GF(p)$ from (4). Hence $N(f(x,y) = 0)$ is $p$, a contradiction.
Thus an equation $\varphi_f(t) = 0$ has two or three distinct roots. This is a final contradiction, and this complete the proof of (b). □

**Remark**: It follows from the proof of Theorem 11 that if $N = 2$ and g.c.d.$(m-1, p-1) \neq 1$, then $f(x,y)$ is not a bent polynomial.

*Example 6.* Assume that g.c.d.$(p-1,3) = 1$. Then $f(x,y) = ax^4 + bx^3y$ $(b \neq 0)$ are bent polynomials. (This is a special example of those given by Kumar, Scholt and Welch ([10]).)

## Acknowledgement

The author wishes to thank Prof. Satoshi Yoshiara for useful comments during the preparation of this article and would like to thank the referee for several useful comments.

## References

1. Blokhuis A., Jungnickel D. and Schmidt B. Proof of the prime power conjecture for projective planes of order $n$ with abelian collineation groups. To appear in Proc.AMS.
2. Coulter R.S. and Matthews R.W.(1997) Planar Functions and Planes of Lenz-Barlotti Class II. Designs, Codes and Cryptography 10:167–184.
3. Dembowski P. and Ostrom T.G.(1968) Planes of order $n$ with colleneation groups of order $n^2$. Math. Z. 103:239–258.
4. Fung C.I.,Siu M.K. and Ma S.L.(1990) On arrays with small off-phase binary autocorrelation. Ars Comb 29A:189–192.
5. Ganley M.J(1976) On a paper of Dembowski and Ostrom. Arch. Math 27:93–98.
6. Gluck D.(1990) A note permutation polynomials and finite geometries. Discrete Math. 80:97–100.
7. Hiramine Y.(1989) A conjecture on affine planes of prime oeder. J. Combin. Theory Ser. A 52:44–50.
8. Hiramine Y.(1991) Facter sets associated with regular collineation groups. J. Algebra 142:414–423.
9. Hiramine Y.(1992) Planar functions and related group algebras. J. Algebra 152:135–145.
10. Kumar P.V.K.,Scholt A. and Welch R.(1985) Generalized bent functions and their properties. J. Combin. Theory, Ser A 40:90–107.
11. Leung K.H.,Ma S.L. and Tan A.V. Planar functions from $Z_n$ to $Z_n$. Preprint.
12. Lidl R. and Niederreiter H.(1984) Finite Fields. Cambridge University Press, Cambridge/London/New York.

13. Ma S.L.(1996) Planar functions,relative difference sets and character theory. J. Algebra 185:342–356.
14. Ma S.L. and Pott A.(1995) Relative difference sets, planar functions and generalized Hadamard matrices. J. Algebra 175:505–525.
15. Nakagawa N.(1993) The non-existence of right cyclic planar functions of degree $p^n$ for $n \leq 2$. J. Combin. Theory Ser A 63:55–64.
16. Nakagawa N.(1997) Left Cyclic Planar Functions Of Degree $p^n$. Utilitas Mathematica 51:89–96.
17. Ronayi L. and Szonyi T.(1989) Planar functions over finite fields. Combinatorica 9:315–320.

# Cryptanalysis of the Sakazaki-Okamoto-Mambo ID-based Key Distribution System over Elliptic Curves

Minghua Qu[1], Doug Stinson[2], and Scott Vanstone[1,2]

[1] Certicom Research, Canada
[2] Department of C&O, University of Waterloo, Canada

**Abstract.** In 1997, H. Sakazaki, E. Okamato and M. Mambo [6] proposed an ID-based key distribution system on an elliptic curve over $\mathbb{Z}_n$. We will cryptanalyze the scheme and demonstrate that when the hashed ID length is about 160 bits, the scheme is insecure. To be specific, after requesting a small number of keys from the Center, our attack allows a new valid key to be constructed without any further interaction with the Center.

## 1 Introduction

In 1986, E. Okamoto [5] proposed an ID-based key distribution system (KDS) whose security depends on the difficulty of factoring an integer that is the product of two large primes, as in the RSA public-key cryptosystem. However, this scheme cannot be constructed on an elliptic curve over $\mathbb{Z}_n$ in a straightforward way because the point corresponding to a user's identity may not be defined on the elliptic curve. As a solution to this problem, Sakazaki-Okamoto-Mambo [6] proposed an ID-based KDS on an elliptic curve over $\mathbb{Z}_n$. The proposed scheme can be also constructed on the ring $\mathbb{Z}_n$.

We will show that some homomorphism-like properties hold in the Sakazaki-Okamoto-Mambo' scheme, and use them to cryptanalyze the scheme. We will demonstrate that, when the hashed ID length is about 160 bits, one can forge a private key $S_{I_i}$ corresponding to some identity $I_i$. Hence the Sakazaki-Okamoto-Mambo scheme with a 160-bit hash function should be considered insecure.

This paper is organized as following: Section 2 describes the Sakazaki-Okamoto-Mambo KDS scheme. Section 3 will discuss the security of the scheme and present and analyze some attacks. Section 4 concludes the paper with some brief comments.

## 2 The Sakazaki-Okamoto-Mambo Scheme

### 2.1 Elliptic Curves over $\mathbb{Z}_n$

For a detailed discussion of elliptic curves over $\mathbb{Z}_n$, see Koblitz [3]. Here we just provide enough information to describe the Sakazaki-Okamoto-Mambo scheme.

Let $n$ be a product of two distinct primes $p$ and $q$ each greater than 3. Let $a, b \in \mathbb{Z}_n$ be such that $\gcd(4a^3 + 27b^2, n) = 1$. An *elliptic curve* over $\mathbb{Z}_n$ with parameters $a$ and $b$ is defined as the set of points

$$\{(x, y) \in \mathbb{Z}_n \times \mathbb{Z}_n : y^2 \equiv x^3 + ax + b \pmod{n}\} \cup \{\mathcal{O}\},$$

where $\mathcal{O}$ is a special point called the *point at infinity* . This elliptic curve is denoted $E_n(a, b)$. Suppose that $G \in E_n(a, b)$ is a *base point* having order

$$k = \text{lcm}(\#E_p(a, b), \#E_q(a, b)).$$

Note that $E_p(a, b)$ and $E_q(a, b)$ denote the corresponding elliptic curves defined over $\mathbb{Z}_p$ and $\mathbb{Z}_q$, and $\#E$ denotes the number of points in an elliptic curve $E$. Such a base point $G$ exists provided that $E_p(a, b)$ and $E_q(a, b)$ are both cyclic groups.

For the key establishment scheme described in this paper, one can select the elliptic curve parameters as follows. First random primes $p$ and $q$ of the same bitlength are generated so that factoring $n = pq$ is intractable (e.g., $p$ and $q$ are each 512 bits in length). Then, elliptic curves $E_p(a_1, b_1)$ over $\mathbb{Z}_p$ and $E_q(a_2, b_2)$ over $\mathbb{Z}_q$ are selected so that $N_1 = \#E_p(a_1, b_1)$ and $N_2 = \#E_q(a_2, b_2)$ are distinct primes (e.g., using Schoof's algorithm [7] and it many enhancements). Let $G_1 = (x_1, y_1) \in E_p(a_1, b_1)$, $G_2 = (x_2, y_2) \in E_q(a_2, b_2)$ be points of order $N_1$, $N_2$, respectively. Finally, one computes $k = N_1 N_2$ and $a, b, x, y \in \mathbb{Z}_n$ satisfying:

$$\begin{cases} a \equiv a_1 \pmod{p} \\ a \equiv a_2 \pmod{q} \end{cases} \qquad \begin{cases} b \equiv b_1 \pmod{p} \\ b \equiv b_2 \pmod{q} \end{cases}$$

and

$$\begin{cases} x \equiv x_1 \pmod{p} \\ x \equiv x_2 \pmod{q} \end{cases} \qquad \begin{cases} y \equiv y_1 \pmod{p} \\ y \equiv y_2 \pmod{q}. \end{cases}$$

Then $G = (x, y)$ is a point of order $k$ on $E_n(a, b)$. Given $n, a, b$ and $G$, computing $k$ is assumed to be intractable without knowledge of the prime factors of $n$.

### 2.2 The Sakazaki-Okamoto-Mambo Scheme over an Elliptic Curve

**Set-up Phase** The Center publishes the parameters of an elliptic curve $E_n(a, b)$, and a base point $G$, as described in Section 2.1. The Center has private key consisting of $k, p$ and $q$.

**Issuing a Private Key to a User** Suppose the Center wants to issue a private key to a user $i$. Let $I_i = h(ID_i)$, where $h$ is a public hash function

such as SHA-1 [4], and $ID_i$ is user $i$'s public identifying information. We call $I_i$ a *hashed identity*. Suppose that $\gcd(I_i, k) = 1$. The Center computes

$$D_i = {I_i}^{-1} \bmod k$$

and

$$S_{I_i} = -D_i G.$$

Hence, it follows that

$$I_i S_{I_i} + G = \mathcal{O}.$$

The Center transmits $(I_i, S_{I_i})$ to user $i$ over a secure (secret and authentic) channel. $S_{I_i}$ is user $i$'s private key, and $I_i$ is his public key.

**Key Exchange Scheme** Suppose Alice and Bob want to establish a common key. Define $[1, n-1] = \{1, \ldots, n-1\}$. Let $I_A$, $I_B$ be Alice's and Bob's public keys, and let $S_{I_A}$ and $S_{I_B}$ be their private keys.

First, Alice randomly chooses an integer $r_A \in [1, n-1]$, computes the elliptic curve point

$$C_A = S_{I_A} + r_A I_B G$$

over $E_n(a, b)$, and sends it to Bob. Similarly, Bob randomly chooses an integer $r_B \in [1, n-1]$, computes

$$C_B = S_{I_B} + r_B I_A G$$

over $E_n(a, b)$, and sends it to Alice.

Then Alice computes

$$K_{AB} = r_A(I_B C_B + G)$$

over $E_n(a, b)$, and Bob computes

$$K_{BA} = r_B(I_A C_A + G)$$

over $E_n(a, b)$. Obviously

$$K_{AB} = K_{BA} = r_A r_B I_A I_B G.$$

Note that the above scheme can also be described over $\mathbb{Z}_n$.

## 3 Cryptanalysis of the Sakazaki-Okamoto-Mambo Scheme

In this section, we will investigate a weakness of the Sakazaki-Okamoto-Mambo' scheme. We will concentrate on the private keys distributed by the Center, and provide methods to forge a private key $S_I$ corresponding to a public key $I$, where $I$ is a hashed identity.

### 3.1 Homomorphism-like Properties of the Sakazaki-Okamoto-Mambo Scheme

In the following, we assume that the modulus $k$ is unknown. All inverses are defined modulo $k$. For any positive integer $x$, define

$$S_x = -x^{-1}G.$$

($S_x$ is the private key corresponding to public key $x$.)

**Lemma 31** *Let $z = xy$ where $x, y$ and $z$ are positive integers. Suppose that $S_z = -z^{-1}G$. Then $S_x = yS_z$ and $S_y = xS_z$.*

*Proof.* Clearly we have

$$xyz^{-1} \equiv 1 \pmod{k},$$

so it follows that

$$-x^{-1} \equiv -yz^{-1} \pmod{k}$$

and

$$-y^{-1} \equiv -xz^{-1} \pmod{k}.$$

Hence,

$$S_x = -x^{-1}G = -yz^{-1}G = yS_z$$

and

$$S_y = -y^{-1}G = -xz^{-1}G = xS_z.$$

□

**Lemma 32** *Suppose that $\gcd(x, y) = 1$, $S_x = -x^{-1}G$ and $S_y = -y^{-1}G$. Then $S_{xy} = k_1S_y + k_2S_x$, where $k_1$ and $k_2$ are integers that can be computed efficiently, given $x$ and $y$.*

*Proof.* Since $\gcd(x, y) = 1$, the extended Eulcidean algorithm can be used to find integers $k_1$ and $k_2$ such that

$$k_1x + k_2y = 1.$$

It follows that

$$-(xy)^{-1} \equiv -k_1y^{-1} - k_2x^{-1} \pmod{k}.$$

Hence,

$$S_{xy} = -(xy)^{-1}G = -k_1y^{-1}G - k_2x^{-1}G = k_1S_y + k_2S_x.$$

□

### 3.2 Attacks on the Sakazaki-Okamoto-Mambo Scheme

Here is the basic idea of the attacks. If we know enough public keys $I_i$ and their corresponding private keys $S_{I_i}$, then we can construct a database

$$DB := \{(x, S_x)\}$$

for small prime integers $x$, using Lemma 31. For a given public key $I$ (i.e., a hashed identity), suppose that $I$ can be factored as

$$I = x_1 x_2 \dots x_u,$$

where $\gcd(x_i, x_j) = 1$ for all $i \neq j$ and $(x_i, S_{x_i}) \in DB$ for all $i$. Then we can compute the private key

$$S_I = -I^{-1}G$$

using Lemma 32.

We now present two attacks on the scheme that use this idea. The first is an attack on a specific pre-chosen identity. The second attack is more general, but less efficient. We will suppose that the length of a hashed identity, say $I$, is 160 bits. Let $t$ be a positive integer. A positive integer $m$ is *$t$-smooth* if all the prime divisors of $m$ are less than $t$. (Typically we will choose $t = 2^{40}$.)

Algorithm 1 is a forgery of a private key $S_I$ corresponding to a specific public key $I$ (where $I$ is the hash value of the identity information of a user $i$).

**Algorithm 1**

1. Find a $t$-smooth hashed identity $I = p_1 p_2 \dots p_u$, where the $p_i$'s are distinct primes.
2. Find a set of hashed identities $I_1, \dots, I_v$ such that, for every $i$ with $1 \leq i \leq u$, there exists an $I_j$ with $1 \leq j \leq v$ such that $p_i | I_j$. (Clearly we can assume $v \leq u$.)
3. For every $j$ with $1 \leq j \leq v$, obtain a private key $S_{I_j}$ corresponding to public key $I_j$ by interacting with the Center.
4. For every $i$ with $1 \leq i \leq u$, compute $S_{p_i}$ using Lemma 31.
5. Construct $S_I$ from the $u$ pairs $(p_i, S_{p_i})$ by repeatedly applying Lemma 32.

In Algorithm 1, we build a database that allows us to forge a specific secret key. Algorithm 2 consructs a large database that will allow various secret keys to be forged. More precisely, a secret key can be forged using Algorithm 2 for a hashed identity $I$ whenever $I$ is $t$-smooth and square-free.

**Algorithm 2**

1. Find a set of hashed identities $I_1, \dots, I_w$ such that, for every prime $p < t$, there exists an $I_j$ with $1 \leq j \leq w$ such that $p | I_j$.

2. For $1 \leq j \leq w$, obtain a private key $S_{I_j}$ corresponding to public key $I_j$ by interacting with the Center.
3. For all primes $p < t$, compute $S_p$ using Lemma 31.
4. Let $I = p_1 p_2 \ldots p_u$ be a $t$-smooth hashed identity, where the $p_i$'s are distinct primes.
5. Construct $S_I$ from the $u$ pairs $(p_i, S_{p_i})$ by repeated applying Lemma 32.

### 3.3 Analysis of the Complexity of the Attacks

In this section, we analyze the complexity of the attacks. First, we need some results on smoothness probabilities. Let $\Psi(x,t)$ denote the number of integers in the interval $[1,x]$ which are $t$-smooth. The notation "log" is used to denote a logarithm to the base $e$. The following result can be found in [1, p. 234].

**Theorem 33** *For $x \geq 4$ and $2 \leq t \leq x$, it holds that $\Psi(x,t) > x^{1-\log\log x/\log t}$.*

If we take $t = x^\alpha$, where $0 < \alpha < 1/2$, then $\Psi(x,t) > x/(\log x)^{1/\alpha}$. Then the probability that a random integer in $[1,x]$ is $t$-smooth is at least $1/(\log x)^{1/\alpha}$. When $x = 2^{160}$ and $t = 2^{40}$, we have $\alpha = 1/4$, and the probability is at least

$$\frac{1}{(160\log 2)^4} = \frac{1}{1.5 \times 10^8} > \frac{1}{2^{28}} \qquad (*).$$

(In practice, however, the probability is much larger than this. In fact, when $1/2 \leq \alpha \leq 1$, the probability is close to $1 + \log\alpha$; see [2, p. 383].)

We first analyze Algorithm 1.

- Suppose we attempt to construct $I$ in step 1 by choosing random identities, hashing them and testing them to see if they are $t$-smooth. We should find a suitable $I$ after $2^{28}$ trials. Assuming that $I = p_1 p_2 \ldots p_u$ is square-free, we proceed to step 2.
- In step 2, we might choose random identities, hash them and test them for divisibility by the $p_i$'s. The probability that a random integer is divisible by $p_i$ is $1/p_i$, so it will take about $p_i$ trials to find a hashed identity divisible by $p_i$, for each $i$. The total number of trials will be about $p_1 + \ldots + p_u$. It is not hard to see that the number of trials is maximized when $u = 4$ and $p_1, p_2, p_3, p_4 \approx 2^{40}$. The number of trials in the worst case is therefore expected to be about $4 \times 2^{40} \approx 2^{42}$.
- In step 3, we require $u$ interactions with the Center to obtain the $S_{p_i}$'s, $1 \leq i \leq u$. In the worst case, we will have $u = 30$, because the product of the first 31 primes exceeds $2^{160}$.
- Finally, step 4 can be done quickly using $u - 1$ applications of Lemma 32.

In practice, the most time-consuming step is probably step 1. This is because the values $I$ in step 1 need to be checked for divisibility by all the primes up

to $2^{40}$. In step 2, we are only testing for divisibility by the $p_i$'s determined in step 1.

This attack is sufficient to cast doubt on the security of the Sakazaki-Okamoto-Mambo scheme if the length of ahashed identity is 160 bits.

Algorithm 2 can be analyzed in a similar fashion. Let $\pi(x)$ denote the number of primes that are less than $x$. (By the prime number theorem, $\pi(x) \approx x/\log x$.) Unfortunately, in step 2 of Algorithm 2, we need to construct a database of $\pi(2^{40})$ keys. This is so large that it is not really practical.

## 4 Summary

The attack presented in Algorithm 1 is at least close to being practical in the case where a hashed identity is 160 bits in length. After requesting a small number of (private) keys from the Center, our attack allows a new valid key to be constructed without any further interaction with the Center. This shows that it is not sufficient for the hash function to be "secure" in order for the Sakazaki-Okamoto-Mambo scheme to be secure. Our attack also works (although is not as effective) if the bitlength of the hash function output is large, e.g., 512 bits.

## References

1. E. Bach and J. Shallit. *Algorithmic Number Theory. Volume 1: Efficient Algorithms*, MIT Press, 1996.
2. D. Knuth. *The Art of Computer Programming, Volume 2, Seminumerical Algorithms* (Third Edition), Addison-Wesley, 1998.
3. N. Koblitz. *A Course in Number Theory and Cryptography (Second Edition)*, Springer-Verlag, 1994.
4. National Institute of Standards and Technology. *Secure Hash Standard (SHS)*, FIPS Publication 180-1, 1995.
5. E. Okamoto. Key distribution systems based on identification information, *Crypto'87* 194–202
6. H. Sakazaki, E. Okamato and M. Mambo. ID-based key distribution system over an elliptic curve, *Contemporary Mathematics* **225** (1999), 215–223 (Fourth International Conference on Finite Fields).
7. R. Schoof. Elliptic curves over finite fields and the computation of square roots mod $p$, *Mathematics of Computation*, **44** (1985), 483–494.

# Differential and Linear Distributions of Substitution Boxes for Symmetric-Key Cryptosystems

Peter Roelse

Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, E-mail: p.l.a.roelse@tue.nl

**Abstract.** In many secret-key cryptosystems substitution boxes are the only non-linear component, and provide resistance against both differential and linear cryptanalyis. In this paper the notions of differential and linear distribution of the mapping defined by a substitution box are introduced. These distributions contain considerable information about its resistance against linear and differential cryptanalysis. With a computer search the differential and linear distribution of four-bit permutations with an optimal resistance against the attacks mentioned above were determined. In particular, this shows that Almost Perfect Nonlinear (APN) permutations on four bits do not exist. The substitution-boxes used in the AES finalist Serpent and the construction used for the S-box of AES are compared with these optimal four-bit permutations. In addition, identities on the elements of the differential and linear distribution of a mapping are presented. These relations are used to explain the close connection between the optimal distributions of four-bit permutations that were found by the computer search.

## 1 Introduction

Substitution boxes (S-boxes) are an important building block for symmetric-key cryptosystems. They can be used in non-linear feedback functions of streamciphers, as well as in the round function of iterated block ciphers. Two popular constructions used for iterated block ciphers are Feistel ciphers, e.g. DES, and substitution-linear transformation networks, e.g. Serpent [1] and AES [9].

Two attack methods that can be applied to a wide range of block ciphers are differential and linear cryptanalysis [3], [13]. Although differential and linear cryptanalysis often require an impractically large amount of chosen and known plain-ciphertext pairs respectively, it is good practice to design block ciphers that are resistant against these attacks.

In the examples mentioned above, the round function consists of three different layers; a key addition layer, an S-box layer and a linear transformation layer. The S-box layer consists of a number of S-boxes operating in parallel on the input data, where each S-box provides a 'local' resistance against linear and differential cryptanalysis. In practice, S-boxes defining permutations are of particular interest, especially on either four or eight bits.

The paper is organized as follows. Section 2 contains the definitions and notations used in the paper and introduces the notions of differential and linear distribution of a mapping. Section 3.1 contains the results of the computer search. The differential and linear distribution of permutations on four bits providing an optimal resistance against differential and linear cryptanalysis are presented and compared with the distributions of the S-boxes used in Serpent and AES. Finally, Sect. 4 contains identities on the elements of the differential and linear distribution of a mapping, which are used to explain the close connection between the distributions of the optimal permutations on four bits that were found by the computer search.

## 2 Basic Concepts

In this section some notations and definitions will be given that will be used throughout the paper. The notions of differential and linear distribution of a mapping from $m$ to $m$ bits, each of which represents an S-box, will be introduced in Sect. 2.1 and Sect. 2.2 respectively. In Sect. 2.3, some algebraic manipulations on a mapping that do not change its differential and linear distribution are discussed.

### 2.1 Differential Distribution of a Mapping

The numbers in the following definition are commonly used for measuring the resistance of a mapping against differential cryptanalysis (see e.g. [4]).

**Definition 1.** Let $m \geq 1$. For a mapping $f : \mathbf{F}_2^m \to \mathbf{F}_2^m$ and $a, b \in \mathbf{F}_2^m$, let the numbers $D_f(a, b)$ be defined as

$$D_f(a, b) := \#\{x \in \mathbf{F}_2^m \mid f(x) + f(x + a) = b\}.$$

The mapping $f$ is called *differentially d-uniform* if $D_f(a, b) \leq d$ for all $0 \neq a \in \mathbf{F}_2^m$ and $b \in \mathbf{F}_2^m$.

The table containing the values of $D_f(a, b)$ for all $a, b \in \mathbf{F}_2^m$ is usually called the XOR Distribution Table of $f$ [3]. Note that given an input difference $a$ and assuming a uniform distribution of the input for $f$, the corresponding probability that the output difference equals $b$ is given by $D_f(a, b)/2^m$ for all $a, b \in \mathbf{F}_2^m$. As differential cryptanalysis tries to exploit non-trivial values of $a$, i.e. $a \neq 0$, and values of $b$ (also called differential characteristics and denoted by $a \to b$), with a probability which is as high as possible, the objective for the designer of symmetric-key cryptosystems is to find mappings with a value for $d$ that is as small as possible. It is easily seen that $d \geq 2$, as all the input differences occur in pairs. Mappings for which equality holds are called *almost perfect nonlinear* (APN), as introduced in [14]. For odd values of $m$, classes of APN permutations on $\mathbf{F}_2^m$ are known. However, it is not known whether APN

permutations exist for even values of $m$. The following definition introduces the notion of differential distribution, which will be used for distinguishing mappings that have the same value for $d$.

**Definition 2.** For a mapping $f : \mathbf{F}_2^m \to \mathbf{F}_2^m$, let $D_i^f$ be the number of times $D_f(a,b)$ equals $i$ for all $a, b \in \mathbf{F}_2^m$ with $(a,b) \neq (0,0)$. The vector $D^f := (D_{2^m}^f, D_{2^m-2}^f, \ldots D_2^f, D_0^f)$ is called the *differential distribution* of $f$.

The following two relations on the elements of this distribution follow directly from counting the total number entries in the XOR Distribution Table (1) and from the fact that $\sum_{b \in \mathbf{F}_2^m} D_f(a,b) = 2^m$ for all $a \in \mathbf{F}_2^m$ (2).

$$\sum_{i=0}^{2^{m-1}} D_{2i}^f = 2^{2m} - 1\,, \tag{1}$$

$$\sum_{i=1}^{2^{m-1}} 2iD_{2i}^f = 2^m(2^m - 1)\,. \tag{2}$$

Assume that the differential distributions for all mappings from $m$ to $m$ bits are ordered lexicographically, i.e. for $f, g : \mathbf{F}_2^m \to \mathbf{F}_2^m$, $D^f < D^g$ if and only if $D_i^f < D_i^g$ for the largest value for $0 \leq i \leq 2^m$ ($i$ even) for which $D_i^f \neq D_i^g$. As each of the differential characteristics with a high probability can potentially be exploited in differential cryptanalysis, it is desirable that the differential distribution is as small as possible. The mappings with the smallest possible differential distribution are called *optimal* w.r.t. resistance against differential cryptanalysis.

## 2.2 Linear Distribution of a Mapping

The numbers in the following definition can be used to measure the resistance of a mapping against linear cryptanalysis (see e.g. [4]), and is similar to Definition 1. The operation $\cdot : \mathbf{F}_2^m \times \mathbf{F}_2^m \to \mathbf{F}_2$ denotes the inner product on the vector space $\mathbf{F}_2^m$.

**Definition 3.** Let $m \geq 1$. For a mapping $f : \mathbf{F}_2^m \to \mathbf{F}_2^m$ and $a, b \in \mathbf{F}_2^m$, let the numbers $L_f(a,b)$ be defined as

$$L_f(a,b) := |\#\{x \in \mathbf{F}_2^m \mid a \cdot x = b \cdot f(x)\} - 2^{m-1}|.$$

The mapping $f$ is called *non-linearly $l$-uniform* if $L_f(a,b) \leq l$ for all $a \in \mathbf{F}_2^m$ and $0 \neq b \in \mathbf{F}_2^m$.

The table containing the values of $L_f(a,b)$ for all $a, b \in \mathbf{F}_2^m$ is usually called the Linear Approximation Table of $f$ [2]. Notice that $a$ and $b$ define a linear relation on the input and output bits of $f$ respectively. Assuming a uniform distribution of the input for $f$, linear cryptanalysis tries to exploit non-trivial

equations (i.e. $b \neq 0$) of the form $a \cdot x = b \cdot f(x)$ (also called linear characteristics and denoted by $a \rightarrow b$) which hold with probability having a distance to $1/2$ that is as large as possible. As the distance to $1/2$ is given by the number $L_f(a,b)/2^{m-1}$, the objective for the designer of symmetric-key cryptosystems is to find mappings with a value for $l$ that is as small as possible. For any mapping $f : \mathbf{F}_2^m \rightarrow \mathbf{F}_2^m$, $l \geq 2^{(m-1)/2}$ [8]. Mappings for which equality holds are called *almost bent* (AB), as introduced in [14]. For odd values of $m$, classes of AB permutations on $\mathbf{F}_2^m$ are known. It is clear that AB mappings can only exists for odd values of $m$. For even $m$ the conjectured lower bound is $2^{m/2}$.

The following definition introduces the notion of linear distribution, which will be used for distinguishing mappings that have the same value for $l$.

**Definition 4.** For a mapping $f : \mathbf{F}_2^m \rightarrow \mathbf{F}_2^m$, let $L_i^f$ be the number of times $L_f(a,b)$ equals $i$ for all $a, b \in \mathbf{F}_2^m$ with $(a,b) \neq (0,0)$. The vector $L^f := (L_{2^{m-1}}^f, L_{2^{m-1}-1}^f, \ldots, L_1^f, L_0^f)$ is called the *linear distribution* of $f$.

Assume that the linear distributions for all mappings from $m$ to $m$ bits are ordered lexicographically, i.e. for $f, g : \mathbf{F}_2^m \rightarrow \mathbf{F}_2^m$, $L^f < L^g$ if and only if $L_i^f < L_i^g$ for the largest value for $0 \leq i \leq 2^{m-1}$ for which $L_i^f \neq L_i^g$. As each of the linear characteristics with a high probability can potentially be exploited in linear cryptanalysis, it is desirable that the linear distribution is as small as possible. The mappings with the smallest possible linear distribution are called *optimal* w.r.t. resistance against linear cryptanalysis.

### 2.3 Algebraic Manipulations on Mappings

Different algebraic methods are known for constructing different S-boxes from one mapping $f$, such that each of these S-boxes has the same value for $d$ and $l$ as $f$ (see e.g. [16]). Of particular interest for this paper are compositions of a mapping $f$ with affine bijective mappings. The following lemma shows that each of these mappings has the same distributions as $f$.

**Lemma 1.** *For a mapping $f : \mathbf{F}_2^m \rightarrow \mathbf{F}_2^m$ and $C, D \in GL(m, \mathbf{F}_2)$ and $c, d \in \mathbf{F}_2^m$, let $g : \mathbf{F}_2^m \rightarrow \mathbf{F}_2^m$ be defined by*

$$xC + c \rightarrow f(x)D + d$$

*for all $x \in \mathbf{F}_2^m$. Then $D^g = D^f$ and $L^g = L^f$.*

*Proof.* For the differential distributions, notice that $D_g(a,b) := \#\{x \in \mathbf{F}_2^m \mid f(xC+c)D + d + f(xC+c+a)D + d = b\} = \#\{y \in \mathbf{F}_2^m \mid f(y) + f(y+a) = bD^{-1}\} = D_f(a, bD^{-1})$, since $C$ and $D$ are invertible matrices. From this it follows that $D^g = D^f$.

For the linear distributions, note that $L_g(a,b) = |\#\{x \in \mathbf{F}_2^m \mid a \cdot (xC+c) = b \cdot (f(x)D+d)\} - 2^{m-1}| = |\#\{x \in \mathbf{F}_2^m \mid aC^T \cdot x + a \cdot c = bD^T \cdot f(x) + b \cdot d\} - 2^{m-1}| = L_f(aC^T, bD^T)$. As $C$ and $D$ are invertible matrices, this defines a bijective mapping between the entries of the linear approximation tables of $g$ and $f$. From this it follows that $L^g = L^f$.

Note that in general not every choice for the matrices $C, D$ and the vectors $c, d$ will lead to a different mapping. For example if $f(x) = 0$ for all values of $x$, only the $2^m$ constant mappings can be constructed by using this lemma. Notice also that for a bijective mapping $f : \mathbf{F}_2^m \to \mathbf{F}_2^m$, its inverse $f^{-1}$ will have the same differential and linear distribution as $f$, since $D_f(a,b) = D_{f^{-1}}(b,a)$ and $L_f(a,b) = L_{f^{-1}}(b,a)$ for all $a, b \in \mathbf{F}_2^m$.

## 3 Permutations

In this section S-boxes defining permutations are discussed. These are of particular interest for practical applications. In substitution-linear transformation ciphers, the S-box layer has to be one-to-one mapping to assure unique decryption. In Feistel ciphers, an invertible round function (for every fixed choice of the round key) avoids attacks based on the non-uniformity of the round function. Taking the cryptographic strength, the implementation complexity and the structure of computer CPUs into account, a natural choice for an S-box layer is to use either a number of permutations on either four or eight bits in parallel, like e.g. in Serpent and AES.

### 3.1 Optimal Permutations on Four Bits

In the following example the differential and linear distribution of one particular permutation $f : \mathbf{F}_2^4 \to \mathbf{F}_2^4$ is given.

*Example 1.* Let the mapping $f : \mathbf{F}_2^4 \to \mathbf{F}_2^4$ be defined by the following table, where the entries are given in hexadecimal notation:

| $x$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | a | b | c | d | e | f |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $f(x)$ | 0 | 1 | 9 | e | d | b | 7 | 6 | f | 2 | c | 5 | a | 4 | 3 | 8 |

The values of the numbers $D_f(a,b)$ and $L_f(a,b)$ are given in Table 1 and Table 2 respectively. From these tables it can be seen that $f$ is differentially and non-linearly 4-uniform. Moreover, the non-zero elements of the differential and linear distribution of $f$ are given by $D_4^f = 15$, $D_2^f = 90$, $D_0^f = 150$ and $L_4^f = 30$, $L_2^f = 120$, $L_0^f = 105$ respectively.

As the number of four-bit permutations equals $16! (> 2^{44})$, performing an exhaustive search over all possible permutations in order to identify the ones with an optimal resistance against differential and linear cryptanalysis is a computationally time-consuming task. However, the amount of work can be reduced considerably by the following observations.

**Definition 5.** Two permutations $f, g : \mathbf{F}_2^m \to \mathbf{F}_2^m$, are called equivalent if there exists an $A \in GL(m, \mathbf{F}_2)$ and an $a \in \mathbf{F}_2^m$ such that

$$g(x) = f(x)A + a,$$

for all $x \in \mathbf{F}_2^m$.

**Table 1.** XOR Distribution Table for $f$ from Example 1.

| $a\backslash b$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | a | b | c | d | e | f |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 4 | 0 | 0 | 0 | 0 | 2 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 2 | 0 |
| 2 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 2 | 0 | 4 | 2 | 0 | 2 | 2 | 0 | 2 |
| 3 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 2 | 2 | 0 | 2 | 2 | 2 | 0 | 4 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 0 | 2 | 0 | 2 | 0 | 0 | 4 | 2 | 2 |
| 5 | 0 | 0 | 0 | 0 | 2 | 0 | 2 | 0 | 2 | 2 | 0 | 4 | 2 | 0 | 0 | 2 |
| 6 | 0 | 2 | 0 | 0 | 2 | 2 | 2 | 4 | 0 | 0 | 2 | 0 | 2 | 0 | 0 | 0 |
| 7 | 0 | 2 | 2 | 2 | 0 | 0 | 4 | 2 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| 8 | 0 | 0 | 0 | 2 | 2 | 2 | 0 | 2 | 0 | 0 | 0 | 2 | 0 | 0 | 2 | 4 |
| 9 | 0 | 2 | 4 | 0 | 0 | 2 | 0 | 0 | 0 | 2 | 0 | 0 | 2 | 0 | 2 | 2 |
| a | 0 | 0 | 2 | 2 | 2 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 4 | 2 | 2 | 0 |
| b | 0 | 2 | 0 | 2 | 0 | 4 | 0 | 0 | 2 | 0 | 0 | 2 | 2 | 2 | 0 | 0 |
| c | 0 | 0 | 2 | 2 | 0 | 2 | 2 | 0 | 0 | 2 | 4 | 2 | 0 | 0 | 0 | 0 |
| d | 0 | 2 | 2 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 2 | 2 | 0 | 2 | 0 | 2 |
| e | 0 | 2 | 0 | 4 | 2 | 0 | 0 | 0 | 2 | 2 | 2 | 0 | 0 | 0 | 2 | 0 |
| f | 0 | 0 | 2 | 0 | 2 | 2 | 0 | 2 | 4 | 2 | 0 | 0 | 0 | 2 | 0 | 0 |

**Table 2.** Linear Approximation Table for $f$ from Example 1.

| $a\backslash b$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | a | b | c | d | e | f |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 0 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 2 | 0 | 2 | 4 | 2 | 0 | 2 | 4 |
| 2 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 2 | 0 | 4 | 4 | 0 | 2 | 2 | 2 | 2 |
| 3 | 0 | 2 | 0 | 2 | 2 | 4 | 2 | 4 | 2 | 0 | 2 | 0 | 0 | 2 | 0 | 2 |
| 4 | 0 | 0 | 2 | 2 | 0 | 0 | 2 | 2 | 0 | 0 | 2 | 2 | 4 | 4 | 2 | 2 |
| 5 | 0 | 2 | 2 | 4 | 0 | 2 | 2 | 0 | 2 | 4 | 0 | 2 | 2 | 0 | 0 | 2 |
| 6 | 0 | 0 | 2 | 2 | 2 | 2 | 4 | 0 | 4 | 0 | 2 | 2 | 2 | 2 | 0 | 0 |
| 7 | 0 | 2 | 2 | 4 | 2 | 0 | 0 | 2 | 2 | 0 | 4 | 2 | 0 | 2 | 2 | 0 |
| 8 | 0 | 2 | 0 | 2 | 0 | 2 | 4 | 2 | 0 | 2 | 0 | 2 | 0 | 2 | 4 | 2 |
| 9 | 0 | 0 | 4 | 0 | 0 | 4 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 |
| a | 0 | 2 | 4 | 2 | 2 | 0 | 2 | 4 | 0 | 2 | 0 | 2 | 2 | 0 | 2 | 0 |
| b | 0 | 4 | 0 | 0 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 4 | 0 | 0 | 0 |
| c | 0 | 2 | 2 | 0 | 4 | 2 | 2 | 0 | 0 | 2 | 2 | 4 | 0 | 2 | 2 | 0 |
| d | 0 | 0 | 2 | 2 | 4 | 0 | 2 | 2 | 2 | 2 | 0 | 0 | 2 | 2 | 0 | 4 |
| e | 0 | 2 | 2 | 0 | 2 | 0 | 0 | 2 | 4 | 2 | 2 | 0 | 2 | 0 | 4 | 2 |
| f | 0 | 4 | 2 | 2 | 2 | 2 | 0 | 0 | 2 | 2 | 0 | 0 | 0 | 4 | 2 | 2 |

By Lemma 1, equivalent permutations have the same differential and linear distributions. Moreover, the classes of equivalent permutations partition the set of all possible permutations. Each of these equivalence classes contains a number of permutations that only depends on $m$, and can be represented by one particular element in this class, as shown in the next theorem. In the following, let $\{x_0 := 0, x_1, x_2, \ldots, x_{2^m-1}\}$ be the set containing all elements of $\mathbf{F}_2^m$.

**Theorem 1.** *For each equivalence class, it holds that: (i) the number of elements in an equivalence class equals*

$$2^m \#GL(m, \mathbf{F}_2) = 2^m \prod_{k=0}^{m-1} (2^m - 2^k),$$

*and (ii) each equivalence class contains a unique permutation $h$ for which $h(0) = 0$ and the matrix $\left(h(x_1)^T h(x_2)^T \dots h(x_{2^m-1})^T\right)$ is in (full) row echelon form.*

*Proof.* (i) Note that all mappings in an equivalence class are permutations. Now consider two permutations $g$ and $h$ in the equivalence class of $f$, i.e. $g(x) = f(x)A + a$ and $h(x) = f(x)B + b$ for all $x \in \mathbf{F}_2^m$ and for some $A, B \in GL(m, \mathbf{F}_2)$ and $a, b \in \mathbf{F}_2^m$. Suppose $g(x) = h(x)$ for all $x \in \mathbf{F}_2^m$. This implies that $f(x)A + a = f(x)B + b$ for all $x \in \mathbf{F}_2^m$. As $f$ is a permutation, one $x$ will be mapped to zero, from which it follows that $a = b$. In addition, there will exist $m$ images of $f$ that are linearly independent over $\mathbf{F}_2$, implying that $A = B$. So all possible choices for $A$ and $a$ lead to different permutations.

(ii) Consider the equivalence class of which $f$ is an element. Define the permutation $g(x) := f(x) + f(0)$ for all $x \in \mathbf{F}_2^m$. Then clearly $f$ and $g$ are equivalent and $g(0) = 0$. Note that every choice for $A \in GL(m, \mathbf{F}_2)$ corresponds to one of the $\#GL(m, \mathbf{F}_2)$ permutations $h(x) := g(x)A$ in the equivalence class with $h(0) = 0$. Now find the unique $A \in GL(m, \mathbf{F}_2)$ such that the matrix $A^T \left(g(x_1)^T g(x_2)^T \dots g(x_{2^m-1})^T\right)$ is in (full) row echelon form, i.e. $A$ represents the elementary row operations (i.e. the Gaussian elimination) needed for this transformation. Define $h(x) := g(x)A$ for all $x \in \mathbf{F}_2^m$.

The computer search can now be restricted to computing the differential and linear distributions of the $16!/322560 < 2^{26}$ equivalence class representatives $h$. It turns out that for $m = 4$, permutations with an optimal resistance against differential and linear cryptanalysis are differentially 4-uniform and non-linearly 4-uniform. In particular, this implies that APN-permutations do not exist for $m = 4$. The possible differential and linear distributions for all classes of permutations with $d = l = 4$ are listed in Table 3.

**Table 3.** Differential and linear distributions of permutations on $\mathbf{F}_2^4$ with $d = l = 4$.

| #classes | $D_4^f$ | $D_2^f$ | $D_0^f$ | $L_4^f$ | $L_2^f$ | $L_0^f$ |
|---|---|---|---|---|---|---|
| 709632 | 15 | 90 | 150 | 30 | 120 | 105 |
| 1290240 | 18 | 84 | 153 | 32 | 112 | 111 |
| 322560 | 24 | 72 | 159 | 36 | 96 | 123 |

From this table, it can be seen that only three different differential and linear distributions occur for permutations on $\mathbf{F}_2^4$ with $d = l = 4$. In addition, the permutations with an optimal resistance against differential cryptanalysis

also provide an optimal resistance against linear cryptanalysis, and their differential and linear distribution coincide with the ones of the mapping from Example 1.

*Remark 1 (Serpent S-boxes).* In the AES finalist Serpent, eight different permutations on $\mathbf{F}_2^4$ are used. The first published version of Serpent makes use of the permutations on four bits as used in the DES S-boxes. In the AES proposal, these were replaced by new ones, offering an improvement in the security of the cipher. All of these new permutations are differentially 4-uniform and non-linearly 4-uniform. It is easily checked that four of them have a differential and linear distribution that equals the second one given in Table 3 and the remaining four having the distributions equal to the third one in this table, implying that they are not optimal in the sense of the definition given in Sect. 2. However, additional design criteria were used for the Serpent S-boxes (see also [1]). One of these is an 'avalanche' criterium, which states that a one bit input difference may not cause a one bit output difference. A computer search confirmed that optimal four-bit permutations satisfying this additional criterium do not exist.

### 3.2 A Construction Based on Finite Fields

For $m > 4$, a computer search like that in the previous subsection is computationally infeasible. Therefore construction methods are desirable in this case. In [15], a construction method for S-boxes based on a mapping in the finite field $\mathbf{F}_{2^m}$ is given. By selecting a basis of $\mathbf{F}_{2^m}$ over $\mathbf{F}_2$ and representing the field as an m-dimensional vector space over $\mathbf{F}_2$, this mapping defines a mapping from $\mathbf{F}_2^m$ into itself. Note that the differential and linear distributions of this mapping do not depend on the particular choice for the basis.

**Theorem 2 (Nyberg).** *The 'inversion' mapping $f : \mathbf{F}_{2^m} \to \mathbf{F}_{2^m}$ defined by*

$$f(x) := x^{2^m-2}$$

*is differentially 2-uniform (i.e. APN) if $m$ is odd and differentially 4-uniform if $m$ is even. The mapping is non-linearly $2^{m/2}$-uniform.*

The value for the non-linearity in this theorem is an upper bound, however, small values of $m$ indicate that this is the smallest possible value for $l$ if $m$ is even. The proof for the non-linearity is based on the Carlitz-Uchiyama bound for exponential sums [7], see also [6] for the relationship between the non-linearity and the number of points on a hyperelliptic curve.

If $m$ is even, then the non-zero elements of the differential distribution of the 'inversion' mapping are given by $D_4^f = 2^m - 1$, $D_2^f = 2^{2m-1} - 2^{m+1} - 2^{m-1} + 2$ and $D_0^f = 2^{2m-1} + 2^m + 2^{m-1} - 2$. This follows directly from the proof of Nyberg's theorem given in [15]. In this proof it is shown that for $a \neq 0$ (implying that also $b \neq 0$, as $f$ is a permutation), the equation

$f(x) + f(x + a) = b$ has exactly four solutions in $\mathbf{F}_{2^m}$ if and only if $a = b^{-1}$ and at most two solutions otherwise. This implies that each non-trivial row and column in the XOR Distribution Table of $f$ will contain exactly one four (i.e. $D_4^f = 2^m - 1$) and all other entries will be zero or two. The values of $D_2^f$ and $D_0^f$ follow from (2) and (1) respectively.

*Remark 2 (AES S-box).* AES uses one permutation on $\mathbf{F}_2^8$. This permutation is based on Nyberg's construction method. This 'inversion' mapping in the finite field $\mathbf{F}_{2^8}$ is composed with an invertible affine mapping on the output, i.e. the S-box is equivalent to the 'inversion' mapping in the sense of Definition 5. The reason for the affine transformation on the output is to avoid a simple mathematical description of the mapping over $\mathbf{F}_{2^8}$, which could potentially be exploited in cryptanalysis, e.g. by applying an interpolation attack. From the discussion above, it follows that for $m = 4$ this construction would yield an S-box with an optimal resistance against both differential and linear cryptanalysis (the permutation from Example 1 is based on the inversion mapping in the field $\mathbf{F}_2[x]/(x^4 + x + 1)$).

## 4 Connections between the Distributions

This section contains connections between the differential and linear distributions of a mapping. In Sect. 4.1 three more relations on the elements of these distributions will be given (in addition to (1) and (2)). In Sect. 4.2, the special cases of APN mappings and differentially 4-uniform mappings will be discussed, and the close connection between the distributions in Table 3 will be explained by using the relations on the elements of their distributions.

### 4.1 Identities on the Differential and Linear Distributions

The identities on the elements of the differential and linear distribution of a mapping will be derived by using results from binary linear error-correcting codes. Such a code $\mathcal{C}$ of block length $n$ is a linear subspace of $\mathbf{F}_2^n$. If the dimension of this subspace is $k$ and the minimum Hamming distance between any two distinct codewords equals $d$, then $\mathcal{C}$ is called an $[n, k, d]$ code. A linear code can be represented by a generator matrix $G \in \mathbf{F}_2^{k \times n}$, for which the rows form a basis for $\mathcal{C}$, i.e. $\mathcal{C} = \{mG \mid m \in \mathbf{F}_2^k\}$. Alternatively, the code can be represented by a parity check matrix $H \in \mathbf{F}_2^{(n-k) \times n}$ of rank $n - k$, i.e. $\mathcal{C} = \{x \in \mathbf{F}_2^n \mid Hx^T = 0\}$. The code of length $n$ and dimension $n - k$ for which $H$ is a generator matrix is called the dual code of $\mathcal{C}$ and is denoted by $\mathcal{C}^\perp$. Let $A_i$ denote the number of words of Hamming weight $i$ $(0 \leq i \leq n)$, then $(A_0, A_1, \ldots, A_n)$ is called the weight distribution of $\mathcal{C}$. The polynomial $A(z) = \sum_{i=0}^{n} A_i z^i$ is called the weight enumerator of $\mathcal{C}$. The weight distribution and enumerator of $\mathcal{C}^\perp$ will be denoted by $(B_0, B_1, \ldots, B_n)$ and $B(z)$ respectively. The connection between $A(z)$ and $B(z)$ is given by the

MacWilliams identity $B(z) = 2^{-k}(1+z)^n A((1-z)/(1+z))$. Related to the MacWilliams identity are the Pless power moment identities [17]. For more detailed information about the theory of error-correcting codes, the reader is referred to [10], [11] or [12].

As in Sect. 3, let $\{x_0 := 0, x_1, x_2, \ldots, x_{2^m-1}\}$ be the set containing all elements of $\mathbf{F}_2^m$. For a mapping $f : \mathbf{F}_2^m \to \mathbf{F}_2^m$ with $f(0) = 0$ define the binary $[n = 2^m, m+1 \leq k \leq 2m+1, d]$ code $\mathcal{C}_f$ by the generator matrix

$$G_f := \begin{pmatrix} 1 & 1 & & \ldots & 1 \\ 0 & x_1^T & x_2^T & \ldots & x_{2^m-1}^T \\ 0 & f(x_1)^T & f(x_2)^T & \ldots & f(x_{2^m-1})^T \end{pmatrix}.$$

The following lemma describes the connection between the linear distribution of $f$ and the weight distribution of $\mathcal{C}_f$.

**Lemma 2.** *The weight distribution of $\mathcal{C}_f$ equals*

$$2^{k-2m-1} \cdot (2^{2m-k+1}, L^f_{2^{m-1}-1}, \ldots, L^f_2, L^f_1, 2L^f_0, L^f_1, L^f_2, \ldots, L^f_{2^{m-1}-1}, 2^{2m-k+1})$$

*and $L^f_{2^{m-1}} = 2^{2m-k+1} - 1$.*

*Proof.* For $a, b \in \mathbf{F}_2^m$, define the codewords $c_{a,b}$ and $c^*_{a,b}$ as $c_{a,b} := (0, a, b)G_f$ and $c^*_{a,b} := (1, a, b)G_f$ respectively, i.e.

$$\begin{aligned} c_{a,b} &= (0, a \cdot x_1 + b \cdot f(x_1), \ldots, a \cdot x_{2m-1} + b \cdot f(x_{2m-1}))\,, \\ c^*_{a,b} &= (1, 1 + a \cdot x_1 + b \cdot f(x_1), \ldots, 1 + a \cdot x_{2m-1} + b \cdot f(x_{2m-1}))\,. \end{aligned}$$

If $w_H(c)$ denotes the Hamming weight of $c$, then from Def. 3 it follows that

$$L_f(a,b) := \mid \#\{i \mid a \cdot x_i + b \cdot f(x_i) = 0\} - 2^{m-1} \mid = \mid 2^{m-1} - w_H(c_{a,b}) \mid .$$

and $L_f(a,b) = \mid w_H(c^*_{a,b}) - 2^{m-1} \mid$. Further, note that $c_{a,b} \neq c^*_{a,b}$, as they differ in the first coordinate and that $w_H(c_{a,b}) = 2^m - w_H(c^*_{a,b})$. If $k = 2m+1$, then every choice for $a$ and $b$ corresponds to two unique codewords $c_{a,b}$ and $c^*_{a,b}$, one of which has Hamming weight $2^{m-1} - L_f(a,b)$ and one with Hamming weight $2^{m-1} + L_f(a,b)$. This implies that the weight distribution of $\mathcal{C}_f$ equals

$$(1, L^f_{2^{m-1}-1}, \ldots, L^f_2, L^f_1, 2L^f_0, L^f_1, L^f_2, \ldots, L^f_{2^{m-1}-1}, 1)$$

and that $L^f_{2^m} = 0$, as only the trivial choice $a = b = 0$ leads to a codeword of weight zero and a word of weight one, but this choice is excluded in the definition of the linear distribution. If $k < 2m+1$, then each codeword in $\mathcal{C}_f$ is counted $2^{2m+1-k}$ times as $a$ and $b$ range over all possible elements in $\mathbf{F}_2^m$.

Note that if $f$ is a permutation with $m > 1$, as in Example 1, all rows in the generator matrix $G_f$ will have even weight, implying that all codewords will be of even weight. In particular, this implies that $L^f_i$ will be zero if $i$ is odd.

Moreover, for permutations both $(0,a,0)G_f$ and $(0,0,b)G_f$ will be elements of a simplex code (if the first zero coordinate is deleted), implying that not only $L_f(a,0)=0$ for $a\neq 0$ but also $L_f(0,b)=0$ for $b\neq 0$.

If the mappings $f$ and $g$ are related as in Lemma 1, then it follows that

$$G_g=\begin{pmatrix}1 & 0 & 0\\ c^T & C^T & 0\\ d^T & 0 & D^T\end{pmatrix}\cdot G_f,$$

implying that $G_f$ and $G_g$ generate the same code. Note that the fact that $f$ and $g$ have the same linear distribution, as shown in Lemma 1, also follows directly from Lemma 2, as $\mathcal{C}_f$ and $\mathcal{C}_g$ have the same weight distribution.

The following relation on the elements of the linear distribution follows directly from counting the number of entries in the Linear Approximation Table or, alternatively, from counting the number of codewords in $\mathcal{C}_f$:

$$\sum_{i=0}^{2^{m-1}} L_i^f = 2^{2m}-1. \tag{3}$$

In the following theorem a second relation on the elements of the linear distribution is given.

**Theorem 3.** *The elements of* $L^f:=(L^f_{2^{m-1}},L^f_{2^{m-1}-1},\ldots,L^f_1,L^f_0)$ *of a mapping* $f:\mathbf{F}_2^m\to\mathbf{F}_2^m$ *satisfy the equation*

$$\sum_{i=1}^{2^{m-1}} i^2 L_i^f = 2^{2m-2}(2^m-1). \tag{4}$$

*Proof.* If $(A_0,A_1,\ldots,A_{2^m})$ denotes the weight distribution of the $[2^m,k]$ code $\mathcal{C}_f$, then the second Pless power moment on this distribution is given by [17]

$$\sum_{i=0}^{2^m} i^2A_i = 2^{m+k-2}(2^m+1) \Leftrightarrow \sum_{i=1}^{2^m-1} i^2A_i = 2^{m+k-2}(2^m+1)-2^{2m},$$

as $A_{2^m}=1$. On the other hand,

$$2^{2m-k+1}\sum_{i=1}^{2^m-1} i^2A_i = \sum_{i=0}^{2^{m-1}-1}((2^{m-1}-i)^2+(2^{m-1}+i)^2)L_i^f =$$

$$2^{2m-1}\sum_{i=0}^{2^{m-1}-1} L_i^f + 2\sum_{i=1}^{2^{m-1}-1} i^2L_i^f = 2^{2m-1}(2^{2m}-2^{2m-k+1})+2\sum_{i=1}^{2^{m-1}-1} i^2L_i^f,$$

by (3) and the fact that $L^f_{2^{m-1}}=2^{2m-k+1}-1$. Combining these two equations gives

$$\sum_{i=1}^{2^{m-1}-1} i^2L_i^f = 2^{3m-2}-2^{4m-k-1},$$

the result now follows by using $L^f_{2^{m-1}}=2^{2m-k+1}-1$.

The connection between the linear and differential distributions follows from observing the dual code of $\mathcal{C}_f$. Note that all codewords in $\mathcal{C}_f^{\perp}$ have even weight, as the all-one vector is in $\mathcal{C}_f$. Moreover, the minimum distance of the dual code is at least four, as all columns in $G_f$ are distinct and non-zero. Of particular interest are the codewords of weight four in the dual code. Recall that $x_0 = f(x_0) = 0$ and let $0 \leq i < j < k < l \leq 2^m - 1$, then each codeword of weight four in $\mathcal{C}_f^{\perp}$ corresponds to

$$\begin{aligned} x_i + x_j &= x_k + x_l\,, \\ f(x_i) + f(x_j) &= f(x_k) + f(x_l)\,. \end{aligned}$$

Define $x_i + x_j =: a$ and $f(x_i) + f(x_j) =: b$, then it follows that $a \neq 0$ and $D_f(a, b) \geq 4$, as $x_i, x_j, x_k$ and $x_l$ are four different solutions to the equation $f(x) + f(x + a) = b$. These observations can be found in [5], where a slightly different definition of the corresponding code is used, i.e. the first row and column of $G_f$ are deleted, so words of weight three and four of the dual code have to be considered there. With their definition of the code, the authors show that the mapping is APN if and only if the minimum distance of the dual code equals five. From their theorem, it follows directly that with the definition of the code used in this paper, the mapping is APN if and only if the minimum distance of $\mathcal{C}_f^{\perp}$ equals six.

A relation between the elements of the differential and linear distribution of a mapping can be obtained by observing the number of codewords of weight four in $\mathcal{C}_f^{\perp}$. The following lemma describes the connection between the linear distribution of $f$ and the number of codewords of weight four in $\mathcal{C}_f^{\perp}$.

**Lemma 3.** *Let $f : \mathbf{F}_2^m \to \mathbf{F}_2^m$ be a mapping with linear distribution $L^f := (L^f_{2^{m-1}}, L^f_{2^{m-1}-1}, \ldots, L^f_1, L^f_0)$ and let $B_i$ denote the number of codewords of weight $i$ in $\mathcal{C}_f^{\perp}$, then*

$$\sum_{i=1}^{2^{m-1}} i^4 L_i^f = 3 \cdot 2^{2m-1} B_4 + 2^{3m-3}(2^m - 1).$$

*Proof.* Let $(A_0, A_1, \ldots, A_{2^m})$ denote the weight distribution of the $[2^m, k]$ code $\mathcal{C}_f$. As $A_{2^m} = 1$ and $B_3 = 0$, the fourth Pless power moment [17] gives

$$\sum_{i=1}^{2^m-1} i^4 A_i = 2^{k-4}(2^m(2^m+1)(2^{2m} + 5 \cdot 2^m - 2) + 24B_4) - 2^{4m}.$$

By using a similar technique as in the proof of Theorem 3, one obtains

$$2^{2m-k+1} \sum_{i=1}^{2^m-1} i^4 A_i = 2^{4m-3} \sum_{i=0}^{2^{m-1}-1} L_i^f + 3 \cdot 2^{2m} \sum_{i=1}^{2^{m-1}-1} i^2 L_i^f + 2 \sum_{i=1}^{2^{m-1}-1} i^4 L_i^f.$$

The result can now be obtained from combining these two equations and using Eqs. 3, 4 and the fact that $L^f_{2^{m-1}} = 2^{2m-k+1} - 1$.

With this lemma, the following statement follows from counting the number of codewords of weight four in $\mathcal{C}_f^{\perp}$ from the differential distribution.

**Theorem 4.** *Let $f : \mathbf{F}_2^m \to \mathbf{F}_2^m$ be a mapping with differential distribution $D^f$ and linear distribution $L^f$, then the elements of these distributions satisfy*

$$\sum_{i=1}^{2^{m-1}} i^4 L_i^f = 2^{2m-1} \sum_{i=2}^{2^{m-1}} \binom{i}{2} D_{2i}^f + 2^{3m-3}(2^m - 1). \tag{5}$$

*Proof.* By Lemma 3, it is sufficient to show that $3B_4 = \sum_{i=2}^{2^{m-1}} \binom{i}{2} D_{2i}^f$. For all $0 \neq a, b \in \mathbf{F}_2^m$ with $D_f(a,b) =: 2i \geq 2$, denote the $2i$ distinct solutions to the equation $f(x) + f(x+a) = b$ by $x_{j_0}, x_{j_1}, x_{j_2}, x_{j_3} \ldots x_{j_{2i-1}}$. Without loss of generality, assume that $x_{j_{2k}} = x_{j_{2k+1}} + a$ with $j_{2k} < j_{2k+1}$ for $k = 0, 1, \ldots, 2i-2$, i.e.

$$x_{j_0} + x_{j_1} = x_{j_2} + x_{j_3} = \ldots = x_{j_{2i-2}} + x_{j_{2i-1}} = a, \tag{6}$$

$$f(x_{j_0}) + f(x_{j_1}) = f(x_{j_2}) + f(x_{j_3}) = \ldots = f(x_{j_{2i-2}}) + f(x_{j_{2i-1}}) = b. \tag{7}$$

For $0 \neq a, b \in \mathbf{F}_2^m$, define the set $\mathcal{S}_f(a,b)$ of ordered pairs of indices as

$$\mathcal{S}_f(a,b) := \{(j_0, j_1), (j_2, j_3), \ldots, (j_{2i-2}, j_{2i-1})\},$$

and define $S_f(a,b) := \emptyset$ if $D_f(a,b) = 0$ for $0 \neq a, b \in \mathbf{F}_2^m$. From this definition, it is clear that $\#\mathcal{S}_f(a,b) = i$ for all $0 \neq a, b \in \mathbf{F}_2^m$ with $D_f(a,b) = 2i$. For each $0 \neq a, b \in \mathbf{F}_2^m$ with $\#\mathcal{S}_f(a,b) \geq 2$, consider all $\binom{i}{2}$ sets consisting of two distinct elements of $\mathcal{S}_f(a,b)$. From (6) and (7), it follows that any ordered pair $(k,l)$ with $0 \leq k < l \leq 2^m - 1$ can be an element of at most one set $S_f(a,b)$ with $0 \neq a, b \in \mathbf{F}_2^m$, implying that all these sets of two elements for all $0 \neq a, b \in \mathbf{F}_2^m$ are distinct. Moreover, as $D_f(0,b) = 0$ for $a \neq 0$, it follows that the number of such sets equals

$$\sum_{i=2}^{2^{m-1}} \binom{i}{2} \#\{(a,b) \mid \#S_f(a,b) = i, 0 \neq a, b \in \mathbf{F}_2^m\} = \sum_{i=2}^{2^{m-1}} \binom{i}{2} D_{2i}^f. \tag{8}$$

Let the positions of the codewords of $\mathcal{C}_f^{\perp}$ be numbered from zero to $2^m - 1$ from left to right and consider the codewords of weight four in this code. For each of these codewords, let the four positions with a one be given by $0 \leq k_0 < k_1 < k_2 < k_3 \leq 2^m$, and define the three $(= \binom{4}{2}/2)$ possible sets of ordered pairs of indices as

$$\mathcal{T}_1(c) := \{(k_0, k_1), (k_2, k_3)\},$$

$$\mathcal{T}_2(c) := \{(k_0, k_2), (k_1, k_3)\},$$

$$\mathcal{T}_3(c) := \{(k_0, k_3), (k_1, k_2)\}.$$

From this definition it is clear that these $3B_4$ sets are all distinct. Moreover, let $\mathcal{T} =: \{(l_0, l_1), (l_2, l_3)\}$ denote such a set and let $a, b \in \mathbf{F}_2^m$ be defined as $a := x_{l_0} + x_{l_1}$ and $b := f(x_{l_0}) + f(x_{l_1})$. Then it follows that $a \neq 0$ and that $\mathcal{T} \subset S_f(a, b)$, as this set corresponds to a unique codeword of weight four in $\mathcal{C}_f^\perp$ with ones in the positions $l_0, l_1, l_2$ and $l_3$, and from the definition of $\mathcal{C}_f^\perp$ it follows that $x_{l_2} + x_{l_3} = x_{l_0} + x_{l_1} = a$ and $f(x_{l_2}) + f(x_{l_3}) = f(x_{l_0}) + f(x_{l_1}) = b$. On the other hand, every set consisting of two distinct elements of $\#\mathcal{S}_f(a, b)$ with $0 \neq a, b \in \mathbf{F}_2^m$ and $\#\mathcal{S}_f(a, b) \geq 2$ corresponds to a unique $c \in \mathcal{C}_f^\perp$ with $w_H(c) = 4$ by (6), (7) and the definition of $\mathcal{C}_f^\perp$, and consequently equals exactly one of the sets $\mathcal{T}_1(c), \mathcal{T}_1(c)$ or $\mathcal{T}_1(c)$. From this one-to-one correspondence and (8), it follows that $\sum_{i=2}^{2^{m-1}} \binom{i}{2} D_{2i}^f = 3B_4$.

### 4.2 APN and Differentially 4-Uniform Mappings

Equations (1), (2), (3), (4) and (5) define five relations on the elements of the differential and linear distribution of a mapping $f$. In this section the special cases for APN and differentially 4-uniform mappings will be discussed. It is well-known that APN mappings always exist, e.g. the mapping $f : \mathbf{F}_{2^m} \to \mathbf{F}_{2^m}$ defined by $f(x) = x^3$ (see [15]). However, for even values of $m$, no APN permutations are known (note that the mapping $f(x) = x^3$ is a three-to-one mapping in this case). Moreover, the results from Sect. 3 show that APN permutations on four bits do not exist. Permutations on an even number of bits are of primary interest in practice and APN permutations might not exists for any even value of $m$. However, by Nyberg's construction (Theorem 2), differentially 4-uniform permutations exist for all even values of $m$. The following corollary describes these two special cases of Theorem 4.

**Corollary 1.** *Let $f : \mathbf{F}_2^m \to \mathbf{F}_2^m$ be a mapping with differential distribution $D^f$ and linear distribution $L^f$.*

1. *The mapping $f$ is differentially 2-uniform (APN) if and only if*

$$L_{2^{m-1}}^f = 0 \quad \textit{and} \quad \sum_{i=1}^{2^{m-1}-1} i^4 L_i^f = 2^{3m-3}(2^m - 1).$$

2. *The mapping $f$ is differentially 4-uniform if and only if*

$$\sum_{i=1}^{2^{m-1}} i^4 L_i^f = 2^{2m-1} D_4^f + 2^{3m-3}(2^m - 1). \tag{9}$$

*Proof.* The results follow immediately from Theorem 4, except for the fact that $L_{2^{m-1}}^f = 0$ if $f$ is APN. Note that if $f$ is APN, then the dimension of $\mathcal{C}_f$ equals $2m + 1$ (i.e. $L_{2^{m-1}}^f = 0$), as otherwise the dual code would have parameters $[2^m, \geq 2^m - 2m, \geq 6]$. Such a code cannot exist, as shortening a

$[2^m, 2^m - 2m, 6]$ code on the first coordinate with respect to zero would imply the existence of a $[2^m - 1, \geq 2^m - 2m - 1, 6]$-code, which cannot exist (see the proof of Theorem 5 in [5]).

Notice that by (1) and (2), the non-zero elements of the differential distribution of an APN mapping are given by

$$D_2^f = 2^{m-1}(2^m - 1), \ D_0^f = 2^{2m-1} + 2^{m-1} - 1.$$

For differentially 4-uniform mappings, $D_4^f$ can be computed by using (9). The values of $D_2^f$ and $D_0^f$ follow from (1) and (2):

$$D_2^f = 2^{m-1}(2^m - 1) - 2D_4^f, \ D_0^f = 2^{2m} - D_4^f - D_2^f - 1.$$

In particular, this implies that the differential distribution of such mappings is determined completely by its linear distribution.

*Example 2.* Consider the differential and linear distributions of the differentially and non-linearly 4-uniform permutations on $\mathbf{F}_2^4$ as given in Table 3. In this case, (4),(3),(9),(2) and (1) reduce to

$$\begin{aligned} L_2^f &= 240 - 4L_4^f, \\ L_0^f &= 255 - L_4^f - L_2^f, \\ D_4^f &= 2L_4^f + \tfrac{1}{8}L_2^f - 60, \\ D_2^f &= 120 - 2D_4^f, \\ D_0^f &= 256 - D_4^f - D_2^f - 1. \end{aligned}$$

Using these equations and the fact that all these numbers should be non-negative and that also the coefficients of the weight enumerator of $\mathcal{C}_f^\perp$ should be non-negative, the following solutions are found:

$$D_4^f = 3i, \ D_2^f = 120 - 6i, \ D_0^f = 135 + 3i,$$

$$L_4^f = 20 + 2i, \ L_2 = 160 - 8i, \ L_0^f = 75 + 6i,$$

for $0 \leq i \leq 12$. Note that the distributions for $i = 0$ would correspond to an APN permutation, which does not exist by Table 3. From this table it follows that from the other 12 distributions, only the three distributions corresponding to $i = 5, 6$ and 8 occur for permutations. Although the relations do not show why the other solutions cannot occur for permutations, they do explain the close connection between the differential and linear distributions of differentially and non-linearly 4-uniform permutations on $\mathbf{F}_2^4$.

## Acknowledgement

The author is grateful to Berry Schoenmakers for his valuable remarks.

## References


1. Anderson, R., Biham, E. and Knudsen, L. (1998). Serpent: A Proposal for the Advanced Encryption Standard, AES CD-1: Documentation, National Institute of Standards and Technology, Information Technology Laboratory.
2. Biham, E. (1995). On Matsui's Linear Cryptanalysis, Proceedings EUROCRYPT '94 (Ed. De Santis, A.), LNCS 950, Springer-Verlag, pp. 341–355.
3. Biham, E. and Shamir, A. (1991). Differential Cryptanalysis of DES-like Cryptosystems, J. Cryptology **4**.
4. Canteaut, A., Charpin, P. and Dobbertin, H. (1999). A new characterization of almost bent functions, Fast Software Encryption '99 (Ed. Knudsen, L.), LNCS 1636, Springer-Verlag, pp. 186–200.
5. Carlet, C., Charpin, P. and Zinoviev, V. (1998). Codes, Bent Functions and Permutations Suitable for DES-like Cryptosystems, Designs, Codes and Cryptography **15** (2), pp. 125–156.
6. Cheon, J.H., Chee, S. and Park, C. (1999). S-boxes with Controllable Non-Linearity, Advances in Cryptology - Proceedings EUROCRYPT '99 (Ed. Stern, J.), LNCS 1592, pp. 286–294.
7. Carlitz, L. and Uchiyama, S. (1957). Bounds for Exponential Sums, Duke Math. J. **24**, pp. 37–41.
8. Chabaud, F. and Vaudenay, S. (1995). Links between Differential and Linear Cryptanalysis, Advances in Cryptology - Proceedings EUROCRYPT '94 (Ed. De Santis, A.), LNCS 950, Springer-Verlag, pp. 356–365.
9. Daemen, J. and Rijmen, V. (1998). AES proposal: Rijndael, AES CD-1: Documentation, National Institute of Standards and Technology, Information Technology Laboratory.
10. Lidl, R. and Niederreiter, H. (1997). Finite Fields (Second edition), Encyclopedia of Mathematics and its Applications 20. Cambridge: Cambridge University Press.
11. Lint, J.H. van (1992). Introduction to Coding Theory (Second edition), Graduate Texts in Mathematics 86. Springer-Verlag.
12. MacWilliams, F.J. and Sloane, N.J.A. (1977). The Theory of Error-Correcting Codes, North Holland Mathematical Library Vol. 16, North-Holland.
13. Matsui, M. (1994). Linear Cryptanalysis Method for DES Cipher, Advances in Cryptology - Proceedings EUROCRYPT '93 (Ed. Helleseth, T.), LNCS 765, Springer-Verlag, pp. 386–397.
14. Nyberg, K. and Knudsen, L. (1993). Provable Security against Differential Cryptanalysis, Advances in Cryptology - Proceedings CRYPTO '93 (Ed. Stinson, D.R.), LNCS 740, Springer-Verlag, pp. 566–574.
15. Nyberg, K. (1994). Differentially Uniform Mappings for Cryptography, Advances in Cryptology – Proceedings EUROCRYPT '93 (Ed. Helleseth, T.), LNCS 765, Springer-Verlag, pp. 55–76.
16. Nyberg, K. (1994). S-Boxes and Round Functions with Controllable Linearity and Differential Uniformity, Fast Software Encryption, Second International Workshop, LNCS 1008, Springer-Verlag, pp. 111–130.
17. Pless, V. (1963). Power Moment Identities on Weight Distributions in Error-Correcting Codes, Info. and Control, Vol. 6, pp. 147–152.

# Exponential Sums and Lattice Reduction: Applications to Cryptography

Igor E. Shparlinski*

Department of Computing, Macquarie University
Sydney, NSW 2109, Australia
igor@ics.mq.edu.au http://www.comp.mq.edu.au/~igor/

**Abstract.** We describe a rather surprising, yet powerful, combination of two famous number theoretic techniques: bounds of *exponential sums* and *lattice reduction* algorithms. This combination has led to a number of cryptographic applications, helping to make rigorous several heuristic approaches and provides a two edge sword to *defend and attack*. It can be used prove important security results and also to create powerful attacks. The examples of the first group include results about the bit security of the Diffie-Hellman key exchange system, of the Shamir message passing scheme and of the XTR and LUC cryptosystems. The examples of the second group include attacks on the Digital Signature Scheme and its modifications which are provably insecure under certain conditions.

## 1 Introduction and Notation

In this paper we describe how a rather unusual combination of two celebrated number theoretic techniques, namely, bounds of *exponential sums* and *lattice reduction* algorithms, provides a powerful cryptographic tool. It can be applied to both proving several security results and designing new attacks.

For example, it has been used to prove certain bit security results for the Diffie-Hellman key exchange system, for the Shamir message passing scheme and for the XTR and LUC cryptosystems. It has also been used to design provably successful attacks on the Digital Signature Scheme and its modifications, including the Nyberg–Rueppel scheme, which are provably insecure under certain conditions.

Here we explain how these two techniques get together, outline several important applications applications and discuss some open problems on exponential sums which arise in this context and which need to be solved before any further progress in this area can be achieved.

Let $p$ denote a prime number and let $\mathbb{F}_p$ denote the finite field of $p$ elements. For integers $s$ and $m \geq 1$ we denote by $\lfloor s \rfloor_m$ the remainder of $s$ on division by $m$. For a prime $p$ and $\ell > 0$ we denote by $\mathrm{MSB}_{\ell,p}(x)$ any integer $u$ such that

$$\left| \lfloor x \rfloor_p - u \right| \leq p/2^{\ell+1}. \tag{1}$$

* Work supported in part by the Australian Research Council.

Roughly speaking, $\mathrm{MSB}_{\ell,p}(x)$ gives $\ell$ most significant bits of $x$ however this definition is more flexible and suits better our purposes. In particular we remark that $\ell$ in the inequality (1) need not be an integer.

Throughout this paper $\log z$ denotes the binary logarithm of $z > 0$.

The implied constants in symbols '$O$' may occasionally, where obvious, depend on the small positive parameters $\varepsilon$ and are absolute otherwise.

## 2 Hidden Number Problem and Lattices

We start with a certain algorithmic problem, introduced in 1996 by Boneh and Venkatesan [5,6], which seemingly has nothing in common with exponential sums. Namely we consider the following

HIDDEN NUMBER PROBLEM, **HNP**: *Recover a number $\alpha \in \mathbb{F}_p$ such that for many known random $t \in \mathbb{F}_p^*$ we are given $\mathrm{MSB}_{\ell,p}(\alpha t)$ for some $\ell > 0$.*

The paper [5] also contains a polynomial time algorithm to solve this problem (with $\ell$ of order $\log^{1/2} p$). The most important ingredient of this algorithm is lattice reduction.

We briefly review a few results and definitions. For general references on lattice theory and its important cryptographic applications, we refer to the recent surveys [34,35].

Let $\{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$ be a set of linearly independent vectors in $\mathbb{R}^s$. The set of vectors

$$L = \left\{ \sum_{i=1}^{s} n_i \mathbf{b}_i \mid n_i \in \mathbb{Z} \right\},$$

is called an $s$-dimensional full rank lattice. The set $\{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$ is called a *basis* of $L$, and $L$ is said to be spanned by $\{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$.

One of the most fundamental problems in this area is the *closest vector problem*, **CVP**: given a basis of a lattice $L$ in $\mathbb{R}^s$ and a target vector $\mathbf{u} \in \mathbb{R}^s$, find a lattice vector $\mathbf{v} \in L$ which minimizes the Euclidean norm $\|\mathbf{u} - \mathbf{v}\|$ among all lattice vectors. It is well know that **CVP** is **NP**-hard (see [34,35] for references). However, its approximate version [2] admits a polynomial time algorithm which goes back to the lattice basis reduction algorithm of Lenstra, Lenstra and Lovász [22].

It has been remarked in Section 2.1 of [29] and then in Section 2.4 of [34] and Section 2.4 of [35] that the following statement holds which is somewhat stronger than that usually used in the literature.

**Theorem 1.** *There exists a polynomial time algorithm which, given an $s$-dimensional full rank lattice $L$ and a vector $\mathbf{r} \in \mathbb{R}^s$, finds a lattice vector $\mathbf{v}$ satisfying the inequality*

$$\|\mathbf{v} - \mathbf{r}\| \le 2^{O\left(s \log^2 \log s / \log s\right)} \min \{\|\mathbf{z} - \mathbf{r}\|, \quad \mathbf{z} \in L\}.$$

*Proof.* The statement is a combination of Schnorr's modification [38] of the lattice basis reduction algorithm of Lenstra, Lenstra and Lovász [22] with a result of Kannan [16] about reduction of the **CVP** to the approximate shortest vector problem.

One can also use a probabilistic analogue [1] of Theorem 1 which gives a slightly better constant.

We are now prepared to sketch the main ideas of [5] to solve the **HNP**. Let $d \ge 1$ be integer. Given $t_i$, $u_i = \mathrm{MSB}_{\ell,p}(\alpha t_i)$, $i = 1, \ldots, d$, we build the lattice $\mathcal{L}(p, \ell, t_1, \ldots, t_d)$ spanned by the rows of the matrix:

$$\begin{pmatrix} p & 0 & \ldots & 0 & 0 \\ 0 & p & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & 0 & \ldots & p & 0 \\ t_1 & t_2 & \ldots & t_d & 1/2^{\ell+1} \end{pmatrix}.$$

and notice

$$\mathbf{w} = (\lfloor \alpha t_1 \rfloor_p, \ldots, \lfloor \alpha t_d \rfloor_p, \alpha/2^{\ell+1}) \in \mathcal{L}(p, \ell, t_1, \ldots, t_d).$$

This vector is very close to the *known* vector $\mathbf{u} = (u_1, \ldots, u_d, 0)$ (at the distance of order $p2^{-\ell}$). Thus applying one of the lattice reduction algorithms one can *hope* to recover $\mathbf{v}$ and thus $\alpha$. In order to make this algorithm *rigorous* one needs to show that (for almost all choices of $t_1, \ldots, t_d \in \mathbb{F}_p$ there is no other lattice vector which is close to $\mathbf{u}$. Namely, taking into account the "stretching" factor in the algorithm of Lemma 1, we have to show that there are very few $d$-tuples $(t_1, \ldots, t_d) \in \mathbb{F}_p^d$ for which the lattice $\mathcal{L}(p, \ell, t_1, \ldots, t_d)$ has a vector $\mathbf{v} \ne \mathbf{w}$ and such that

$$\|\mathbf{v} - \mathbf{u}\| \le p2^{-\ell} \exp\left(O\left(\frac{d \log^2 \log d}{\log d}\right)\right).$$

The last inequality implies that

$$\|\mathbf{v} - \mathbf{w}\| \le p2^{-\ell} \exp\left(O\left(\frac{d \log^2 \log d}{\log d}\right)\right) \tag{2}$$

which is our main tool.

Any vector $\mathbf{v} \in \mathcal{L}(p, \ell, t_1, \ldots, t_d)$ is of the form

$$\mathbf{w} = \left(\beta t_1 - \lambda_1 p, \ldots, \beta t_d - \lambda_d p, \beta/2^{\ell+1}\right),$$

with some integers $\beta$ and $\lambda_1, \ldots, \lambda_d$. Thus (2) implies that for all $i = 1, \ldots, d$ we have

$$(\alpha - \beta) t_i \equiv y_i \pmod{p} \tag{3}$$

for some $y_i \in [-h, h]$ where

$$h = p2^{-\ell} \exp\left(O\left(\frac{d \log^2 \log d}{\log d}\right)\right).$$

The probability

$$\Pr_{y \in \mathbb{F}_p} [\gamma t = y \pmod p \quad y \in [-h, h]] \le \frac{2h+1}{p} \tag{4}$$

for any $\gamma \neq 0$.

Therefore the probability $P$ that (3) holds for all $i = 1, \ldots, d$ and at least one $\beta \neq \alpha$, is at most

$$P \le (p-1)\left(\frac{2h+1}{p}\right)^d \le p(3h/p)^d = 2^{-\ell d} \exp\left(O\left(\frac{d^2 \log^2 \log d}{\log d}\right)\right).$$

Thus if

$$\ell = \left\lceil C \frac{\log^{1/2} p \log\log\log p}{\log\log p} \right\rceil \qquad \text{and} \qquad d = 2\left\lceil \frac{\log p}{\ell} \right\rceil$$

with some absolute constant $C > 0$ then the lattice reduction algorithm returns $\mathbf{v}$ with probability exponentially close to 1.

## 3 Extended Hidden Number Problem, Lattices and Exponential Sums

It has turned out that for many applications, including some results about the so-called bit security of Diffie-Hellman, Shamir, XTR and some other cryptosystems [12,13,27,40–42] and rigorous results on attacks (following the heuristic arguments of [15,31]) on the DSA and DSA-like signature schemes [10,32,33], the condition that $t$ is selected uniformly at random from $\mathbb{F}_p$ is too restrictive.

It has been systematically exploited in the papers [10,12,13,27,32,33,40–42] that the method of [5] can be extended to the case where $t$ is selected from a sequence $\mathcal{T}$ having some uniformity of distribution property.

Accordingly, we consider the following:

EXTENDED HIDDEN NUMBER PROBLEM, **EHNP**: *Recover a number $\alpha \in \mathbb{F}_p$ such that for many known random $t \in \mathcal{T}$ we are given $\mathrm{MSB}_{\ell,p}(\alpha t)$ for some $\ell > 0$.*

If $\mathcal{T} = \mathbb{F}_p$ then rather simple counting arguments of Section 2 show that the number of $d$-tuples $(t_1, \ldots, t_d) \in \mathbb{F}_p^d$ for which the algorithm of Lemma 1

returns a false vector is exponentially small. However for other sequences $\mathcal{T}$ one needs a result about the uniformity of distribution of $\mathcal{T}$.

In the quantitative form which is based on best known lattice reduction algorithms [1,2,16,17,22,34,35,38] this has been obtained in [32].

Recall that the *discrepancy* of an $N$-element sequence $\Gamma = \{\gamma_1, \ldots, \gamma_N\}$ of elements of the interval $[0, 1]$ is defined as

$$\mathcal{D}(\Gamma) = \sup_{J \subseteq [0,1]} \left| \frac{A(J, N)}{N} - |J| \right|,$$

where the supremum is extended over all subintervals $J$ of $[0, 1]$, $|J|$ is the length of $J$, and $A(J, N)$ denotes the number of points $\gamma_n$ in $J$ for $0 \le n \le N - 1$.

We say that a finite sequence $\mathcal{T}$ of integers is *$\Delta$-homogeneously distributed modulo $p$* if for any integer $a$, with $\gcd(a, p) = 1$ the discrepancy of the sequence $\{\lfloor at \rfloor_p / p\}_{t \in \mathcal{T}}$ is at most $\Delta$.

In this case the arguments of Section 2 go through with only one change, namely (4) becomes

$$\Pr_{y \in \mathcal{T}} [\gamma t = y \pmod p \quad y \in [-h, h]] \le \frac{2h + 1}{p} + \Delta.$$

This leads to the following result from [32] which extends the algorithm of [5] to the **EHNP** with a general sequence $\mathcal{T}$.

**Theorem 2.** *For a prime $p$, define $\ell = \lceil \log^{1/2} p \rceil + \lceil \log \log p \rceil$, and $d = 2 \left\lceil \log^{1/2} p \right\rceil$. Let $\mathcal{T}$ be a $2^{-\log^{1/2} p}$-homogeneously distributed modulo $p$ sequence of integer numbers. There exists a deterministic polynomial time algorithm $\mathcal{A}$ such that for any fixed integer $\alpha$ in the interval $[0, p - 1]$, given a prime $p$ and $2d$ integers*

$$t_i \qquad \text{and} \qquad u_i = \mathrm{MSB}_{\ell, p}(\alpha t_i), \qquad i = 1, \ldots, d,$$

*its output satisfies for sufficiently large $p$*

$$\Pr_{t_1, \ldots, t_d \in \mathcal{T}} [\mathcal{A}(p, t_1, \ldots, t_d; u_1, \ldots, u_d) = \alpha] \ge 1 - 2^{-(\log p)^{1/2} \log \log p}$$

*if $t_1, \ldots, t_d$ are chosen uniformly and independently at random from the elements of $\mathcal{T}$.*

It follows from Corollary 3.11 of [36], that $\mathcal{T}$ is $\Delta$-homogeneously distributed modulo $p$ with

$$\Delta = O\left( \frac{\log p}{\#\mathcal{T}} \max_{c=1,\ldots,p-1} \left| \sum_{t \in \mathcal{T}} \exp(2\pi i c t / p)) \right| \right). \tag{5}$$

Therefore, in order to apply this result one can establish the uniformity of distribution of various sequences of $\mathcal{T}$ arising in cryptographic applications and thus one needs to estimate *exponential sums* with elements of $\mathcal{T}$. Thus bounds of exponential sums enter the problem. It has turned out that in some cases relevant exponential sums are well studied in number theory, and thus the corresponding cryptographic result follows immediately, for example, see Section 4. On the other hand, in some case the exponential sums are of very unusual structure which has no meaningful number theoretic interpretations and thus they have required special treatment, for example, see Section 5.

## 4 Bit Security of the Diffie–Hellman Secret Key

We recall the problem which underlies the Diffie–Hellman key exchange system: given an element $g$ of order $\tau$ modulo $p$, find an efficient algorithm to recover Diffie–Hellman secret key $K = \lfloor g^{xy} \rfloor_p$ from $\lfloor g^x \rfloor_p$ and $\lfloor g^y \rfloor_p$.

Typically, either $\tau = p - 1$ (thus $g$ is a primitive root) or $\tau = q$, a large prime divisor of $p - 1$.

The size of $p$ and $\tau$ is determined by the present state of art in the *discrete logarithm problem*. Typically, $p$ is at least about 500 bits, $\tau$ is at least about 160 bits.

However after the common DH key $K = \lfloor g^{xy} \rfloor_p$ is established, only a small portion of bits of $K$ will be used as a common key for some pre-agreed *private* key cryptosystem.

Thus a natural question arises: *Assume that finding $K$ is infeasible, is it still infeasible to find certain bits of $K$?*

In 1996, Boneh and Venkatesan [5] found very elegant links between the **EHNP** and the above problem.

Indeed, assume there is an efficient algorithm to find $\ell$ most significant bits of $\lfloor g^{xy} \rfloor_p$ from from $X = \lfloor g^x \rfloor_p$ and $Y = \lfloor g^y \rfloor_p$. Then, given $A = \lfloor g^a \rfloor_p$ and $B = \lfloor g^b \rfloor_p$ one can select a random $u \in [0, \tau - 1]$ one can apply the above algorithm to $A$ and $U = \lfloor Bg^u \rfloor_p$ getting

$$\mathrm{MSB}_{\ell,p}\left(g^{a(b+u)}\right) = \mathrm{MSB}_{\ell,p}\left(\alpha g_a^u\right)$$

where $\alpha = \lfloor g^{ab} \rfloor_p$ and $g_a = g^a$. Thus we have a special case of the **EHNP**. Unfortunately the paper [5] has a minor gap in the proof of Theorem 2 of that paper. It is claimed that if $g$ is a primitive root (that is, if $\tau = p - 1$) then the obtained problem is exactly the **HNP**. However, this is true only if $g_a$ is a primitive root as well, thus if $\gcd(a, p - 1) = 1$.

To fix this gap and to extend the result to the case of $\tau < p - 1$, Gonzalez Vasco and Shparlinski [12] have used the following bounds of exponential sums from [21].

**Theorem 3.** *Let $Q$ be a sufficiently large integer. The following statement holds with $\vartheta = 1/3$ for all primes $p \in [Q, 2Q]$, and with $\vartheta = 0$ for all primes $p \in [Q, 2Q]$ except at most $Q^{5/6+\varepsilon}$ of them. For any $\varepsilon > 0$ there exists $\delta > 0$ such that for any $g \in \mathbb{F}_p$ of multiplicative order $T \ge p^{\vartheta+\varepsilon}$ the bound*

$$\max_{\gcd(a,p)=1} \left| \sum_{x=0}^{T-1} \exp\left(2\pi i a g^x / p\right) \right| = O\left(Tp^{-\delta}\right)$$

*holds.*

Using (5) we see that under the conditions of Theorem 3 the sequence $g^x$, $x = 0, \ldots, T-1$, is $p^{-\delta}$-homogeneously distributed modulo $p$.

Combining this result with the above arguments and Theorem 2, one can obtain the following statement about the bit security of the Diffie–Hellman secret key.

For each integer $\ell \ge 1$ define the oracle $\mathcal{DH}_\ell$ as an 'black box' which given the values of $X = \lfloor g^x \rfloor_p$ and $Y = \lfloor g^y \rfloor_p$ outputs the value of $\mathrm{MSB}_{\ell,p}(g^{xy})$.

**Theorem 4.** *Let $Q$ be a sufficiently large integer. The following statement holds with $\vartheta = 1/3$ for all primes $p \in [Q, 2Q]$, and with $\vartheta = 0$ for all primes $p \in [Q, 2Q]$ except at most $Q^{5/6+\varepsilon}$ of them. Let $k = \left\lceil \log^{1/2} p \right\rceil + \lceil \log\log p \rceil$. For any $\varepsilon > 0$, sufficiently large $p$ and any element $g \in \mathbb{F}_p^*$ of multiplicative order $T \ge p^{\vartheta+\varepsilon}$, there exists a probabilistic polynomial time algorithm which for any pair $(a, b) \in [0, T-1]^2$, given the values of $A = \lfloor g^a \rfloor_p$ and $B = \lfloor g^b \rfloor_p$, makes $O\left(\log^{1/2} p\right)$ calls of the oracle $\mathcal{DH}_k$ and computes $\lfloor g^{ab} \rfloor_p$ correctly with probability $1 + O\left(2^{-\log^{1/2} p}\right)$.*

## 5 Attack on the Digital Signature Algorithm

On the other hand, in some cases the corresponding exponential sums are new and require a separate study. For example, in [32] the sequence arising in the attack on the Digital Signature Algorithm (DSA) has been studied. We recall the DSA settings. Assume that $q$ and $p$ are primes with $q | p-1$ and that $g \in \mathbb{F}_p$ is a fixed element of multiplicative order $q$. Let $\mathcal{M}$ be the set of messages to be signed and let $h : \mathcal{M} \to \mathbb{F}_q$ be an arbitrary hash-function. They all (that is, $p$, $q$, $g$, $\mathcal{M}$, $h$) are *publicly* known.

The *secret key* is an element $\alpha \in \mathbb{F}_q^*$ which is known only to the signer.

To sign a message $\mu \in \mathcal{M}$, the signer chooses a random integer $k \in \mathbb{F}_q^*$ usually called the *nonce*, and which must be kept secret. We define the following two elements of $\mathbb{F}_q$:

$$r(k) = \left\lfloor \lfloor g^k \rfloor_p \right\rfloor_q, \qquad s(k, \mu) = \left\lfloor k^{-1} \left(h(\mu) + \alpha r(k)\right) \right\rfloor_q.$$

The pair $(r(k), s(k,\mu))$ is the *DSA signature* of the message $\mu$ with a nonce $k$.

The attack on the DSA which has been developed in [31] (and which simplifies and improves the attack from [15]) is based on the solving the **HNP** with the sequence

$$t(k,\mu) = \left\lfloor 2^{-\ell} r(k) s(k,\mu)^{-1} \right\rfloor_q, \qquad (k,\mu) \in \mathcal{S}, \tag{6}$$

where $\mathcal{S}$ is the set of pairs $(k,\mu) \in [1, q-1] \times \mathcal{M}$ with $s(k,\mu) \neq 0$.

Denote by $W$ the number of solutions of the equation $h(\mu_1) = h(\mu_2)$, $\mu_1, \mu_2 \in \mathcal{M}$. Thus $W/|\mathcal{M}|^2$ is probability of collision and expected to be of order $q^{-1}$ for any practically usable hash function.

In [33] the heuristic results of [31] have been made rigorous by proving the following statement.

**Theorem 5.** *Let $Q$ be a sufficiently large integer. The following statement holds with $\vartheta = 1/3$ for all primes $p \in [Q, 2Q]$, and with $\vartheta = 0$ for all primes $p \in [Q, 2Q]$ except at most $Q^{5/6+\varepsilon}$ of them. For any $\varepsilon > 0$ there exists $\delta > 0$ such that for any $g \in \mathbb{F}_p$ of multiplicative order $q \geq p^{\vartheta+\varepsilon}$ the sequence (6) the bound*

$$\max_{\gcd(a,q)=1} \left| \sum_{(k,\mu)\in\mathcal{S}} \exp\left(2\pi i a t(k,\mu)/q\right) \right| = O\left(W^{1/2} q^{3/2-\delta}\right)$$

*holds.*

Using (5) we see that under the conditions of Theorem 5 the sequence (6) is $q^{-\delta/3}$-homogeneously distributed modulo $q$ provided that

$$W \leq \frac{(\#\mathcal{M})^2}{q^{1-\delta}}. \tag{7}$$

This result is based on a combination of the bounds of exponential sums with exponential functions from [21] given in Theorem 3, with the *Weil* bound, see [28] and the Vinogradov method of estimates of double sums. As we have mentioned, the inequality (7) usually holds in the stronger form $W = O\left(|\mathcal{M}|^2/q\right)$.

Then the above arguments together with Theorem 2 imply the following statement.

For an integer $\ell$ we define the oracle $\mathcal{DSA}_\ell$ which, for any given DSA signature $(r(k), s(k,\mu))$, $k \in [0, q-1]$, $\mu \in \mathcal{M}$, returns the $\ell$ least significant bits of $k$.

**Theorem 6.** *Let $Q$ be a sufficiently large integer. The following statement holds with $\vartheta = 1/3$ for all primes $p \in [Q, 2Q]$, and with $\vartheta = 0$ for all primes $p \in [Q, 2Q]$ except at most $Q^{5/6+\varepsilon}$ of them. For any $\varepsilon > 0$ there exists $\delta > 0$*

*such that for any element $g \in \mathbb{F}_p$ of multiplicative order $q$, where $q \ge p^{\vartheta+\varepsilon}$ is prime, and any hash function $h$ satisfying (7), given an oracle $\mathcal{DSA}_\ell$ with $\ell = \left\lceil \log^{1/2} q \right\rceil + \lceil \log\log q \rceil$, there exists a probabilistic polynomial time algorithm to recover the DSA secret key $\alpha$, from $O\left(\log^{1/2} q\right)$ signatures $(r(k), s(k,\mu))$ with $k \in [0, q-1]$ and $\mu \in \mathcal{M}$ selected independently and uniformly at random. The probability of success is at least $1 - 2^{-(\log\log q)\log^{1/2} q}$.*

## 6 Other Applications and Open Questions

The method of the proof of Theorem 4 can be used to establish the bit security of several other exponentiation based cryptographic algorithms. Several such schemes, including the *ElGamal cryptosystem* (see Section 8.4 in [30]) and the *Shamir message passing scheme* (see Protocol 12.22 of [30]), have been outlined in [5,6]. As yet another example we also mention the *Matsumoto–Takachima–Imai key-agreement protocol*, see Section 12.6 of [30]. In fact the treatment of the Shamir message passing scheme in [5] has the same gap as the treatment of the Diffie-Hellman scheme. Accordingly, using exponential sums this gap has been fixed in [12].

In [27] several results on the recently introduced in [23,24] the *XTR cryptosystem* and on the *LUC cryptosystem*, see [3,43]. These results are approximately the same strength as those known for the Diffie-Hellman scheme (however apply only elements of relatively large multiplicative order that those in Theorem 4). These results are also based on bounds of exponential sums, however instead of Theorem 3, a certain bound of exponential sums conjectured by Deligne in [9] (and proved in some special case). In the full generality it has been proved by Katz [18] (the proof follows from Theorem 4.1.1 of [18] after some standard transformations). In fact, for the cases relevant to the XTR and LUC cryptosystems, simpler and more explicit statements are given in [25] and in Chapter 6 of [26]. To be more precise, the approach of [27] makes use of the bound the bounds of the exponential sums

$$\max_{\gamma \in \mathbb{F}_{p^m}^*} \left| \sum_{t \in \mathcal{G}} \exp\left(2\pi i \mathrm{Tr}\left(\gamma t\right)/p\right) \right| \le p^{(m-m/s)/2}$$

where $\mathrm{Tr}(z) = z + z^p + \ldots + z^{p^5}$ is the trace of $z \in \mathbb{F}_{p^m}$ in $\mathbb{F}_p$ and $\mathcal{G}$ is a subgroup of the group of the elements $z \in \mathbb{F}_{p^m}$ with

$$z^{1+p^{m/s}+\ldots+p^{m-m/s}} = 1.$$

which holds for any divisor $s$ of $m$.

The case of the *XTR cryptosystem* corresponds to $m = 6$ and $s = 2$. The case of the *LUC cryptosystem* corresponds to $m = 2$ and $s = 2$.

The result of Theorem 6 has been extended to other DSA-like signature schemes, including the *elliptic curve* version of DSA in [10,33]. In particular, the bound of [20] provides an analogue of Theorem 3 for exponential sums over an orbit generated by a point on an elliptic curve, see [33]. However some interesting questions still remain open. For example, for the *Nyberg–Rueppel* signature scheme the range of $p$ and $q$ in which the results of [10] are nontrivial are narrower than in practical applications. It is shown in [10] that the attack designed in that paper on the Nyberg–Rueppel signature scheme can be reduced to **EHNP** with the sequence of multipliers

$$r(k,\mu) = \left\lfloor \left\lfloor h(\mu) g^k \right\rfloor_p \right\rfloor_q, \qquad (k,\mu) \in [1, q-1] \times \mathcal{M}.$$

Unfortunately it is not clear how to estimate the exponential sums

$$\sum_{\mu \in \mathcal{M}} \sum_{k \in \mathbb{F}_q^*} \exp\left(2\pi i c r(k,\mu)\right), \qquad c \in [1, q-1],$$

and obtaining such a bound is an interesting open question. Using a rather indirect method, it has been shown in [10] that the sequence $r(k,\mu)$ is $2^{-\log^{1/2} q}$ homogeneously distributed modulo $q$, provided that

$$W \le \frac{(\#\mathcal{M})^2 q^{3-\delta}}{p^3}$$

for some $\delta > 0$. We remark that in the settings of the Nyberg–Rueppel signature scheme it is natural to assume that $h$ is bijective, that is, $W = \#\mathcal{M}$. Also, if the message set $\mathcal{M}$ is "dense" (that is, $\#\mathcal{M}$ is of order $p$) then the above result holds for $q \ge p^{2/3+\delta}$. It would be very interesting to lower this bound.

The results and ideas of [32] have recently been used in [7] to design an attack on another DSA-based cryptosystem. It is shown in [7] that in the above cryptosystem there is a way to extract all necessary information from the protocol itself, thus no additional "leakage" is assumed. In fact, Theorem 5 allows us to make the attack of [7] rigorously proved and also to extend it to other small subgroups of $\mathbb{F}_p^*$ (not only those with a power of 2 elements as in [7]).

Yet another modifications of the **HNP** has recently been introduced in [14]. Namely, that paper introduces the following

HIDDEN NUMBER PROBLEM WITH HIDDEN MULTIPLIER, **HNP-HM**: *Recover a number $\alpha \in \mathbb{F}_p$ such that for many unknown random $t \in \mathcal{T}$ we are given* $\mathrm{MSB}_{\ell,p}(\alpha t)$, $\mathrm{MSB}_{\ell,p}(t)$ *and* $\mathrm{MSB}_{\ell,p}(\alpha)$ *for some $\ell > 0$.*

In the case $\mathcal{T} = \mathbb{F}_p^*$ and $\ell \ge (4/5 + \varepsilon) \log p$ the paper [14] provides a polynomial time algorithm for the **HNP-HM**. In fact it also works in more

general residue rings (which is important for applications to [37]). As one can see this result is substantially weaker than those known for **HNP** and **EHNP** where one can take $\ell$ of order $\log^{1/2} p$. However, using exponential sums, it has been shown in [14] that indeed for **HNP-HM** to have a unique solution the value of $\ell$ must be very large. Namely for $\ell \le (1/2 + \varepsilon) \log p$ there can be exponentially many possibilities for $\alpha$.

The aforementioned algorithm has been used in [14] to establish a certain bit security result for the "timed-release crypto" introduced by Rivest, Shamir and Wagner [37] and also to design a "correcting" algorithm for noisy exponentiation black-boxes.

It is an interesting and challenging problem to study **HNP-HM** for more general sequences $\mathcal{T}$, in particular for subgroups of $\mathbb{F}_p^*$.

In the case $\mathcal{T} = \mathbb{F}_p^*$ the paper [6] provides a *non-uniform* polynomial time algorithm for the **HNP** which works with $\ell = O(\log \log p)$. We recall that non-uniformity means that the algorithm exists but to actually design this algorithm one may need exponential time (thus such algorithms are of rather limited value). Nevertheless it would be of interest to extend this result to subgroups of $\mathbb{F}_p^*$. In order to get such a generalisation one needs an analogue of Lemma 2.4 for subgroups and this seems to be a rather feasible task taking into account the bounds of exponential sums of Theorem 3.

Finally, several more modifications of the **HNP** have been considered in the papers [4,11,19,39,44]. However they are of more algebraic than geometric nature and lattices have not been involved in their study.

## Acknowledgement

The author thanks Phong Nguyen for a careful reading of the manuscript and fruitful discussion.

## References

1. M. Ajtai, R. Kumar and D. Sivakumar, 'A sieve algorithm for the shortest lattice vector problem', *Proc. 33rd ACM Symp. on Theory of Comput.*, Crete, Greece, July 6-8, 2001, 601–610.
2. L. Babai, 'On Lovász' lattice reduction and the nearest lattice point problem', *Combinatorica*, **6** (1986), 1–13.
3. D. Bleichenbacher, W. Bosma and A. K. Lenstra, 'Some remarks on Lucas-based Cryptograph', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **963** (1995), 386–396.
4. D. Boneh and I. E. Shparlinski, 'On the unpredictability of bits of the elliptic curve Diffie–Hellman scheme', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **2139** (2001), 201–212.
5. D. Boneh and R. Venkatesan, 'Hardness of computing the most significant bits of secret keys in Diffie–Hellman and related schemes', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1109** (1996), 129–142.

6. D. Boneh and R. Venkatesan, 'Rounding in lattices and its cryptographic applications', *Proc. 8th Annual ACM-SIAM Symp. on Discr. Algorithms*, ACM, NY, 1997, 675–681.
7. D. R. L. Brown and A. J. Menezes, 'A small subgroup attack on a key agreement protocol of Arazi', *Research Report CORR 2001-50*, Faculty of Math., Univ. Waterloo, Waterloo, 2001, 1–5.
8. R. Canetti, J. B. Friedlander, S. Konyagin, M. Larsen, D. Lieman and I. E. Shparlinski, 'On the statistical properties of Diffie–Hellman distributions', *Israel J. Math.*, **120** (2000), 23–46.
9. P. Deligne, *Cohomologie 'etale (SGA* $4\frac{1}{2}$*)*, Lect. Notes in Math., Springer-Verlag, Berlin, **569** (1977).
10. E. El Mahassni, P. Q. Nguyen and I. E. Shparlinski, 'The insecurity of Nyberg–Rueppel and other DSA-like signature schemes with partially known nonces', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **2146** (2001), 97–109.
11. M. I. González Vasco, M. Näslund and I. E. Shparlinski, 'The hidden number problem in extension fields and its applications', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **2286** (2002), 105–117.
12. M. I. González Vasco and I. E. Shparlinski, 'On the security of Diffie–Hellman bits', *Proc. Workshop on Cryptography and Computational Number Theory*, Singapore 1999, Birkhäuser, 2001, 257–268.
13. M. I. González Vasco and I. E. Shparlinski, 'Security of the most significant bits of the Shamir message passing scheme', *Math. Comp.*, **71** (2002), 333–342.
14. N. A. Howgrave-Graham, P. Q. Nguyen and I. E. Shparlinski, 'Hidden number problem with hidden multipliers, timed-release crypto and noisy exponentiation', *Math. Comp.*, (to appear).
15. N. A. Howgrave-Graham and N. P. Smart, 'Lattice attacks on digital signature schemes', *Designs, Codes and Cryptography*, **23** (2001), 283–290.
16. R. Kannan, 'Algorithmic geometry of numbers', *Annual Review of Comp. Sci.*, **2** (1987), 231–267.
17. R. Kannan, 'Minkowski's convex body theorem and integer programming', *Math. of Oper. Research*, **12** (1987), 231–267.
18. N. M. Katz, *Gauss sums, Kloosterman sums, and monodromy groups*, Ann. of Math. Studies, **116**, Princeton Univ. Press, 1988.
19. E. Kiltz, 'A primitive for proving the security of every bit and about universal hash functions & hard core bits', *Preprint*, 2001, 1–19.
20. D. R. Kohel and I. E. Shparlinski, 'Exponential sums and group generators for elliptic curves over finite fields', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1838** (2000), 395–404.
21. S. V. Konyagin and I. Shparlinski, *Character sums with exponential functions and their applications*, Cambridge Univ. Press, Cambridge, 1999.
22. A. K. Lenstra, H. W. Lenstra and L. Lovász, 'Factoring polynomials with rational coefficients', *Mathematische Annalen*, **261** (1982), 515–534.
23. A. K. Lenstra and E. R. Verheul, 'The XTR public key system', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1880** (2000), 1–19.
24. A. K. Lenstra and E. R. Verheul, 'Key improvements to XTR', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1976** (2000), 220–233.
25. W.-C. W. Li,'Character sums and abelian Ramanujan graphs', *J. Number Theory*, **41** (1992), 199–217.
26. W.-C. W. Li, *Number theory with applications*, World Scientific, Singapore, 1996.

27. W.-C. W. Li, M. Näslund and I. E. Shparlinski, 'The hidden number problem with the trace and bit security of XTR and LUC', *Proc. Crypto'02*, Santa Barbara, 2002, Lect. Notes in Comp. Sci., Springer-Verlag, Berlin, (to appear).
28. R. Lidl and H. Niederreiter, *Finite fields*, Cambridge University Press, Cambridge, 1997.
29. D. Micciancio, 'On the hardness of the shortest vector problem', *PhD Thesis*, MIT, 1998.
30. A. J. Menezes, P. C. van Oorschot and S. A. Vanstone, *Handbook of Applied Cryptography*, CRC Press, Boca Raton, FL, 1996.
31. P. Q. Nguyen, 'The dark side of the hidden number problem: Lattice attacks on DSA', *Proc. Workshop on Cryptography and Computational Number Theory*, Singapore 1999, Birkhäuser, 2001, 321–330.
32. P. Q. Nguyen and I. E. Shparlinski, 'The insecurity of the Digital Signature Algorithm with partially known nonces', *J. Cryptology* (to appear).
33. P. Q. Nguyen and I. E. Shparlinski, 'The insecurity of the elliptic curve Digital Signature Algorithm with partially known nonces', *Designs, Codes and Cryptography*, (to appear).
34. P. Q. Nguyen and J. Stern, 'Lattice reduction in cryptology: An update', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1838** (2000), 85–112.
35. P. Q. Nguyen and J. Stern, 'The two faces of lattices in cryptology', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **2146** (2001), 146–180.
36. H. Niederreiter, *Random number generation and quasi–Monte Carlo methods*, SIAM, Philadelphia, 1992.
37. R. L. Rivest, A. Shamir and D. A. Wagner, 'Time-lock puzzles and timed-release crypto', *Preprint*, 1996, 1–9.
38. C. P. Schnorr, 'A hierarchy of polynomial time basis reduction algorithms', *Theor. Comp. Sci.*, **53** (1987), 201–224.
39. I. E. Shparlinski, 'Security of polynomial transformations of the Diffie–Hellman key', *Cryptology ePrint Archive, Report 2000/023*, 2000, 1–9.
40. I. E. Shparlinski, 'Sparse polynomial approximation in finite fields', *Proc. 33rd ACM Symp. on Theory of Comput.*, Crete, Greece, July 6-8, 2001, 209–215.
41. I. E. Shparlinski, 'On the generalised hidden number problem and bit security of XTR', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **2227** (2001), 268–277.
42. I. E. Shparlinski, 'Security of most significant bits of $g^{x^2}$', *Inform. Proc. Letters*, **83** (2002), 109–113.
43. P. J. Smith and C. T. Skinner, 'A public-key cryptosystem and a digital signature system based on the Lucas function analogue to discrete logarithms', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **917** (1995), 357-364.
44. E. R. Verheul, 'Certificates of recoverability with scalable recovery agent security', *Lect. Notes in Comp. Sci.*, Springer-Verlag, Berlin, **1751** (2000), 258–275.

# An Alternate Construction of the Berlekamp Subalgebra

Greg Stein

City University of New York,
300 Jay Street,
Brooklyn, NY 22201
United States

**Abstract.** We give an alternate construction of the Berlekamp subalgebra for cyclotomic polynomials over finite fields. This gives a new, deterministic, polynomial time reduction of factoring these polynomials to knowing the trace of an appropriate root of unity.

In an earlier paper [16] we demonstrated a very constructive technique, using the theory of cyclotomy, for factoring the $r^{th}$ cyclotomic polynomial, $\Phi_r(x)$, over $\mathbb{F}_p$, $r$ and $p$ prime, which was both deterministic and polynomial in time $O(r \log p)$ given the traces of this polynomial (i.e. the traces of the $r^{th}$ roots of unity over $\mathbb{F}_p$). Here we give an algebraic approach, also using ideas from cyclotomy, which will yield a factorization of $\Phi_r(x)$ given a single one of its traces and, in most cases, this will be a complete factorization.

The technique presented here is to factor out an ideal generated by quadratic elements from a polynomial ring in a number of indeterminates over $\mathbb{F}_p$ as well as an explicit isomorphism from this ring onto a subalgebra of $\mathbb{F}_p[x]/\Phi_r(x)$, to whit, the Berlekamp subalgebra. The determination of any zero-divisors in this ring therefore immediately yield a factorization of $\Phi_r(x)$. In the event that we happen to know the trace of some primitive $r^{th}$ root of unity it will turn out that we will immediately know zero-divisors in this ring.

Let $p$ and $r$ be distinct primes, $d$ the multiplicative order of $p$ in $\mathbb{F}_r$, $m = (r-1)/d$, $\alpha$ a generator of $\mathbb{F}_r^*$, and $\zeta$ a primitive $r^{th}$ root of unity in an appropriate extension field of $\mathbb{F}_p$. Recall that $\Phi_r(x)$ factors into the product of $m$ $d^{th}$ degree polynomials, irreducible over $\mathbb{F}_p$, $\Phi_r(x) = \prod_{i=0}^{m-1} g_i(x)$ where

$$g_i(x) = \prod_{k=0}^{d-1} \left( x - \zeta^{\alpha^i p^k} \right) \tag{1}$$

and that the trace of $g_i(x)$ is

$$t_i = \sum_{j=0}^{d-1} \zeta^{\alpha^i p^j} \in \mathbb{F}_p, i \in \mathbb{Z}/m\mathbb{Z}. \tag{2}$$

The theory of cyclotomy gives us the relations

$$t_i t_{i+k} = \sum_{h=0}^{m-1} \left[(k,h) - d\theta_k\right] t_{i+h} \tag{3}$$

where $(k,h)$ is the cyclotomic number defined as the number of solutions to $x+1=y$ where $x$ and $y$ are in the cyclotomic classes $H_k$ and $H_h$ with

$$H_i = \left\{\alpha^i, \alpha^{m+i}, \alpha^{2m+i}, \cdots, \alpha^{(d-1)m+i}\right\} \subseteq \mathbb{F}_r^* \tag{4}$$

and

$$\theta_k = \begin{cases} 1, \ d \text{ even}, \ i = 0 \\ 1, \ d \text{ odd}, \ i = m/2 \\ 0, \text{ otherwise} \end{cases} \tag{5}$$

For further information on cyclotomy and cyclotomic numbers see [2], [3], [16] or [17].

Recall that the Chinese Remainder Theorem gives an isomorphism between $\mathbb{F}_p[x]/\Phi_r(x)$ and the direct sum of splitting fields for $\Phi_r(x)$

$$\varphi : \mathbb{F}_p[x]/\Phi(x) \cong \mathbb{F}_p[x]/g_0(x) \oplus \cdots \oplus \mathbb{F}_p[x]/g_{m-1}(x) \tag{6}$$

by

$$f(x) + (\Phi(x)) \mapsto (f(x) + (g_0(x)), \cdots, f(x) + (g_{m-1}(x))). \tag{7}$$

In order to minimize confusion let us agree to represent an element of $\mathbb{F}_p[x]/\Phi(x)$ by the unique coset representative modulo $\Phi_r(x)$ which is a polynomial of degree less than $r-1$. Similarly, let us agree to represent an element of $\mathbb{F}_p[x]/g_0(x) \oplus \cdots \oplus \mathbb{F}_p[x]/g_{m-1}(x)$ by an $m$-tuple of polynomials whose $i^{th}$ entry is the unique coset representative modulo $g_i(x)$ which is a polynomial of degree less than $d$.

Recalling the relations from the product formula, (3), let $\{T_i \mid i \in \mathbb{Z}/m\mathbb{Z}\}$ be indeterminates and define $R \subset \mathbb{F}_p[T_0, \cdots, T_{m-1}]$ to be the ideal generated by the set

$$\{T_i T_{i+k} - \sum_{h=0}^{m-1} [(k,h) - d\theta_k] T_{i+h} \mid i, j \in \mathbb{Z}/m\mathbb{Z}\} \cup \{1 + T_0 + \cdots + T_{m-1}\} \tag{8}$$

where $(k,h)$ and $\theta_k$ are as defined above. Now we define the quotient ring

$$\mathbb{F}_p[\mathbb{T}] = \mathbb{F}_p[T_0, \cdots, T_{m-1}]/R. \tag{9}$$

We may think of $\mathbb{F}_p[\mathbb{T}]$ as the $m$-dimensional $\mathbb{F}_p$-algebra of homogeneous first degree polynomials in $T_0, \cdots, T_{m-1}$ where the ring action is given by the relations in $R$, relations that are designed to mimic the relations given by the product rule for the periods in the theory of cyclotomy.

Let $\mathbf{K}$ be any splitting field for $\Phi_r(x)$ over $\mathbb{F}_p$ and define the polynomials

$$T_i(x) = \sum_{j=0}^{d-1} x^{\alpha^i p^j} \in \mathbf{K}[x]. \tag{10}$$

For $i, k \in \mathbb{Z}/m\mathbb{Z}$ we compute, for any $l \in \mathbb{Z}/d\mathbb{Z}$,

$$\begin{aligned} T_i(\zeta^{\alpha^k p^l}) &= \sum_{j=0}^{d-1} \left(\zeta^{\alpha^k p^l}\right)^{\alpha^i p^j} \\ &= \sum_{j=0}^{d-1} \left(\zeta^{\alpha^{k+i} p^{l+j}}\right) \\ &= \sum_{j=0}^{d-1} \left(\zeta^{\alpha^{k+i} p^j}\right) = t_{i+k}. \end{aligned}$$

Since the polynomial $T_i(x) - t_{i+k} \in \mathbb{F}_p[x]$ vanishes on all of the roots of $g_k(x) \in \mathbb{F}_p[x]$ there exists a polynomial $h_k(x) \in \mathbb{F}_p[x]$ so that

$$T_i(x) = g_k(x)h_k(x) + t_{i+k}. \tag{11}$$

Now let $\overline{T_i(x)}$ be the projection of $T_i(x) \in \mathbb{F}_p[x]$ into $\mathbb{F}_p[x]/\Phi_r(x)$ and let $\varphi$ be the isomorphism in (6). Then, for $i \in \mathbb{Z}/m\mathbb{Z}$ we have

$$\varphi(\overline{T_i(x)}) = (t_i, t_{i+1}, \cdots, t_{i-1}) \tag{12}$$

where it is understood that the $j^{th}$ entry of the $m$-tuple is a coset representative mod $g_j(x)$.

We have the same relations among the $\overline{T_i(x)}$ as we do among the $T_i \in \mathbb{F}_p[\mathbb{T}]$. Given $\overline{T_i(x)}$ and $\overline{T_j(x)}$ note that

$$\begin{aligned} &\varphi\left(\overline{T_i(x)T_{i+k}(x)}\right) \\ &= (t_i t_{i+k}, t_{i+1}t_{i+1+k}, \cdots, t_{i-1}t_{i-1+k}) \\ &= \left(\sum_{h=0}^{m-1} [(k,h) - d\theta_k] t_{i+h}, \sum_{h=0}^{m-1} [(k,h) - d\theta_k] t_{i+1+h}, \cdots, \right. \\ &\qquad \left. \sum_{h=0}^{m-1} [(k,h) - d\theta_k] t_{i-1+h}\right) \\ &= \sum_{h=0}^{m-1} \left([(k,h) - d\theta_k]\,(t_{i+h}, t_{i+1+h}, \cdots, t_{i-1+h})\right) \\ &= \sum_{h=0}^{m-1} \left([(k,h) - d\theta_k]\varphi\left(\overline{T_{i+h}(x)}\right)\right) \end{aligned}$$

hence, since $\varphi$ is an isomorphism,

$$\overline{T_i(x)T_{i+k}(x)} = \sum_{h=0}^{m-1} \left( [(k,h) - d\theta_k] \overline{T_{i+h}(x)} \right). \tag{13}$$

Now define $\psi : \mathbb{F}_p[\mathbb{T}] \to \mathbb{F}_p[x]/\Phi_r(x)$ by

$$\psi : \sum_{i=0}^{m-1} \alpha_i T_i \mapsto \sum_{i=0}^{m-1} \alpha_i \overline{T_i(x)}. \tag{14}$$

**Lemma 1.** *$\psi$ is an injective $\mathbb{F}_p$-algebra homomorphism.*

*Proof.* Define the homomorphism $\gamma : \mathbb{F}_p[T_0, \cdots, T_{m-1}] \to \mathbb{F}_p[x]/\Phi(x)$ by $T_i \mapsto \overline{T_i(x)}$ and note that (8) and (13) above show that $\varphi(R) = 0$, so $\gamma$ factors through $\mathbb{F}_p[\mathbb{T}]$ via $\psi$, hence $\psi$ is an $\mathbb{F}_p$-algebra homomorphism. To see that $\psi$ is injective let us stray momentarily from our agreement regarding the representation of elements in $\mathbb{F}_p[x]/\Phi_r(x)$ and note that, modulo $\Phi_r(x)$, any polynomial has a unique representation as an $r - 1^{st}$ degree polynomial without a constant term. Specifically, since $x^r \equiv 1 \bmod \Phi_r(x)$, we have, for $\beta_1 \equiv \beta_2 \bmod r$, that $x^{\beta_1} \equiv x^{\beta_2} \bmod \Phi_r(x)$ and that $1 \equiv -x - x^2 - \cdots - x^{r-1} \bmod \Phi_r(x)$. Now note that the $T_i(x) = \sum_{j=0}^{d-1} x^{\alpha^i p^j}$ where, for $0 \le j \le d-1$, $\alpha^i p^j \neq 0 \in \mathbb{F}_r$ [since the order of $p \bmod r$ is $d$]. Therefore, using the representation above, we write

$$\overline{T_i(x)} = \sum_{j=0}^{d-1} x^{\overline{\alpha^i p^j}} + (\Phi_r(x)) \tag{15}$$

where $\overline{\alpha^i p^j}$ is the least residue of $\alpha^i p^j \mod r$. If we now remark that, for $0 \le i < k \le m-1$, $\alpha^i p^{l_1} \not\equiv \alpha^k p^{l_2} \bmod r$ for any $l_1$ and $l_2$, then there can be no non-trivial relations among the $\overline{T_i(x)}$. That is, $\overline{T_0(x)}, \cdots, \overline{T_{m-1}(x)}$ have distinct representations as polynomials of degree less than $r$, no constant term, and that no two of these representations have terms of equal degree, so the only linear combinations of these elements which equals zero is the trivial one. Therefore

$$\psi\left( \sum_{i=0}^{m-1} \alpha_i T_i \right) = \sum_{i=0}^{m-1} \alpha_i \overline{T_i(x)} = 0 \Leftrightarrow \alpha_i = 0 \ \forall i \in \mathbb{Z}/m\mathbb{Z} \tag{16}$$

and so $\psi$ is injective. □

It is interesting to remark at this point that a direct sum of $m$ copies of $\mathbb{F}_p$, call it $\mathcal{B}$, also known as the *Berlekamp subalgebra*[1] sits inside of $\mathbb{F}_p[x]/g_0(x) \oplus$

[1] See, for example, Section 2.4 of [9]

$\cdots \oplus \mathbb{F}_p[x]/g_{m-1}(x)$ in a natural way and that $\mathrm{Im}(\varphi\psi)$ is contained in this sum. As $\psi$ is injective and $\#(\mathbb{F}_p[\mathbb{T}]) = \#(\mathbb{F}_p \oplus \cdots \oplus \mathbb{F}_p) = p^m$, finite, it follows that $\varphi\psi$ is an isomorphism from $\mathbb{F}_p[\mathbb{T}]$ to $\mathcal{B}$. Keeping in mind the one-to-one correspondence of idempotents in $\mathbb{F}_p[\mathbb{T}]$ and those in $\mathcal{B}$, we see that any idempotent in $\mathbb{F}_p[\mathbb{T}]$ will give a non-trivial factorization of $\Phi_r(x)$.

Suppose that we happen to know one of the traces, $t_i$. Then

$$\varphi\left(\overline{T_0(x)} - t_i\right) = (t_0 - t_i, \cdots, t_{m-1} - t_i) \in \mathcal{B}$$

has a zero in the $i^{th}$ position. This element will be zero if, and only if, all of the $t_i$ are equal.

**Lemma 2.** *Given distinct primes, $r$ and $p$, the traces of the irreducible factors of $\Phi_r(x)$ cannot all be the same.*

*Proof.* As remarked above, if the traces were all the same then $\overline{T_0(x)} - t_i = 0 \in \mathbb{F}_p[x]/\Phi_r(x)$. Recalling the representation of elements in $\mathbb{F}_p[x]/\Phi_r(x)$ discussed in the proof of Lemma 1 we may represent $\overline{T_0(x)} - t_i$ by

$$\begin{aligned}\overline{T_0(x)} - t_i &= \sum_{j=0}^{d-1} x^{\overline{p^j}} + t_i \sum_{j=1}^{r-1} x^j \\ &= \sum_{j=1}^{r-1} s_j x^j\end{aligned}$$

where $\overline{p^j}$ is the least residue of $p^j \bmod r$ and

$$s_j = \begin{cases} 1 + t_i & \text{if } j \in H_0 \\ t_i, & \text{otherwise} \end{cases} \tag{17}$$

This element will be 0 if, and only if $s_j = 0$ for all $0 \leq j \leq r-1$. As some $j$ are in $H_0$ and some are not, this would simultaneously force $t_i = 1$ and $t_i = 0$. □

So $\overline{T_0(x)} - t_i$ will be a non-trivial zero divisor. It should further be noted that

$$\left(\varphi\left(\overline{T_0(x)} - t_i\right)\right)^{p-1} = \left((t_0 - t_i)^{p-1}, \cdots, (t_{m-1} - t_i)^{p-1}\right) \tag{18}$$

where, since we are working in characteristic $p$, $(t_j - t_i)^{p-1} = 0$ or 1 as $t_j = t_i$ or not, respectively. As $\left(\overline{T_0(x)} - t_i\right)^{p-1}$ can be computed in $\log p$ multiplications it follows that we will have found a nontrivial idempotent of $\mathcal{B}$ in polynomial time. Whether working with the idempotent or the zero divisor we will have a nontrivial factorization of $\Phi_r(x)$.

It is a well known fact[2] that if $f(x)$ is any polynomial over any finite field $\mathbb{F}_q$, and if $b(x)$ is an element of Berlekamp subalgebra of $\mathbb{F}_q[x]/f(x)$, then we have

$$f(x) = \prod_{a \in \mathbb{F}_q} \gcd\left(f(x), b(x) - a\right) \tag{19}$$

Noting that $\overline{T_0(x)}$ is an element of the Berlekamp sub-algebra of $\mathbb{F}_q[x]/\Phi_r(x)$ we have

$$\Phi_r(x) = \prod_{a \in \mathbb{F}_p} \gcd\left(\Phi_r(x), \overline{T_0(x)} - a\right) \tag{20}$$

If we now observe that $\gcd\left(\Phi_r(x), \overline{T_0(x)} - a\right) = 1$ for a$\notin \{t_0, \cdots, t_{m-1}\}$ and that $\gcd\left(\Phi_r(x), \overline{T_0(x)} - t_i\right)$ is a proper, non-trivial factor of $\Phi_r(x)$ for $i = 0, \cdots, m-1$, (20) becomes

$$\Phi_r(x) = \prod_{i=0}^{m-1} \gcd\left(\Phi_r(x), \overline{T_0(x)} - t_i\right) \tag{21}$$

Noting that these gcd calculations can be done in time $O\left(r \log^2 r\right)$ [1] we see that if we know all of the traces of the irreducible factors of $\Phi_r(x)$ then we may compute the complete factorization of $\Phi_r(x)$ deterministically in time $O\left(r^2 \log^2 r\right)$.

Also note that for any $i$, and $j$, $\overline{T_j(x)} - t_i$ will also be a zero divisor, so we may get other non trivial factorizations of $\Phi_r(x)$. In particular, if $t_i$ is distinct from all of the other traces then we will have a complete factorization of $\Phi_r(x)$ via

$$\Phi_r(x) = \prod_{j=0}^{m-1} \gcd\left(\Phi_r(x), \overline{T_j(x)} - t_i\right) \tag{22}$$

Please note that it may be the case that not all of the $t_i$ are distinct [15]. In this case although we will get a nontrivial factorization, it need not be a complete factorization.

## Summary

We may think of the $\mathbb{F}_p$- algebra $\mathbb{F}_p[\mathbb{T}]$ as the set of all sums of the form $\sum_{i=0}^{m-1} \alpha_i T_i$, $\alpha_i \in T$, and where multiplication is given by

$$T_i T_{i+k} = \sum_{h=0}^{m-1} \left[(k,h) - d\theta_k\right] T_{i+h}$$

[2] See, for example. Section 2.4 of [9]

This ring is $\mathbb{F}_p$- algebra isomorphic to the Berlekamp subalgebra of $\mathbb{F}_p[x]/\Phi_r(x)$, therefore finding zerodivisors in this ring is equivalent to finding factors of $\Phi_r(x)$. If $t$ is any trace of of $\Phi_r(x)$ then $T_i - t$ will be a zerodivisor in $\mathbb{F}_p[\mathbb{T}]$ and if, as is usually the case when $p$ is large, $t$ is distinct from all the other traces of $\Phi_r(x)$, then $\{T_i - t \mid 0 \leq i \leq m-1\}$ corresponds to the complete set of irreducible factors of $\Phi_r(x)$. Once $t$ is known, the computations involves $O(r \log^2 r)$ operations in $\mathbb{F}_p$ to compute each factor.

## References


1. D. Bini and V. Pan, **Polynomial and Matrix Computations**, Birkhauser, Boston, (1994).
2. L.D. Baumert and W.H. Mills, Uniform cyclotomy. *Journal of Number Theory*, [**14**], (1982), 67–82.
3. L.E. Dickson, Cyclotomy, higher congruences, and Waring's problem. *Amer. J. Math.*, [**57**], (1935), 391–424.
4. D. Jungnickel, **Finite Fields**, Wissenschaftsverlag, Mannheim, (1993).
5. D. Knuth, **The Art of Computer Programming**, volume 2, *Semi-numerical algorithms*, 2nd ed., Addison-Wesley, Reading, (1981).
6. H. W. Lenstra, Jr., Finding isomorphims between finite fields. *Math. Comp.*, [**56**], (1991), 329–347.
7. R. Lidl and H. Niederreiter, **Finite Fields**. Encyclopedia of Mathematics and its Applications, v. 20, Addison-Wesley, Reading, 1983.
8. R. Lidl and H. Niederreiter, **Introduction to Finite Fields and their Applications**, revised edition, Cambridge University Press, Cambridge, 1994.
9. A. J. Menezes, ed., **Applications of Finite Fields**, Kluwer, Boston, (1993).
10. G. Myerson, Period polynomials and Gauss sums for finite fields. *Acta Arith.*, [**39**], (1981), 251–264.
11. R. Schoof, Elliptic curves over finite fields and the computation of square roots mod$p$. *Math. Comp.* [**44**], (1985), pp. 483–494.
12. V. Shoup, New algorithms for finding irreducible polynomials over finite fields. *Math. Comp.* [**54**], (1990), pp. 435–447.
13. C. Small, **Arithmetic of Finite Fields**, Marcel Dekker, New York, (1991).
14. G. Stein, Factoring cyclotomic polynomials over large finite fields. *Finite Fields and Applications*, London Mathematical Society Lecture Note Series #233, S. Cohen & R. Niederrieter, eds., Cambridge University Press, pp. 349–354 (1996).
15. G. Stein, Traces of roots of unity over finite fields. *Finite fields: theory. applications, and algorithms* (Waterloo, ON, 1997) 113-121, Contemp. Math. 225, Amer. Math Soc., Providence RI 1999.
16. G. Stein Using the theory of cyclotomy to factor cyclotomic polynomials over finite fields. *Math. Comp.* [70] (2001), no. 235, 1237–1251.
17. T. Storer, **Cyclotomy and Difference Sets**, Markham, Chicago, 1967.

# On the $F_p$-Linearity of the Generalized Gray Map Image of a $Z_{p^{k+1}}$-Linear Code

H. Tapia-Recillas[1] and G. Vega[2]

[1] Dpto. Matemáticas, UAM-I, México, D.F., *htr@xanum.uam.mx*
[2] Dirección General de Servicios de Cómputo Académico (DGSCA), UNAM, 04510 México, D.F., MÉXICO, *gerardov@servidor.unam.mx*

**Abstract.** A necessary and sufficient condition for the generalized Gray map image $G(D)$ of a $Z_{p^{k+1}}$-linear code $D$ to be $F_p$-linear is given for any prime $p$ and any integer $k \geq 1$. If $p = 2$ and the linear code is assumed to be cyclic, a necessary condition for $G(D)$ to be linear is also given in terms of the generators of the ideal of $D$. Some examples to illustrate the results are given.

## 1 Introduction

In [3] the Gray map on finite chain rings is introduced and shown to be an isometry which generalizes the Gray map given in [1]. Also in [3] this map is used to demostrate the existence of a non-linear ternary $(36, 3^{12}, 15)$ code. In [8] necessary and sufficient conditions are given for the image, $G(D)$, under the generalized Gray map $G$ defined on the ring $Z_{2^{k+1}}$, of a $Z_{2^{k+1}}$-code $D$ (not neccesarily linear) to be linear. In particular, if the code $D$ is linear a characterization is provided for the image $G(D)$ to be linear. For the case of the ring $Z_{p^{k+1}}$, where $p$ is a prime and $k$ is an integer $\geq 1$, a necessary and sufficient condition for the Gray map image of a $Z_{p^{k+1}}$-linear code to be $F_p$-linear is provided in this note. These conditions are similar to those given in [4] for the case $p = 2$ and $k = 1$, i.e., for the usual Gray map over the ring $Z_4$. For the case $p = 2$, assuming the linear code $D$ is also cyclic, a necessary condition for the Gray map image $G(D)$ to be linear, involving the generators of the ideal associated to the code, is also presented. Some examples to illustrate the results are given.

## 2 Some Basic Results

Some concepts that will be used in the paper are introduced in this section. In particular the definition of the Gray map as given in [3] as well as other basic results are recalled.

Let $F = F_q$ be a finite field with $q = p^s$ elements, where $p$ is a prime and $s$ a positive integer. Let $u, v \in F^q$ where $u$ lists all the elements of the field $F$ and $v$ is the all-1 vector. For a positive integer $k$ let:

$$c_i = (v + \delta_{i,0}(u - v)) \otimes \cdots \otimes (v + \delta_{i,k-1}(u - v))$$

for $i = 0, 1, ..., k$, where $\delta$ stands for the Kronecker delta and " $\otimes$ " is the tensor product ([6]). It is easy to see that the vectors $c_i$ generate a $[q^k, k+1, (q-1)q^{k-1}]$-linear code $C$ over the field $F$ (cf. [3]).

Let $p$ and $k$ be as above and let $R = Z_{p^{k+1}}$ be the ring of integers modulo $p^{k+1}$ with residue field $F_p$. Any element $a$ of the ring $R$ can be written in its $p$-adic expansion as: $a = r_k(a)p^k + r_{k-1}(a)p^{k-1} + \cdots + r_1(a)p + r_0(a)$ where $r_i(a) \in T$ and $T \subseteq R$ is a set of representatives of the elements of $F_p$. The generalized *Gray map* on $R$ is defined as (cf. [3]):

$$G : R \longrightarrow C, \quad G(a) = r_k(a)c_k + r_{k-1}(a)c_{k-1} + \cdots + r_1(a)c_1 + r_0(a)c_0$$

The Gray map is bijective on $C$ and it can be extended coordinatewise to the cartesian product of $R$ as:

$$G : R^n \longrightarrow F_p^{np^k}, \quad G(\underline{a}) = (G(a_1), ..., G(a_n))$$

where $\underline{a} = (a_1, ..., a_n) \in R^n$.

Observe that the above extension of $G$ is the same as the one given in [3], i.e., $G(\underline{a}) = \sum_{i=0}^{k} a^{(i)} \otimes c_i$ where the $a^{(i)}$'s stand for the $p$-adic components of the element $\underline{a} \in R^n$.

We recall that the homogeneous weight on $R = Z_{p^{k+1}}$ is defined as:

$$wt_{\text{hom}}(a) = \begin{cases} 0 & \text{if } a = 0 \\ p^k & \text{if } a \in p^k R \setminus \{0\} \\ (p-1)p^{k-1} & \text{otherwise} \end{cases}, \qquad \forall\, a \in Z_{p^{k+1}}.$$

Let $d_{\text{hom}}$ and $d_H$ be the metrics induced by the homogeneous and Hamming weights on $R^n$ and $F_p^{np^k}$ respectively. In ([3]) it is shown that the Gray map:

$$G : (R^n, d_{\text{hom}}) \longrightarrow (F_p^{np^k}, d_H)$$

is an isometry whose image is the Generalized Reed-Muller code $\text{GRM}(1, k)$ over the field $F_p$.

## 3 The Characterization

In this section some properties of the Gray map $G$ related to the addition on the ring $R = Z_{p^{k+1}}$ are described. A necessary and sufficient condition for the image $G(D)$ of a $Z_{p^{k+1}}$-linear code $D$ to be linear is also given.

For any $a \in R$ let $a = r_k(a)p^k + r_{k-1}(a)p^{k-1} + \cdots + r_1(a)p + r_0(a)$ be its $p$-adic expansion. For any two elements $a, b$ of $R$ let:

$$a \oplus b = \sum_{i=0}^{k} p^i \rho_i(a, b), \qquad Q(a, b) = \sum_{i=0}^{k} p^i Q_i(a, b)$$

where, as an integer, $r_i(a)+r_i(b) = pQ_i(a,b)+\rho_i(a,b)$ and $0 \leq \rho_i(a,b) \leq p-1$, for all $i = 0, 1, 2, ..., k$.

Observe that since $0 \leq r_i(a), r_i(b) \leq p-1$ and $0 \leq \rho_i(a,b) \leq p-1$, then $Q_i(a,b) = 0, 1$. Also, if $p = 2$, $\rho_i(a,b) = r_i(a) \oplus r_i(b)$, the addition on the binary field $F_2$ and $Q(a,b) = a \odot b = \sum_{j=0}^{k} r_i(a)r_i(b)$, (see also [9] and [8]).

*Proposition 1.* Let "+" be the addition on the ring $Z_{p^{k+1}}$. Then for any elements $a, b$ of $Z_{p^{k+1}}$ we have:

$$a + b = (a \oplus b) + pQ(a,b).$$

*Proof.* The proof is by induction on $k$. Clearly the Proposition is true when $k = 1$. Suppose that the relation is true for $k \geq 1$, and let $a$ and $b$ be in $\mathbb{Z}_{p^{k+2}}$. By the induction hypothesis we have:

$$(a - r_0(a) + b - r_0(b))/p = \sum_{i=1}^{k+1} p^{i-1}[(\rho_i(a,b) + pQ_i(a,b)].$$

This relation is equivalent to

$$a + b = [r_0(a) + r_0(b)] + \sum_{i=1}^{k+1} p^i[(\rho_i(a,b) + pQ_i(a,b)]$$

which directly gives the proof, since $r_0(a) + r_0(b) = pQ_0(a,b) + \rho_0(a,b)$.

*Proposition 2.* With the same notation as above:

$$G(a \oplus b) = G(a) + G(b)$$

where the operation on the right-hand side is performed in $F_p^{p^k}$.

*Proof.* From the definition of $(a \oplus b)$ it follows that $G(a \oplus b) = \sum_{i=0}^{k} \rho_i(a,b)c_k$. On the other hand $G(a)+G(b) = [r_k(a)+r_k(b)]c_k+[r_{k-1}(a)+r_{k-1}(b)]c_{k-1}+ \cdots +[r_1(a)+r_1(b)]c_1+[r_0(a)+r_0(b)]c_0$. By reduction modulo $p$ on the brackets of this last expression the claim follows.

The next result is an immediate consequence of the previous Propositions.

*Corollary 3.* Let $G$ be the Gray map defined on the ring $Z_{p^{k+1}}$ as described above. Then for any elements $a, b$ in $Z_{p^{k+1}}$:

$$G(a) + G(b) = G(a + b - pQ(a,b))$$

where $pQ(a,b)$ is the integer that appears in the expression for $a + b$ in Proposition 1.

The operation " + " on the ring $R$ is extended, coordinatewise, to the cartesian product $R^n$ in the obvious way: if $\underline{a} = (a_1, ..., a_n)$, $\underline{b} = (b_1, ..., b_n) \in R^n$ then $\underline{a} + \underline{b} = (a_1 + b_1, ..., a_n + b_n)$.

From Corollary 3 it is easy to see that the Gray map on the cartesian product of the ring $R$ satisfies the following:

*Proposition 4.* For any elements $\underline{a} = (a_1, ..., a_n)$, $\underline{b} = (b_1, ..., b_n)$ of $R^n$:

$$G(\underline{a}) + G(\underline{b}) = G(\underline{a} + \underline{b} - p\underline{Q}(\underline{a}, \underline{b}))$$

where $\underline{Q}(\underline{a}, \underline{b}) = (Q(a_1, b_1), ..., Q(a_n, b_n))$.

Recall that a $R$-linear code of length $n$ means a $R$-submodule of $R^n$. We now have:

*Theorem 1.* Let $G$ be the generalized Gray isometry on $R^n$ and let $D$ be a $R$-linear code of $R^n$. Then:

$$G(D) \text{ is } F_p - \text{linear } \textit{if and only if } p\underline{Q}(\underline{a}, \underline{b}) \in D$$

for all $\underline{a}$, $\underline{b} \in D$

*Proof.* If $G(D)$ is a linear code of $R^n$ it follows from Proposition 4 that $G(\underline{a} + \underline{b} - p\underline{Q}(\underline{a}, \underline{b})) \in G(D)$ which implies that $\underline{a} + \underline{b} - p\underline{Q}(\underline{a}, \underline{b}) \in D$. Since $D$ is linear, $p\underline{Q}(\underline{a}, \underline{b}) \in D$. It is easy to see that if $p\underline{Q}(\underline{a}, \underline{b}) \in D$ for all $\underline{a}, \underline{b} \in D$, then $2\underline{a} - p\underline{Q}(\underline{a}, \underline{a}) \in D$. Hence $\alpha G(\underline{a}) \in G(D)$ for any $\alpha \in F_p$. The rest of the proof follows from Proposition 4.

## 4 Linear Cyclic Codes

In this section we assume further that the linear code $D$ is cyclic. A necessary condition on the generators of the code $D$ for its generalized Gray map image, $G(D)$, to be linear is given when $p = 2$. If $k = 1$ the condition is also sufficient and in this case the result is the one presented in ([9], Proposition 16) for codes over $Z_4$.

Let $a \in Z_{p^{k+1}}$ and let $a = r_0(a) + r_1(a)p + \cdots + r_k(a)p^k$ be its $p$-adic expansion. The reduction modulo $p$ of the element $a$ is just $\tilde{a} = r_0(a) \in F_p$. If $a(x) = a_0 + a_1x + \cdots + a_mx^m$ is an element of the polynomial ring $Z_{p^{k+1}}[x]$, its reduction modulo $p$ is $\tilde{a}(x) = \tilde{a}_0 + \tilde{a}_1x + \cdots + \tilde{a}_mx^m \in F_p[x]$. Assume now that $n$ is such that $(p, n) = 1$. Let $R_n = Z_{p^{k+1}}[x]/(x^n - 1)$ and for any element $\alpha(x) \in R_n$, $< \alpha(x) >$ will denote the ideal of $R_n$ generated by $\alpha(x)$.

Let $D$ be a $Z_{p^{k+1}}$-linear cyclic code. According to ([5], Corollary 3.5), polynomials $f_0, f_1, ..., f_k$ exist in $R_n$ such that $f_k | f_{k-1} | \cdots | f_1 | f_0 | (x^n - 1)$ and $D =< f_0, pf_1, ..., p^k f_k >$, the ideal generated by these polynomials.

*Observations.* From the above definitions it is easy to see that:

1. If $a(x) \in R_n$ then $p^k a(x) = p^k \tilde{a}(x)$ in $R_n$.
2. If $\tilde{u}(x)$, $\tilde{v}(x) \in F_p[x]/(x^n - 1)$ and $p^k \tilde{u}(x) = p^k \tilde{v}(x)$ in $R_n$, then $\tilde{u}(x) = \tilde{v}(x)$.

*Lemma 1.* Let $D \subseteq R_n$ be a linear cyclic code as described above. Then $D \cap < p >=< pf_k >$.

*Proof.* Let $u \in D \cap < p >$, i.e., $u = ps$ for some $s \in R_n$ and $u = m_0 f_0 + pm_1 f_1 + \cdots + p^k m_k f_k$ where $m_i \in R_n$. Since $f_k | f_i$ then $f_i = q_i f_k$ with $q_i \in R_n$ for $i = 0, 1, ..., k-1$, and hence $u = m_0 q_0 f_k + pf_k T$ where $T \in R_n$. Also since $u = ps$, $\tilde{u} = 0 = \tilde{m}_0 \tilde{q}_0 \tilde{f}_k$ and $\tilde{f}_k \neq 0$, then $\tilde{m}_0 \tilde{q}_0 = 0$, i.e., $m_0 q_0 = pt$ for some $t \in R_n$. Thus $u = ptf_k + pf_k T$ which implies that $u \in < pf_k >$. The other contention is obvious.

For the rest of this section we take $p = 2$. Let $u, v \in Z_{2^{k+1}}$ and let $u = r_0(u) + r_1(u)2 + \cdots + r_k(u)2^k$ and $v = r_0(v) + r_1(v)2 + \cdots + r_k(v)2^k$ be their 2-adic expansion, respectively. The operation $u \odot v = r_0(u)r_0(v) + r_1(u)r_1(v)2 + \cdots + r_k(u)r_k(v)2^k$ as defined above (cf. [9], [8]) is extended coordinatewise to the cartesian product of the ring $Z_{2^{k+1}}$ and to the polynomial ring $Z_{2^{k+1}}[x]$ in the obvious way: if $u(x) = u_0 + u_1 x + \cdots + u_r x^r$ and $v(x) = v_0 + v_1 x + \cdots + v_r x^r$ then $u(x) \odot v(x) = (u_0 \odot v_0) + (u_1 \odot v_1)x + \cdots + (u_r \odot v_r)x^r$. Also, if $A$ and $B$ are subsets of $Z_{2^{k+1}}[x]$ then $A \odot B = \{a \odot b : a \in A,\ b \in B\}$.

Let $n$ be a positive odd integer and let $D =< f_0, 2f_1, ..., 2^k f_k > \subseteq R_n$ be a linear cyclic code where the polynomials $f_i$ are such that $f_k | f_{k-1} | \cdots | f_1 | f_0 | (x^n - 1)$. Let $G$ denote the generalized Gray map induced on the ring $R_n$.

*Proposition 5.* With the notation as above assume that the generalized Gray map image $G(D)$ of the linear cyclic code $D$ is a binary linear code. Then

$$< \tilde{f}_i(x) > \odot < \tilde{f}_j(x) > \subseteq < \tilde{f}_k(x) >, \forall i, j = 0, 1, ..., k$$

where this last relation is performed in the ring $F_2[x]/(x^n - 1)$.

*Proof.* From Theorem 1 above it follows that $G(D)$ is linear if and only if $2\underline{Q}(\underline{a}, \underline{b}) \in D$ for all $\underline{a},\ \underline{b} \in D$. For the case $p = 2$ the element $\underline{Q}(\underline{a}, \underline{b})$ is just $(\underline{a} \odot \underline{b})$, (cf. §3). In terms of the ideal $D$, this condition is equivalent to: $2(a(x) \odot b(x)) \in D$ for all $a(x),\ b(x) \in D$. Since $2(a(x) \odot b(x)) \in D$, by Lemma 1 it follows that $2(a(x) \odot b(x)) \in D \cap < 2 >=< 2f_k(x) >$ and, therefore, $2^k(a(x) \odot b(x)) = 2^k t(x) f_k(x)$ for some $t(x) \in R_n$. Thus by the above observations $2^k(\tilde{a}(x) \odot \tilde{b}(x)) = 2^k \tilde{t}(x) \tilde{f}_k(x)$ which implies that $\tilde{a}(x) \odot \tilde{b}(x) = \tilde{t}(x)\tilde{f}_k(x)$ for all $a(x)$ and $b(x)$ in $D$. In particular, it follows that $\tilde{f}_i(x) \odot \tilde{f}_j(x) \in < \tilde{f}_k(x) >$ for all $i, j = 0, 1, ..., k$, proving the Proposition.

If $k = 1$ and it is assumed that $\tilde{f}_i(x) \odot \tilde{f}_j(x) \in < \tilde{f}_k(x) >$ for all $i, j = 0, 1, ..., k$, then from the above observations it follows that $2(f_i(x) \odot f_j(x)) \in D$ and hence $2(a(x) \odot b(x)) \in D$ for all $a(x), b(x)$ in $D$, and from Theorem 1 it follows that $G(D)$ is linear. In this case the result is Proposition 16 of [9].

## 5 Examples

In this section two examples are provided in which the previous results are used, especially Theorem 1.

**Example 1.** In [3] the authors use the Gray map $G$ on the ring $Z_{3^2}$ to show that the Gray map image $G(D) \subset Z_3^{36}$ of a $Z_{3^2}$-linear code $D \subset Z_{3^2}^{12}$ is a non-linear $(36, 3^{12}, 15)$ ternary code. The linear code $D$ is the extension by a parity check of the code $E$ which is the lift to $Z_{3^2}[x]/(x^{11}-1)$ of the ternary cyclic $[11, 6, 5]$ Golay code. The code $E$ is the cyclic code generated by the polynomial $e(x) = x^5 + 7x^4 + 8x^3 + x^2 + 6x + 8 \in Z_{3^2}[x]/(x^{11}-1)$, which is the Hensel lift to $Z_{3^2}[x]$ of the generator polynomial $f(x) = x^5 + x^4 + 2x^3 + x^2 + 2 \in Z_{3^2}[x]$ of the ternary Golay code. The authors prove that the image $G(D)$ of the linear code $D$ is not $Z_3$-linear by showing that the elements $\underline{a} = (0, 0, 0, 0, 0, 1, 7, 8, 1, 6, 8, 5)$ and $\underline{b} = (0, 0, 0, 0, 1, 7, 8, 1, 6, 8, 0, 5)$ of $D$ are such that $G(\underline{a} + \underline{b}) - G(\underline{a}) - G(\underline{b}) \notin G(D)$.

We now use the characterization given in Theorem 1 above to show that $G(D)$ is not a linear code. In fact from a direct calculation it is readily seen that for the elements $\underline{a}$ and $\underline{b}$ as given above, $3\underline{Q}(\underline{a}, \underline{b}) = (0, 0, 0, 0, 0, 0, 3, 3, 0, 0, 0, 3)$, which obviously is not an element of the code $D$ since the polynomial $3x^4 + 3x^3$ is not a multiple of the generator of the code $E$.

**Example 2.** In [2] the authors take the binary Golay $[23, 12, 7]$ code generated by the polynomial $x^{11} + x^9 + x^7 + x^6 + x^5 + x + 1$ and Hensel-lift it to the polynomial $x^{11} + 2x^{10} - x^9 + 4x^8 + 3x^7 + 3x^6 - x^5 + 2x^4 + 4x^3 + 4x^2 + x - 1$ over $Z_8$, generating a $[23, 12]$ code over $Z_8$. Extending this code by a parity check results in a $Z_8$-linear $[24, 12]$ code $M_8$. The authors prove that the image $G(M_8)$ is not a $Z_2$-linear code by showing that the elements $\underline{v} = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 2, 7, 4, 3, 3, 7, 2, 4, 4, 1, 7, 3)$ and $\underline{w} = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 2, 7, 4, 3, 3, 7, 2, 4, 4, 1, 7, 0, 3)$ of $M_8$ are such that $G(\underline{v} + \underline{w}) - G(\underline{v}) - G(\underline{w}) \notin G(M_8)$.

This result can also be proved by means of Theorem 1 above. In fact, it is easy to see that for the two elements given above, $2\underline{Q}(\underline{v}, \underline{w}) = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 4, 0, 0, 6, 6, 4, 4, 0, 0, 4, 0, 6)$, which is not in $M_8$ since the sum of its entries does not vanish modulo 8. Hence, by Theorem 1, $G(M_8)$ is not linear.

## Acknowledgement

This research was carried out while the first author was on sabbatical leave at DGSCA, UNAM. Research partially supported by CONACYT grant No. L007 and SNI, Mexico.

## References

1. C. Carlet, "$Z_{2^k}$-linear Codes", IEEE *Trans. Inform. Theory*, vol. 44, pp. 1543–1547, July, 1998.
2. I. Duursma, M. Greferath, S.N. Litsyn and S.E. Schmidt, "A $Z_8$-Linear Lift of the Binary Golay Code and a Nonlinear Binary $(96, 3^{37}, 24)$-Code", IEEE *Trans. Inform. Theory*, vol. 47, pp. 1596–1598, May., 2001.

3. M. Greferath and S.E. Schmidt, "Gray isometries for Finite Chain rings and a Nonlinear Ternary $(36, 3^{12}, 15)$ Code", IEEE *Trans. Inform. Theory*, vol. 45, pp. 2522–2524, Nov., 1999.
4. A. R. Hammons, Jr., P. V. Kumar, A. R. Calderbank, N. J. A. Sloane, and P. Solé, "The $Z_4$-linearity of Kerdock, Preparata, Goethals, and related codes," *IEEE Trans. Inform. Theory*, vol. 40, pp. 301–319, 1994.
5. P. Kanwar and S. R. López-Parmouth, "Cyclic Codes over the Integers Modulo $P^m$". *Finite Fields and their Applications* **3**, 334–352 (1997).
6. F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes.* Amsterdam, The Netherlands: North-Holland, 1977.
7. H. Tapia-Recillas and G. Vega, "A Generalization of Negacyclic Codes." *Proc. International Workshop on Coding and Cryptography, WCC 2001*, (D. Augot, Ed.), Paris, France, pp. 519–529, 2001.
8. G. Vega and H. Tapia-Recillas, "On the $Z_{p^k}$-Linear and Quaternary Codes". Submitted for publication.
9. J. Wolfmann, "Binary images of cyclic codes over $Z_4$". *IEEE, Trans. Inform. Theory*, vol. 47, No.5, pp. 1773–1779, 2001.

# Construction of Modular Curves and Computation of Their Cardinality over $\mathbb{F}_p$

Cédric Tavernier

Projet codes, Bâtiment 10, INRIA Rocquencourt 78150 Le Chesnay, France

**Abstract.** Following [3], and in using several results, we describe an algorithm which compute with a level N given the cardinality over $\mathbb{F}_p$ of the Jacobian of elliptic curves and hyperelliptic curves of genus 2 which come from $X_0(N)$. We will also sketch how to get a plane affine model for these curves.

## 1 Introduction

Elliptic curves are used for electronic signature. A required condition to have a secure cryptosystem is to have $\#Jac(C(\mathbb{F}_p))$ nearly prime. It is known that the computation over $\mathbb{F}_p$ of elliptic curves and hyperelliptique curves of genus two is a difficult problem. Several algorithms (Schoof (1985), Atkin, Elkies, Sato, Pila, Huang) exist with polynomial complexity in $Log(p)$. These methods consist in computing the Frobenius action on $l$-torsion points. This gives the cardinality modulo $l$ (CRT construction). A new way : G. Frey and M. Müller (1998), used $X_0(N)$ and newforms to compute the cardinalities of jacobian of elliptic and hyperelliptic modular curves over $\mathbb{F}_p$.

In section two we will give some results and definitions about $X_0(N)$. The curves $X_0(N)$ has a structure of Riemann surface compact and it is a curve with rational coefficients, so we will study the space of holomorphic differentials $\Omega^1(X_0(N))$ of $X_0(N)$. In fact $\Omega^1(X_0(N))$ is isomorphic to the space of modular forms which are vanishing on cusps of $X_0(N)$ and this space is called space of cusp-forms.

In the third section we will introduce the Hecke algebra. The Hecke algebra is generated by some operators called the Hecke operators and the Atkin-Lehner operators. We will see how this algebra acts on the modular curves $X_0(N)$ and further more on its Jacobian and on its homology. In consequence, we will give some definitions and results about a sub-space of the cusp-forms which is called the space of new-forms.

In the fourth section we will study the first homology group $H_1(X_0(N), \mathbb{Z})$ and the relative homology $H_1(X_0(N), cusps, \mathbb{Z})$ and we will see that there is a correspondence between the homology group and cusp-forms. An important problem is to give a representation for the elements of the homology groups and we want a representation which can be easily computed. Thus we will study two methods, one using the theory of modular symbols and one using the Manin-symbols. We will present the algorithms which permit to convert

Modular-symbols into Manin-symbols and conversely. With these theories we will be able to restrict to new-forms.

In the fifth section we will summarize some results about modular Abelian varieties. We know that for a level N given the new-forms are in correspondence with Abelian varieties of conductor $N^g$ where g is the dimension of these Abelian varieties. In particular we are interested in computing the cardinality over $\mathbb{F}_p$ of these Abelian varieties, so we will give some results about L-series and Abelian varieties.

In the sixth section we will describe the algorithm to compute the cardinality $\mathbb{F}_p$ of Abelian varieties coming from new-forms, and more specially we will give a method to restrict to Abelian varieties of dimension one, that is to say elliptic curves, and Abelian varieties of dimension two which are sometimes Jacobian of modular hyperelliptic curves of genus two.

In the seventh section we will sketch some possible algorithms to obtain an affine model of curve $C$ such that the the Jacobian of $C$ is isogeneous to $A_f$, we will apply some methods due to [1] for genus one and [10] for the genus two.

## 2 About Modular Curves $X_0(N)$

Here, we give some definitions and results about $X_0(N)$.

We know that $SL_2(\mathbb{Z})$ is generated by $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $R = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$.

For a positive integer $N$, we consider the group denoted by

$$\Gamma_0(N) = \left\{ \alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \mid c \equiv 0 \ mod \ N \right\}.$$

which is the Hecke subgroup of $SL_2(\mathbb{Z})$ of level $N$. The groups $SL_2(\mathbb{Z})$ and $\Gamma_0(N)$ act on the upper half plane $\mathbb{H} = \{z \in \mathbb{C} \mid \Im(z) > 0\}$ by homographic transformations given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z \longmapsto \frac{az+b}{cz+d}.$$

We denote the orbits of this action by $Y_0(N) = \Gamma_0(N)\backslash\mathbb{H}$. The quotient $Y_0(N)$ is equipped with a complex analytic structure which comes from $\pi : \mathbb{H} \longrightarrow \Gamma_0(N)\backslash\mathbb{H}$. We compactify $Y_0(N)$ by adjoining the set of cusps $\mathbb{Q} \cup \{\infty\}$. We denote $\mathbb{H}^* = \mathbb{H} \cup \mathbb{Q} \cup \{\infty\}$ and we denote

$$X_0(N) = \Gamma_0(N)\backslash\mathbb{H}^*,$$

the modular curve of $\Gamma_0(N)$. So $X_0(N)$ is a Riemann surface compact and it can be seen as a projective algebraic curve defined over $\mathbb{C}$.

**Definition 1** *A modular form for $\Gamma_0(N)$ of weight 2 is a function $f : \mathbb{H} \longrightarrow \mathbb{C}$ such that*

1. *$f$ is holomorphic on $\mathbb{H}$;*
2. *for any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$, $f(\gamma z) = (cz+d)^2 f(z)$;*
3. *$f$ is holomorphic at the cusps.*

*We denote this space by $M_2(N)$. If a modular form $f$ vanishes at the cusps, then $f$ is called a cusp-form and we denote the space of cusp-forms of weight two, $S_2(N)$.*

**Proposition 1** *Let $\pi$ be the quotient map $\mathbb{H}^* \longrightarrow \Gamma_0(N)\backslash\mathbb{H}^*$, and for any holomorphic differential $\omega$ on $\Gamma_0(N)\backslash\mathbb{H}^*$, set $\pi^*\omega = f dz$. Then $\omega \longmapsto f$ is an isomorphism from the space of holomorphic differentials $\Omega^1(X_0(N))$ on $\Gamma_0(N)\backslash\mathbb{H}^*$ to $S_2(N)$. The dimension of $S_2(N)$ as a complex vector space is equal to the genus of the curve $X_0(N)$.*

## 3 The Hecke Algebra $\mathbb{T}_N$

Let $\tau \in \mathbb{H}$. We denote by $E$ the elliptic curve $\mathbb{C}/L$ where $L = \mathbb{Z} + \mathbb{Z}\tau$. Let $C_N$ be a cyclic subgroup of order $N$ of $E[N]$, the group of $N$-torsion points. Then, by $P = (E, C_N)_\sim$ we denote the isomorphism class of a pair $(E, C_N)$. Using the modular interpretation of the points on $X_0(N)_\mathbb{Z}$ we can define the Atkin-Lehner operators which are also called Atkin-Lehner involutions [2] and are denoted $W_n$ . Let $n$ be a positive divisor of $N$ such that $\gcd(n, N/n) = 1$, then the action of the $n$-th Atkin-Lehner operator is given by

$$W_n(P) = (E/C_n, (C[n] \times C_{N/n})/C_n)_\sim.$$

Also using the modular interpretation we define the Hecke operator [2]. Let $n$ not dividing $N$, then the $n$-th Hecke operator is denoted $T_n$ and its action is given by

$$T_n(P) = \sum_G (E/G, (C[n] \times C_{N/n})/C_n)_\sim,$$

where $G$ runs thought the set of subgroups of order $n$ of $E$ that have trivial intersection with $(C[n] \times C_{N/n})$. As a consequence, $W_n$ and $T_n$ act on

1. the Jacobian variety $J_0(N)$ of $X_0(N)$;
2. the space of cusp forms $S_2(N)(\mathbb{Z})$;
3. the homology group $H_1(X_0(N), \mathbb{Z})$.

**Definition 2** *The Hecke algebra $\mathbb{T}_N$ of level $N$ is the $\mathbb{Z}$-sub-algebra of the endomorphism ring $End_\mathbb{Z}(\Omega^1(X_0(N))_\mathbb{Z})$ generated by*

$$W_n \text{ with } n|N,\ \gcd(n, N/n) = 1 \text{ and } T_k \text{ with } \gcd(k, N) = 1.$$

*The Hecke algebra is commutative.*

**Theorem 1** *The operators $T_n$ and $W_n$ have the following properties:*

1. $T_{nm} = T_n T_m$ *if* $\gcd(m,n) = 1$;
2. $T_p T_{p^r} = T_{p^{r+1}} + pT_{p^{r-1}}$ *if $p$ prime doesn't divide $N$;*
3. $T_p T_{p^r} = T_p^r$, $r \geq 1$, *if $p$ divides $N$.*

**Definition 3** *A cusp-form $f \in S_2(N)$ is a Hecke-eigenform, if $f$ satisfies*

$$T(f) = \lambda_T \cdot f \text{ for all } T \in \mathbb{T}_N;$$

*where $\lambda_T$ is the Hecke-eigenvalues with respect to $T$. We denote*

$$E_{\lambda_T} = \{f \in S_2(N) \mid T(f) = \lambda_T \cdot f\},$$

*the $\lambda_T$-eigenspace where $T \in \mathbb{T}_N$ is fixed. Now we define the space of old forms of $S_2(N)$ as*

$$S_2^{old} = \left\langle g(dz) \mid g(z) \in S_2(M) \text{ with } M|N; M \neq N; d|\frac{N}{M} \right\rangle.$$

**Definition 4** *The orthogonal complement of $S_2^{old}$ with respect to the Petersson inner product:*

$$\langle f, g \rangle = \int_{X_0(N)} f(z)\overline{g(z)}dxdy \quad \text{with } f, g \in S_2(N),\ z = x + iy,$$

*is denoted by $S_2^{new}(N)$ and is called space of new-forms. a cusp-form $f$ is a new-form if and only if $f(z) = q + \sum_{n\geq 2} a_n q^n$ and $f$ is a Hecke-eigenform.*

**Theorem 2** ***(Atkin-Lehner (1970))***. *$S_2^{new}(N)$ is stable under all operators $T_n$, and so $S_2^{new}(N)$ decomposes into a direct sum of orthogonal subspaces $X_i$,*

$$S_2^{new}(N) = \oplus X_i$$

*each of which is a simultaneous eigenspace for all $T_n$ with* $\gcd(n, N) = 1$. *The $T_p$ for $p|N$ stabilize each $X_i$ over $\mathbb{C}$. The spaces $X_i$ in the above decomposition all have dimension* 1 *over $\mathbb{C}$.*

It is known that $Hom(S_2(N), \mathbb{C})$ is a free $\mathbb{T}_N \otimes \mathbb{C}$-module of rank one and $\mathbb{T}_N$ is a free $\mathbb{Z}$-module of rank equal to the genus of $X_0(N)$.

**Proposition 2** ***(Merel (1994))***. *Let $R$ be a commutative ring and let $\psi \in Hom(\mathbb{T}_N, R)$, then*

$$\sum_{n=1}^{\infty} \psi(T_n) q^n \in S_2(N)(R).$$

We will use this property to compute the Fourier expansion of cusp-forms. Since $\mathbb{T}_N$ is a free $\mathbb{Z}$-module of finite rank acting on $S_2(N)$ we get

**Lemma 1** *Let* $f = q + \sum_{n=2}^{\infty} a_n q^n \in S_2^{new}(N)$ *be a Hecke eigenform and* $T \in \mathbb{T}_N$. *Then the eigenvalue* $\lambda_T$ *is a totally real integral algebraic integer and the field*

$$K_f = \mathbb{Q}(\lambda_T \mid T \in \mathbb{T}_N)$$

*is a finite extension of* $\mathbb{Q}$.

## 4 Hecke Theory on Modular Symbols

Let us consider $H_1(X_0(N), \mathbb{Z}) = \mathrm{AB}(\Pi^1(X_0(N), z))$ which is the Abelian group obtained by taking as generators all closed paths on $X_0(N)$, and by factoring out by the relation that two clothed paths are equivalent if one can be continuously deformed into the other. Let $\alpha,\ \beta \in \mathbb{H}^*$ be points equivalent under the action of $\Gamma_0(N)$, so that $\beta = M(\alpha)$ for some $M \in \Gamma_0(N)$, then any smooth path from $\alpha$ to $\beta$ dètermines an integral homology class in $H_1(X_0(N), \mathbb{Z})$ which only depends only on $\alpha$ and $\beta$ ($\mathbb{H}^*$ is simply connected). We denote this homology class by the modular symbol $\{\alpha, \beta\}$. Conversely every integral homology class $\gamma \in H_1(X_0(N), \mathbb{Z})$ can be represented by such a modular symbol $\{\alpha, \beta\}$.

**Proposition 3** *Let* $\alpha, \beta, \gamma \in \mathbb{H}^*$, *and let* $M \in \Gamma_0(N)$. *Then*

1. $\{\alpha, \alpha\} = 0$;
2. $\{\alpha, \beta\} + \{\beta, \gamma\} + \{\gamma, \alpha\}$;
3. $\{M\alpha, M\beta\} = \{\alpha, \beta\}$;

**Corollary 1** *The map* $M \mapsto \{\alpha, M\alpha\}$ *is a surjective group morphism* $\Gamma_0(N) \longrightarrow H_1(X_0(N), \mathbb{Z})$, *which is independent of* $\alpha \in \mathbb{H}^*$.

One considers $H_1(X_0(N), cusp, \mathbb{Z})$, the relative homology of $X_0(N)$ with respect to the set of the cusps. In particular we can see that $H_1(X_0(N), \mathbb{Z})$ is a subgroup of $H_1(X_0(N), cusp, \mathbb{Z})$ because we can take as an element of $H_1(X_0(N), \mathbb{Z})$, a linear combination of elements $\{\alpha, M\alpha\}$ with $\alpha \in \mathbb{Q} \cup \{\infty\}$. We denote by $\mathbb{Z}^{\nu_\infty}$ the set of the cusps $\Gamma_0(N)\backslash\mathbb{Q} \cup \{\infty\}$. A modular symbol $\{\alpha, \beta\}$ is an element of $H_1(X_0(N), cusp, \mathbb{Z})$, where $\alpha,\ \beta$ are cusps. For $\alpha \in \mathbb{Q} \cup \{\infty\}$, we denote by $[\alpha]$ its image in $\Gamma_0(N)\backslash\mathbb{Q} \cup \{\infty\}$. Later we will study more precisely the correspondence.

**Proposition 4** ***(Eichler and Shimura)*** *One has the exact sequence*

$$\begin{array}{ccccccccc} & & & & & \overset{\delta}{} & & \overset{\theta}{} & \\ 0 \to & H_1(X_0(N), \mathbb{Z}) & \to & H_1(X_0(N), cusp, \mathbb{Z}) & \to & & \mathbb{Z}^{\nu_\infty} & \to \mathbb{Z} \to 0 & \\ & \{\alpha, M\alpha\} & \mapsto & \{\alpha, M\alpha\} & & & & & \\ & & & \{\alpha, \beta\} & & \mapsto [\alpha] - [\beta] & & & \\ & & & & & & \lambda[\alpha] & \mapsto \lambda & \end{array} \tag{1}$$

Now we give some recalls: the projective line over $\mathbb{Z}/N\mathbb{Z}$ is defined by

$$\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z}) = \{(c,d) \in (\mathbb{Z}/N\mathbb{Z})^2 \mid \gcd(c,d,N) = 1\}/\sim,$$

where $(c,d) \sim (c',d')$ iff $cd' \equiv c'd \mod N$. We can show that the map

$$\Gamma_0(N)\backslash SL_2(\mathbb{Z}) \longrightarrow \mathbb{P}^1(\mathbb{Z}/N\mathbb{Z}),\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto (c:d) \mod N$$

is a bijection between the right coset $\Gamma_0(N)\backslash SL_2(\mathbb{Z})$ and the projective line $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$. The elements of $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$ are called Manin-symbols.

**Theorem 3** ***(Manin 1972)*** *$H_1(X_0(N), cusp, \mathbb{Z})$ is a free $\mathbb{Z}$-module and it rank is equal to $2g(X_0(N)) + \nu_\infty(N) - 1$. It is generated by the modular symbols*

$$\{\{M(0), M(\infty)\} | M \in \Gamma_0(N)\backslash SL_2(\mathbb{Z})\};$$

*and we have the isomorphism*

$$\mathbb{Z}[\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})]/\langle u + uS, u + uR + uR^2 | u \in \mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})\rangle \cong H_1(X_0(N), cusp, \mathbb{Z}).$$

Let $i$ be the following involution which acts on $\mathbb{H}^*$, on the Manin-symbols and the modular symbols by the following relations:

$$i(z) = -\overline{z},\quad i((c,d)) = (c,d)\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},\quad i(\{\alpha,\beta\}) = \{-\alpha,-\beta\}.$$

Restricting (1) to invariant elements under the involution, we get:

$$0 \longrightarrow H_1(X_0(N),\mathbb{Z})_+ \longrightarrow H_1(X_0(N), cusp, \mathbb{Z})_+ \overset{\delta_+}{\longrightarrow} \mathbb{Z}_+^{\nu_\infty}$$

If we want to construct a basis of $H_1(X_0(N),\mathbb{Z})_+$, we have to construct the matrix of $\delta_+$, thus a basis of $H_1(X_0(N), cusp, \mathbb{Z})_+$ and of $\mathbb{Z}_+^{\nu_\infty}$. The construction is similar if we want to construct a basis of $H_1(X_0(N),\mathbb{Z})$, we just have to omit the involution action.

To find a basis of $H_1(X_0(N),\mathbb{Z})_+$, we use the following relations:

1. $(c:d) + (c:d)S = (c:d) + (-d:c) = 0$;
2. $(c:d) + (c:d)R + (c:d)R^2 = (c:d) + (c+d:-c) + (d:-c-d) = 0$;
3. $(c:d) - i((c:d)) = (c:d) - (-c:d) = 0$.

These formulas give us the relations between the elements of the representative system of $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$, then we just have to index-link the elements of this representative system in a canonic basis, then we can construct our $\mathbb{Z}$-module quotient. It is similarly to obtain a $\mathbb{Z}$-module basis of $\mathbb{Z}_+^{\nu_\infty}$. We have the following equivalence:

1. $i([\alpha]) = [\alpha]$ et $[\alpha] \equiv [\beta] \Longleftrightarrow \alpha = \pm\beta \mod \Gamma_0(N)$ ;
2. For $j = 1, 2$, let $\alpha_j = p_j/q_j$, be equivalent cusps written in lowest terms. Then $s_1 q_2 \equiv \pm s_2 q_1 \ mod\ gcd(q_1 q_2, N)$ where $s_j$ satisfies $p_j s_j \equiv 1 \mod q_j$.

Now we present some results about the correspondence between homology and cusp-forms.

**Proposition 5** ***(Merel 1994)*** *We have isomorphisms*

1. $H_1(X_0(N), cusp, \mathbb{Z}) \cong \varepsilon is(\Gamma_0(N)) \oplus S_2(N) \oplus \overline{S_2(N)}$;
2. $H_1(X_0(N), \mathbb{Z}) \cong S_2(N) \oplus \overline{S_2(N)}$ *and* $H_1(X_0(N), \mathbb{Z})_+ \cong S_2(N)$;
3. $\dim H_1(X_0(N), \mathbb{Z})_+ = \dim H_1(X_0(N), \mathbb{Z})_- = g(X_0(N))$;

*where* $\overline{S_2(N)}$ *is the anti-holomorphic space of cusp-forms,* $\varepsilon is(\Gamma_0(N))$ *is a space of modular forms which is called space of Eisenstein series, and we noted* $g(X_0(N))$ *as the genus of* $X_0(N)$.

We are going to describe the action of Hecke algebra on Manin-symbols and modular symbols:

**Proposition 6** *For* $p$ *prime and* $p \nmid N$, *if* $\alpha$, $\beta$ *are cusps, we have:*

$$T_p(\{\alpha, \beta\}) = \{p\alpha, p\beta\} + \sum_{k=0}^{p-1} \left\{ \frac{\alpha + k}{p}, \frac{\beta + k}{p} \right\}.$$

*If* $p^a || N$, *then let* $W_p = \begin{pmatrix} p^a x & y \\ Nz & p^a t \end{pmatrix}$, *with* $x,\ y,\ z,\ t \in \mathbb{Z}$, $\det(W_p) = p^a$. *Then*

$$T_{p^a}(\{\alpha, \beta\}) = \left\{ \frac{p^a x\alpha + y}{Nz\alpha + p^a t}, \frac{p^a x\beta + y}{Nz\beta + p^a t} \right\}.$$

To compute the matrix of Hecke operators acting on cusp-forms we need to be able to convert the modular symbols into Manin symbols [1].

If $(c : d) \in \mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$, with the Bezout lemma we can find $a, b \in \mathbb{Z}$ such that $\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = 1$, so we are be able to convert a Manin symbol into modular symbol:

$$(c : d) \longrightarrow M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \longrightarrow \{M(0), M(\infty)\} = \left\{ \frac{a}{c}, \frac{b}{d} \right\}.$$

If we give us a modular symbol $\left\{\frac{a}{c}, \frac{b}{d}\right\}$, we have the following algorithm:

$$\left\{ \frac{a}{c}, \frac{b}{d} \right\} = \left\{ \frac{a}{c}, 0 \right\} + \left\{ 0, \frac{b}{d} \right\} \text{ and we note } \left\{ 0, \frac{b}{d} \right\} = \{0, t\}$$

Let $[a_1, \ldots, a_n]$ be the simple continued fraction expansion of $t$, i.e.

$$a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}$$

If we note $C_k = [a_1, \ldots, a_k]$ then the numerator $p_k$ and the denominator $q_k$ of $C_k$ satisfy the equations (For $i = 3, 4, \ldots, k$)

$$p_i = a_i p_{i-1} + p_{i-2},\ p_{-1} = 0,\ p_0 = 1,\ p_1 = a_1,\ p_2 = a_1 a_2 + 1,$$
$$q_i = a_i q_{i-1} + q_{i-2},\ q_{-1} = 1,\ q_0 = 0,\ q_1 = 1,\ \ q_2 = a_2.$$

where $t = \frac{p_n}{q_n}$ and we know that $p_i q_{i-1} - p_{i-1} q_i = (-1)^i$, where $i \geq 0$. So we obtain that

$$\{0, t\} = \sum_{i=0}^{i=n} \{M_i(0), M_i(\infty)\} \text{ with } M_i = \begin{pmatrix} (-1)^{i-1} p_i & p_{i-1} \\ (-1)^{i-1} q_i & q_{i-1} \end{pmatrix}.$$

We present another method: the Hecke algebra can act directly on Manin symbols, in such manner continued fractions are not needed.

**Definition 5** *Let* $M_n = \{v \in M^{2\times 2}(\mathbb{Z}) \mid \det(v) = n\}$ *and*

$$(c:d)M = \begin{cases} 0 & \text{if } (c:d)M \notin \mathbb{P}^1(\mathbb{Z}/N\mathbb{Z}) \\ (c:d)M & \text{if } (c:d)M \in \mathbb{P}^1(\mathbb{Z}/N\mathbb{Z}). \end{cases}$$

*For any integer* $n \in \mathbb{Z}$ *we will say that the element* $\Theta_n = \sum_{M \in M_n} u_M M \in \mathbb{Z}[M_n]$ *satisfies the condition* $C_n$ *if*

$$\sum_{M \in M_n} u_M (M(\infty) - M(0)) = (\infty) - (0).$$

**Theorem 4** ***(Merel 1994)*** *If* $\Theta_n$ *satisfies the condition* $C_n$ *then we have the following formula for the action of Hecke and Atkin-Lehner operators on Manin symbols:*

$$T_n((c:d)) = \sum_{M \in M_n} u_M (c:d)M \text{ for } \gcd(n, N) = 1,$$

$$W_n((c:d)) = \sum_{M \in M_n,\ (c:d)M \equiv (0,0) \mod n} u_M \epsilon_n(gM) \text{ for } n|N,$$

*where* $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$ *and* $\epsilon_n(gM)$ *is the unique element of* $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$ *congruent to* $(1,0)gM \mod n$ *and to* $(0,1)gM \equiv (c:d)M \mod N/n$.

**Theorem 5** ***(Merel 1994)*** *The element*

$$\sum_{a>b\geq 0,\ d>c\geq 0,\ ad-bc=n} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{Z}[M_n] \tag{2}$$

*satisfies the condition* $C_n$.

Now we give a very useful result which gives us an algorithm to restrict to new-forms which are in correspondence with elliptic curves and hyperelliptic curves of genus two.

**Theorem 6** ***(Merel 1994)*** *Let $x \in \mathbb{Z}[\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})]$ and $\Theta = \sum_{M \in M_1} u_M M$ which satisfies the condition $C_1$ and let*

$$\epsilon_1 : S_2(N) \longrightarrow S_2(N/n), \quad \epsilon_1(x) = \sum u_M x M \tag{3}$$

*for $n$ dividing $N$ and where the sum is restricted to the matrices $M$ such that $xM \in \mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$. Then $x$ belongs to $S_2^{new}(N)$ if and only if $x$ and $W_N(x)$ belong to the kernel of $\epsilon_1$ for all divisors $n$ of $N$.*

Using (2), (3), it is easy to see that the sum $\Theta$ is restricted to the matrix identity. With these results we now are able to construct a basis of new-forms which are in correspondence with Abelian varieties of genus one or two. We are going to see this correspondence in the following section.

## 5 New-forms and Abelian Varieties

**Theorem 7** *Let $f = q + \sum_{n=2}^{\infty} a_n q^n$ be a Hecke eigenform and let $K_f = \mathbb{Q}(a_n \mid n \in \mathbb{N})$ be the field generated by the Fourier coefficients of $f$. Then there exists an Abelian sub-variety $A_f$ of $J_0(N)$ and an isomorphism $\theta$ from $K_f$ to $End(J_0(N)) \otimes \mathbb{Q}$ with the properties:*

1. $\dim(A_f) = [K_f : \mathbb{Q}] = d$;
2. *If* $\gcd(n, N) = 1$, *then $\theta(a_n)$ coincides with the restriction of $T_n$ to $A_f$;*
3. *The conductor $N(A_f)$ is equal to $N^d$, where $d = \dim(A_f)$.*

*Moreover the pair $(A_f, \theta)$ is unique and $A_f$ is a simple Abelian variety defined over $\mathbb{Q}$.*

### 5.1 L-series and Applications

**Case of elliptic curves** Recall that for an elliptic curves $E$ over $\mathbb{Q}$, we define

$$L(E, s) = \prod_{p \text{ good}} \frac{1}{1 - a_p p^{-s} + p^{1-s}} \cdot \prod_{p \text{ bad}} \frac{1}{1 - a_p p^{-s}} = \sum a_n n^{-s}$$

where

$$a_p = \begin{cases} p + 1 - N_p & \text{if } p \text{ good;} \\ 1 & \text{if } p \text{ split nodal;} \\ -1 & \text{if } p \text{ nonsplit nodal;} \\ 0 & \text{if } p \text{ cuspidal.} \end{cases} \qquad \text{and } N_p = \#E(\mathbb{F}_p).$$

Recall that to a new-form $f$ we can associate a Dirichlet series which admits an Euler product [7]

$$L(f,s) = \prod_{\gcd(p,N)=1} \frac{1}{1-a_p p^{-s}+p^{1-s}} \cdot \prod_{p|N} \frac{1}{1-a_p p^{-s}} = \sum a_n n^{-s}$$

**Theorem 8** ***(Eichler-Shimura)*** *Let $f = q + \sum_{n=2}^{\infty} a_n q^n$ a new-form with $a_n \in \mathbb{Z}$ for all $n \geq 0$. Then there exists an elliptic curve $E_f$ of conductor $N$ such that $L(f,s) = L(E,s)$.*

In fact we know now that all elliptic curves are modular, that is to say that all elliptic curves of conductor $N$ are simple factors of the Jacobian $J_0(N)$.

**Case of Abelian variety of genus 2** Let $f = q + \sum_{n=2}^{\infty} a_n q^n$ be a Hecke eigenform, with $K_f(a_n \mid n \in \mathbb{Z})$ being a quadratic extension of $\mathbb{Q}$. Let $I_f = \{Id, \sigma\}$ be the set of distinct embedding of $K_f$ into $\mathbb{C}$, then we define the L-series of $f$ in $p$ by

$$L_p(f,s) = \begin{cases} 1 - a_p s + p s^2 & \text{if } p \text{ doesn't divide } N, \\ 1 - a_p s & \text{if } p \text{ divide } N. \end{cases}$$

**Theorem 9** *Let $L_p(A_f, s)$ be the L-series of $A_f$ in $p$. Then, for a $p$ prime not dividing $N$, we have the following properties:*

1. $L_p(A_f, s) = \prod_{\sigma \in I_f} L_p(f^\sigma, s)$;
2. $L_p(A_f, 1) = \#(A_f \otimes \mathbb{F}_p)$.

In particular we have the following formula

$$L_p(A_f, 1) = (1 + p + a_p)(1 + p + \sigma(a_p)) = \chi_p^f(p+1),$$

where $\chi_p^f$ is the minimal polynomial of $T_p$ acting on $f$.

**Remark 1** *The same properties hold if $K_f(a_n \mid n \in \mathbb{Z})$ is an extension of $\mathbb{Q}$ of greater degree but we are just interested by elliptic curves and hyperelliptic curves of genus 2.*

## 6 Steps to Compute the Cardinality over $\mathbb{F}_p$

We are going to summarize by the following points how to compute the cardinality of elliptic curves or Abelian variety over $\mathbb{F}_p$:

- First we construct a representative system of $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$, one can take all elements $(d,i)$ with $d$ dividing $N$ and $\gcd(d,i) = 1$, then we have to choose a representant in the coset $(d,i)$ where $d$ is fixed because two elements $(d,i)$ and $(d,j)$ are equivalent if and only if $i - j \equiv 0 \mod N/d$;

- Secondly we find a Manin-symbol basis of $H_1(X_0(N), cusp, \mathbb{Z})_+$, for this, the relations that we have seen before are essential:
  1. $(c:d)+(-d:c)=0$;
  2. $(c:d)+(c+d:-c)+(d:-c-d)=0$;
  3. $(c:d)-(-c:d)=0$.

  We just index-link the representative system of $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$ by the elements of a canonical basis, then we just recognize the relations seen above in this canonical basis. After we quotient a free $\mathbb{Z}$-module of rank $\#(\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z}))$ index-linked by the canonical basis by the relation seen above. We obtain in fact a morphism $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z}) \to H_1(X_0(N), cusp, \mathbb{Z})_+$
- Similarly, we construct a $\mathbb{Z}$-module basis of $\mathbb{Z}_+^{\nu_\infty}$. To do this:
  1. First we extract a representative system of cusps which come from $H_1(X_0(N), cusp, \mathbb{Z})_+$, we saw that it is possible because we can convert a Manin-symbol into modular symbol. So we will obtain two cusps for each modular symbol.
  2. Then we use the equivalent properties of cusps: [1]
     (a) $i([\alpha]) = [\alpha]$ et $[\alpha] \equiv [\beta] \iff \alpha = \pm\beta \bmod \Gamma_0(N)$ ;
     (b) For $j = 1, 2$, let $\alpha_j = p_j/q_j$, be equivalent cusps written in lowest terms. Then $s_1 q_2 \equiv \pm s_2 q_1 \bmod gcd(q_1 q_2, N)$ where $s_j$ satisfies $p_j s_j \equiv 1 \bmod q_j$.

  We also obtain a morphism cusps $\longrightarrow \mathbb{Z}_+^{\nu_\infty}$
- Now we are able to construct the matrix of $\delta_+$ because we have a basis of $\mathbb{Z}_+^{\nu_\infty}$ and $H_1(X_0(N), cusp, \mathbb{Z})_+$ , with the extended Euclidean algorithm we just convert some Manin-symbols into modular symbols and extract the two cusps of each modular symbol.
- To obtain a Manin-symbol basis of $S_2(N)$, we just compute the kernel of $\delta_+$. Thus we obtain the vector basis. We get the Manin-symbol basis in looking the index-linking that we choose for the representative system of $\mathbb{P}^1(\mathbb{Z}/N\mathbb{Z})$.
- Our goal now is to restrict to basis of new-forms which are in correspondence with Abelian varieties of genus one or two. Thus we choose the smaller prime $p$ not dividing $N$, and we compute the matrix of the $p$-th Hecke operator $T_p$ acting on $S_2(N)$. Then we compute the characteristic polynomial of $T_p$ and we extract a basis of eigenvectors which corresponds to irreducible factors of degree one or two of the characteristic polynomial of $T_p$.
- For each eigenvector we verify if it is an element of $S_2^{new}(N)$ with the map (3) $\epsilon_1 : S_2(N) \longrightarrow S_2(N/n)$ with $n \mid N$. We keep only the elements which belong to $S_2^{new}(N)$. So we get a Manin-symbol basis of $S_2^{new}(N)$ which are in correspondence with these Abelian varieties of genus one or two. Of course sometimes there doesn't exist modular Abelian varieties with level $N$ given. We are just interested by the best cases, where there is at least an Abelian variety of genus one or two.
- Now we would like compute the Fourier coefficients of these new-forms in order to get the cardinality of these varieties over $\mathbb{F}_p$. In fact the

eigenvalues of the $p$-th Hecke operator acting on a element of the basis of $S_2^{new}(N)$ is equal to the $p$-th coefficient of the new-form which is in bijection with this element. So to compute the series of a new-form we just compute the eigenvalues of the Hecke algebra acting on the element of the basis of $S_2^{new}(N)$.

In this example we programmed this algorithm with Magma:

Here is a representative system of $\mathbb{P}^1(\mathbb{Z}/33\mathbb{Z})$. We choose this natural order given by magma as index-linking for the canonic basis:

```
>RepresSyst(33);

[
    [ 1, 0 ],[ 1, 1 ],[ 1, 2 ],[ 1, 3 ],[ 1, 4 ],[ 1, 5 ],..
    ,[ 1, 32 ],[ 3, 14 ],[ 3, 11 ],[ 3, 20 ],..,[ 3, 1 ],
    [ 11, 3 ],[ 11, 2 ],[ 11, 1 ],[ 0, 1 ]
]
```

Now here is a Manin-symbol basis of $S_2(33)$. We know that the genus of $X_0(N)$ is equal to 3, thus we get 3 vectors. We also take this natural order to index-link the basis of $S_2(33)$.

```
>S2base(33);

  [              [               [
    -[ 3, 5 ]        [ 3, 5 ]      -[ 3, 4 ]
    -[ 11, 3 ]     +[ 3, 4 ]       -[ 11, 3 ]
    +[ 11, 1 ]     -[ 3, 1 ]       +[ 11, 1 ]
  ],               +[ 11, 3 ]    ]
                   -[ 11, 1 ]
                ],
```

The smaller prime $p$ not dividing $N$ is 2. Thus we compute the action of the 2-th Hecke operator on $S_2(33)$, because we want to extract the new-forms which interest us.

```
> HeckeAction(2,33);

      [ 0  2  1]
A := [ 0 -2  0]
      [ 2  2 -1]
```

```
> CharcPolyHecke(2,33);

[
  <x - 1, 1>,
  <x + 2, 2>
]
```

We see that there are two eigenspaces, one is associated to the eigenvalue 1 and the other is associated to the eigenvalue $-2$. We need the eigenvectors of this eigenspaces:

```
> Eigenspace(A,-2);
Echelonized basis:
( 1  0 -1)
( 0  1  0)
```

```
> Eigenspace(A,1);
Echelonized basis:
(2 2 1)
```

We search now the elements of $S_2^{new}(N)$. The degree of the irreducible factors of the characteristic polynomial of $T_2$ is one. Thus if $S_2^{new}(N) \neq 0$, the new-forms are in correspondence with elliptic curves (up to isogeny) of conductor equal to 33. We verify that the 1-eigenvector satisfies the condition of the theorem 6 (3), that is to say that Eigenspace(A,1) has to belong to the kernel of the map $\epsilon_1$ for the divisors 3 and 11 of 33.

```
> Epsilon1(A,1,3);
(0)
```

```
> Epsilon1(A,1,11);
(0)
```

```
> Epsilon1(W33(A,1),3);
(0)
```

```
> Epsilon1(W33(A,1),11);
(0)
```

Epsilon1(A,i,3), for $i = 1, -2$ is always equal to 0 because $\dim(S_2(3)) = 0$ whereas $\dim(S_2(11)) = 1$. We verify that the other eigenvectors which are associated to the eigenvalue $-2$ do not belong to $S_2^{new}(N)$:

```
> Epsilon1(A,-2,3);
(0)
(0)
```

```
> Epsilon1(A,-2,11);
 (-4)
 (1)
```

Therefore, the eigenvector which is associated to the eigenvalue 1 of the Hecke operator $T_2$ belongs to $S_2^{new}(N)$. So the 1-eigenspace is in fact a new-form for which we can compute its Fourier coefficients. We see that the elements which is in correspondence with the eigenvalues $-2$ belong to because $\dim(E_{-2}) = 2$.

We have two ways to compute the Fourier coefficients, we can apply the $p$-th Hecke operator directly on the 1-eigenvector of Manin-symbols, using the Manin and Merel results, or we can transform these Manin-symbols into modular symbols and then we use the continued fraction method. In practice the continued fraction method is more easier to implement.

$N = 33$ : $\text{genus}(X_0(33)) = 3$, we can get a new-form associated to a elliptic curve: [9]

$$f(z) = q + q^2 - q^3 - q^4 - 2q^5 - q^6 + 4q^7 - 3q^8 + q^9 - 2q^{10} + q^{11} + q^{12} - 2q^{13} \ldots$$

This elliptic curve admits for minimal model $E : y^2 + xy = x^3 + x^2 - 11x$ [1]

## 7 Construction of Modular Curves from New-forms

First we summarize the results of Shimura [4]. Let $f(z)$ a new-form of weight two and $\omega(f) = 2\pi i f(z)dz$ be the associated differential.

Let $I_f = \{\sigma_1, \ldots, \sigma_d\}$ be all the distinct embeddings of $K_f = \mathbb{Q}(a_1, \ldots)$ into $\mathbb{C}$ which is the field generated by the coefficients of $f$. Let $\{f^{\sigma_1}, \ldots f^{\sigma_d}\}$ be the complete set of new-forms conjugate to $f$ over $\mathbb{Q}$. There exists an Abelian variety $A_f$ rational over $\mathbb{Q}$ (see section 4 theorem 7) such that the space of differential 1-forms $\Omega^1(A_f)$ is isomorphic to $\sum_{\sigma \in I_f} \mathbb{C}\omega(f^\sigma)$. Let $\mathbf{f} = (f^{\sigma_1}, \ldots, f^{\sigma_d})^t$ and $\omega(\mathbf{f}) = (\omega(f^{\sigma_1}), \ldots, \omega(f^{\sigma_d}))^t$. Then the image of $H_1(X_0(N), \mathbb{Z})$ by the map

$$H_1(X_0(N), \mathbb{Z}) \longrightarrow \mathbb{C}^d;\ \gamma \longmapsto \int_\gamma \omega(\mathbf{f}) = \left( \int_\gamma \omega(f^{\sigma_1}), \ldots, \int_\gamma \omega(f^{\sigma_d}) \right)^t$$

q is a free $\mathbb{Z}$-module of rank $2d$. It is a lattice $\Lambda_f$ in $\mathbb{C}^d$ and we get

$$A_f \cong \mathbb{C}^d / \Lambda_f.$$

When $d = 1$ we have an elliptic curve and in this case it is possible to get a minimal model of elliptic curve $C$ such that $C \cong A_f$. See [1].

When $d = 2$, sometimes we can get a model of hyperelliptic curve of genus two $C$ such that $Jac(C) \cong A_f$. This model can be obtained if the period matrix of $A_f$ satisfies certain conditions. See [10].

## 8 Conclusion

With this method it is possible to construct a general family of elliptic curves because we know that all elliptic curves are modular. In fact for a level $N$ given we are able to construct up to isogeny all the elliptic curves and in fact, without constructing these elliptic curves, we can give the number of elliptic curves over $\mathbb{Q}$ (up to isogeny) with given conductor $N$. The complexity of this algorithm is polynomial in $N$, so if the level $N$ is not too large we can get a great number of Abelian varieties. We have the Shimura-Taniyama conjecture which asserts that any Abelian variety $A$ with real multiplication, both defined over $\mathbb{Q}$, is isogenous to a factor of $J_0(N)$ for a suitable $N$. So we just can say that with this algorithm we can compute the number of modular Abelian varieties of genus two and conductor $N^2$ with level $N$ given. We just interested by Abelian varieties of genus one or two because the hyperelliptic curves of genus two and elliptic curves may give good cryptosystems. An important problem in cryptography is to compute of the cardinality of the Jacobian of these curves over $\mathbb{F}_p$.

Computing the cardinality over $\mathbb{F}_p$ with this algorithm is not possible if $p$ is too large: we see that if we choose the method using continued fraction we need to compute $p$ continued fractions on fractions of number very closed to $p$. This computing need about $\mathcal{O}(p \log(p))$ arithmetic operations. The method which uses the matrices sum acting on the Manin-symbols is not better because we know (see [5], [3]) that these families have a cardinal very closed to

$p\log(p)$ and it is not easy in practice to construct this sum. So we can't use these methods in cryptography.

A possible improvement would be to find the matrix of the $p$-th Hecke operator with p large. It would be interesting if we find for example a method to compute the $p$-th Hecke operator action modulo some small prime numbers $l_i$ with a polynomial complexity who depends of $l_i$. (CRT)

## References


1. John Cremona. *Arithmetic of modular elliptique curves.* Cambridge University Press, 1992.
2. Bas Edixhoven. The modular curves $X_0(N)$. In *Trieste, ICTP, Summer school on elliptic curves*, 1997.
3. Gerhard Frey and Michael Müller. Arithmetic of modular curves and application. In Matzat, B.H; Greuel, G.M.; Hiss, G. editor, *Algorithmic Algebra and Number Theory*, Springer-Verlag, 1999.
4. G. Shimura. *Introduction to the arithmetic theory of automorphic Functions*, Princeton University Press, 1971.
5. Loïc Merel. Universal fourier expansions of modular forms. *Lecture Notes in Mathematics*, 1994.
6. Jean-Francois Mestre. Construction de courbes de genre 2 à partir de leur modules. *Effective Methods in Algebraic Geometry*, 1991.
7. Joseph Milne. Elliptic curves. Available on `http://www.jmilne.org/math/CourseNotes/math679.html`, 1996.
8. Jean-Pierre Serre. *Cours d'arithmétique.* Presses Univ. France, 1970.
9. William A. Stein. The modular forms database. Available on `http://modular.fas.harvard.edu/Tables/index.html`, 1999.
10. Xiangdong Wang. 2-dimensional simple factors of $J_0(N)$. *Manuscripta Mathematica*, 1995.
11. Xiangdong Wang. The Hecke operators on the cusp-forms of $\Gamma_0(N)$. In G. Frey, editor, *On Artin's Conjecture for Odd 2-dimensional Representations*, number 1585 in Lecture notes in Mathematics, pages 59–94. Springer-Verlag, 1995.

# Asymptotic Properties of Global Fields

M. A. Tsfasman*

Institut de Mathématiques de Luminy, Independent University of Moscow,
and Institute for Information Transmission Problems.
E-mail: tsfasman@iml.univ-mrs.fr.

**Abstract.** The main object of our study is an "infinite" global field, i.e., an infinite algebraic extension either of $\mathbb{Q}$ or of $\mathbb{F}_r(t)$. In order to understand such fields we study sequences of usual global fields, both number and function, with growing discriminant (respectively, genus). We manage to generalize the Odlyzko–Serre bounds and the Brauer–Siegel theorem. This leads to asymptotic bounds on the ratio $\log hR/\log\sqrt{|D|}$ valid without the standard assumption $n/\log\sqrt{|D|}\to 0$, thus including, in particular, the case of unramified towers. Then we produce examples of class field towers, showing that this assumption is indeed necessary for the Brauer–Siegel theorem to hold. To understand what is going on, we introduce zeta-functions of infinite global fields, and study measures corresponding to limit distributions of zeroes of usual zeta functions.

**Definitions.** Let $K$ be a global field, i.e., either a number field of degree $n$ and absolute value of the discriminant $D$ with $r_1$ real and $r_2$ pairs of complex embeddings, or a function field over $\mathbb{F}_r$; in the function field case we set $r_1 = r_2 = 0$ and $g = genus(K)$; in the number field case, $g = \log\sqrt{|D|}$. By $h$ we denote the class-number of $K$ (which equals the number of $\mathbb{F}_r$-rational points on the Jacobian of $K$ in the function field case); $R$ denotes the regulator of $K$ in the number field case and equals 1 in the function field case. We write log for $\log_e$ in the number field case, and for $\log_r$ in the function field case over $\mathbb{F}_r$; the Euler constant is denoted by $\gamma$. For a prime power $q$ let $N_q(K) := |\{v \in P(K) : Norm(v) = q\}|$ , where $P(K)$ is the set of non-archimedean places of $K$.

We call a sequence $\{K_i\}$ of global fields a *family* if $g(K_i)\to\infty$ for $i\to\infty$, in the function field case $r$ being fixed. We call a family $\{K_i\}$ *asymptotically exact* if and only if there exist all the limits (for any prime power $q$)

$$\phi_q := \lim_{i\to\infty}\frac{N_q(K_i)}{g(K_i)}\ , \quad \phi_{\mathbb{R}} := \lim_{i\to\infty}\frac{r_1(K_i)}{g(K_i)}\ , \quad \phi_{\mathbb{C}} := \lim_{i\to\infty}\frac{r_2(K_i)}{g(K_i)}\ .$$

A simple diagonal argument shows that every family $\{K_i\}$ contains an asymptotically exact subfamily.

By an infinite global field we mean an infinite algebraic extension either of $\mathbb{Q}$ or of $\mathbb{F}_r(t)$.

* Supported in part by RBRF 99-01-01204. This lecture is based on my joint work with Serge Vlăduţ.

**Lemma 1.** *Let $\mathcal{K}$ be an infinite global field. Any such field is the union of a tower of embedded global fields, $\mathcal{K} = \cup K_i$. Any embedded tower is asymptotically exact, and moreover, the corresponding limits do not depend on the choice of the tower, but only on the field $\mathcal{K}$ itself.*

**Basic Inequality.** To simplify the exposition, in what follows we assume the Generalized Riemann Hypothesis. Without it, the results for number fields become weaker (see Remark 1).

First we prove

**Theorem 1 (Basic Inequality).** *For any infinite global field (and for any asymptotically exact family of global fields) one has*

$$\sum_q \frac{\phi_q \log q}{\sqrt{q}-1} + \phi_{\mathbb{R}}(\log 2\sqrt{2\pi} + \frac{\pi}{4} + \frac{\gamma}{2}) + \phi_{\mathbb{C}}(\log 8\pi + \gamma) \leq 1,$$

*the sum being taken over all prime powers $q$.*

In the number field case this result generalizes the Odlyzko–Serre inequalities for the discriminant. In the function field case it becomes

$$\sum_{m=1}^{\infty} \frac{m\beta_m}{r^{m/2}-1} \leq 1,$$

and generalizes the Drinfeld–Vlăduţ theorem. Here $\beta_m = \phi_{r^m}$ is the ratio of the number of places of degree $m$ to the genus.

**Generalized Brauer–Siegel Theorem.** Our next concern is the class number.

**Lemma 2.** *For any infinite global field $\mathcal{K} = \cup K_i$ (and for any asymptotically exact family of global fields $K_i$) the limit*

$$BS(\mathcal{K}) = \lim_{i\to\infty} \frac{\log h(K_i)R(K_i)}{g(K_i)}$$

*exists. Moreover, for an infinite global field it does not depend on the choice of the tower, but only on the field $\mathcal{K}$ itself.*

Then we prove

**Theorem 2 (Generalized Brauer–Siegel Theorem).** *For any infinite global field (and for any asymptotically exact family of global fields) one has*

$$BS(\mathcal{K}) = 1 + \sum_q \phi_q \log \frac{q}{q-1} - \phi_{\mathbb{R}} \log 2 - \phi_{\mathbb{C}} \log 2\pi \ ,$$

*the sum being taken over all prime powers $q$.*

In the function field case it simplifies to

$$BS(\mathcal{K}) = 1 + \sum_{m=1}^{\infty} \beta_m \log \frac{r^m}{r^m - 1}.$$

**Bounds.** Next question is that of the possible asymptotic behaviour of the Brauer–Siegel ratio $BS(\mathcal{K})$. For the number field case we have

**Theorem 3 (Bounds).** *For any family of number fields*

$$BS_{low} \le \liminf_{i\to\infty} \frac{\log h(K_i)R(K_i)}{g(K_i)} \le \limsup_{i\to\infty} \frac{\log h(K_i)R(K_i)}{g(K_i)} \le BS_{up} \quad ,$$

*where*

$$BS_{low} = 1 - \frac{\log 2\pi}{\gamma + \log 8\pi} \approx 0.5165...,$$

*and*

$$BS_{up} = 1 + \frac{\log \frac{3}{2} + \log \frac{5}{4} + \log \frac{7}{6}}{\frac{\gamma}{2} + \frac{\pi}{4} + \log 2\sqrt{2\pi} + \frac{\log 2}{\sqrt{2}-1} + \frac{\log 3}{\sqrt{3}-1} + \frac{\log 5}{\sqrt{5}-1} + \frac{\log 7}{\sqrt{7}-1}} \approx 1.0938... .$$

We see that the possible values for the ratio $BS(\mathcal{K})$ lie in the interval $(0.5165..., 1.0938...)$.

Note that in the statement of the classical Brauer–Siegel theorem there is the assumption $n/\log\sqrt{|D|} \to 0$, which is equivalent to $\phi_{\mathbb{R}} = \phi_{\mathbb{C}} = 0$ (and hence, $\phi_q = 0$ for any $q$, since $\sum_{m=1}^{\infty} m\phi_{p^m} \le \phi_{\mathbb{R}} + 2\phi_{\mathbb{C}}$ for any prime $p$). In this case, Theorem 2 implies $BS = 1$.

The analogue of Theorem 3 for the function field case is

$$1 \le \liminf_{i\to\infty} \frac{\log_r h(K_i)}{g(K_i)} \le \limsup_{i\to\infty} \frac{\log_r h(K_i)}{g(K_i)} \le 1 + (\sqrt{r} - 1)\log_r \frac{r}{r-1} \quad .$$

**Examples.** Having in mind the classical value 1 of the Brauer–Siegel theorem itself, it is natural to ask whether in our more general setting there exist examples when the ratio differs from 1, or not. They do exist.

**Theorem 4.** *Consider the field $\mathcal{K}_1$ which is the union of the 2-class field tower over $\mathbb{Q}(\cos\frac{2\pi}{11}, \sqrt{2}, \sqrt{-23})$. This field is infinite and*

$$0.5939... \le BS(\mathcal{K}_1) \le 0.6025....$$

*For the field $\mathcal{K}_2$ which is the union of such 2-class field tower over*

$$\mathbb{Q}(\sqrt{11\cdot 13\cdot 17\cdot 19\cdot 23\cdot 29\cdot 31\cdot 37\cdot 41\cdot 43\cdot 47\cdot 53\cdot 59\cdot 61\cdot 67})$$

*that nine prime ideals lying over 2, 3, 5, 7 and 71 split completely, we have*

$$1.0602... \leq BS(\mathcal{K}_1) \leq 1.0798....$$

This shows that the condition $n/\log|D| \to 0$ in the Brauer–Siegel theorem is indeed indispensable for the ratio to be 1. Corresponding examples for the function field case are provided by modular curves.

**Zeta functions.** To understand these results we introduce a limit zeta-function, roughly, the limit of $g$-th roots of classical zeta-functions.

The *limit zeta function* of an asymptotically exact family is defined by the product

$$\zeta_\phi^0(s) := \prod_q (1 - q^{-s})^{-\phi_q} ,$$

$q$ running over all prime powers. The "corrected" zeta function is defined as

$$\zeta_\phi(s) := e^s 2^{-\phi_{\mathbb{R}}} \pi^{-s\phi_{\mathbb{R}}/2} (2\pi)^{-s\phi_{\mathbb{C}}} \Gamma(\frac{s}{2})^{\phi_{\mathbb{R}}} \Gamma(s)^{\phi_{\mathbb{C}}} \prod_q (1 - q^{-s})^{-\phi_q}.$$

**Lemma 3.** *The product defining zeta functions* $\zeta_\phi^0(s)$ *and* $\zeta_\phi(s)$ *absolutely converges for* $\mathrm{Re}\,(s) \geq \frac{1}{2}$.

Note that $\zeta_\phi(s)$ depends only on $\phi$'s and does not depend on the particular sequence of global fields. In particular, we have defined $\zeta_{\mathcal{K}}(s)$ for any infinite global field $\mathcal{K}$.

Let now

$$\xi_\phi(s) := (\log \zeta_\phi)' = \zeta_\phi'/\zeta_\phi =$$

$$1 - \frac{\phi_{\mathbb{R}}}{2} \log \pi - \phi_{\mathbb{C}} \log 2\pi + \frac{1}{2}\phi_{\mathbb{R}} \psi(\frac{s}{2}) + \phi_{\mathbb{C}} \psi(s) - \sum_q \phi_q \frac{\log q}{q^s - 1}$$

where $\psi(s) = \frac{\Gamma'}{\Gamma}(s)$ .

Then one can express the Basic Inequality (Theorem 1) as

$$\xi_\phi(1/2) \geq 0 ,$$

and the generalized Brauer–Siegel theorem (Theorem 2) either as

$$\lim_{i\to\infty} \frac{\log h_i R_i}{g_i} = \log \zeta_\phi(1),$$

or as

$$\lim_{i\to\infty} \frac{\log \kappa_i}{g_i} = \log \zeta_\phi^0(1),$$

$\kappa_i$ being the residue of $\zeta_{K_i}(s)$ at 1.

**Measures.** Let $\mathcal{K} = \cup K_j$ be an infinite number field. For each $K_j$ we define the measure

$$\Delta_{K_j} := \frac{1}{2g_{K_j}} \sum_{\zeta_{K_j}(\rho)=0} \delta_{t(\rho)} ,$$

where $t(\rho) = (\rho - \frac{1}{2})/i$, and $\rho$ runs over all non-trivial zeros of the zeta-function $\zeta_{K_j}(s)$. Because of the GRH $t(\rho)$ is real, and $\Delta_{K_j}$ is a discrete measure on $\mathbb{R}$. Moreover, $\Delta_{K_j}$ is a measure of slow growth.

Now we are ready to express the limit distribution of zeta zeros in terms of the parameters $(\phi_{\mathbb{R}}, \phi_{\mathbb{C}}, \phi_q)$ of the field $\mathcal{K}$.

**Theorem 5.** *In the space of measures of slow growth on $\mathbb{R}$ the limit*

$$\Delta_{\mathcal{K}} := \lim_{j\to\infty} \Delta_{K_j}$$

*exists. Moreover, the measure $\Delta_{\mathcal{K}}$ has a continuous density $M_{\mathcal{K}}$,*

$$M_{\mathcal{K}}(t) = \mathrm{Re}\left(\xi_{\mathcal{K}}\left(\frac{1}{2} + it\right)\right) =$$

$$1 - \sum_q \phi_q h_q(t) \log q + \frac{1}{2}\phi_{\mathbb{R}} \mathrm{Re}\,\psi\left(\frac{1}{4} + \frac{it}{2}\right) + \phi_{\mathbb{C}} \mathrm{Re}\,\psi\left(\frac{1}{2} + it\right)$$

$$-\frac{\phi_{\mathbb{R}}}{2} \log \pi - \phi_{\mathbb{C}} \log 2\pi,$$

*where*

$$h_q(t) = \frac{\sqrt{q}\cos(t \log q) - 1}{q + 1 - 2\sqrt{q}\cos(t \log q)}, \quad \psi(s) = \frac{\Gamma'}{\Gamma}(s).$$

**Corollary (Asymptotic Explicit Formula).** *We have*

$$\sum_q \frac{\phi_q \log q}{\sqrt{q} - 1} + \phi_{\mathbb{R}}(\log \sqrt{8\pi} + \frac{\pi}{4} + \frac{\gamma}{2}) + \phi_{\mathbb{C}}(\log 8\pi + \gamma) = 1 - M_{\mathcal{K}}(0) .$$

In other words, we have transformed the basic inequality into equality, showing that the deficiency, i.e., the difference between 1 and the left hand side, equals

$$\delta_{\mathcal{K}} = \xi_\phi(\frac{1}{2}) = M_{\mathcal{K}}(0) \geq 0.$$

In the function field case, the corresponding statement is also true. The Asymptotic Explicit Formula gives the asymptotic distribution law for Frobenius angles for asymptotically exact families of function fields, or, which is the same, the limit distribution law for zeroes of their zeta-functions. Let

$\mathcal{K} = \cup K_j$ be an infinite global field. For a zero $\rho$ of the zeta-function $\zeta_{K_j}(s)$ define $t(\rho)$ as

$$t(\rho) := \frac{\rho - \frac{1}{2}}{\pi i} .$$

Clearly, $t(\rho)$ is a real number defined modulo 2, and we suppose that $t(\rho) \in (-1, 1]$ which determines it uniquely.

Let

$$\Delta_j := \frac{1}{g_j} \sum_{\zeta_{K_j}(\rho)=0} \delta_{t(\rho)} ,$$

where $\delta_{t(\rho)}$ is, as usual, the Dirac measure supported at $t(\rho)$ . Then $\Delta_j$ is a measure of total mass 2 on $\mathbb{R}/2\mathbb{Z}$, and $\Delta_j$ is symmetric with respect to $t \mapsto -t$. Points of $\mathbb{R}/2\mathbb{Z}$ are given by their representatives in $(-1, 1]$.

In the function field case in the weak topology on the space of measures on $\mathbb{R}/2\mathbb{Z}$ the limit

$$\Delta(\mathcal{K}) := \lim_{j\to\infty} \Delta_j$$

exists. The measure $\Delta(\mathcal{K})$ has a continuous density $M_{\mathcal{K}}$ ,

$$M_{\mathcal{K}}(t) = \mathrm{Re}\,(\xi_\beta(\frac{1}{2} + \frac{i\pi}{\log r}t)) = 1 - \sum_{m=1}^{\infty} m\beta_m h_m(t)$$

for

$$h_m(t) = \frac{q^{m/2}\cos(\pi m t) - 1}{q^m + 1 - 2q^{m/2}\cos(\pi m t)} ,$$

which depends only on the field $\mathcal{K}$ and we have

$$\delta_{\mathcal{K}} = \xi_{\mathcal{K}}(\frac{1}{2}) = 1 - \sum_{m=1}^{\infty} \frac{m\beta_m}{r^{m/2} - 1} = M_{\mathcal{K}}(0) .$$

**Remark 1 (Unconditional results)**. What can we prove with no GRH at hand? Instead of the Basic Inequality we get two weaker ones

$$2\sum_q \phi_q \log q \sum_{m=1}^{\infty} (q^m + 1)^{-1} + \phi_{\mathbb{R}}(\gamma/2 + 1/2 + \log 2\sqrt{\pi}) + \phi_{\mathbb{C}}(\gamma + \log 4\pi) \le 1$$

and

$$\sum_q \frac{\phi_q \log q}{q - 1} + (\gamma/2 + \log 2\sqrt{\pi})\phi_{\mathbb{R}} + (\gamma + \log 2\pi)\phi_{\mathbb{C}} \le 1.$$

We can prove the existence of $BS(\mathcal{K})$ and the Generalized Brauer–Siegel Theorem only for the case of a normal infinite number field $\mathcal{K}$. (For any

asymptotically exact family of number fields we still have the "less than or equal to" inequality for the upper limit of the ratio.) The bounds weaken to

$$(0.4087\ldots, 1.1588\ldots).$$

In our examples the possible range for $BS(\mathcal{K})$ is then, respectively,

$$(0.5939\ldots, 0.6235\ldots) \text{ and } (1.0602\ldots, 1.0921\ldots).$$

We have no unconditional results for zero distributions.

**Remark 2 (Proofs)**. The proofs of Theorems 1 and 2 and of their unconditional counterparts are those of analytic number theory. We use Guinand–Weil and Lagarias–Odlyzko explicit formulas. Theorem 3 is treated as a linear programming problem, the proof being rather cumbersome but elementary. To produce the examples like those of Theorem 4 we turn to the algebraic number theory, using the class field towers technique and decomposition considerations. The zeta-function theory and Theorem 5 need some analytic technique again. It is comparatively easy to guess what the results should look like, and even to prove them in the function field case. The subtle part is to prove different convergencies in the number field case.

**Conclusion.** An infinite global field has a very strange set of invariants, which can be expressed either as an infinite sequence of non-negative real numbers $\{\phi_q(\mathcal{K})\}$, or as a limit zeta function $\zeta_{\mathcal{K}}(s)$, or as a limit zero measure $\Delta_{\mathcal{K}}$. This makes me expect that there exists a yet unknown to us non-trivial theory of infinite global fields.

## References

M.A. Tsfasman, S.G. Vlăduţ. *Asymptotic Properties of Zeta-Functions.* J. Math. Sciences, 1997, v. 84, n. 5, pp. 1445–1467.

M.A. Tsfasman, S.G. Vlăduţ. *Infinite Global Fields and the Generalized Brauer–Siegel Theorem.* Moscow Math. J., 2002, v. 2, n. 2.

# Author Index