Series on Coding Theory and Cryptology – Vol. 5

Editors

Edgar Martínez-Moro
Carlos Munuera
Diego Ruano

ADVANCES IN ALGEBRAIC GEOMETRY CODES

ADVANCES IN ALGEBRAIC GEOMETRY CODES

Series on Coding Theory and Cryptology

Editors: Harald Niederreiter *(National University of Singapore, Singapore)* and San Ling *(Nanyang Technological University, Singapore)*

Published

Vol. 1 Basics of Contemporary Cryptography for IT Practitioners
by B. Ryabko and A. Fionov

Vol. 2 Codes for Error Detection
by T. Kløve

Vol. 3 Advances in Coding Theory and Cryptography
eds. T. Shaska et al.

Vol. 4 Coding and Cryptology
eds. Yongqing Li et al.

Vol. 5 Advances in Algebraic Geometry Codes
eds. E. Martínez-Moro, C. Munuera and D. Ruano

Series on Coding Theory and Cryptology – Vol. 5

ADVANCES IN ALGEBRAIC GEOMETRY CODES

Editors

Edgar Martínez-Moro
Carlos Munuera
Universidad de Valladolid, Spain

Diego Ruano
Technical University of Denmark, Denmark

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data

Advances in algebraic geometry codes / edited by Edgar Martínez-Moro, Carlos Munuera & Diego Ruano.

p. cm. -- (Series on coding theory and cryptology ; v. 5)

Includes bibliographical references.

ISBN-13: 978-981-279-400-0 (hardcover : alk. paper)

ISBN-10: 981-279-400-X (hardcover : alk. paper)

1. Coding theory. 2. Geometry, Algebraic. 3. Error-correcting codes (Information theory).

I. Martínez-Moro, Edgar. II. Munuera, Carlos. III. Ruano, Diego.

QA268.A378 2008

005.7'2--dc22

2008035038

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

Printed in Singapore.

Preface

Error-correcting codes are used to achieve a reliable transmission of information through noisy channels. Due to their importance for many applications they became a meeting point between mathematics, computer science and engineering. All error-correcting codes are constructed using mathematical tools but, perhaps, the most deep and fascinating links between (classical) mathematics and codes can be found in Algebraic Geometry Codes.

The theory of Algebraic Geometry codes started over thirty years ago with the works of V.D. Goppa. Nowadays this theory is both a ripe subject and an exciting research field. At the same time, it has impelled research in different mathematical areas, as for example curves over finite fields.

In this book we try to provide the fundamentals, the 'state of the art' and the 'state of research', of this field. It consists of twelve chapters written by some of the most renowned specialists worldwide, each of them devoted to one of the main leading topics in this subject. These chapters are mostly self-contained and have been designed to be read independently.

We hope that this book will be useful for students and researchers in algebraic geometry and coding theory, as well as for computer scientists and engineers interested in information transmission.

We want to thank all the authors for their contribution to this volume. It was their efforts which made the publication of this book possible. Also we want to thank World Scientific and E. H. Chionh for their continuous support and excellent editorial job.

C. Munuera and *E. Martínez-Moro*
Dept. of Applied Mathematics,
University of Valladolid

D. Ruano
Department of Mathematics,
Technical University of Denmark

This page intentionally left blank

Contents

Chapter 1

Algebraic Geometry Codes: General Theory

Iwan M. Duursma

Department of Mathematics,
University of Illinois at Urbana-Champaign,
duursma@math.uiuc.edu

This chapter describes some of the basic properties of geometric Goppa codes, including relations to other families of codes, bounds for the parameters, and sufficient conditions for efficient error correction. Special attention is given to recent results on two-point codes from Hermitian curves and to applications for secret sharing.

Contents

Introduction

Geometric Goppa codes became famous when Tsfasman, Vladuts and Zink showed that infinite families of such codes can be constructed that exceed the Gilbert-Varshamov bound. An important step towards actual application of the codes came when Justesen, Larsen, Jensen, Havemose and Høholdt gave an efficient decoding algorithm for a special class of curves. Many curves have since then been proposed and studied for the construction of geometric Goppa codes. Decoding algorithms can now correct any geometric Goppa code up to half its designed minimum distance and improvements in their implementation continue to be made. Several new applications have been proposed that use special features of geometric Goppa codes. This chapter presents basic properties of geometric Goppa codes. The material is divided over four sections, with results on linear codes, cyclic codes, Reed-Muller codes, and geometric Goppa codes.

1.1. Linear codes and the affine line

Let $\mathbb{F}$ be a finite field. A $\mathbb{F}$-*linear code* C of *length* n is a linear subspace of $\mathbb{F}^n$. For $x, y \in \mathbb{F}^n$, the *Hamming distance* of x and y is

$$d(x,y) = |\{i : x_i \neq y_i, i = 1, 2, \ldots, n\}|.$$

The *minimum distance* of a nontrivial code C is

$$d(C) = \min\{d(x,y) : x, y \in C, x \neq y\}.$$

The dimension k of a code and the minimum distance d satisfy the Singleton bound,

$$k + d \leq n + 1.$$

Codes that attain the upper bound are called *maximum distance separable (MDS)*. An example is the code

$$C(< k, \mathbb{F}) = \{\ (f(a_1), f(a_2), \ldots, f(a_q))\ :\ f \in \mathbb{F}[x]_{<k}\ \},$$

for a fixed ordering $(a_1, a_2, \ldots, a_q)$ of the elements in $\mathbb{F}$. The code $C(< k, \mathbb{F})$ is a special case of an extended cyclic code, a Reed-Muller code, and a geometric Goppa code. Those three families of codes are the subject of the next three sections. In this section we describe a number of properties that are important for all three families but that actually hold for much larger

classes of codes if not for all linear codes.

The *dual code* $C^\perp$ of a code C is the maximal subspace of $\mathbb{F}^n$ that is orthogonal to C with respect to the standard inner product. A code is nondegenerate if neither the code nor its dual has a coordinate where all words are zero. The dual of the code $C(< k, \mathbb{F})$ is the code $C(< q-k, \mathbb{F})$, since $\sum_{x\in\mathbb{F}} x^i = 0$, for $i = 0, 1, \ldots, q-2$.

The *Singleton defect* or the *genus* of a code is $g(C) = n + 1 - k - d$. The dual of a MDS code is again MDS, but in general a code and its dual may have different genera. Every subset of k coordinates in an MDS code carries full information about the codeword. For a general code the *MDS discrepancy* or the *information defect* is the minimal m such that every subset of k coordinates contains at least $k - m$ information symbols. The parameter m is the same for a code and its dual and is at most the genus of a code.

1.1.1. *Dimension and infinite families*

A code C of type $[n, k, d]$ is *optimal* if it has maximal dimension for given length and minimum distance. For a family $\{[n_i, k_i, d_i]\}$ of optimal codes of increasing length with $\lim d_i/n_i = \delta$, define $\alpha(\delta) = \limsup k_i/n_i$. For an optimal code, each of its q^{n-k} cosets in $\mathbb{F}^n$ contains at least one vector y with $d(y, 0) < d$. The lower bound

$$q^{n-k} \le |\{y \in \mathbb{F}^n : d(y, 0) < d\}|$$

for the dimension of an optimal code is called the *Gilbert-Varshamov bound.* For $0 \le \delta \le \theta = (q-1)/q$,

$$\frac{1}{n} \log |\{y \in \mathbb{F}^n : d(y, 0) < \delta n\}| = H_q(\delta) + o(1),$$

where $H_q(x) = x \log_q(q-1) - x \log_q x - (1-x) \log_q(1-x)$, for $0 < x \le \theta$.

Theorem 1.1. *(asymptotic Gilbert-Varshamov bound) For an infinite family of optimal codes with relative distance $d/n = \delta$,*

$$\alpha(\delta) \ge 1 - H_q(\delta), \quad \textit{for } 0 < \delta \le \theta.$$

Lemma 1.2. *For a q-ary linear code with $k > m+1$ and $n-k > m(q^{m+1}-1)/(q-1) - (m+1)$ the information defect is at least m.*

Proof. Divide the coordinates in a subset of $k-(m+1)$ independent coordinates and its complement of $n-k+(m+1) > m(q^{m+1}-1)/(q-1)$ coordinates. In the subcode with zeros in the $k-(m+1)$ coordinates there is a block of size at least $m+1$ in which the nonzero coordinates are essentially repeated. Together the $k-(m+1)+(m+1)$ coordinates contain only $k-m$ information symbols. □

For families of $q-$ary linear codes with k and $n-k$ going to infinity, the information defect (and therefore also the genus) goes to infinity. As a consequence we obtain the following upper bound for the dimension of an optimal code.

Theorem 1.3. *(asymptotic Plotkin bound) For an infinite family of codes with $k, n-k \to \infty$, we have $d \leq \theta(n-k)$ as $n \longrightarrow \infty$, or*

$$\alpha(\delta) \leq 1 - \delta/\theta, \quad \text{for } 0 < \delta \leq \theta.$$

1.1.2. *Duality and differentials*

Let C be a linear code of length n. Omitting the i-th coordinate produces the *punctured code* $P_i(C)$ of length $n-1$. The *shortened code* $S_i(C)$ is the subcode of $P_i(C)$ of words with omitted i-th coordinate equal to zero. In general $P(C)^\perp = S(C^\perp)$. For a subset $\mathcal{P} = \{a_1, a_2, \ldots, a_n\}$ of the field $\mathbb{F}$, define a code

$$C(< k, \mathcal{P}) = \{\ (f(a_1), f(a_2), \ldots, f(a_n))\ :\ f \in \mathbb{F}[x]_{<k}\ \}.$$

The code $C(< k, \mathcal{P})$ is a punctured version of the code $C(< k, \mathbb{F})$. The dual code is a shortened version of the code $C(< q-k, \mathbb{F})$.

$$C(< k, \mathcal{P})^\perp = \{\ (f(a_1), f(a_2), \ldots, f(a_n))\ :\ f \in \mathbb{F}[x]_{<q-k}, f|_{\mathbb{F}-\mathcal{P}} = 0\ \}.$$

Let $p(x) = (x-a_1)(x-a_2)\cdots(x-a_n)$ and let $x^q - x = p(x)r(x)$. Then $-1 = p'(a_i)r(a_i)$, for $i = 1, 2, \ldots, n$. With f of the form $f = rh$,

$$C(< k, \mathcal{P})^\perp = \{\ (\frac{h(a_1)}{p'(a_1)}, \frac{h(a_2)}{p'(a_2)}, \ldots, \frac{h(a_n)}{p'(a_n)})\ :\ h \in \mathbb{F}[x]_{<n-k}\ \}.$$

We give a description of the dual code using differentials. For a polynomial $h \in \mathbb{F}[x]$ of degree $\deg h < n$, let

$$\omega = \frac{h}{p}\,dx = (\frac{c_1}{x-a_1} + \frac{c_2}{x-a_2} + \cdots + \frac{c_n}{x-a_n})\,dx$$

be a differential with at most simple poles and with residues $c_1, c_2, \ldots, c_n$ at $x = a_1, a_2, \ldots, a_n$, respectively. Then $h(a_i) = c_i p'(a_i)$ and

$$C(< k, \mathcal{P})^{\perp} = \{(\mathrm{res}_{a_1}(\omega), \mathrm{res}_{a_2}(\omega), \ldots, \mathrm{res}_{a_n}(\omega)) \ : \\ \omega = \frac{h}{p}\, dx, \ \deg h < n - k\}.$$

1.1.3. *Minimum distance*

Several useful inequalities exist for the parameters of codes A, B, C with

$$\sum_i a_i b_i c_i = 0, \quad \text{for all } a \in A, b \in B, c \in C.$$

Let $a * b = (a_1 b_1, \cdots, a_n b_n)$ denote the Hadamard product or coordinate-wise product of two vectors a and b. The relation between A, B, C can be formulated as $A * B = \{a * b : a \in A, b \in B\} \subset C^{\perp}$. Such decompositions of the dual code under the Hadamard product form the basis for several bounds for the minimum distance.

Theorem 1.4. *(Roos bound for linear codes) For a linear code C, and for linear codes A and B with $A * B \subset C^{\perp}$,*

$$g(A) < d(B^{\perp}) - 1 \ \Rightarrow \ d(C) \geq k(A) + d(B^{\perp}) - 1.$$

Proof. It is enough to show that for every subset I of $(k(A) - 1) + (d(B^{\perp}) - 1)$ positions there exists a word $a * b \in A * B$ with precisely one nonzero coordinate in those positions. First choose $a \in A$ with zeros in $k(A) - 1$ positions of I. Then choose $b \in B$ with a single nonzero coordinate in the remaining $d(B^{\perp}) - 1$ positions such that the nonzero coordinate appears in a position where a is nonzero. This is possible since a has no more then $n - d(A) < k(A) - 1 + d(B^{\perp}) - 1$ zeros. □

Theorem 1.5. *(Symmetric Roos bound for linear codes) For a linear code C, and for linear codes A and B with $A * B \subset C^{\perp}$,*

$$g(A) < k(B) \text{ and } g(B) < k(A) \\ \Rightarrow \ d(C) \leq g(A) + g(B) \text{ or } d(C) \geq k(A) + k(B).$$

The two versions of the Roos bound can be used in combination, with different choices for A and B, to produce stronger results.

Theorem 1.6. *(Shift bound or Coset bound) Let C be a linear code and let $C_1 \subset C$ be a maximal subcode. If there exist vectors $a_1, \ldots, a_w$ and*

$b_1, \ldots, b_w$ such that

$$\begin{cases} a_i * b_j \in C^\perp & \text{for } i+j \leq w, \\ a_i * b_j \in C_1^\perp \backslash C^\perp & \text{for } i+j = w+1, \end{cases}$$

then words in $C \backslash C_1$ have weight at least w.

Proof. For all $c \in C \backslash C_1$ and $a * b \in C_1^\perp \backslash C^\perp$, $\sum_i a_i b_i c_i \neq 0$. Thus it suffices to show the existence, for any choice of $w-1$ coordinates, of a vector $a * b \in C_1^\perp \backslash C^\perp$ that vanishes in those coordinates. The conditions show that the vectors $a_1, \ldots, a_w$ are linearly independent, and there exists a nonzero linear combination a of the vectors $a_1, \ldots, a_w$ vanishing at $w-1$ given coordinates coordinates. If i is maximal such that a_i has a nonzero coefficient in the linear combination a then $a * b_{w+1-i} \in C_1^\perp \backslash C^\perp$ and vanishes in the $w-1$ coordinates. □

Theorem 1.7. *(Iterated coset bound) Repeated application of the coset bound to a sequence $C_r \subset \cdots \subset C_1 \subset C_0 = C$ gives the lower bound $d(C) \geq \min\{d_1, d_2, \ldots, d_r, d(C_r)\}$, where $d(C_{i-1}/C_i) \geq d_i$ is obtained with the coset bound.*

1.1.4. *Error correction*

Let A, B, and C be nondegenerate linear codes such that

$$\sum_i a_i b_i c_i = 0, \quad \text{for all } a \in A, b \in B, c \in C.$$

If $k(A) > t$ and $d(B^\perp) > t$ then (A, B) is called a t-error-locating pair for C. For a given error-locating pair the error positions in a received word can be located by solving a suitable system of linear equations.

Theorem 1.8. *Let (A, B) be a t-error-locating pair for C. For $c \in C$ and for a vector e of weight at most t, let $y = c + e$. Every vector $a \in A$ with $a * y \perp b$ for all $b \in B$ has the property $a * e = 0$.*

An error-locating pair for C is called error-correcting if moreover $d(A) + d(C) > n$. For a given error-correcting pair a codeword can be recovered from the zeros in an error locating vector $a \in A$ by solving a second suitable system of linear equations.

Theorem 1.9. *Let (A, B) be a t-error-correcting pair for C. For $c \in C$ and for a vector e of weight at most t, let $y = c + e$. Let $a \in A$ have the*

property $a * e = 0$. Then $c \in C$ is the unique solution to the system of equations $c \in C$ and $a * c = a * y$.

The key equation $a * y \perp b$ for all $b \in B$ amounts to a linear system of $\dim(B)$ equations in $\dim(A)$ unknowns. A different formulation gives a key equation with n linear equations in $\dim(A) + \dim(B^{\perp})$ unknowns.

Theorem 1.10. *For $c \in C$ and for a vector e of weight at most t, let $y = c + e$. For every pair of vectors $a \in A, \hat{b} \in B^{\perp}$ with $a * y = \hat{b}$, the vector c is the unique solution to the system of equations $c \in C$ and $a * c = \hat{b}$.*

In general, the decoding is not completed with $c \in C$ since c is merely an encoding of the relevant information symbols. In such cases it may be better to bypass the computation of c and to solve directly for the information symbols. The t-error-correcting code $C(< q - 2t, \mathbb{F})$ has a t-error-correcting pair $(A = C(\leq t, \mathbb{F}), B = C(< t, \mathbb{F})$. The key equation for an error-locating vector is: determine $g(x) \in \mathbb{F}[x]_{\leq t}$ such that

$$\sum_i y_i g(x_i) h(x_i) = 0, \quad \text{for all } h \in \mathbb{F}[x]_{<t}.$$

When t errors occur, the solution for $g(x)$ is the unique polynomial that vanishes in those positions. The second key equations is: determine $g(x) \in \mathbb{F}[x]_{\leq t}$ and $\hat{h}(x) \in \mathbb{F}[x]_{<q-t}$ such that

$$y_i g(x_i) = \hat{h}_i(x_i), \quad \text{for } i = 1, 2, \ldots, n.$$

When t errors occur, the solution is the pair $(g(x), f(x)g(x))$ where $c_i = f(x_i)$ for $i = 1, 2 \ldots, n$. In general, the information symbols are the coefficients of f. The key equation with n equations generalizes to list decoding. List decoding produces a list of bounded size ℓ that contains all code words that are within distance t of the received word.

Theorem 1.11. *For a code $C(< k, \{x_1, \ldots, x_n\})$ and a received vector y, let*

$$Q(x, y) = \sum_{i=0}^{\ell} g_i(x) y^i, \quad \deg g_i < n - t - i(k-1),$$

be a nonzero polynomial such that $Q(x_i, y_i) = 0$ for $i = 1, 2, \ldots, n$. Then $y - f(x)$ divides $Q(x, y)$ for all f with $y_i = f(x_i)$ in at least $n - t$ positions.

Let E_I be the subspace of $\mathbb{F}^n$ generated by unit vectors e_i with $i \in I$, for $I \subset \{1, \ldots, n\}$. For an error vector $e \in E_I$, we reformulate the sufficient conditions for error correction in terms of I. Let $\bar{I} = \{1, \ldots, n\} \backslash I$.

Theorem 1.12. *For a linear code C, let A and B be linear codes with $A * B \subset C^\perp$, such that for all $a \in A$ and $c \in C$, $a * c = 0$ if and only if $a = 0$ or $c = 0$. Let $y = c + e$, with $c \in C$ and $e \in E_I$. If I is such that*

$$A \cap E_{\bar{I}} \neq 0 \quad \text{and} \quad B^\perp \cap E_I = 0$$

then there exists a nonzero vector $a \in A$ with $a * y \perp b$ for all $b \in B$. And for any such a, $c \in C$ is the unique solution to the system of equations $c \in C$ and $a * c = a * y$.

1.1.5. *Linear secret sharing schemes*

An *ideal $\mathbb{F}-$linear secret sharing scheme* (LSSS) $\Sigma = \Sigma_0(\Pi)$ on the set of *players* $\mathcal{P} = \{1, 2, \ldots, n\}$ is a sequence $\Pi = (\pi_0, \pi_1, \ldots, \pi_n)$ of surjective linear mappings $\pi_i \colon E \to \mathbb{F}$, where E is a vector space of finite dimension over $\mathbb{F}$. For a given $s \in \mathbb{F}$ and for a randomly chosen $x \in E$ with $\pi_0(x) = s$, the values $\pi_1(x), \ldots, \pi_n(x)$ form a collection of *shares* for the *secret value* $\pi_0(x)$. A subset $A \subset \mathcal{P}$ is *qualified* or *accepted by* Σ if the players in A can determine the secret value uniquely from their shares. Otherwise A is *unqualified* or *rejected by* Σ.

Lemma 1.13. *A subset $A \subset \mathcal{P}$ is unqualified if and only if there exists $x \in E$ with $\pi_0(x) = 1$ and $\pi_i(x) = 0$ for all $i \in A$.*

For a LSSS $\Sigma = \Sigma_0(\Pi)$, let $\hat{C} = \{(\pi_1(x), \ldots, \pi_n(x), \pi_0(x)) : x \in E\}$ be the linear code of length $n + 1$ with shares in the first n positions and secret value in the last position. Let C denote the punctured code $\{(\pi_1(x), \ldots, \pi_n(x)) : x \in E\}$ and let C^0 denote the shortened code $\{(\pi_1(x), \ldots, \pi_n(x)) : x \in E, \pi_0(x) = 0\}$.

Theorem 1.14. *(Rejection bound) Let $\Sigma = \Sigma(\hat{C})$. If there exist vectors $a_0, \ldots, a_t \in \mathbb{F}^n$ and $b_0, \ldots, b_t \in \mathbb{F}^n$ such that*

$$\begin{cases} a_i * b_j \in C^0 & \text{for } i + j < t. \\ a_i * b_j \in C \backslash C^0 & \text{for } i + j = t. \end{cases}$$

then any subset $A \subset \mathcal{P}$ of size at most t is rejected by Σ.

Proof. A subset of players can not recover the secret s if and only if there exists a vector in $C \backslash C^0$ that is zero in their positions. The conditions show that the vectors $a_0, \ldots, a_t$ are independent. For a given set of t players there exists a nonzero linear combination a of the vectors $a_0, \ldots, a_t$ that vanishes at their coordinates. If i is maximal such that a_i has a nonzero

coefficient in the linear combination a then $a * b_{t-i} \in C \backslash C^0$ and vanishes in the t coordinates. □

A LSSS Σ is *nondegenerate* if the secret can be reconstructed as a linear combination of all the shares. That is, there exist $r_1, \ldots, r_n \in \mathbb{F}$ such that

$$\pi_0(x) = \sum_i r_i \pi_i(x), \quad \text{for all } x \in E.$$

The same values reconstruct the sum $\pi_0(x) + \pi_0(y)$ of two secrets from the pairwise sums $\pi_i(x) + \pi_i(y)$ of their shares. We call Σ *additive in $n-t$ positions* if for any subset $A \subset \mathcal{P}$ of size t there exists a choice for $r_1, \ldots, r_n \in \mathbb{F}$ with $r_i = 0$ for $i \in A$.

Proposition 1.15. *For a given LSSS $\Sigma(\hat{C})$, let $\Sigma(\hat{D})$ be the scheme defined with the dual code $\hat{D}$ of $\hat{C}$. Then $\Sigma(\hat{C})$ is additive in $n-t$ positions if and only if $\Sigma(\hat{D})$ rejects all subsets $A \subset \mathcal{P}$ of size t.*

To implement secure protocols for multiparty computations that involve addition and multiplication, a stronger property is needed. A LSSS Σ is *multiplicative* if the product $\pi_0(x) \cdot \pi_0(y)$ of two secrets can be reconstructed as a linear combination of the pairwise products $\pi_i(x) \cdot \pi_i(y)$ of the shares, i.e. if there exist $r_1, \ldots, r_n \in \mathbb{F}$ such that

$$\pi_0(x)\pi_0(y) = \sum_i r_i \pi_i(x)\pi_(y), \quad \text{for all } x, y \in E.$$

We call Σ *multiplicative in $n-t$ positions* if for any subset $A \subset \mathcal{P}$ of size t there exists a choice for $r_1, \ldots, r_n \in \mathbb{F}$ with $r_i = 0$ for $i \in A$. A LSSS Σ is called *strongly multiplicative* if for any unqualified subset $A \subset \mathcal{P}$ there exists a choice for $r_1, \ldots, r_n \in \mathbb{F}$ with $r_i = 0$ for $i \in A$.

Proposition 1.16. *For a given LSSS $\Sigma(\hat{C})$, let $\Sigma(\hat{B})$ be the scheme defined with the maximal code $\hat{B}$ that is orthogonal to $\hat{C} * \hat{C}$. Then $\Sigma(\hat{C})$ is multiplicative in $n-t$ positions if and only if $\Sigma(\hat{B})$ rejects all subsets $A \subset \mathcal{P}$ of size t. And $\Sigma(\hat{C})$ is strongly multiplicative if and only if $\Sigma(\hat{B})$ rejects all unqualified subsets $A \subset \mathcal{P}$.*

A LSSS $\Sigma(\hat{C})$ that is multiplicative in $n-t$ positions (or that is strongly multiplicative) has a decomposition $\hat{D} \supset \hat{C} * \hat{B}$ of the dual code $\hat{D}$. This decomposition can be used to apply error correction as in the previous section to recover the secret in the presence of corrupted shares. The following theorem outlines a dedicated secret reconstruction procedure that avoids correcting corrupted shares and instead computes the secret directly. For

geometric Goppa codes the theorem is a way to recover the value $f(P_0)$ from possibly erroneous values $f(P_1), \ldots, f(P_n)$. Since the point P_0 can be chosen arbitrarily, the function f can be recovered completely and the theorem provides a way to decode geometric Goppa codes up to half their designed minimum distance (Theorem 1.58 in Section 1.4.7).

Theorem 1.17. *(Secret reconstruction) Let* $\Sigma = \Sigma_0(\Pi), \Sigma' = \Sigma_0(\Pi'), \Sigma'' = \Sigma_0(\Pi'')$ *be LSSSs such that* $\sum_i \pi_i(x)\pi'(y)\pi''(z) = 0$, *for all* $x \in E, y \in E', z \in E''$. *For a possibly corrupted vector of shares* $(s_1, \ldots, s_n)$ *for* Σ, *let* $(y, z) \in E' \times E''$ *be such that* $\pi'_0(y) = \pi''_0(z) = 1$ *and*

$$0 = \sum_i s_i \pi'_i(y)\pi''_i(z_0) \quad \forall z_0 \in E'' \text{ with } \pi''_0(z_0) = 0,$$

$$0 = \sum_i s_i \pi'_i(y_0)\pi''_i(z) \quad \forall y_0 \in E' \text{ with } \pi'_0(y_0) = 0.$$

If the corrupted shares are contained in a subset A *that is rejected by both* Σ' *and* Σ'' *then such a pair* (y, z) *exists and the secret for the uncorrupted vector of shares is*

$$s = -\sum_i s_i \pi'(y)\pi''(z).$$

If either Σ' *or* Σ'' *rejects* A *but not both then a pair* (y, z) *may not exist. If it exists then the formula for the secret produces the correct value for* s.

Proof. Assume that A is rejected by Σ''. Then $z = z_1$ with $\pi''_0(z_1) = 1$ and $\pi''_i(z_1) = 0$ for $i \in A$ gives a solution for z. An arbitrary $z \in E''$ with $\pi''_0(z) = 1$ is of the form $z = z_0 + z_1$ with $\pi''_0(z_0) = 0$. For a solution y to the first equation and for an arbitrary $z \in E''$ with $\pi''_0(z) = 1$,

$$\sum_i s_i \pi'_i(y)\pi''_i(z) = \sum_i s_i \pi'(y)\pi''(z_1) = \sum_i \pi_i(x)\pi'(y)\pi''(z_1) = -s.$$

This clearly implies the claims in the theorem. □

The choices that are made for y and z in general need not vanish in the corrupted shares. In general, the secret is reconstructed without obtaining information about corrupted players. Clearly the two equations reduce to a single equation when $\Sigma' = \Sigma''$.

A LSSS $\Sigma = \Sigma_0(\Pi)$ with $\sum_i \pi_i(x)\pi_i(y)\pi_i(z) = 0$ for all $x, y, z \in E$ is called trilinear. Such a scheme is strongly multiplicative and can reconstruct the secret efficiently whenever the corrupted shares are con-

tained in an unqualified subset. A trilinear scheme that rejects all subsets of size t is multiplicative in $n - t$ positions. The Shamir LSSS $\Sigma_0(\leq t, \{a_1, \ldots, a_n, a_0\})$ is the scheme $\Sigma_0(\Pi)$, where $\Pi : \mathbb{F}[x]_{\leq t} \longrightarrow \mathbb{F}^{n+1}$, $f \to (f(a_1), \ldots, f(a_n), f(a_0))$.

Theorem 1.18. *The Shamir LSSS* $\Sigma_0(\leq t, (a_1, \ldots, a_n, a_0))$ *rejects all subsets of size* t *or less and accepts all subsets of size* $t+1$ *or more. For* $3t < n$*, the scheme is trilinear.*

Proof. For $p = (x - a_0)(x - a_1) \cdots (x - a_n)$, and for $r_i = p'(a_i)$,

$$\sum_{i=0}^{n} r_i f(a_i) g(a_i) h(a_i) = 0, \forall f, g, h \in \mathbb{F}_{\leq t}[x].$$

□

1.1.6. *Weight distributions and codes over extension fields*

The weight distribution of a linear code C of length n is the vector $(A_0, A_1, \ldots, A_n)$, where A_i is the number of words of weight i in C. For a q-ary code the weight enumerator $A(x, y)$ and the projective weight enumerator $\bar{A}(x, y)$ are defined by

$$A(x, y) = \sum_{i=0}^{n} A_i x^{n-i} y^i = x^n + (q-1)\bar{A}(x, y).$$

For the code $C(< k, \{a_1, \ldots, a_n\})$ we can describe the projective weight enumerator in terms of the zeta function of the affine line. The latter is a generating function for the number of monic polynomials of a given degree, with Euler product factorization

$$(1 - qT)^{-1} = \prod_{f \text{monic, irr}} (1 - T^{\deg f})^{-1}.$$

The number of monic polynomials of degree less than k that vanish in precisely $n - i$ elements of $\{a_1, \ldots, a_n\}$ becomes

$$\bar{A}_i = [T^{k-1}] \frac{\binom{n}{i} T^{n-i} (1 - T)^i}{(1 - T)(1 - qT)}$$

and

$$\bar{A}(x, y) = [T^{k-1}] \frac{(xT + y(1 - T))^n}{(1 - T)(1 - qT)}.$$

For a given coordinate, the weight enumerator of a code can be described recursively in terms of the punctured code and the shortened code at the given coordinate.

$$A(x,y) = \begin{cases} xS(x,y), & \text{if } j \text{ is a loop} \\ (x+(q-1)y)P(x,y), & \text{if } j \text{ is a bridge} \\ yP(x,y)+(x-y)S(x,y), & \text{otherwise.} \end{cases}$$

A coordinate is called a loop if shortening preserves the dimension and a bridge if puncturing lowers the dimension. A code is nondegenerate if it has no loops or bridges. An invariant that satisfies a recursion of the above type is called a Tutte-Grothendieck invariant. By continuing the recursion it is clear that there exist polynomials $T(x,y)$, called the Tutte polynomial, and $W(x,y) = T(x+1,y+1)$, called the Whitney polynomial, such that

$$\frac{A(x,y)}{(x-y)^k y^{n-k}} \;=\; T(\frac{x+(q-1)y}{x-y}, \frac{x}{y}) \;=\; W(\frac{qy}{x-y}, \frac{x-y}{y}).$$

The recursive procedure, and thus the polynomials T and W, remains the same if the q-ary code is extended to a code with coefficients in an extension field of size q^m. The weight enumerator $A^{(m)}$ of the q^m-ary code is

$$\frac{A^{(m)}(x,y)}{(x-y)^k y^{n-k}} \;=\; W(\frac{q^m y}{x-y}, \frac{x-y}{y}).$$

For a weight enumerator $A(x,y)$, let

$$P(x,y) = \frac{1}{n}(\frac{\partial}{\partial_x} + \frac{\partial}{\partial_y})A(x,y), \qquad S(x,y) = \frac{1}{n}(\frac{\partial}{\partial_x})A(x,y)$$

be the average punctured and shortened weight enumerator, respectively. They clearly satisfy the recursion type $A(x,y) = yP(x,y) + (x-y)S(x,y)$ of a nondegenerate code. Let $a_w = A_w/\binom{n}{w}$, for $w = 0,1,\ldots,n$. Define the normalized weight enumerator as

$$a(t) = \frac{1}{q-1}(a_d + a_{d+1}t + \cdots + a_n t^{n-d})$$

Theorem 1.19. *The expression*

$$a(t)(1+t)^{d+1} \pmod{t^{n-d+1}}$$

is invariant under puncturing and averaging or shortening and averaging. For the q-ary code $C(<k,\{a_1,\ldots,a_n\}$ the expression agrees with the evaluation of $1/(1-T)(1-qT)$ at $T = t/(1+t)$.

1.2. Cyclic codes and classical Goppa codes

A $\mathbb{F}$-linear code C of length n with coordinates $\{0, 1, \ldots, n-1\}$ is cyclic if, after identifying words $c = (c_0, c_1, \ldots, c_n)$ with polynomials $c(x) = c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}$, the code is an ideal in the ring $R = \mathbb{F}[x]/(x^n - 1)$. The ring R is a principal ideal domain. Polynomials $g(x)$ and $\gcd(x^n - 1, g(x))$ generate the same ideal in R and cyclic codes of length n correspond one-to-one to factors $g(x)$ of $x^n - 1$. If $\gcd(\mathrm{char}\,\mathbb{F}, n) = 1$, then $x^n - 1$ factors over $\mathbb{F}$ as a product of distinct irreducible polynomials. For a factorization $x^n - 1 = f_1 \cdots f_t$ into t irreducible factors, there are 2^t cyclic codes of length n over $\mathbb{F}$.

The code C with generating polynomial $g(x)|x^n - 1$ is determined by the irreducible factors in $g(x)$ or by the zeros of $g(x)$ in an algebraic closure $\overline{\mathbb{F}}$ of $\mathbb{F}$. Let $\alpha \in \overline{\mathbb{F}}$ be a primitive n-th root of unity. For $i \in \mathbb{Z}/n\mathbb{Z}$, let $m_i(x)$ be the minimal polynomial of α^i over $\mathbb{F}$. If $g(x) = \mathrm{lcm}\{m_i(x) : i \in I\}$ then I is called a *defining set* for C. The maximal defining set for C is the set $\{i \in \mathbb{Z}/n\mathbb{Z} : g(\alpha^i) = 0\}$. The dual code of a cyclic code with maximal defining set I is cyclic with maximal defining set $I^* = \mathbb{Z}/n\mathbb{Z} \setminus -I$, where we use $\sum_{k=0}^{n-1}(\alpha^{i+j})^k = 0$, for all $i, j \in \mathbb{Z}/n\mathbb{Z}$ with $i + j \neq 0$, Thus, the dual code of the code generated by $g(x)$ is the code generated by $h(x) = (x^n - 1)/g^*(x)$, where $g^*(x)$ is the reciprocal polynomial of $g(x)$.

1.2.1. *Reed-Solomon and BCH codes*

Of particular interest among cyclic codes are *BCH codes*, that are defined with a defining set of the form $I = \{b + 1, b + 2, \ldots, b + \delta - 1\}$. A BCH code over a field of q elements is called *primitive* if the length $n = q^m - 1$, for $m \geq 1$. A Reed-Solomon code is a primitive BCH code of length $n = q^m - 1$ over the field of q^m elements. For the given defining set, a Reed-Solomon code has parameters $[q^m, q^m + 1 - \delta, \delta]$. Primitive BCH codes in general have a maximal defining set that is larger than I. They are subcodes of Reed-Solomon codes and have minimum distance $d \geq \delta$. A lower bound for the dimension is $k \geq n - m(\delta - 1)$, with an improvement $k \geq n - m(q-1)\lceil(\delta-1)/q\rceil$ when $b = 0$. BCH codes are an important way to construct long codes over a given finite field such that both the minimum distance and the dimension have lower bounds. However asymptotically BCH codes are not good. For an infinite family of BCH codes of increasing length, either the relative distance d/n or the information rate k/n goes to zero as n goes to infinity.

Theorem 1.20. *The Reed-Solomon code of length $n = q^m - 1$ over the field of q^m elements with defining set $I = \{b+1, b+2, \ldots, b+\delta-1\}$ has as codewords the vectors $(f(\alpha^0), f(\alpha^1), \ldots, f(\alpha^{n-1}))$, for $f \in L = \langle x^{-b}, \ldots, x^{-1}, 1, x, \ldots, x^a \rangle$, where a is such that $a + b = n - \delta$. The BCH code over the field of q elements with the same length and defining set is a subcode of the Reed-Solomon code.*

The space L is the vector space of rational functions in x with pole order at most a at ∞, pole order at most b at 0, and no other poles. In the terminology of the next section the Reed-Solomon code over the field $\mathbb{F}$ with defining set $I = \{b+1, b+2, \ldots, b+\delta-1\}$ is a two-point code $C_L(aP_\infty + bP_0, \mathbb{F}^*)$. When $b = 0$ the code $C_L(aP_\infty, \mathbb{F})$ is called a one-point code. These are the codes $C(\leq a, \mathbb{F})$ that were used as a main example in the previous section. BCH codes with $b = 0$, i.e. subfield subcodes of one-point codes, are called *narrow sense*.

To apply the theorems in Section 1.1.3 to cyclic codes requires a decomposition of their defining set. We illustrate this for the dual of the two-error correcting BCH code of length $n = 15$ over $\mathbb{F}_4$. The BCH code has defining set $\{1, 2, 3, 4\}$ and complete defining set $I = \{1, 2, 3, 4, 8, 12\}$. The dual code has complete defining set $I^* = \{0, 1, 2, 4, 5, 6, 8, 9, 10\}$. The code and its dual are of type $[15, 9, 5]$ and $[15, 6, 8]$, respectively. For the decomposition $I^* \supset \{0, 1, 2, 4, 5, 6\} + \{0, 4\}$, Theorem 1.4 gives $d \geq 8$. For the decomposition $I^* \supset \{0, 2, 4\} + \{0, 2, 4, 6\}$, Theorem 1.4 only gives $d \geq 7$. On the other hand this decomposition can be used with Theorem 1.9 to correct any three errors. For the decomposition $I^* \supset \{0, 1, 4, 5\} + \{0, 1, 4, 5\}$, Theorem 1.5 gives $d \leq 4$ or $d \geq 8$. The pair meets the conditions of Theorem 1.9 for correcting three errors, so the possibility $d \leq 4$ is easily excluded. Of the two decompositions that correct any three errors, the second has the property that the codes A and B can be defined over $\mathbb{F}_4$, while in the first case decoding takes place over the field $\mathbb{F}_{16}$.

1.2.2. *Classical Goppa codes*

The family of classical Goppa codes includes as subfamily the BCH codes but is large enough to contain infinite families of codes of increasing length that attain the asymptotic Gilbert-Varshamov lower bound for the dimension of optimal codes.

Let $\alpha_1, \cdots, \alpha_n \subset \mathbb{F}$ be distinct field elements and let $g(x) \in \mathbb{F}[x]$ be a monic polynomial that is relatively prime to $p(x) = (x - \alpha_1) \cdots (x - \alpha_n)$.

The classical Goppa code defined with the polynomial $g(x)$ is the set of all words $c = (c_1, \cdots, c_n) \in \mathbb{F}^n$ with

$$\frac{c_1}{x-\alpha_1}\,dx + \cdots + \frac{c_n}{x-\alpha_n}\,dx \;=\; \frac{h(x)}{p(x)}\,dx, \text{ for } g(x)|h(x).$$

The polynomial h is of degree at most $n-1$. It vanishes at the zeros of a word $c = (c_1, \cdots, c_n)$. Since $g|h$, we have $n - d \leq \deg h - \deg g \leq n - 1 - t$ and $d \geq t + 1$.

Theorem 1.21. *Let C be the classical Goppa code over $\mathbb{F}_{q^m}$ defined with relatively prime polynomials $g(x)$ and $p(x)|x^{q^m} - x$. The dual code $C^\perp$ of C is obtained by evaluation of functions in $L = \langle h/g : \deg h < \deg g \rangle$.*

Proof. As in Section 1.1.2,

$$C = \langle\, (\frac{g(\alpha_1)\alpha_1^i}{p'(\alpha_1)}, \cdots, \frac{g(\alpha_n)\alpha_n^i}{p'(\alpha_n)}) \,:\, i = 0, 1, \ldots, n-t-1 \,\rangle,$$

$$C^\perp = \langle\, (\frac{\alpha_1^j}{g(\alpha_1)}, \cdots, \frac{\alpha_n^j}{g(\alpha_n)}) \,:\, j = 0, 1, \ldots, t-1 \,\rangle.$$

□

In the terminology of geometric Goppa codes, C is defined by evaluating residues of differentials $\omega \in \Omega(G - P_\infty)$ and $C^\perp$ by evaluating values of functions $f \in L(G - P_\infty)$, where G is the divisor of zeros of $g(x)$. A classical Goppa code over the subfield $\mathbb{F}$ of size q is a subfield subcode of the code C. The Reed-Solomon code of length $n = q^m - 1$ with $I = \{1, 2, \ldots, \delta - 1\}$ has a dual code that is defined by the evaluation of functions $f \in L((\delta-1)P_\infty - P_0) = \langle x, \ldots, x^{\delta-1} \rangle$. To realize a Reed-Solomon code as a classical Goppa code we evaluate instead the functions $f \in L((\delta-1)P_0 - P_\infty) = \langle x^{1-\delta}, \ldots, x^{-1} \rangle$. A value in position α for the Reed-Solomon code appears with the different evaluation in position α^{-1}. The rearranged Reed-Solomon code is a classical Goppa code with divisor $G = (\delta-1)P_0$ and polynomial $g(x) = x^{\delta-1}$.

Let $V(d-1) = |\{y \in \mathbb{F}^n : d(y, 0) < d\}|$ be the number of words in a closed ball of Hamming radius $d-1$. Recall that the Gilbert-Varshamov bound shows that for given n and d, there exist codes with $q^{n-k} \leq V(d-1)$.

Theorem 1.22. *For a given length $n = q^m$ and minimum distance d, there exist irreducible polynomials $g(x)$ over $\mathbb{F}_{q^m}$ such that the classical Goppa code defined with $g(x)$ has minimum distance at least d and dimension attaining the asymptotic Gilbert-Varshamov bound.*

Proof. Let t denote the degree of $g(x)$. The number of irreducible polynomials of degree t over q^m is at least $(q^{mt} - d(t)q^{mt/2})/t$, where $d(t)$ is the number of divisors of t. For a word $(c_1, \ldots, c_n)$ of weight at most $d-1$, $h(x)$ has at least $n - d + 1$ zeros in common with $p(x)$, and the cofactor of degree at most $d-2$ contains no more than d/t irreducible factors of degree t. Thus, for $d/t \cdot V(d-1) < (q^{mt} - d(t)q^{mt/2})/t$, there exist classical Goppa codes with polynomial $g(x)$ of degree t and minimum distance d. Since $q^{n-k} \leq q^{mt}$, there exist classical Goppa codes with

$$q^{n-k} \cdot (1 - d(t)q^{-mt/2}) \leq dV(d-1).$$

After taking logarithms and dividing by n, the factors $(1 - d(t)q^{-mt/2})$ and d are absorbed in $o(1)$ as n goes to infinity. □

1.2.3. *Dual BCH codes*

The RS code of length $n = q^m - 1$ with defining set $I = \{1, 2, \ldots, \delta - 1\}$ has parameters $[q^m - 1, q^m - \delta, \delta]$ over the field $\mathbb{F}_{q^m}$. The BCH code over the subfield $\mathbb{F}_q$ with the same length and defining set is the subcode of the RS code with coefficients in $\mathbb{F}_q$. The dual of a BCH code is again cyclic but it is in general not a BCH code. With Delsarte's theorem it can be described as the trace of the dual RS code.

Theorem 1.23. *(Delsarte's Theorem) Let C be a linear code of length n over $\mathbb{F}_{q^m}$ with dual code $C^\perp$. For the subfield $\mathbb{F} = \mathbb{F}_q$ and for the trace map* $\mathrm{Tr}(x) = x + x^q + \cdots + x^{q^{m-1}}$,

$$(C \cap \mathbb{F}^n)^\perp = \mathrm{Tr}(C^\perp).$$

The extended RS code of length $n = q^m$ is the code $C(\leq q^m - \delta, \mathbb{F}_{q^m})$, with dual code $C(\leq \delta - 1, \mathbb{F}_{q^m})$. The weights of nonconstant codewords in the dual of the extended BCH code can be estimated with the Hasse-Weil bound.

Theorem 1.24. *For a polynomial $f \in \mathbb{F}_{q^m}[x]$, let $N(f)$ denote the number of zeros in $(\mathrm{Tr}(f(\alpha)) : \alpha \in \mathbb{F}_{q^m})$. If f is of degree at most $\delta - 1$ and not of the form $a(y^q - y) + b$, for $a \in \mathbb{F}_q, b \in \mathbb{F}_{q^m}$, then*

$$|q \cdot N(f) - q^m| \leq (\delta - 2)(q-1)q^{m/2}.$$

The bound compares the number $q \cdot N(f)$ of solutions (x, y) for the equation $y^q - y = f(x)$ with the number q^m of points on the affine line.

The weight distribution of a dual BCH code describes the number of rational points on curves of the form $y^q - y = f(x)$, for f of bounded degree.

The RS code and the BCH code describe linear relations among the vectors $(\alpha, \alpha^2, \ldots, \alpha^{\delta-1}) \in \mathbb{F}_{q^m}^{\delta-1}$, for $\alpha \in \mathbb{F}_{q^m}^*$. With class field theory the vectors have a natural interpretation as reduced Frobenius automorphisms inside a ray class group of conductor δ. This interpretation will be used in two directions. Weil's theorem on $L-$series for ray class fields gives estimates for the weight distribution of BCH codes. And BCH codes describe the relations between elements in ray class groups that determine the properties of quotient fields of the ray class field with many rational points.

Let $\mathbb{F}$ be a finite field of size $q = p^m$, for a prime p. Let $Q_p(\alpha)$ be a cyclotomic extension of the p-adic numbers with α a primitive n-th roots of unity for $n = p^m - 1$ and let $Z_p[\alpha]$ be the ring of integers in $Q_p(\alpha)$. For a positive integer e, let R_e be the finite ring $Z_p[\alpha]/(p^e)$. So that $|R_e| = q^e$.

For a fixed positive integer δ, let $I = \{1, 2, \ldots, \delta - 1\}$ and let $I^* = \{i \in I : \gcd(p, i) = 1\}$. For $i \in I^*$, let e_i be the unique integer with $ip^{e_i - 1} < \delta \leq ip^{e_i}$. So that $\sum_{i \in I^*} e_i = \delta - 1$.

Theorem 1.25. *(Class field theory) Let $\mathbb{F}$ be a finite field. For every non-negative integer δ, there exists a unique maximal abelian extension $K/\mathbb{F}(x)$, called the* ray class field extension *of conductor δ, for which all characters have conductor at most $\delta(x)_\infty$ and in which $(x)_0$ splits completely. The extension is finite of degree $q^{\delta-1}$ with Galois group*

$$\mathrm{Gal}(K/\mathbb{F}(x)) \simeq (\mathbb{F}[T]/T^\delta)^*/\mathbb{F}^* \simeq \oplus_{i \in I^*} R_{e_i}.$$

For $\alpha \in \mathbb{F}$, let $(K/\mathbb{F}(x), \alpha) \in \mathrm{Gal}(K/\mathbb{F}(x)$ denote the Frobenius automorphism. Under the isomorphisms

$$(K/\mathbb{F}(x), \alpha) \leftrightarrow (1 + \alpha T) \leftrightarrow (\alpha^i : i \in I^*).$$

If H is the subgroup generated by the Frobenius elements for $\alpha \in \mathbb{F}$, then the fixed field $K^H/\mathbb{F}$ defines an extension with group G/H in which ∞ is completely ramified and in which $x = a$ splits completely, for all $a \in \mathbb{F}$.

The set of all relations $(c_\alpha \in \mathbb{Z}/p^e\mathbb{Z} : \alpha \in \mathbb{F}^*)$ with $\sum_\alpha c_\alpha F_\alpha = 0$ defines a cyclic code modulo p^e. The Frobenius element F_α, for $\alpha \in \mathbb{F}^a st$, can be represented by the column vector $h_\alpha = (p^{e-e_i}\alpha^i : i \in I^*) \in R_e^{|I^*|}$. The code

has a generator polynomial $g(x) = g_0(x) + pg_1(x) + \cdots + p^{e-1}g_{e-1}$ with $g_{e-1}|\cdots|g_1|g_0|x^n - 1$, such that, for $i \in I^*$, $g_j(\alpha^i) = 0$ if and only if $j < e_i$ if and only if $ip^j < \delta$. The extended cyclic code $C(p^m, \delta)$ modulo p^e is the set of all relations $(c_\alpha \in \mathbb{Z}/p^e\mathbb{Z} : \alpha \in F)$ with $\sum_\alpha c_\alpha F_\alpha = 0$ and moreover $\sum_\alpha c_\alpha = 0$. The code $C(2^m, 3)$ is defined modulo 4. It is known as the quaternary Preparata code and its dual as the quaternary Kerdock code. The reduction of the code $C(p^m, \delta)$ modulo p is the extended primitive BCH code with designed minimum distance δ.

Theorem 1.26. *(Weil bound) Let χ be a character for $K/\mathbb{F}(x)$ of conductor δ, and let F_α denote the Frobenius element in $G = \mathrm{Gal}(K/\mathbb{F}(x))$ that corresponds to $x = \alpha$. Then*

$$|\sum_{\alpha \in F} \chi(F_\alpha)| \leq (\delta - 2)\sqrt{q}.$$

The theorem applies to characters of characteristic p^e and is more general then Theorem .... The latter follows by writing $q \cdot N(f) - q^m = \sum_{\beta^{q-1}=1} \sum_{\alpha^{q^m-1}=1} \chi(\mathrm{Tr}(\beta f(\alpha))$. In general, for a polynomial $f = \sum_{i \in I^*} p^{e-e_i} \sum_{ip^j < \delta} f_{ip^j} x^{ip^j} \in R_e[x]$, for a trace map $\mathrm{Tr} : R_e \to \mathbb{Z}/p^e\mathbb{Z}$, and for a nontrivial character $\chi : \mathbb{Z}/p^e\mathbb{Z} \to \mathbb{C}$,

$$|\sum_{\alpha^{q^m} - \alpha = 0} \chi(\mathrm{Tr}(f(\alpha))| \leq (\delta - 2)\sqrt{q}.$$

Let $\Delta = \{\, i' \in I^* : \exists i \in I^* \,|\, i < i',$ and $i' \equiv i \cdot q^j \pmod{n} \,\}$. For $i' \in \Delta$ with witness i, a relation $\sum_j c_j \alpha^{ij} = 0 \in R_{e_i}$ implies that $\sum_j \alpha^{i'j} = 0 \in R_{e'_i}$. Therefore, the group G/H has size at least $\prod_{i' \in \Delta} |R_{e'_i}|$.

Theorem 1.27.

(1) For $q = r^2$, let $\delta = r + 2$ and $I = \{1, 2 \ldots, r + 1\}$. Then $|G/H| \geq r$.
(2) For $q_0 = 2^s, q = 2^{2s+1}$, let $\delta = 2q_0 + 2$ and $I = \{1, 2, \ldots, 2q_0 + 1\}$. Then $2q_0 + 1 \in I'$, and $|G/H| \geq |R_1| = q$.
(3) For $q_0 = 3^s, q = 3^{2s+1}$, let $\delta = 3q_0 + 3$ and $I = \{1, 2, \ldots, 3q_0 + 2\}$. Then $3q_0 + 1, 3q_0 + 2 \in I'$ and $|G/H| \geq |R_1 \times R_1| = q^2$.

Proof. (1) The elements α^{r+1} span the subfield $\mathbb{F}_r$ of $R_1 = \mathbb{F}_q$. And $|\mathbb{F}_q/\mathbb{F}_r| = r$. (2) $2q_0 + 1 \equiv 2q_0(q_0 + 1) \pmod{q-1}$. (3) $3q_0 + 1 \equiv 3q_0(q_0 + 1) \pmod{q-1}$ and $3q_0 + 2 \equiv 3q_0(2q_0 + 1) \pmod{q-1}$. □

1.3. Reed-Muller codes

For geometric Goppa codes defined by the evaluation of functions in points $X = \{P_1, \ldots, P_n\}$ that form an ideal-theoretic complete intersection, the main properties can be established without the usual tools for algebraic curves. The dimension is given by the Hilbert function of X (instead of the Riemann-Roch theorem for curves), the dual code is of the same explicit form as the code itself (instead of defined in terms of differentials), and the code and its dual are related by the a-invariant of X (instead of the residue theorem for curves). Codes defined on complete intersections generalize the affine Reed-Muller codes and are part of the larger class of evaluation codes.

For a finite field $\mathbb{F}$, let $A = \mathbb{F}[x_0, x_1, ..., x_m] = \bigoplus_{\nu \geq 0} A(\nu)$ be the graded ring of polynomials in $m+1$ variables with homogeneous components $A(\nu)$, and let $X = \{P_1, ..., P_n\} \subseteq \mathbb{P}^m(\mathbb{F})$ be a set of n distinct points. For a positive integer ν, the $\mathbb{F}$-linear code $C(\nu, X)$ of length n is the image of the homogeneous component $A(\nu)$ after evaluation on X. That is, for a given choice of representatives for $P_1, \ldots, P_n$, the code $C(\nu, X) = \alpha(A(\nu))$, for the $\mathbb{F}$-linear evaluation map

$$\alpha : A \longrightarrow \mathbb{F}^n, \qquad \alpha(f) = (f(P_1), ..., f(P_n)).$$

For the field $\mathbb{F}$ of two elements, the *binary Reed-Muller code* $RM(\nu, m)$ is defined as the code $C(\nu, X)$ with $X = \{(1 : x_1 : \cdots : x_m) : x_i \in \mathbb{F}\}$. Replacing the binary field with an arbitrary finite field yields the class of *a*ffine or *g*eneralized Reed-Muller codes $GRM(\nu, m)$. Evaluation of $A(\nu)$ on a complete set X of representatives for the points of projective m-space over $\mathbb{F}$ yields the class of *projective Reed-Muller codes* $PRM(m, r)$.

Let $I_X = \bigoplus_{\nu \geq 0} I_X(\nu) \subseteq A$ be the vanishing ideal of X. Then the code $C(\nu, X)$ is isomorphic to $S(\nu)/I_X(\nu)$ and its dimension is $H_X(\nu)$, where H_X is the Hilbert function of I_X. If the ideal I_X is a complete intersection, that is if $I_X = (f_1, \ldots, f_m)$ such that f_i is not a zero divisor in $\mathbb{F}[x_0, x_1, \ldots, x_m]/(f_1, \ldots, f_{i-1})$, then the Hilbert function is completely determined by the multi-degree $(\nu_1, \ldots, \nu_m)$ of I_X. Moreover, duality of codes can be described in terms of the a-invariant $(\nu_1 + \cdots + \nu_m) - m - 1$ of X.

Theorem 1.28. *Let X be an ideal-theoretic complete intersection X of multi-degree $(\nu_1, ..., \nu_n)$ with ideal $I_X = (f_1, ..., f_m)$. Let $a_X = (\nu_1 + \cdots +$*

$\nu_m) - (m+1)$ *be the a-invariant of* I_X*. The Hilbert function* $H_X(\nu)$ *of* I_X *is*

$$\binom{m+\nu}{\nu} - \sum_i \binom{m+\nu-\nu_i}{\nu-\nu_i} + \sum_{i<j} \binom{m+\nu-(\nu_i+\nu_j)}{\nu-(\nu_i+\nu_j)} + \cdots + (-1)^m \binom{m+\nu-(\nu_1+\cdots+\nu_m)}{\nu-(\nu_1+\cdots+\nu_m)}$$

For $0 \le \nu \le a_X$, $H_X(\nu) + H_X(a_X - \nu) = n$.

The generalized Reed-Muller code $GRM(\nu, m)$ is defined with $X = \mathbb{P}^m(\mathbb{F})\backslash(x_0 = 0)$. It has vanishing ideal $I = (x_1^q - x_1 x_0^{q-1}, \ldots, x_m^q - x_m x_0^{q-1})$ with multi-degree $(\nu_1, \ldots, \nu_m) = (q, \ldots, q)$ and a-invariant $mq - (m+1)$. The code $GRM(\nu, m)$ has dual code $GRM(qm - m - 1 - \nu, m)$. The set $\mathbb{P}^m(\mathbb{F})$ of all points in projective $m-$space is in general not a complete intersection. Complete intersections in $\mathbb{P}^2(\mathbb{F})$ are described by the Bezout theorem.

Theorem 1.29. *(Bezout) The ideal generated by two polynomials* $f_1, f_2 \in \mathbb{F}[x_0, x_1, x_2]$ *with no common factors is a complete intersection. Over the algebraic closure of* $\mathbb{F}$ *the intersection of the curves* $f_1 = 0$ *and* $f_2 = 0$ *contains* $\deg f_1 \cdot \deg f_2$ *points counted with multiplicities.*

Examples of complete intersections in $\mathbb{P}^2$ are: (1) the projective line with multi-degree $(1, q+1)$ and $|X| = q+1, a_X = q-1$. (2a) the rational points on the Hermitian curve with multi-degree $(r+1, r^2-r+1)$ and $|X| = r^3+1, a_X = q-1$. (2b) the subset of rational points with multi-degree (r, r^2) and $|X| = r^3, a_X = r^2+r-3$. (3) the Klein curve with multi-degree $(4, 6)$ and $|X| = 24, a_X = 7$.

Theorem 1.30. *Let* $C(\nu, X)$ *be defined on the intersection* $X = \{P_1, ..., P_n\} \subseteq \mathbb{P}^2(\mathbb{F})$ *of two curves* $f_1 = 0$ *and* $f_2 = 0$ *with no common component. Let* $\nu_1 = \deg f_1$ *and* $\nu_2 = \deg f_2$*. The Hilbert polynomial* $H_X(\nu)$ *of* I_X *is*

$$\binom{2+\nu}{\nu} - \binom{2+\nu-\nu_1}{\nu-\nu_1} - \binom{2+\nu-\nu_2}{\nu-\nu_2} + \binom{2+\nu-(\nu_1+\nu_2)}{\nu-(\nu_1+\nu_2)}.$$

For $0 \le \nu \le a_X$, $H_X(\nu) + H_X(a_X - \nu) = n$.

Thus, when X is the set of r^3+1 rational points of the Hermitian curve of degree $r+1$ then, for any $0 \le \nu \le r^2 - 1$, the codes $C(\nu, X)$ and

$C(q-1-\nu, X)$ are dual to each other. This can also be seen as follows.

For codes $C(< k, \mathbb{F})$ on the affine line, duality of $C(< k, \mathbb{F})$ and $C(< q-k, \mathbb{F})$ amounts to the property $\sum_{x\in\mathbb{F}} x^i = 0$, for $i = 0, 1, \ldots, q-2$. If we extend the summation to points on the projective line, we have $\sum_{(x:y)} x^i y^{q-1-i} = 0$, for $i = 0, 1, \ldots, q-1$. The cases $i > 0$ reduce to the affine line $x = 1$ and the cases $i < q-1$ to the affine line $y = 1$. Note that the total degree $q-1$ of $x^i y^{q-1-i}$ makes the summation independent of a choice of representative for the projective points. That the a-invariant for the projective line and the Hermitian curve is $q-1$ in both cases corresponds to the fact that the rational points of the Hermitian curve form a codeword in the code spanned by lines [6], [50].

1.4. Geometric Goppa codes

Geometric Goppa codes use algebraic curves for their construction. Similar to codes on the affine line (Section 1.1), they can be defined in two different ways, by evaluating functions or by computing residues of differentials. In combination with well known theorems for algebraic curves, the definitions immediately reveal the following important properties of geometric Goppa codes:

- An explicit geometric description of both a code and its dual.
- Good lower bounds for the dimension, the minimum distance, and the dual minimum distance of a code.
- Expressions for code parameters in terms of invariants of algebraic curves.
- A multiplicative structure on codes.

Following are some important results for geometric Goppa codes that crucially depend on these properties:

- Constructions of polynomial complexity for asymptotically good codes.
- Efficient algebraic decoding.
- Applications to secret sharing and efficient multi-party computation.

In this section, we first give the definitions and the main properties of geometric Goppa codes (Sections 1.4.1, 1.4.2), followed by a summary of curves that have been used for their construction (Section 1.4.3). One-point codes and two-point codes are discussed in Sections 1.4.4 and 1.4.5. Finally, we present results on error correction (Section 1.4.6), secret sharing (Section 1.4.7), and weight distributions (Section 1.4.8).

1.4.1. *Curves and linear codes*

An *algebraic curve* $\mathcal{X}/\mathbb{F}$ is defined as an algebraic variety (i.e. an irreducible algebraic set) of dimension one over the field $\mathbb{F}$. The field of rational functions is denoted by $\mathbb{F}(\mathcal{X})$, the module of rational differentials by $\Omega(\mathcal{X})$. Among all curves with function field $\mathbb{F}(\mathcal{X})$ there is up to isomorphism a unique nonsingular projective curve. We define the codes in terms of the function field $\mathbb{F}(\mathcal{X})$ of $\mathcal{X}$. The geometric properties that we establish for codes hold for codes that are defined with the unique nonsingular projective model of $\mathcal{X}$. Function fields of algebraic curves over a finite field can be characterized as finite separable extensions $K/\mathbb{F}(x)$.

Points on a curve $\mathcal{X}$ are identified with places of the function field, rational points with places of degree one. Let t denote a generator of the maximal ideal of a place. For a rational function f, define the divisor $(f) = \sum \nu_t(f)P$, where P runs over all places and ν_t denotes the discrete valuation at P. For a divisor E, define

$$L(E) \;=\; \{\, f \in \mathbb{F}(\mathcal{X})^* \,:\, (f) + E \geq 0 \,\} \cup \{\, 0 \,\}$$

as the linear space of rational functions with pole divisor bounded by E.

Definition 1.31. Let $D = P_1 + P_2 + \cdots + P_n$, for distinct rational points $P_1, P_2, \ldots, P_n$, and let G be a divisor with support disjoint from D. The code $C_L(D, G)$ is the image of the linear map

$$\alpha_L \;:\; L(G) \quad\longrightarrow\quad \mathbb{F}^n, \qquad f \mapsto (\, f(P_1), \ldots, f(P_n)\,).$$

The map establishes an isomorphism $L(G)/L(G-D) \simeq C_L(D,G)$.

In general, $\dim L(G) \leq \deg G + 1$. To estimate the dimension of a code we need a lower bound for $L(G)$.

Theorem 1.32. *(Riemann) There exists a minimal constant $g \geq 0$ depending only on $\mathcal{X}$, such that $\dim L(G) \geq \deg G + 1 - g$. Moreover, for every divisor G of degree $\deg G > 2g - 2$, $\dim L(G) = \deg G + 1 - g$. The parameter g is called the genus of the curve $\mathcal{X}$.*

Theorem 1.33. *(code parameters) For $2g - 2 < \deg G < n$, the code $C_L(G, D)$ has dimension $k = \deg G + 1 - g$ and minimum distance $d \geq n - \deg G$. The dual code $C_L(G, D)$ has dimension $k^\perp = n - (\deg G + 1 - g)$ and minimum distance $d^\perp \geq \deg G - (2g - 2)$. In particular,*

$$n + 1 - g \leq k + d, k^\perp + d^\perp \leq n + 1.$$

Proof. The only part remaining is to show that $d^\perp \geq \deg G - (2g-2)$. For any $\tau < d^\perp$ positions $Q_1, \ldots, Q_\tau$, $\dim L(G) - \dim L(G - Q_1 \cdots - Q_\tau) = \tau$ and the encoding map α_L is surjective on the τ positions. □

A divisor is called principal if it is the divisor of a function. The relation $E_1 \sim E_2$ if and only if $E_1 - E_2$ is principal defines an equivalence relation on divisors.

Theorem 1.34. *(Approximation theorem) For a divisor E and a finite set of places S, there exists a divisor E' that is linearly equivalent to E and that has support outside S.*

In many cases it is attractive to define codes where D and G have one or more rational points in common. For the construction of such codes one may replace G with an equivalent divisor using the approximation theorem. However, the following theorem gives an important geometric property of algebraic curves that makes the construction of such codes straightforward without replacing the divisor G.

Theorem 1.35. *For a nonsingular curve $\mathcal{X}$ and for rational functions $(f_0, f_1, \ldots, f_m)$, the rational map $(f_0 : f_1 : \cdots : f_m) : \mathcal{X} \longrightarrow \mathbb{P}^m$ is a morphism (is defined everywhere).*

In case the divisors D and G have a rational point P in common, the evaluation map α_L in the definition of $C_L(G, D)$ is modified at the coordinate $\alpha_{L,P}$. For a given local parameter t at P, and for $i = \mathrm{ord}_P(G)$,

$$\alpha_{L,P} \ : \ L(G) \ \longrightarrow \ \mathbb{F}, \quad f \mapsto (t^i f)(P).$$

The bounds in Theorem 1.33 are based on properties of the geometric embedding of points in projective space and remain valid for the modified construction.

The Klein curve is defined by the equation $X^3Y + Y^3Z + Z^3X = 0$. Define a divisor $\Delta = (0:0:1) + (0:1:0) + (1:0:0)$. A monomial $X^aY^bZ^c$ intersects the curve with multiplicities

$$(X^3Y + Y^3Z + Z^3X = 0) \cap (X^aY^bZ^c = 0) = \\ (3a+b)(0:0:1) + (3b+c)(1:0:0) + (3c+a)(0:1:0).$$

We find a basis $\langle X^2Y/XYZ, Y^2Z/XYZ, Z^2X/XYZ, XYZ/XYZ \rangle$ for $L(2\Delta)$. Over the field of eight elements, the curve has 24 rational points. The given basis does not evaluate in the three points of Δ. An option

is to define the code on the remaining 21 points or to replace 2Δ with an equivalent divisor that has support in an extension field of $\mathbb{F}_8$. The straightforward solution suggested by the theorem is to embed the points as images of the morphism $(X^2Y : Y^2Z : Z^2X : XYZ)$. The morphism sends $(0:0:1) \mapsto (0:1:0:0)$, $(1:0:0) \mapsto (0:0:1:0)$, and $(0:1:0) \mapsto (1:0:0:0)$. There exist no 6 distinct rational points with $Q_1 + \cdots + Q_6 \sim 2\Delta$ and the code $C_L(2\Delta, D)$ is of type $[24, 4, 19]$. The distance is an arithmetic peculiarity of the configuration of flexpoints on the Klein curve that can be explained in terms of the large automorphism group of the curve but not with any of the theorems in this chapter.

Not all properties of codes are preserved by the modified construction: For divisors $G_1 \leq G_2$ that have supports disjoint from D, the code $C_L(D, G_1)$ is a subcode of the code $C_L(D, G_2)$. When $G_2 - G_1$ is not disjoint from D this is in general no longer true.

1.4.2. *Duality and differentials*

For a differential ω, define the divisor $(\omega) = \sum \nu_t(\omega)P$, where P runs over all places and $\nu_t(fdt) = \nu_t(f)$. The rational differentials $\Omega(\mathcal{X})$ form a free $\mathbb{F}(\mathcal{X})$ module of rank one. The divisor class of a differential is called the canonical divisor class, K denotes a divisor representing the class. For a divisor E, define the linear space of rational differentials

$$\Omega(E) \;=\; \{\, \omega \in \Omega(\mathcal{X})^* \,:\, (\omega) \geq E \,\} \cup \{\, 0 \,\}.$$

Definition 1.36. Let $D = P_1 + P_2 + \cdots + P_n$, for distinct rational points $P_1, P_2, \ldots, P_n$, and let G be a divisor with support disjoint from D. The code $C_\Omega(D, G)$ is the image of the linear map

$$\alpha_\Omega \;:\; \Omega(G - D) \;\longrightarrow\; \mathbb{F}^n, \qquad \omega \mapsto (\,\mathrm{res}_{P_1}(\omega), \ldots, \mathrm{res}_{P_n}(\omega)\,).$$

The map establishes an isomorphism $\Omega(G-D)/\Omega(G) \simeq C_\Omega(D, G)$.

In case the divisors D and G have a rational point P in common, the evaluation map α_Ω is modified at the coordinate $\alpha_{\Omega,P}$. For a given local parameter t at P, and for $i = \mathrm{ord}_P(G)$,

$$\alpha_{\Omega,P} \;:\; \Omega(G - D) \;\longrightarrow\; \mathbb{F}, \quad \omega \mapsto \mathrm{res}_P(t^{-i}\omega).$$

Theorem 1.37. *(Residue theorem) The summation over all places of the residues of a differential is well-defined and equal to zero.*

Each differential ω induces a natural isomorphism

$$L((\omega) - E) \quad \xrightarrow{\sim} \quad \Omega(E), \qquad f \mapsto f\omega.$$

If η is a differential with a simple pole at P and residue $\mathrm{res}_P(\eta) = 1$, and if f is a function with no pole at P, then $\mathrm{res}_P(f\eta) = f(P)$.

Lemma 1.38. *If η is a differential with simple poles at $P_1, P_2, \ldots, P_n$ and residues equal to 1 at those points then*

$$C_\Omega(G, D) = C_L((\eta) + D - G, D).$$

Proof. For $f \in L((\eta) + D - G)$, $f(P) = \mathrm{res}_P(f\eta)$, where the differential $\omega = f\eta$ has divisor $(\omega) = (f) + (\eta) \geq G - D$. □

Lemma 1.39. *Let f be a nonzero rational function. The differential df/f has at most simple poles and the residue at P is $\mathrm{res}_P(df/f) = \mathrm{ord}_P(f)$.*

Theorem 1.40. *(Riemann-Roch) The dimensions of $L(E)$ and $L(K-E) \simeq \Omega(E)$ are related by*

$$\dim L(E) - \dim L(K - E) = \deg(E) + 1 - g.$$

Together, the Residue theorem and the Riemann-Roch theorem imply that $C_\Omega(G, D)$ is the dual code of $C_L(D, G)$.

Theorem 1.41. *The codes $C_L(D, G)$ and $C_\Omega(G, D)$ are dual codes.*

As the dual of $C_L(D, G)$, the code $C_\Omega(D, G)$ has minimum distance at least $\deg G - (2g - 2)$ (Theorem 1.33).

Theorem 1.42. *(Symmetric floor bound) Let $G = A + B + Z$, for $Z \geq 0$ such that $L(A + Z) = L(A)$ and $L(B + Z) = L(B)$. For D with $D \cap Z = 0$, a nonzero word in $C_\Omega(D, G)$ has weight at least $\deg G - (2g - 2) + \deg Z$.*

Proof. Suppose that $c \in C_\Omega(D, G)$ is nonzero in the positions $Q = Q_1 + \cdots + Q_d$, so that there exists $E \geq 0$ with $K + Q \sim A + B + Z + E$. With the Riemann-Roch theorem,

$$\dim L(A + E) - \dim L(B + Z - Q) = \deg(A + E) + 1 - g,$$
$$\dim L(A + Z) - \dim L(B + E - Q) = \deg(A + Z) + 1 - g.$$

It follows that

$$\deg E - \deg Z = l(A + E) - l(A) + l(B + E - Q) - l(B - Q) \geq 0.$$

Finally, $\deg E \geq \deg Z$ gives $d \geq \deg G - (2g - 2) + \deg Z$. □

The divisor K satisfies: $\deg K = 2g-2$ and $l(K) = g$. The genus g of a nonsingular plane curve of degree m satisfies $g = (m-1)(m-2)/2$. For a plane curve, let the divisor L denote the intersection divisor of a line with the curve, then $K = (m-2)L$ represents the canonical class.

For divisors $G + G' \sim D$, the codes $C_L(G, D)$ and $C_L(G', D)$ are in general not dual codes, unless $g = 1$. The two codes have the same number of words of designed distance. Namely G is equivalent to a sum of rational points Q if and only if G' is equivalent to the sum of rational points Q', where $Q + Q' = D$. If one code has distance greater than the designed distance then the other code as well. With the Klein curve over $\mathbb{F}_8$, and for $G = 2\Delta$, we found a code of type $[24, 4, 19]$. In this case, $D \sim 8\Delta$ and the code with $G' = 6\Delta$ is of type $[24, 16, 7]$. This is the best known three-error-correcting code of length 24 over $\mathbb{F}_8$. Its weight distribution is given in Table 1.3.

Theorem 1.43. *(Clifford's theorem) For a divisor E such that both $L(E)$ and $\Omega(E)$ are nontrivial,*

$$\dim L(E) \leq \frac{\deg(E)}{2} + 1.$$

1.4.3. *Families of curves*

The first step towards good geometric Goppa codes over a field $\mathbb{F}_q$ is the search for curves $\mathcal{X}/\mathbb{F}_q$ that have many rational points for a given genus. For a given curve $\mathcal{X}/\mathbb{F}_q$ of genus g with N rational points, we can construct $\mathbb{F}_q$-linear codes of length N of any dimension $0 \leq k \leq N$ such that $k + d \geq N + 1 - g$.
The class of Deligne-Lusztig varieties was defined for the purpose of studying representations of algebraic groups. The class contains three families of irreducible curves (Table 1.1). Curves in each family have the maximal number of rational points for their genus and they have large automorphism groups. In each case, the automorphism group is of order $N(N-1)(q-1)$ and acts $2-$transitively on the set of rational points. The curve of unitary type was already known as the Hermitian curve. Another much studied curve is the Klein curve, or the modular curve $X(7)$. From its definition as a modular curve it follows that it is a nonsingular quartic with automorphism group the simple group $PSL(2, 7)$ of order 168. Klein found the model $X^3Y + Y^3Z + Z^3X = 0$ for the unique curve with these properties. Over the field of eight elements it has the maximal number of 24 rational

points for a curve of genus 3.

Table 1.1. Deligne-Lusztig curves $X/\mathbb{F}_q$.

Type	Unitary	Suzuki	Ree
X	$y^r + y = x^{r+1}$.	$y^q + y = x^{q_0}(x^q + x)$.	$y^q - y = x^{q_0}(x^q - x)$, $z^q - z = x^{q_0}(y^q - y)$.
q	$q = r^2$	$q = 2q_0^2 \geq 8$.	$q = 3q_0^2 \geq 27$.
g	$r(r-1)/2$	$q_0(q-1)$	$\frac{3}{2}q_0(q-1)(q+q_0+1)$
N	r^3+1	q^2+1	q^3+1
conductor	$r+2$	$2q_0+2$	$3q_0+3$

Serre initiated the construction of curves with many points using class field theory. This has been a very successful method to show that certain pairs (N, g) occur as the number of rational points and the genus of a curve. The actual construction of the curves is in general not straightforward. Lauter uses class field theory to show the existence of curves with the parameters of the Deligne-Lusztig curves. In those cases there is a connection between class field theory and BCH codes (Theorem 1.27).

For asymptotic results we need families of curves of increasing genus such that $\liminf N_i/g_i > 0$ as $g_i \to \infty$. For any given family $\limsup N_i/g_i \leq \sqrt{q} - 1$ (Drinfeld-Vladuts bound). Asymptotic results were first obtained by Tsfasman, Vladuts and Zink. They use families of modular curves over $\mathbb{F}_q$ to attain the best possible $\liminf N_i/g_i = \ell - 1$, for $q = \ell^2$. In subsequent papers polynomial constructions were given for the codes from these curves. Garcia and Stichtenoth presented several constructions for optimal towers that have a short and explicit recursive definition (Table 1.2).

Table 1.2. Recursively defined towers of function fields ($F_1 = \mathbb{F}_q(x_1)$).

(A)	$F_{n+1} = F_n(z_{n+1})$	$z_{n+1}^\ell + z_{n+1} = x_n^{\ell+1}$, $x_n = z_n/x_{n-1}$	$q = \ell^2$
(B)	$F_{n+1} = F_n(x_{n+1})$	$x_{i+1}^\ell + x_{i+1} = x_i^\ell/(x_i^{\ell-1}+1)$	$q = \ell^2$
(C)	$F_{n+1} = F_n(x_{n+1})$	$x_{i+1}^m + (x_i+1)^m = 1$	$m\|(q-1)/(p-1)$
(D)	$F_{n+1} = F_n(x_{n+1})$	$x_{i+1}^{\ell-1} + (x_i+1)^{\ell-1} = 1$	$q = \ell^2$

The towers (A) and (B) are wildly ramified while the towers (C) and (D) are tamely ramified. An efficient construction of codes in the tower (A) is given in [70]. In [78], codes are constructed with the field F_3 in

the tower (B). The towers (A) and (B) correspond to Drinfeld modular curves [27] and the towers (C) and (D) to classical modular curves [25]. Examples are the modular towers $X_0(3^n)$ in char $= 2$ and $X_0(2^n)$ in char $= 3$. Klein gives solutions for the modular equations $J_2(j(\tau), j(2\tau)) = 0$ and $J_3(j(\tau), j(3\tau)) = 0$ in terms of resolvents,

$$J_2(\psi_2(\eta), \psi_2(\eta_0)) = 0, \quad \text{for } \psi_2(\eta) = 64\frac{(\eta+3)^3}{(\eta-1)^2}, \ (\eta-1)(\eta_0-1) = 1,$$
$$J_3(\psi_3(\eta), \psi_3(\eta_0)) = 0, \quad \text{for } \psi_3(\eta) = 27\frac{\eta(\eta+8)^3}{(eta-1)^3}, \ (\eta-1)(\eta_0-1) = 1,$$

such that $\psi_2(z'^2) = \psi_2(z^2)$ for all symmetries of the triangle $\{1, -1, \infty\}$, and $\psi_3(z'^3) = \psi_3(z^3)$ for all symmetries of the tetrahedron $\{1, \omega, \omega^2, \infty\}$. In particular,

$$\psi_2(z'^2) = \psi_2(z^2), \quad \text{for } z' = \frac{z+3}{z-1} \ (1 \leftrightarrow \infty, -1 \leftrightarrow -1).$$
$$\psi_3(z'^3) = \psi_3(z^3), \quad \text{for } z' = \frac{z+2}{z-1} \ (1 \leftrightarrow \infty, \omega \leftrightarrow \omega^2).$$

The modular equation is symmetric in its two arguments and so is the equation $(\eta-1)(\eta_0-1) = 1$ in the variables η, η_0. In the $z-$plane, a recursive formula for the modular tower can be achieved by rotating z before adjoining z_0 (as described in Cohn, Iteration and the icosahedron).

$$z' = \frac{z+3}{z-1}, \quad (z'^2 - 1)(z_0^2 - 1) = 1.$$
$$z' = \frac{z+2}{z-1}, \quad (z'^3 - 1)(z_0^3 - 1) = 1.$$

In the variables $x = -1/z, y = -1/z_0$, the recursive formulas are

$$y^2 + \left(\frac{1+x}{1-3x}\right)^2 = 1.$$
$$y^3 + \left(\frac{1+x}{1-2x}\right)^3 = 1.$$

In char $= 3$ the first tower $X_0(2^n)$ is of type (D), and in char $= 2$ the second tower $X_0(3^n)$ is of type (C).

The equation $F_{n+1} = F_n(x_{n+1}), x_{i+1}^2 + x_{i+1} = x_i + 1 + 1/x_i$ defines an asymptotically good tower over $\mathbb{F}_8$. It has a generalization to arbitrary cubic fields.

1.4.4. *One-point codes*

For a curve with many rational points for a given genus, any choice of divisor G will give a good code. In many cases, once the degree of G has been fixed, a convenient choice is a divisor $G = mP_\infty$ with support at a single point P_∞. The codes $C_L(mP_\infty, D)$ are called one-point codes. It follows from Lemma 1.38 that the dual code of a one-point code is again a one-point code if there exists a differential η with divisor $(2g-2+n)P_\infty - D$ that has residues equal to 1 at the points $P_1, P_2, \ldots, P_n$. In that case

$$C_\Omega(mP_\infty, D) = C_L((2g-2+n-m)P_\infty, D).$$

For the projective line, for the Hermitian curves, and for the Suzuki curves, the dual of a one-point code is again a one-point code. For each of these curves, the divisor D can be chosen to be the set of all rational points minus the point P_∞. For this choice of D, there exists an algebraic function $x \in K$ such that $n = [K : \mathbb{F}(x)] \cdot q$ and $\eta = df/f$ for $f = x^q - x$.. The one-point codes can be extended by including the point P_∞ in D. The modified construction for one-point codes is straightforward and in some cases the longer codes that are obtained in this way have larger automorphism groups.

For the Klein curve $X^3Y + Y^3Z + Z^3X = 0$, the dual of a one-point code is in general not a one-point code. The curve has three points O_0, O_1, O_2 with $XYZ = 0$. Let $K = L$ be the canonical divisor class and let $\Delta = O_0 + O_1 + O_2$. The divisor classes K and 2Δ are invariant under the full automorphism group $PSL(2,7)$. The spaces $L(m(L-\Delta))$ are spanned by monomials. For the Klein curve over $\mathbb{F}_8$, the codes $C_L(m(L-\Delta), D)$ are better than the one-point codes on the same curve, are closed under duality, and have interesting geometric properties.

The space $L(mP_\infty)$ is a subset of the affine ring $R = \cup_{m \geq 0} L(mP_\infty)$ of rational functions with poles only at P_∞. The ring is a finitely generated $\mathbb{F}$-algebra. If $\phi_1, \ldots, \phi_r$ are generators and $m_1, \ldots, m_r$ are their pole orders then the set of all possible pole orders is the semigroup $\Lambda = \mathbb{Z}m_1 + \cdots + \mathbb{Z}m_r \subset \mathbb{Z}$ The complement $\mathbb{Z} \backslash \Lambda$ is finite of size g. Especially when r is small, the ring R can be used for efficient encoding (if the code is a one-point code) or efficient decoding (using a key equation in standard form if the dual code is a one-point code, or a key equation in Welch-Berlekamp form if the code is a one-point code).

The Hermitian curve over a field $\mathbb{F}$ of size q^2 is the curve $X/\mathbb{F} : y^q + y = x^{q+1}$. For every $x \in \mathbb{F}$, there are q solutions for $y \in \mathbb{F}$. Together with the point at infinity $P_\infty = (0:1:0)$ the curve has q^3+1 rational points. Codes from Hermitian curves are among the most studied geometric Goppa codes. The semigroup Λ of nongaps is generated by $\{q, q+1\}$.

Lemma 1.44. *For an integer a, write $a = a_0(q+1) - a_1$ with $0 \le a_1 \le q$. Then a is a nongap if and only if $a_1 \le a_0$.*

For the Suzuki curve, the semigroup of nongaps is generated by $\{q, q + q_0, q + 2q_0, q + 2q_0 + 1\}$, and for the Klein curve by $\{3, 5, 7\}$. For Hermitian one-point codes, the actual minimum distance is completely determined by properties of the nongaps. We give a first proof based on the following lemma.

Lemma 1.45. *For every point $R = (x_0, y_0) \neq P_\infty$, there exists an effective divisor E_R of degree q such that $(y - y_0) = R + E_R - (q+1)P_\infty$ and $E_R \cap P_\infty = 0$.*

Theorem 1.46. *Let $G = K + (a_0(q+1) - a_1)P_\infty$, with $K = (q-2)(q+1)P_\infty$ a canonical divisor. Then*

$$d(C_\Omega(G, D)) = \begin{cases} a_0(q+1) - a_1 & \text{if } a_1 \le a_0 \\ a_0(q+1) - a_0 & \text{if } a_1 > a_0 \end{cases}$$

Proof. Let $Q = Q_1 + \cdots + Q_d$ and assume that there exists a nonzero differential $\omega \in \Omega(G - Q)$ with $(\omega) = G - Q + E, E \ge 0$. Then $Q \sim (a_0(q+1) - a_1)P_\infty + E$. For each point $R \in E$ apply the lemma to find $Q + \sum E_R \sim ((a_0 + \deg E)(q+1) - a_1)P_\infty$. With the first lemma $a_1 \le a_0 + \deg E$. □

In this case, it appears natural to formulate the bound for the code $C_\Omega(G, D)$. The result and the proof depends on G but not on D. Below we repeat the proof for a code $C_L(G^*, D)$ which essentially leads us back to the case of a code $C_\Omega(G, D)$ after making the assumption $D \sim nP_\infty$.

Proof. (second proof) We prove the minimum distance bound for the code $C_L(m^*P_\infty, D)$, where $m^* = n + 2g - 2 - m = n - a_0(q+1) + a_1$. Assume that there exists a nonzero $f \in L(m^*P_\infty - Q')$, with $(f) = Q' + E - m^*P_\infty$, $E \ge 0$. Then the complement $Q = D - Q' \sim (n - m^*)P_\infty + E$. As in the first proof, $Q + \sum E_R \sim ((a_0 + \deg E)(q+1) - a_1)P_\infty$ and $a_1 \le a_0 + \deg E$. □

A special case of the theorem can be obtained with Theorem 1.42. Let $A = B = (a_0(q+1)-q)P_\infty, Z = (q-a_0-1)P_\infty, 1 \le a_0 \le q-1$. Then the code $C_\Omega(G,D)$ with $G = (2a_0-1)(q+1)-a_0$ has $d = d^* + (q-a_0-1)$. This corresponds to the case $G = ((q-2)(q+1) + (2a_0+1-q)(q+1) - a_0)P_\infty$ in the theorem above, with $a_0 \ge 2a_0 + 1 - q$ if and only if $a_0 \le q-1$.

For Hermitian one-point codes of length q^3, the q^3 finite rational points form a complete intersection with coordinate ring

$$\begin{aligned} & \mathbb{F}[x,y]/(y^q + y - x^{q+1}, x^{q^2} - x) \\ = \ & \langle x^i y^j : 0 \le i \le q^2 - 1, 0 \le j \le q-1 \rangle. \end{aligned}$$

For the q^3 monomials $x^i y^j$ in the vector space basis,

$$\sum_{P \in D} x^i y^j = \begin{cases} 1 & \text{if } x^i y^j = x^{q^2-1} y^{q-1} \\ 0 & \text{otherwise} \end{cases}$$

Duality can be stated as

$$\sum_{P \in D} x^i y^j = 0, \quad \text{for } i + j \le q-1, \ (i,j) \ne (0, q-1).$$

The monomials with $i + j \le q-1$, $(i,j) \ne (0,q-1)$ generate but do not form a basis for the coordinate ring. The set of all $q^3 + 1$ rational points is also a complete intersection, with coordinate ring

$$\begin{aligned} & \mathbb{F}[x,y]/(y^q + y - x^{q+1}, x(y^{q^2} - y)/(y^q + y)) \\ = \ & \langle x^i y^j : 0 \le i \le q, 0 \le j \le q^2 - q \rangle. \end{aligned}$$

For the $q^3 + 1$ monomials $x^i y^j$ in the vector space basis,

$$\sum_{P \in D \cup P_\infty} x^i y^j = \begin{cases} 1 & \text{if } x^i y^j = x^q y^{q^2-q} \\ 0 & \text{otherwise.} \end{cases}$$

Duality can be stated as

$$\sum_{P \in D} x^i y^j = 0, \quad \text{for } i + j \le q-1.$$

This is the same duality as that for a summation over all points of the projective line, and indeed follows from that duality since the points on the Hermitian curve form a codeword in the code of the point-line graph of the projective plane [6], [50]. The monomials with $i + j \le q-1$ generate but

do not form a basis for the coordinate ring. The tables give the monomial basis for each of the two coordinate rings when $q = 3$.

$-$	1	y	y^2
1	0	4	8
x	3	7	11
x^2	6	10	14
x^3	9	13	17
x^4	12	16	20
x^5	15	19	23
x^6	18	22	26
x^7	21	25	29
x^8	24	28	32

$-$	1	y	y^2	y^3	y^4	y^5	y^6
1	0	4	8	12	16	20	24
x	3	7	11	15	19	23	27
x^2	6	10	14	18	22	26	30
x^3	9	13	17	21	25	29	33

The coordinate ring $\mathbb{F}[x, y]$ of the Hermitian curve itself is often considered as an $\mathbb{F}[x]$ algebra with free basis $\{1, y, \ldots, y^{q-1}\}$. For one-point codes of full length q^3+1, the ring $\mathbb{F}[x, y]$ may be considered as an $\mathbb{F}[y]$ algebra with free basis $\{1, x, \ldots, x^q\}$.

1.4.5. *Two-point codes*

Let $\mathcal{X}$ be a curve and let P_∞, P_0 be distinct rational points. A two-point code is defined with a divisor $G = aP_\infty + bP_0$. For the rational function field $\mathbb{F}(x)$ let P_∞ be the simple pole of x and P_0 the simple zero of x. Then, for $a + b \geq 0$, $L(aP_\infty + bP_0) = \langle x^{-b}, \ldots, x^a \rangle$. Thus, two-point codes are to one-point codes what BCH codes are to narrow sense BCH codes. The larger class of codes contains some codes that are better without giving up the advantages of efficient encoding and decoding. The subsemigroup $H(P_\infty, P_0)$ of $\mathbb{N} \times \mathbb{N}$ was introduced in 1985 by Joe Harris. It consists of all ordered pairs (a, b) such that there exists a rational function on $\mathcal{X}$ with polar divisor $aP_\infty + bP_0$. It generalizes the subsemigroup $H(P_\infty)$ of $\mathbb{N}$. The complement $G(P_\infty) = \mathbb{N} \backslash H(P_\infty)$ of gaps at P_∞ is of size g. The size of the complement $G(P_\infty, P_0) = \mathbb{N} \times \mathbb{N} \backslash H(P_\infty, P_0)$ does not depend on the genus alone. For the questions that we are interested in we extend $H(P_\infty, P_0)$ to the subsemigroup of $\mathbb{Z} \times \mathbb{Z}$ of nongaps at P_∞ and P_0. Thus

$$H(P_\infty, P_0) = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} : \\ \exists f \in L(aP_\infty + bP_0) | \operatorname{ord}_{P_\infty} = -a, \operatorname{ord}_{P_0} = -b\}.$$

The semigroup is contained in the halfplane $a + b \geq 0$, but not in the first quadrant. The complement $G(P_\infty, P_0) = \mathbb{Z} \times \mathbb{Z} \backslash H(P_\infty, P_0)$ is contained

in the halfplane $a + b \leq 2g - 1$. We extend the definition of the set Γ by Kim [51] to the subsemigroup of the full integer plane.

$$\Gamma(P_\infty, P_0) = \{(a,b) \in H(P_\infty, P_0) : \text{ for given } a,\ b \text{ is minimal with } (a,b) \in H(P_\infty, P_0)\}.$$

Proposition 1.47. *The set Γ is defined as the graph of a function $\sigma : \mathbb{Z} \longrightarrow \mathbb{Z}$. The function σ is a permutation of the integers. If f is a nonzero rational function with $(f) = m(P_0 - P_\infty)$ then σ is determined by its images on a set of representatives for the integers modulo m.*

Proof. The pair (a,b) is in $\Gamma(P_\infty, P_0)$ if and only if $L(aP_\infty + bP_0) \neq L((a-1)P_\infty + bP_0)$ and $L(aP_\infty + (b-1)P_0) = L((a-1)P_\infty + (b-1)P_0)$ if and only if $L(aP_\infty + bP_0) \neq L(aP_\infty + (b-1)P_0)$ and $L((a-1)P_\infty + bP_0) = L((a-1)P_\infty + (b-1)P_0)$. That is, for given b, a is minimal with $(a,b) \in H(P_\infty, P_0)$. Clearly, if $(a,b) \in \Gamma$ then $(a+m, b-m) \in \Gamma$. □

We call the ordered pair $(a,b) \in \Gamma$ a discrepancy pair. A pair of integers (a,b) is a nongap if and only if the discrepancies (a,b') and (a',b) satisfy $b' \leq b$ and $a' \leq a$.

Lemma 1.48. *For two rational points P_∞, P_0 on the Hermitian curve $y^q + y = x^{q+1}$ over $\mathbb{F}_{q^2}$, there exists f with $(f) = (q+1)(P_0 - P_\infty)$. The set of discrepancies*

$$\Gamma(P_\infty, P_0) = \{(a_0(q+1) - a_1, -a_0(q+1) + a_1 q) : a_0 \in \mathbb{Z}, 0 \leq a_1 \leq q\}.$$

Proof. It suffice to consider $a_0 = 0, a_1 = 0, 1, \ldots, q$. The minimal choices correspond to functions $(y/x)^{a_1}$ with order $-a_1$ at P_∞ and qa_1 at $P_0 = (0,0)$. □

Lemma 1.49. *Write $(a,b) \in \mathbb{Z} \times \mathbb{Z}$ as $(a_0(q+1) - a_1, b_0(q+1) - b_1)$ with $a_0, b_0 \in \mathbb{Z}$ and $0 \leq a_1, b_1, \leq q$. Then (a,b) is a nongap if and only if $a_1, b_1 \leq a_0 + b_0$.*

The following result was first obtained, in a different formulation and with a different proof, by Homma and Kim. We state the result as formulated by Beelen [3] and Park [65].

Theorem 1.50. *Let $G = K + aP_\infty + bP_0 \geq K + P_\infty + P_0$, where K is the canonical divisor, and write*

$$a = a_0(q+1) - a_1, \quad 0 \leq a_1 \leq q,$$
$$b = b_0(q+1) - b_1, \quad 0 \leq b_1 \leq q.$$

Let $d = d(C_\Omega(G,D))$ *and let* $d^* = \deg(G) - (2g-2) = a+b$.

(1)	$a_1, b_1 \le a_0 + b_0$,	$d = d^*$.
(2a)	$b_1 \le a_0 + b_0 \le a_1$,	$d = d^* + a_1 - (a_0 + b_0)$.
(2b)	$a_1 \le a_0 + b_0 \le b_1$,	$d = d^* + b_1 - (a_0 + b_0)$.
(3a)	$a_0 + b_0 \le a_1 \le b_1$ *and* $a_1 < q$,	$d = d^* + a_1 + b_1 - 2(a_0 + b_0)$.
(3b)	$a_0 + b_0 \le b_1 \le a_1$ *and* $b_1 < q$	$d = d^* + a_1 + b_1 - 2(a_0 + b_0)$.
(4)	$a_0 + b_0 \le a_1 = b_1 = q$	$d = d^* + q - (a_0 + b_0)$.

Proof. Let $H = (q+1)P_\infty$. For $a_0 + b_0 \le a_0 + b_0 + r < a_1$,

$$G + P_\infty = ((a_0 + b_0 + r)H - a_1 P_\infty) + ((q - 1 - r)H - qP_\infty - b_1 P_0)$$

is the sum of two gaps. For $a_0 + b_0 \le a_0 + b_0 + s < b_1$,

$$G + P_\infty = ((q - 1 - s)H - qP_\infty) + ((a_0 + b_0 + s)H - a_1 P_\infty - b_1 P_0)$$

is the sum of two gaps. Applying the coset bound Theorem 1.6 repeatedly, as in Theorem 1.7, gives a lower bound for the minimum distance that adds the number of pairs of gaps to the designed distance. The first group of pairs adds $a_1 - (a_0 + b_0)$ to cases (2a,3a,4). The second group adds $b_1 - (a_0 + b_0)$ to the case (3a). The cases (2b) and (3b) follow by symmetry. □

1.4.6. *Error correction*

For algebraic decoding it is important to have triples of codes A, B, C with $\sum a_i b_i c_i = 0$ for all $a \in A, b \in B, c \in C$. For a choice of error-locating code $A = C_L(F,D)$, the general formats we use are

(1) $A = C_L(F,D)$ $B = C_L(G-F,D)$ $C = C_\Omega(G,D)$.
(2) $A = C_L(F,D)$ $B = C_\Omega(G+F,D)$ $C = C_L(G,D)$.

The direct application of Theorems 1.8 and 1.9 to geometric Goppa codes is as follows. For $c \in C$, let $y = c + e$ be a received word such that e is nonzero in the error positions $Q = Q_1 + \ldots + Q_t$. For $\dim A > t$, there exists a nonzero $f \in L(F - Q)$, i.e. a nonzero function that vanishes in the error positions. The function f is obtained as a solution to the key equation.

(1) Find $f \in L(F) : \sum_i f(P_i)g(P_i)y_i = 0$, $\forall g \in L(G-F)$.
(2) Find $f \in L(F), h \in L(G+F) : f(P_i)y_i = h(P_i)$, for $i = 1, 2, \ldots, n$.

The key equations produce a nonzero $f \in L(F-Q)$ from which the codeword c can be uniquely decoded if

$$\begin{array}{ll} (1) & L(F-Q) \neq 0 \ \text{ and } \ \Omega(G-F-Q) = 0. \\ (2) & L(F-Q) \neq 0 \ \text{ and } \ L(G+F+Q-D) = 0. \end{array}$$

Two codes $C_\Omega(G,D)$ and $C_L(G^*,D)$ are equal if $G+G^* = (\eta)+D$, for a suitable differential η that depends on D but not on G and G^*. The two key equations are equivalent and lead to algorithms with the same performance. In particular, using the first with G and the second with $G^* = (\eta)+D-G$ leads to similar conditions for decoding the same codes.

Theorem 1.51. *(Basic algorithm) In both key equations, a choice of F with $\deg F = g+t$ will correctly decode a received word with $t \leq (d^*-1)/2 - g/2$ errors.*

Proof.

$$\begin{array}{ll} (1) & \deg(G-F-Q) \ = \ 2g-2+d^*-g-2t > 2g-2. \\ (2) & \deg(G+F+Q-D) \ = \ g+2t-d^* < 0. \end{array}$$

□

If decoding fails with the divisor F because $L(F-Q)=0$ then with little extra computational cost decoding can be attempted with the updated divisor $F+P_\infty$. For this process it is important that

$$\begin{array}{ll} (1) & L(F-Q)=0 \ \Rightarrow \ \Omega(G-F-P_\infty-Q)=0. \\ (2) & L(F-Q)=0 \ \Rightarrow \ L(G+F+P_\infty+Q-D)=0. \end{array}$$

Lemma 1.52. *For a pair of divisors A and B with $\deg B < \dim L(A+B)$*

$$L(B) \neq 0 \ \Rightarrow \ L(A) \neq 0.$$

Proof. Assume $L(B) \neq 0$. Replacing B with an equivalent effective divisor if necessary, $\dim L(A+B) \leq \dim L(A) + \deg B$, and thus $L(A) \neq 0$.

□

Theorem 1.53. *(Modified algorithm) In both key equations, the implications necessary for updating the key equation from a choice F to a choice $F+P_\infty$ hold when*

$$t \leq (d^*-1)/2 + (\dim L(E)-1) - \deg E/2,$$

where (1) $E = K-G+2F+P_\infty$, or (2) $E = G+2F+P_\infty-D$.

With the Riemann-Roch theorem, the defect is the same for E and for $K - E$,

$$\deg(E)/2 - (l(E) - 1) \;=\; \deg(K - E)/2 - (l(K - E) - 1).$$

A divisor is called special if both $L(E) \neq 0$ and $L(K - E) \neq 0$. Clifford's theorem gives that the defect is nonnegative when E is a special divisor.

For the Hermitian curve $K \sim (2g - 2)P_\infty$ and $D \sim nP_\infty$. For one-point codes with odd designed distance $d^* = 2t + 1$, if F goes through $tP_\infty, \ldots, (t+g)P_\infty$ then up to equivalence E goes through (1) $K, \ldots, 2P_\infty, 0$ or (2) $0, 2P_\infty, \ldots, K$.

Theorem 1.54. *The modified algorithm for one-point Hermitian codes from the curve $y^q + y = x^{q+1}$ corrects any number of errors $t \leq (d^* - 1)/2 - q(q-2)/8$.*

For the case $q = 4$, the defect is one and we present an example where an error of size $t = (d^* - 1)/2$ is decoded as an error of size $t + 1$ in the same coset.

Consider $X/\mathbb{F}_{16} : y^4 + y = x^5$. The evaluation of

$$f = x^9y + x^8y + x^8 + x^7y^2 + x^6 + x^5y^3 + x^5 + x^4y^3 + x^4y^2 + x^4 + x^3y^3 + x^3 + x^2y^3 + xy^3 + x + y^3$$

gives a word $c = (c_1, \ldots, c_{23}, 0, \ldots, 0) \in C_L(41P_\infty, D)$ of weight 23. The nonzero positions lie on the lines

$$\ell_1 : x = \alpha^5,\ \ell_2 : x = \alpha^{10},\ \ell_3 : y = (x + 1),\ \ell_4 : y = \alpha^5(x + 1),\ \ell_5 : y = \alpha^{10}(x + 1).$$

Let

$$Q_1 = (\ell_1 - P_\infty) + (\ell_2 - P_\infty) + (\ell_3 - (0, 1)) \sim 13P_\infty - (0, 1).$$
$$Q_2 = \ell_4 + \ell_5 + (0, 1) \sim 10P_\infty + (0, 1).$$

The vanishing ideals for Q_1 and Q_2 are generated by

$$Q_1 : (x^5 + y^4 + y, f_1 = x^2y + \cdots, g_1 = x^6 + \cdots),$$
$$Q_2 : (x^5 + y^4 + y, f_2 = x^2y^2 + \cdots, g_2 = y^3 + \cdots).$$

$-$	1	y	y^2	y^3	y^4	y^5	y^6	$\cdots$
1	$-$	$-$	$-$	$-$	$-$	$+$	$+$	$\cdots$
x	$-$	$-$	$-$	$-$	$+$	$+$	$+$	$\cdots$
x^2	$-$	$+$	$+$	$+$	$+$	$+$	$+$	$\cdots$
x^3	$-$	$+$	$+$	$+$	$+$	$+$	$+$	$\cdots$
x^4	$-$	$+$	$+$	$+$	$+$	$+$	$+$	$\cdots$

$-$	1	y	y^2	y^3	y^4	y^5	y^6	$\cdots$
1	$-$	$-$	$-$	$+$	$+$	$+$	$+$	$\cdots$
x	$-$	$-$	$+$	$+$	$+$	$+$	$+$	$\cdots$
x^2	$-$	$-$	$+$	$+$	$+$	$+$	$+$	$\cdots$
x^3	$-$	$-$	$+$	$+$	$+$	$+$	$+$	$\cdots$
x^4	$-$	$-$	$+$	$+$	$+$	$+$	$+$	$\cdots$

If the word $c = (c_1, \ldots, c_{23}, 0, \ldots, 0)$ is received as $(c_1, \ldots, c_{12}, 0, \ldots, 0)$, with errors in the eleven positions corresponding to Q_2, then the modified algorithm finds the smallest error-locating function f_1 for Q_1 before it finds the smallest error-locating function f_2 for Q_2, and the word is decoded as the allzero word. Among the functions that solve the key equation, there are functions with leading monomial x^2y, x^3y, x^4y that locate Q_1 and functions with leading monomials xy^2, y^3, xy^3, y^4 that locate Q_2. Assuming that the codeword is of the form $s \cdot x^9y + \cdots$, for a given $s \in \mathbb{F}$, we can add one more constraint to the key equation. The functions with leading monomial x^2y, x^3y, x^4y remain valid only when $s = 0$. The functions with leading monomials xy^2, y^3, xy^3, y^4 remain valid only when $s = 1$. None of the seven functions remains valid when $s \neq 0, 1$. The number of errors t is therefore at least 11 if $s = 1$, at least 12 if $s = 0$ and at least 15 if $s \neq 0, 1$. The decoder should therefore first explore the case $s = 1$ which in this case leads to the closest codeword.

	1	y	y^2	y^3	y^4	y^5	y^6	$\cdots$
1	0	1	1	1	0	1	1	$\cdots$
x	1	1	0	1	0	0	$s?$	$\cdots$
x^2	0	0	0	1	1	0	.	$\cdots$
x^3	0	0	0	0	1	.	.	$\cdots$
x^4	0	1	0	1	1	.	.	$\cdots$

	1	y	y^2	y^3	y^4	y^5	y^6	$\cdots$
1	$-$	$-$	$-$	$?$	$?$	$+$	$+$	$\cdots$
x	$-$	$-$	$?$	$?$	$+$	$+$	$+$	$\cdots$
x^2	$-$	$?$	$+$	$+$	$+$	$+$	$+$	$\cdots$
x^3	$-$	$?$	$+$	$+$	$+$	$+$	$+$	$\cdots$
x^4	$-$	$?$	$+$	$+$	$+$	$+$	$+$	$\cdots$

1.4.7. *Secret reconstruction for algebraic-geometric LSSSs*

An ideal $\mathbb{F}$-linear secret sharing scheme $\Sigma = \Sigma_0(\Pi)$ on the set of players $\{1, 2, \ldots, n\}$ is defined as an $\mathbb{F}$-linear map $\Pi : E \longrightarrow \mathbb{F}^{n+1}$. For $x \in E$, the values $\pi_1(x), \ldots, \pi_n(x) \in \mathbb{F}$ are the shares of the secret value $\pi_0(x) \in \mathbb{F}$. We recast the main properties of a linear secret sharing scheme in the language of geometric Goppa codes. We also show that every geometric Goppa code can be decoded up to half the designed distance.

Let $\mathcal{X}/\mathbb{F}$ be a curve. Let $\mathcal{P} = \{P_1, \ldots, P_n\}$ be a set of n rational points and let P_0 be a fixed rational point not in $\mathcal{P}$. For a choice of divisor G, define an algebraic geometric LSSS $\Sigma = \Sigma_0(G, \mathcal{P})$ with the $\mathbb{F}-$linear map $\alpha_L : L(G) \longrightarrow \mathbb{F}^{n+1}$. For $f \in L(G)$, the values $f(P_1), \ldots, f(P_n) \in \mathbb{F}$ are the shares of the secret value $f(P_0) \in \mathbb{F}$.

Lemma 1.55. *For G of degree $\deg G = 2g + t$, the AG-LSSS $\Sigma_0(G, \mathcal{P})$ rejects any subset of size at most t and accepts any subset of size at least $t + 2g + 1$.*

Proof. A subset $A = \{Q_1, \ldots, Q_a\} \subset \mathcal{P}$ is unqualified if and only if $L(G - A) \neq L(G - A - P_0)$. The latter holds for all $a \leq t$ and fails for all $a \geq t + 2g + 1$. □

For a different proof, that uses Riemann's Theorem, let $f_0, \ldots, f_g \in L(G - Q_1 \cdots - Q_t)$ be functions with increasing orders of vanishing at P_0 in the range $\{0, \ldots, 2g\}$. And let $h_0, \ldots, h_g \in L(2gP_0)$ be functions with increasing pole order at P_0 in the range $\{0, \ldots, 2g\}$. By the pigeonhole principle there exist f_i and g_j such that $f_i g_j$ is a unit at P_0.

For a proof that uses Riemann's theorem in combination with Theorem 1.14, let $f_0, f_1, \ldots, f_{g+t} \in L(G)$ be functions with increasing orders of vanishing at P_0 in the range $\{0, \ldots, 2g+t\}$. And let $g_0, \ldots, g_{g+t} \in L((2g+t)P_0)$ be functions with increasing pole order at P_0 in the range $\{0, \ldots, 2g+t\}$. By the pigeonhole principle there exist subsequences $f'_0, \ldots, f'_t$ and $g'_0, \ldots, g'_t$ such that

$$\begin{cases} f'_i * g'_j \in L(G - P_0) & \text{for } i + j < t. \\ f'_i * g'_j \in L(G) \backslash L(G - P_0) & \text{for } i + j = t. \end{cases}$$

Now apply Theorem 1.14. The last proof shows that in special cases the rejection threshold can be higher depending on the vanishing orders at P_0 of the divisor G. The following theorem appears in [9].

Theorem 1.56. *For a divisor G of degree $\deg G = 2g + t$, and for a set of rational points $\mathcal{P}$ of size n, the AG-LSSS $\Sigma_0(G, \mathcal{P})$ is multiplicative in $n - t$ positions (resp. strongly multiplicative) if $3t < n - 4g$ (resp. $3t < n - 6g$).*

Proof. A subset of $n - t$ players can interpolate the product fg of two functions $f, g \in L(G)$ if $2\deg G < n - t$, that is if $3t < n - 4g$. Unqualified subsets for Σ are of size at most $t + 2g$. Strong multiplication is guaranteed if the dual code $C_L(G', D + P_0)$ of $C_L(2G, D + P_0)$ rejects all subsets of size

$t + 2g$. This is the case if $\deg G' = n + 1 + 2g - 2 - 2\deg G \geq 4g + t$, that is if $3t < n - 6g$. □

The theorem shows that for a curve $\mathcal{X}/\mathbb{F}$ of genus g with N rational points, and for $3t+4g < n \leq N-1$, there exist linear secret sharing schemes $\Sigma = \Sigma_0(G, \mathcal{P})$ on n participants such that

- Σ reject all subsets of size t, and
- Σ reconstructs products of secrets from any $n - t$ products of shares.

One of the main results in [9] is that efficient linear secret sharing schemes for an increasing number of participants can be constructed over a small base field using asymptotically good families of curves.

Strong multiplication can be realized with the weaker bound $3t+4g < n$ by choosing the divisor G of degree $2g + t$ such that $\Sigma_0(G, \mathcal{P})$ is trilinear, i.e. such that $C_L(G, D + P_0)$ is essentially orthogonal to $C_L(2G, D + P_0)$. This gives the following generalization of a Shamir secret sharing scheme (Theorem 1.18).

Theorem 1.57. *For a divisor G such that there exists a differential η with $(\eta) = 3G - D - P_0$, the AG-LSSS $\Sigma_0(G, \mathcal{P})$ is trilinear.*

We give such a choice for the Hermitian curve $\mathcal{X}/\mathbb{F}_{16} : Y^4Z + YZ = X^5$. It has 65 rational points that form a complete intersection $X = \mathcal{P}$ with a-invariant $a = q - 1 = 15$. The LSSS $\Sigma(\hat{C})$ defined with the Reed-Muller code $\hat{C} = RM(\nu = 5, X = \mathcal{P})$ is trilinear. The Reed-Muller code is equivalent to a geometric Goppa code defined with a divisor $G \sim 5L$. The curve has parameters $N = 65$ and $g = 6$, the code $\hat{C}$ is of type $[65, 20, 40]$, and the scheme $\Sigma(\hat{C})$ has parameters $n = 64$ and $t = 13$.

For a LSSS $\Sigma_0(G, \mathcal{P})$ with $\deg G \leq n - (2t + 1)$, any two vectors of shares differ in at least $2t + 1$ positions. If at most t shares are corrupted then it is a priori possible to detect the corrupted shares and to determine their correct value. The assumption $4g + 2t = 2\deg G < n - t$ that is used for schemes that are multiplicative in $n - t$ positions corresponds to the much weaker $\deg G \leq n - (2t + 1) - 2g$.

For a LSSS $\Sigma_0(G, \mathcal{P})$ with $\deg G \leq n - (2t + 1) - 2g$, correcting t corrupted shares is straightforward with the key equation in Theorem 1.10. Let $(s_1, \ldots, s_n)$ be a vector of possibly corrupted shares that differs in at

most t positions from the vector $(f(P_1), \ldots, f(P_n))$, for $f \in L(G)$. After choosing a suitable divisor F, we solve for $g \in L(F)$ and $h \in L(G+F)$ such that $g(P_i)s_i = h(P_i)$ for $i = 1, \ldots, n$. The function f is recovered as $f = h/g$. The procedure succeeds if the corrupted positions $Q = Q_1 + \ldots + Q_t$ satisfy

$$L(F - Q) \neq 0 \quad \text{and} \quad L(G + F + Q - D) = 0.$$

The conditions hold for $t + g \leq \deg F \leq t + 2g$. The choice $\deg F = t + g$ gives a key equation with smallest number of variables and this is the most efficient choice. For F of degree $\deg F = t+3g/2$, both conditions hold with $\deg Q = t+g/2$. This choice corrects the largest number of corrupted shares. For $\deg F = t + 2g$, and in particular for $F = G$, only up to t corrupted shares can be corrected but there exists a solution for g with $g(P_0) \neq 0$ and in that case the secret can be recovered as $f(P_0) = h(P_0)/g(P_0)$. The last choice corresponds to the reconstruction procedure in [12] for a general LSSS. The constraint $g(P_0) \neq 0$ is not needed for an AG-LSSS if we evaluate the secret as $f(P_0) = (h/g)(P_0)$.

To correct t corrupted shares in a LSSS $\Sigma = \Sigma_0(G, \mathcal{P})$ with $\deg G \leq n - (2t + 1)$, we use the procedure in Theorem 1.17. The procedure makes use of two schemes $\Sigma' = \Sigma_0(F, \mathcal{P})$ and $\Sigma'' = \Sigma_0(F^*, \mathcal{P})$ such that $C_L(F + F^*, \mathcal{P} + P_0)$ is orthogonal to $C_L(G, \mathcal{P} + P_0)$. Let $f \in L(G)$. If $(s_1, \ldots, s_n)$ is a vector that differs from the vector $(f(P_1), \ldots, f(P_n))$ in the positions $Q = Q_1 + \ldots + Q_t$, then the procedure returns the correct value for $f(P_0)$ if

$$L(F - Q) \neq L(F - Q - P_0) \quad \text{and} \quad L(F^* - Q) \neq L(F^* - Q - P_0).$$

If one of the conditions fails the procedure may not return a value. An incorrect value is returned only if

$$L(F - Q) = L(F - Q - P_0) \quad \text{and} \quad L(F^* - Q) = L(F^* - Q - P_0).$$

Theorem 1.58. *Let $C = C_L(G, \mathcal{P})$ be a geometric Goppa code of length n with divisor G of degree $\deg G = n - (2t + 1)$. Let P_0 be a point not in $\mathcal{P}$. For $f \in L(G)$, let $(s_1, \ldots, s_n)$ be a vector that differs in no more than t positions from the vector $(f(P_1), \ldots, f(P_n))$. Among the values for $f(P_0)$ that are returned by the reconstruction procedure when it is applied with $F = tP_0, \ldots, (t + 2g)P_0$, the correct value for $f(P_0)$ outnumbers any other value.*

Proof. $\dim L((t+2g)P_0 - Q) - \dim L(tP_0 - Q - P_0) = g + 1$. For $F = tP_0, \ldots, (t+2g)P_0$, the condition $L(F-Q) \neq L(F-Q-P_0)$ holds $g+1$ times and fails g times. The matching divisor F^* similarly meets the condition $L(F^* - Q) \neq L(F^* - Q - P_0)$ exactly $g+1$ times and fails it g times. With the pigeonhole principle, both conditions hold, and the correct value is returned, at least once. Moreover, the number of times that an incorrect value is returned is at most the number of times that both conditions fail which is one less than the number of times that both conditions hold. □

1.4.8. *Weight distributions*

Weight distributions of linear codes are in general hard to determine. The extra structure of geometric Goppa codes makes it possible to approach their weight distribution as a distribution problem of effective divisors over divisor classes and to benefit from the group structure on the divisor classes. For a code $C_L(G, D)$ with injective encoding map $L(G) \longrightarrow C_L(G, D)$, words of weight w correspond to functions in $L(G)$ with $n - w$ zeros in D. The correspondence between nonzero words of weight w and effective divisors in the class of G that intersect D in $n - w$ points is $(q-1)$-to-one. Thus, in order to determine weight distributions, we may consider all effective divisors of a given degree that intersect D in a given number of points and their distribution over the finitely many divisor classes of that degree.

The main tools for pursuing the above approach are zeta functions, to study divisor distributions, and Fourier analysis over the finite group of divisor classes of degree zero. In this section we show that the weight distributions of the codes $C_L(G, D)$, where G runs over a full set of inequivalent divisors $G_1, G_2, \cdots, G_h$ of the same degree, have an average weight distribution that depends only on the zeta function of the curve and the degrees of the divisors G and D. The error terms for each individual weight distribution are controlled by the $L-series$ $L(T, \chi)$, where $\chi = \chi_1, \ldots, \chi_h$ is an unramified character of the function field.

Let $K = \mathbb{F}(\mathcal{X})$ be the function field of $\mathcal{X}$ and let $\mathcal{P}_K$ be the set of all the places of K. The *group of divisors* $D(K)$ is the free abelian group generated by the set of places $\mathcal{P}_K$. The principal divisors (f), for a nonzero $f \in K$, form a subgroup $P(K)$ of $D(K)$. The quotient $D(K)/P(K)$ is the *divisor class group* $C(K)$. The group $C(K)$ is finitely generated of the form $\Gamma \times \mathbb{Z}$. The finite torsion subgroup Γ is the group of divisor classes of degree zero.

The set of places $\mathcal{P}_K$ generates the *semigroup of effective divisors* $E(K)$. For a fixed divisor class E of degree one, let

$$L(T) = \sum_{a \geq 0} \sum_{g \in \Gamma} |(g + aE) \cap E(K)| \; X^g T^a$$

be a generating function for the number of effective divisors in the divisor class $g + aE$. Let $\{e_\chi : \chi \in \hat{\Gamma}\}$ be a basis of primitive idempotents for $\mathbb{C}\Gamma$,

$$e_\chi = \frac{1}{|\Gamma|} \sum_{g \in \Gamma} \chi(-g) X^g,$$

so that $X^g e_\chi = \chi(g)\, e_\chi$. Define coordinate functions $L(T, g), L(T, \chi) \in \mathbb{C}[[T]]$ via

$$L(T) = \sum_g L(T, g) X^g = \sum_\chi L(T, \chi) e_\chi.$$

The function $L(T, g)$ is a generating function for the number of effective divisors in the divisor class $g + aE$, for $a \geq 0$. The function $L(T, \chi)$ is a Dirichlet L-series for a Dirichlet character of trivial conductor. For a non-trivial character χ, $L(T, \chi)$ is a polynomial of degree $2g - 2$ with cyclotomic integer coefficients. For the trivial character χ_0, $L(T, \chi_0) = Z(T)$ is the zeta function of the curve. For a rational place P, let $g_P + E$ be the divisor class of P, for $g_P \in \Gamma$. For a subset $\mathcal{P}$ of rational places, define

$$\Lambda(T) = \prod_{P \in \mathcal{P}} (1 + X^{g_P} T) \in \mathbb{C}\Gamma[T],$$

with coordinate functions

$$\Lambda(T) = \sum_g \Lambda(T, g) X^g = \sum_\chi \Lambda(T, \chi) e_\chi.$$

Theorem 1.59. *The distribution over divisor classes of effective divisors that contain precisely a given number of elements from $\mathcal{P}$ is given by*

$$A(U, T) = L(T)\Lambda(U - T) \in \mathbf{C}\Gamma[U](T).$$

The coordinate function $A(U, T, g) \in \mathbf{C}[U][[T]]$ is the generating function for the number of effective divisors in the divisor class $g + (i + j)E$ with precisely i elements of $\mathcal{P}$ in the support.

Proof. The generating function $L(T)$ has an Euler product decomposition. The contribution of $P \in \mathcal{P}$ to $A(U, T)$ is, with $g = g_P$,

$$\frac{1 + X^g(U - T)}{1 - X^g T} = 1 + X^g U + X^{2g} UT + X^{3g} UT^2 + \cdots.$$

Hence the variable U keeps track of the precise number of places $P \in \mathcal{P}$ that contribute to a term of $A(U,T)$. □

To compute weight distributions with the theorem we compute the coordinate functions $A(U,T,\chi) = L(T,\chi)\Lambda(U-T,\chi)$ on the basis of idempotents and apply an inverse Fourier transform to recover coordinate functions $A(U,T,g)$ for $A(U,T)$. The top row in the table below gives the weight distribution for a code of type $[24,16,7]$ over $\mathbb{F}_8$ constructed with the Klein curve. The method outlined here produces the weight distributions for all $2744 = 14^3$ codes of type $[24,16]$ on the Klein curve. For the code and its dual, only the weights below the Singleton bound are listed. Using only the contribution of the trivial character $\chi = \chi_0$ gives the average weight distribution for codes defined with inequivalent divisors of the same degree.

$$\frac{1}{|\Gamma|}\sum_g A(U,T,g) = \frac{1}{|\Gamma|}Z(T)(1+U-T)^n.$$

Table 1.3. Weight distributions for the 2744 distinct $[24,16,\geq 6]$ codes on the Klein quartic over $\mathbb{F}_8$.

	Small weights			Small dual weights		
#	$\bar{A}_6$	$\bar{A}_7$	$\bar{A}_8$	$\bar{A}^\perp_{14}$	$\bar{A}^\perp_{15}$	$\bar{A}^\perp_{16}$
1	0	2520	37620	696	4200	11340
7	52	2184	38643	852	3720	11907
24	35	2170	38709	672	4329	11753
24	56	2138	37968	707	4469	10846
168	38	2167	38642	683	4312	11752
168	60	2131	37896	745	4278	11276
168	47	2190	38106	735	4212	11544
168	53	2136	38340	747	4167	11643
126	52	2104	38430	692	4404	11378
126	40	2176	38280	660	4484	11336
252	60	2060	38537	729	4246	11718
504	48	2140	38288	692	4374	11506
504	49	2154	38336	731	4222	11558
504	46	2165	38348	717	4272	11478
avg	49.1	2144.2	38328.1	714.7	4288.5	11525.2

Computed as an inverse Fourier transform of the unramified $L-$series of the curve.

1.5. Bibliographic notes

There are many textbooks for coding theory, including [4], [48], [55], [57]. The books [5], [36], [44], [49], [54], [62], [68], [71], [72], [75], [77], [79], as well as the survey chapters [10], [42], [45], [47], discuss algebraic geometry codes, each with a distinct approach and emphasis. We give a few more references for the topics discussed in this chapter. Roos bound for the minimum distance [22], Linear secret sharing schemes [12], Weight distributions and codes over extension fields [21], [76], Dual BCH codes [20], [32], [69], Codes from the Klein and Suzuki curves [8], [17], [33], [39], [61], Floor bound [7], [58], [56], Explicit towers [1], [11], [25], [30], [31], [59], [70], [78], One-point codes [29], [52], [80], Two-point codes [2], [3], [46], [51], [60], [65], Error correction [13], [23], [37], [38], [66], Secret reconstruction for algebraic-geometric LSSSs [9], [10], [14], Weight distributions [18].

References

[1] P. Beelen and I. I. Bouw, Asymptotically good towers and differential equations, *Compos. Math.* **141**(6), 1405–1424, (2005).
[2] P. Beelen and N. Tutaş, A generalization of the Weierstrass semigroup, *J. Pure Appl. Algebra.* **207**(2), 243–260, (2006).
[3] P. Beelen, The order bound for general algebraic geometric codes, *Finite Fields Appl.* **13**(3), 665–680, (2007).
[4] J. Bierbrauer, *Introduction to coding theory.* Discrete Mathematics and its Applications (Boca Raton), (Chapman & Hall/CRC, Boca Raton, FL, 2005).
[5] R. E. Blahut, *Algebraic codes on lines, planes, and curves: an engineering approach.* (Cambridge University Press, Cambridge, 2008)
[6] A. Blokhuis, A. Brouwer, and H. Wilbrink, Hermitian unitals are code words, *Discrete Math.* **97**(1-3), 63–68, (1991).
[7] C. Carvalho and F. Torres, On Goppa codes and Weierstrass gaps at several points, *Des. Codes Cryptogr.* **35**(2), 211–225, (2005).
[8] C.-Y. Chen and I. M. Duursma, Geometric Reed-Solomon codes of length 64 and 65 over $\mathbb{F}_8$, *IEEE Trans. Inform. Theory.* **49**(5), 1351–1353, (2003).
[9] H. Chen and R. Cramer. Algebraic geometric secret sharing schemes and secure multi-party computations over small fields. In *CRYPTO*, pp. 521–536, (2006).
[10] H. Chen, Algebraic geometric codes with applications, *Front. Math. China.* **2**(1), 1–11, (2007).
[11] H. Cohn, *Introduction to the construction of class fields.* Cambridge Studies in Advanced Mathematics volume 6 (Cambridge University Press, Cambridge, 1985). Reprint by (Dover Publications Inc., New York, 1994).
[12] R. Cramer, V. Daza, I. Gracia, J. J. Urroz, G. Leander, J. Martí-Farré, and C. Padró. On codes, matroids and secure multi-party computation from

linear secret sharing schemes. In *Advances in cryptology—CRYPTO 2005*, vol. 3621, *Lecture Notes in Comput. Sci.*, pp. 327–343. Springer, Berlin, (2005).

[13] I. M. Duursma, Algebraic decoding using special divisors, *IEEE Trans. Inform. Theory.* **39**(2), 694–698, (1993).

[14] I. M. Duursma, Majority coset decoding, *IEEE Trans. Inform. Theory.* **39** (3), 1067–1070, (1993).

[15] I. M. Duursma, *Decoding codes from curves and cyclic codes.* (Technische Universiteit Eindhoven, Eindhoven, 1993). Dissertation, Technische Universiteit Eindhoven, Eindhoven, 1993.

[16] I. M. Duursma and R. Kötter, Error-locating pairs for cyclic codes, *IEEE Trans. Inform. Theory.* **40**(4), 1108–1121, (1994).

[17] I. M. Duursma, Monomial embeddings of the Klein curve, *Discrete Math.* **208/209**, 235–246, (1999). Combinatorics (Assisi, 1996).

[18] I. M. Duursma, Weight distributions of geometric Goppa codes, *Trans. Amer. Math. Soc.* **351**(9), 3609–3639, (1999).

[19] I. M. Duursma, C. Rentería, and H. Tapia-Recillas, Reed-Muller codes on complete intersections, *Appl. Algebra Engrg. Comm. Comput.* **11**(6), 455–462, (2001).

[20] I. M. Duursma, Preparata codes through lattices, *IEEE Trans. Inform. Theory.* **47**(1), 36–44, (2001).

[21] I. M. Duursma. Combinatorics of the two-variable zeta function. In *Finite fields and applications*, vol. 2948, *Lecture Notes in Comput. Sci.*, pp. 109–136. Springer, Berlin, (2004).

[22] I. M. Duursma and R. Pellikaan, A symmetric Roos bound for linear codes, *J. Combin. Theory Ser. A.* **113**(8), 1677–1688, (2006).

[23] D. Ehrhard. Decoding algebraic-geometric codes by solving a key equation. In *Coding theory and algebraic geometry (Luminy, 1991)*, vol. 1518, *Lecture Notes in Math.*, pp. 18–25. Springer, Berlin, (1992).

[24] D. Ehrhard, Achieving the designed error capacity in decoding algebraic-geometric codes, *IEEE Trans. Inform. Theory.* **39**(3), 743–751, (1993).

[25] N. D. Elkies. Explicit modular towers. In *Proceedings of the Thirty-Fifth Annual Allerton Conference on Communication, Control and Computing (Univ. of Illinois at Urbana-Champaign.*

[26] N. D. Elkies. Excellent codes from modular curves. In *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*, pp. 200–208 (electronic), New York, (2001). ACM.

[27] N. D. Elkies. Explicit towers of Drinfeld modular curves. In *European Congress of Mathematics, Vol. II (Barcelona, 2000)*, vol. 202, *Progr. Math.*, pp. 189–198. Birkhäuser, Basel, (2001).

[28] G. L. Feng and T. R. N. Rao, Decoding algebraic-geometric codes up to the designed minimum distance, *IEEE Trans. Inform. Theory.* **39**(1), 37–45, (1993).

[29] A. García, S. J. Kim, and R. F. Lax, Consecutive Weierstrass gaps and minimum distance of Goppa codes, *J. Pure Appl. Algebra.* **84**(2), 199–207, (1993).

[30] A. García and H. Stichtenoth, A tower of Artin-Schreier extensions of function fields attaining the Drinfeld-Vladuts bound, *Invent. Math.* **121**(1), 211–222, (1995).

[31] A. Garcia and H. Stichtenoth. Explicit towers of function fields over finite fields. In *Topics in geometry, coding theory and cryptography*, vol. 6, *Algebr. Appl.*, pp. 1–58. Springer, Dordrecht, (2007).

[32] G. van der Geer, R. Schoof, and M. van der Vlugt, Weight formulas for ternary Melas codes, *Math. Comp.* **58**(198), 781–792, (1992).

[33] M. Giulietti, G. Korchmáros, and F. Torres, Quotient curves of the Suzuki curve, *Acta Arith.* **122**(3), 245–274, (2006).

[34] V. D. Goppa, Decoding and Diophantine approximations, *Problems of Control and Information Theory/Problemy Upravlenija i Teorii Informacii.* **5** (3), 195–206, (1976).

[35] V. D. Goppa, Codes on algebraic curves, *Dokl. Akad. Nauk SSSR.* **259**(6), 1289–1290, (1981).

[36] V. D. Goppa, *Geometry and codes.* vol. 24, *Mathematics and its Applications (Soviet Series)*, (Kluwer Academic Publishers Group, Dordrecht, 1988). Translated from the Russian by N. G. Shartse.

[37] V. Guruswami and M. Sudan, Improved decoding of Reed-Solomon and algebraic-geometry codes, *IEEE Trans. Inform. Theory.* **45**(6), 1757–1767, (1999).

[38] V. Guruswami and A. C. Patthak, Correlated algebraic-geometric codes: improved list decoding over bounded alphabets, *Math. Comp.* **77**(261), 447–473 (electronic), (2008).

[39] J. P. Hansen and H. Stichtenoth, Group codes on certain algebraic curves with many rational points, *Appl. Algebra Engrg. Comm. Comput.* **1**(1), 67–77, (1990).

[40] J. P. Hansen and J. P. Pedersen, Automorphism groups of Ree type, Deligne-Lusztig curves and function fields, *J. Reine Angew. Math.* **440**, 99–109, (1993).

[41] J. P. Hansen, Linkage and codes on complete intersections, *Appl. Algebra Engrg. Comm. Comput.* **14**(3), 175–185, (2003).

[42] J. W. P. Hirschfeld. Linear codes and algebraic curves. In *Geometrical combinatorics (Milton Keynes, 1984)*, vol. 114, *Res. Notes in Math.*, pp. 35–53. Pitman, Boston, MA, (1984).

[43] J. W. P. Hirschfeld, G. Korchmáros, and F. Torres. The number of points on an algebraic curve over a finite field. In *Surveys in combinatorics 2007*, vol. 346, *London Math. Soc. Lecture Note Ser.*, pp. 175–200. Cambridge Univ. Press, Cambridge, (2007).

[44] J. W. P. Hirschfeld, G. Korchmáros, and F. Torres, *Algebraic curves over a finite field.* (Princeton University Press, Princeton, 2008)

[45] T. Høholdt, J. H. van Lint, and R. Pellikaan. Algebraic geometry of codes. In *Handbook of coding theory, Vol. I, II*, pp. 871–961. North-Holland, Amsterdam, (1998).

[46] M. Homma and S. J. Kim, The complete determination of the minimum distance of two-point codes on a Hermitian curve, *Des. Codes Cryptogr.* **40**

(1), 5–24, (2006).
[47] W.-b. Hu and C.-p. Xing, A survey on algebraic-geometry codes, *Adv. Math. (China).* **35**(6), 641–656, (2006).
[48] W. C. Huffman and V. Pless, *Fundamentals of error-correcting codes.* (Cambridge University Press, Cambridge, 2003).
[49] N. E. Hurt, *Many rational points. Coding theory and algebraic geometry* Mathematics and its Applications, 564. (Kluwer Academic Publishers, Dordrecht, 2003)
[50] J. D. Key, Hermitian varieties as codewords, *Des. Codes Cryptogr.* **1**(3), 255–259, (1991).
[51] S. J. Kim, On the index of the Weierstrass semigroup of a pair of points on a curve, *Arch. Math.* **62**(1), 73–82, (1994).
[52] C. Kirfel and R. Pellikaan, The minimum distance of codes in an array coming from telescopic semigroups, *IEEE Trans. Inform. Theory.* **41**(6, part 1), 1720–1732, (1995). Special issue on algebraic geometry codes.
[53] K. Lauter, Deligne-Lusztig curves as ray class fields, *Manuscripta Math.* **98** (1), 87–96, (1999).
[54] J. H. van Lint and G. van der Geer, *Introduction to coding theory and algebraic geometry.* vol. 12, *DMV Seminar*, (Birkhäuser Verlag, Basel, 1988).
[55] J. H. van Lint, *Introduction to coding theory.* vol. 86, *Graduate Texts in Mathematics*, (Springer-Verlag, Berlin, 1999), third edition.
[56] B. Lundell and J. McCullough, A generalized floor bound for the minimum distance of geometric Goppa codes, *J. Pure Appl. Algebra* **207**(1), 155–164, (2006).
[57] F. J. MacWilliams and N. J. A. Sloane, *The theory of error-correcting codes.* (North-Holland Publishing Co., Amsterdam, 1977). North-Holland Mathematical Library, Vol. 16.
[58] H. Maharaj and G. L. Matthews, On the floor and the ceiling of a divisor, *Finite Fields Appl.* **12**(1), 38–55, (2006).
[59] H. Maharaj. Explicit towers and codes. In *Recent trends in coding theory and its applications*, vol. 41, *AMS/IP Stud. Adv. Math.*, pp. 35–71. Amer. Math. Soc., Providence, RI, (2007).
[60] G. L. Matthews, Weierstrass pairs and minimum distance of Goppa codes, *Des. Codes Cryptogr.* **22**(2), 107–121, (2001).
[61] G. L. Matthews, Codes from the Suzuki function field, *IEEE Trans. Inform. Theory.* **50**(12), 3298–3302, (2004).
[62] C. Moreno, *Algebraic curves over finite fields.* vol. 97, *Cambridge Tracts in Mathematics*, (Cambridge University Press, Cambridge, 1991).
[63] H. Niederreiter and C. P. Xing, Low-discrepancy sequences obtained from algebraic function fields over finite fields, *Acta Arith.* **72**(3), 281–298, (1995).
[64] H. Niederreiter and C. Xing, *Rational points on curves over finite fields: theory and applications.* vol. 285, *London Mathematical Society Lecture Note Series*, (Cambridge University Press, Cambridge, 2001).
[65] S. Park, *Applications of algebraic curves to cryptography*, Thesis (University of Illinois, Urbana, 2007).
[66] F. Parvaresh and A. Vardy. Correcting errors beyond the guruswami-sudan

radius in polynomial time. In *Proceedings of the 46th IEEE Symposium on Foundations of Computer Science (FOCS), 2005.*

[67] J. P. Pedersen. A function field related to the Ree group. In *Coding theory and algebraic geometry (Luminy, 1991)*, vol. 1518, *Lecture Notes in Math.*, pp. 122–131. Springer, Berlin, (1992).

[68] O. Pretzel, *Codes and algebraic curves.* vol. 8, *Oxford Lecture Series in Mathematics and its Applications*, (The Clarendon Press Oxford University Press, New York, 1998).

[69] R. Schoof, Families of curves and weight distributions of codes, *Bull. Amer. Math. Soc. (N.S.).* **32**(2), 171–183, (1995).

[70] K. W. Shum, I. Aleshnikov, P. V. Kumar, H. Stichtenoth, and V. Deolaikar, A low-complexity algorithm for the construction of algebraic-geometric codes better than the Gilbert-Varshamov bound, *IEEE Trans. Inform. Theory.* **47** (6), 2225–2241, (2001).

[71] S. A. Stepanov, *Codes on algebraic curves.* (Kluwer Academic/Plenum Publishers, New York, 1999).

[72] H. Stichtenoth, *Algebraic function fields and codes.* Universitext, (Springer-Verlag, Berlin, 1993).

[73] H. Stichtenoth and C. Xing, Excellent nonlinear codes from algebraic function fields, *IEEE Trans. Inform. Theory.* **51**(11), 4044–4046, (2005).

[74] M. A. Tsfasman, S. G. Vlăduţ, and T. Zink, Modular curves, Shimura curves, and Goppa codes, better than Varshamov-Gilbert bound, *Math. Nachr.* **109**, 21–28, (1982).

[75] M. A. Tsfasman and S. G. Vlăduţ, *Algebraic-geometric codes.* vol. 58, *Mathematics and its Applications (Soviet Series)*, (Kluwer Academic Publishers Group, Dordrecht, 1991). Translated from the Russian by the authors.

[76] M. A. Tsfasman and S. G. Vlăduţ, Geometric approach to higher weights, *IEEE Trans. Inform. Theory.* **41**(6, part 1), 1564–1588, (1995). Special issue on algebraic geometry codes.

[77] M. Tsfasman, S. Vlăduţ, and D. Nogin, *Algebraic geometric codes: basic notions.* vol. 139, *Mathematical Surveys and Monographs*, (American Mathematical Society, Providence, RI, 2007).

[78] C. Voss and T. Høholdt, An explicit construction of a sequence of codes attaining the Tsfasman-Vlăduţ-Zink bound: the first steps, *IEEE Trans. Inform. Theory.* **43**(1), 128–135, (1997).

[79] J. L. Walker, *Codes and curves.* vol. 7, *Student Mathematical Library*, (American Mathematical Society, Providence, RI, 2000). IAS/Park City Mathematical Subseries.

[80] K. Yang and P. V. Kumar. On the true minimum distance of Hermitian codes. In *Coding theory and algebraic geometry (Luminy, 1991)*, vol. 1518, *Lecture Notes in Math.*, pp. 99–107. Springer, Berlin, (1992).

[81] C. Xing, Nonlinear codes from algebraic curves improving the Tsfasman-Vlăduţ-Zink bound, *IEEE Trans. Inform. Theory.* **49**(7), 1653–1657, (2003).

Chapter 2

The Decoding of Algebraic Geometry Codes

Peter Beelen and Tom Høholdt

Department of Mathematics,
Technical University of Denmark,
Matematiktorvet, Building 303S,
DK 2800, Kgs.Lyngby,
{p.beelen,t.hoeholdt}@mat.dtu.dk

Contents

2.1. Introduction

The work on decoding of algebraic geometry codes started in 1986 and in the following 10 years a lot of papers appeared. The paper [11] surveys all the work on decoding until 1995. In this chapter we will present decoding algorithms using recent ideas and methods.

The chapter is organized as follows: In Section 2.2 we present the basic algorithm for decoding a general algebraic geometry code $C_L(D, G)$, this algorithm only decodes error-patterns of weight smaller than $\frac{d-1}{2} - g$ where d is the Goppa bound on the minimum distance of the code and g is the genus of the curve used in the construction. Section 2.3 contains a syndrome

formulation of the basic algorithm, this gives the possibility of correcting $\frac{d-1}{2} - \frac{g}{2}$ errors. In Section 2.4 we introduce and prove the generalized order bound, which improves the Goppa bound on the minimum distance in many cases and in Section 2.4 we use majority voting to give an algorithm that corrects up to $\frac{d_S-1}{2}$ errors where d_S is the order bound on the minimum distance. Section 2.6 contains a list decoding algorithm which gives the possibility of correcting more errors, but then the correct codeword is on a (small) list. In Section 2.7 we give a syndrome formulation of the list decoder in order to drastically reduce the complexity of this algorithm. The sections contain a number of examples illustrating the methods. We have chosen not to include references in the text, but in Section 2.8 we discuss the literature. The chapter ends with a full list of references.

2.2. The basic algorithm

2.2.1. *Decoding*

When an (n, k) code C is used for correcting errors, one of the important problems is the design of a *decoder*. One can consider this as a mapping from $\mathbb{F}_q^n$ into the code C, as an algorithm or sometimes even as a physical device. We will usually see a decoder as a mapping or as an algorithm. One way of stating the objective of the decoder is: for a received vector r, select a codeword c that minimizes $d(r, c)$. This is called *maximum likelihood decoding*. It is clear that if the code is t-error correcting, i.e $t < \frac{d_{min}}{2}$ and $r = c + e$ with $w(e) \leq t$ then the output of such a decoder is c.
It is often difficult to design a maximum likelihood decoder, but if we only want to correct t errors where $t < \frac{d_{min}}{2}$ it is sometimes easier to get a good algorithm.

Definition 2.1. A *minimum distance decoder* is a decoder that, given a received word r, selects the codeword c that satisfies $d(r, c) < \frac{d_{min}}{2}$ if such a codeword exists, and otherwise declares failure.

It is obvious that there can be at most one such codeword.
We will also in the following consider a so-called *list decoder*.

Definition 2.2. Let $0 \leq \tau \leq n$. A τ list decoder is a decoder that, given a received word r, outputs all codewords c such that $d(r, c) \leq \tau$.

Again it is clear that if $\tau < \frac{d_{min}}{2}$ then there is at most one codeword, but for larger τ there could be more, hence the name list decoder. Also,

for such a decoder to be useful, the number of codewords on the list should be small.

2.2.2. *The basic algorithm for decoding of algebraic geometry codes*

Let χ be an algebraic curve i.e. an affine or projective variety of dimension one, which is absolutely irreducible and nonsingular and whose defining equations are (homogeneous) polynomials with coefficients in a finite field $\mathbb{F}$, and let $\mathscr{F}$ denote its function field. Moreover, let G and $D = P_1 + \cdots + P_n$ be $\mathbb{F}$-rational divisors on χ with $\operatorname{Supp} D \cap \operatorname{Supp} G = \varnothing$ and denote by g the genus of the curve χ. Moreover define the functions

$$\begin{aligned} \mathrm{Ev}_D &: L(G) \to \mathbb{F}^n \\ & f \mapsto (f(P_1), \ldots, f(P_n)) \end{aligned}$$

and

$$\begin{aligned} \mathrm{Res}_D &: \Omega(G - D) \to \mathbb{F}^n \\ & \omega \mapsto (\mathrm{res}_{P_1}(\omega), \ldots, \mathrm{res}_{P_n}(\omega)), \end{aligned}$$

that are used to construct the codes $C_L(D, G)$ and $C_\Omega(D, G)$.

Suppose that we wish to use the code $C_L(D, G)$ and that we have received the word $(r_1, \ldots, r_n)$ containing at most t errors. The algorithm below works with a divisor A with $\operatorname{Supp} A \cap \operatorname{Supp} D = \varnothing$ satisfying

(1) $\deg A < n - t$
(2) $\deg A > \frac{n + \deg G}{2} + g - 1$

It can be seen that if $t < \frac{n - \deg G}{2} - g$ then such a divisor A exists, but we will see later in Section 3 that condition (2) above can be relaxed and then we can work with larger t.

The idea of the algorithm is to find a nonzero polynomial $Q(y) \in \mathscr{F}[y]$ such that:

(i) $Q(y) = Q_0 + Q_1 y$ where $Q_0 \in L(A)$ and $Q_1 \in L(A - G)$
(ii) $Q_0(P_j) + r_j Q_1(P_j) = 0, j = 1, \ldots, n$

The polynomial $Q(y)$ is called an *interpolation polynomial*.

Lemma 2.3. *Suppose the transmitted word is $ev_D(f)$ with $f \in L(G)$ and $Q(y)$ satisfy (i) and (ii) then $f = -\frac{Q_0}{Q_1}$.*

Proof. Since $f \in L(G)$ and $Q_1 \in L(A-G)$ we have $fQ_1 \in L(A)$ and therefore $Q(f) \in L(A)$. We also have that $Q_0(P) + f(P)Q_1(P) = 0$ for at least $n-t$ points $P \in \{P_1, \ldots, P_n\}$, so $Q(f) \in L(A - P_{i_1} - \cdots - P_{i_s})$ with $s \geq n-t$. But $\deg(A - P_{i_1} - \cdots - P_{i_s}) < 0$ and therefore $Q(f) = 0$ and the result follows. □

We also get that $Q(y) = Q_1(-f+y)$ and therefore Q_1 must have the error-positions among its zeroes. For this reason Q_1 is called an *error-locator*.

Lemma 2.4. *If the divisor A satisfies condition* (2) *above then there exists a nonzero $Q(y) \in \mathscr{F}[y]$ satisfying (i) and (ii).*

Proof. Let $\{g_1, \ldots, g_{l_0}\}$ be a basis for $L(A)$ and $\{h_1, \ldots, h_{l_1}\}$ a basis for $L(A-G)$. We then write

$$Q_0 = \sum_{i=1}^{l_0} q_{0i} g_i$$

and

$$Q_1 = \sum_{i=1}^{l_1} q_{1i} h_i$$

so (ii) becomes

$$\sum_{i=1}^{l_0} q_{0i} g_i(P_j) + r_j \sum_{i=1}^{l_1} q_{1i} h_i(P_j) = 0, \text{ with } j = 1, \ldots, n.$$

Since $l_0 + l_1 = l(A) + l(A-G) \geq \deg A + \deg(A-G) - 2g + 2 = 2\deg A - \deg G - 2g + 2 > n$ the n linear homogenous equations have more that n unknowns (q_{0i} and q_{1i}) so there is a nonzero solution. □

Based on the considerations above we can now present the so-called *basic algorithm*

Input: A received word $(r_1, r_2, \ldots, r_n)$.
Find a polynomial $Q(y)$ satisfying (i) and (ii).
If $f = -\frac{Q_0}{Q_1} \in L(G)$ **Output**: $\mathrm{Ev}_D(f)$.
Else **Output**: Failure.

2.3. Syndrome formulation of the basic algorithm

In this section we will reformulate the basic algorithm using so-called syndromes. The advantage of this over the description given before is that an interpolation polynomial can be found easier now, since its defining system of linear equations can be divided into two pieces. Also we will be able to show now that the basic algorithm for the code $C_L(D,G)$ can correct up to $t < (n - \deg G - g)/2$ errors improving the previous result

that it can correct up to $(d-1)/2-g$ errors. For future reference we give the following definitions:

$$\mathbf{M}_A := \begin{pmatrix} g_1(P_1) & \dots & g_{l_0}(P_1) \\ \vdots & & \vdots \\ g_1(P_n) & \dots & g_{l_0}(P_n) \end{pmatrix}, \tag{2.1}$$

$$\mathbf{D_r} := \begin{pmatrix} r_1 & & \\ & \ddots & \\ & & r_n \end{pmatrix} \tag{2.2}$$

and

$$\mathbf{M}_{A-G} := \begin{pmatrix} h_1(P_1) & \dots & h_{l_1}(P_1) \\ \vdots & & \vdots \\ h_1(P_n) & \dots & h_{l_1}(P_n) \end{pmatrix} \tag{2.3}$$

The interpolation conditions can then be written as:

$$\mathbf{M}_A \cdot \mathbf{q_0} + \mathbf{D_r}\mathbf{M}_{A-G} \cdot \mathbf{q_1} = \mathbf{0} \tag{2.4}$$

Thus we can find interpolation polynomials using linear algebra techniques on system (2.4). A faster method can be obtained by multiplying system (2.4) from the left with a suitable invertible matrix. We will construct this matrix using differentials on the curve χ.

Lemma 2.5. *Let A be a non-trivial divisor and write $l_0 = l(A)$. Further let $D = P_1 + \cdots + P_n$ as before and suppose that* $\operatorname{Supp} A \cap \operatorname{Supp} D = \varnothing$. *Finally let $\langle \cdot, \cdot \rangle$ denote the standard inner product on $\mathbb{F}^n$. Then there exist differentials $\omega_1, \dots, \omega_n$ such that*

(i) The set $\{\operatorname{Res}_D(\omega_1), \dots, \operatorname{Res}_D(\omega_n)\}$ is a basis for $\mathbb{F}^n$,

(ii) The set $\{\operatorname{Res}_D(\omega_1), \dots, \operatorname{Res}_D(\omega_{n-l_0})\}$ is a basis of the code $C_\Omega(D, A)$,

(iii) For any point $P \in \operatorname{Supp} D$ and $1 \le i \le n$, we have $v_P(\omega_i) \ge -1$,

(iv) For any $\mathbf{c} \in C_L(D, A)$ and $1 \le j \le n - l_0$, we have $\langle \mathbf{c}, \operatorname{Res}_D(\omega_j) \rangle = 0$.

Proof. First of all, we choose some point T outside $\operatorname{Supp} D$ (not necessarily rational). Note that $C_\Omega(D, -T) = \mathbb{F}^n$, since it is the dual of the code $C_L(D, -T)$ and $L(-T) = \{0\}$. So for any $v \in \mathbb{F}^n$, there exists a differential

$\omega \in \Omega(-T-D)$ such that $(\mathrm{Res}_D(\omega)) = v$. Since $\deg A < n$, we see that $\dim\Omega(A-D) \geq \dim C_\Omega(D,A) = n - \dim C_L(D,A) = n - l(A) = n - l_0$. Therefore, starting with a basis $v_1, \ldots, v_{n-l_0}$ of $C_\Omega(D,A)$, we can find differentials $\omega_1, \ldots, \omega_{n-l_0} \in \Omega(A-D)$ such that $v_i = \mathrm{res}(\omega_i)$.

We can complete the set $\{v_1, \ldots, v_{n-l_0}\}$ to a basis of $\mathbb{F}^n$ by adding l_0 suitable vectors to it, say $\{v_{n-l_0+1}, \ldots, v_n\}$. By the above remark we can then find differentials $\omega_{n-l_0+1}, \ldots, \omega_n \in \Omega(-T-D)$ such that $v_j = \mathrm{Res}_D(\omega_j)$ for all j between $n-l_0+1$ and n. This proves items (i), (ii) and (iii).

It is clear that if $j \leq n-l_0$ and $\mathbf{c} \in C_L(D,A)$, then $\langle \mathbf{c}, \mathrm{Res}_D(\omega_j)\rangle = 0$, since $\mathrm{Res}_D(\omega_j) \in C_\Omega(D,A) = C_L(D,A)^\perp$. This proves item (iv), and the lemma follows. □

Definition 2.6. Let G and $D = P_1 + \cdots + P_n$ be divisors defining a code as usual. Given a differential ω, a function h, and a word $\mathbf{r} = (r_1, \ldots, r_n) \in \mathbb{F}^n$, we define the following *syndrome*:

$$s_{\omega,h}(\mathbf{r}) := \langle \mathbf{r}, \mathrm{Res}_D(h\omega)\rangle,$$

where as before $\langle \cdot, \cdot \rangle$ denotes the standard inner product on $\mathbb{F}^n$.

Remark 2.7. The name syndrome is justified in the following sense. If $\omega \in \Omega(A-D)$, $h \in L(A-G)$, and $\mathbf{c} \in C_L(D,G)$, say $\mathbf{c} = \mathrm{Ev}_D(f)$, then

$$\begin{aligned} s_{\omega,h}(\mathbf{c}) &= \langle \mathrm{Ev}_D(f), \mathrm{Res}_D(h\omega)\rangle \\ &= \sum_{i=1}^{n} f(P_i)\ \mathrm{res}_{P_i}(h\omega) \\ &= \sum_{i=1}^{n} \mathrm{res}_{P_i}(fh\omega) \\ &= 0 \end{aligned} \tag{2.5}$$

where the last equality follows from the residue theorem, since the differential $f\,h\,\omega$ cannot have poles outside $\mathrm{Supp}\,D$.

Proposition 2.8. *Let G, D and A be as above, let $\{h_1, \ldots, h_{l_1}\}$ be a basis of $L(A-G)$, and let $\omega_1, \ldots, \omega_{n-l_0} \in \Omega(A-D)$ be differential forms such that $\{\mathrm{Res}_D(\omega_1), \ldots, \mathrm{Res}_D(\omega_{n-l_0})\}$ is a basis of the code $C_\Omega(D,A)$. Then the system in equation (2.4) is equivalent to the following system:*

$$\begin{pmatrix} s_{\omega_1,h_1}(\mathbf{r}) & \dots & s_{\omega_1,h_{l_1}}(\mathbf{r}) \\ \vdots & & \vdots \\ s_{\omega_{n-l_0},h_1}(\mathbf{r}) & \dots & s_{\omega_{n-l_0},h_{l_1}}(\mathbf{r}) \end{pmatrix} \begin{pmatrix} q_{11} \\ \vdots \\ q_{1l_1} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{2.6}$$

That is, for any solution $(q_{11}, \dots, q_{1l_1})$ of system (2.6) there exists exactly one l_0-tuple $(q_{01}, \dots, q_{0l_0})$ such that $(q_{01}, \dots, q_{0l_0}; q_{11}, \dots, q_{1l_1})$ is a solution of system (2.4), and conversely, given a solution $(q_{01}, \dots, q_{0l_0}; q_{11}, \dots, q_{1l_1})$ of system (2.4), the l_1-tuple $(q_{11}, \dots, q_{1l_1})$ is a solution of system (2.6).

Proof. Let $\omega_1, \dots, \omega_n$ be differentials satisfying the properties in Lemma 2.5. From this basis, we define the matrix $\mathbf{H}$ by putting the i-th row of M equal to $\mathrm{Res}_D(\omega_i)$. We will multiply system (2.4) with $\mathbf{H}$ from the left. Note that $\mathbf{H}$ is regular, implying that the multiplied system has exactly the same solutions as the original one. First we investigate the matrix $\mathbf{H}\,\mathbf{M_A}$. Since $\deg A < n$, we see that $\dim C_L(D, A) = l(A) = l_0$. Hence the matrix $\mathbf{M}_A$ (as well as the matrix $\mathbf{H}\,\mathbf{M}_A$) has rank l_0. On the other hand, according to item 4 in Lemma 2.5, the first $n - l_0$ rows of $\mathbf{H}\,\mathbf{M}_A$ are zero. Thus the $l_0 \times l_0$ matrix $\mathbf{B}$ obtained by deleting the first $n - l_0$ rows from $\mathbf{H}\,\mathbf{M}_A$ is regular.

We have now shown that when we multiply system (2.4) from the left by $\mathbf{H}$, we obtain a system of the form:

$$\left(\frac{\mathbf{0}}{\mathbf{B}}\right) \begin{pmatrix} q_{01} \\ \vdots \\ q_{0l_0} \end{pmatrix} + \mathbf{H}\,\mathbf{D_r}\,\mathbf{M}_{A-G} \begin{pmatrix} q_{11} \\ \vdots \\ q_{1l_1} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{2.7}$$

A direct computation shows that the entries of the matrix $\mathbf{H}\,\mathbf{D_r}\,\mathbf{M}_{A-G}$ indeed are syndromes as defined in Definition 2.6. In other words: system (2.6) is nothing but the first $n - l_0$ equations of system (2.7). Since $\mathbf{B}$ is regular, the claim of the proposition now follows. □

We define $\mathbf{S}^{(A)}(\mathbf{r})$ to be the matrix occurring in Proposition 2.8, i.e. we define:

$$\mathbf{S}^{(A)}(\mathbf{r}) := \begin{pmatrix} s_{\omega_1,h_1}(\mathbf{r}) & \dots & s_{\omega_1,h_{l_1}}(\mathbf{r}) \\ \vdots & & \vdots \\ s_{\omega_{n-l_0},h_1}(\mathbf{r}) & \dots & s_{\omega_{n-l_0},h_{l_1}}(\mathbf{r}) \end{pmatrix}. \tag{2.8}$$

Given two matrices $\mathbf{M_1}$ and $\mathbf{M_2}$, we denote by $\mathbf{M_1}|\mathbf{M_2}$ the matrix whose columns are those of $\mathbf{M_1}$ followed by those of $\mathbf{M_2}$. As a bonus of the proof of the previous proposition, we get the following:

Corollary 2.9. *The rank of the matrix* $\mathbf{M}_A|\mathbf{D_r}\mathbf{M}_{A-G}$ *is at most* $l_0 + t$, *were* t *denotes the number of errors in* $\mathbf{r}$.

Proof. In the proof of Proposition we defined a regular matrix H such that

$$\mathbf{H} \cdot (\mathbf{M}_A|\mathbf{D_r}\mathbf{M}_{A-G}) = \left(\begin{array}{c|c} \mathbf{0} & \mathbf{S}^{(A)}(\mathbf{r}) \\ \hline \mathbf{B} & * \end{array}\right).$$

Therefore we see that

$$\operatorname{rank}(\mathbf{M}_A|\mathbf{D_r}\mathbf{M}_{A-G}) = \operatorname{rank}(\mathbf{H} \cdot (\mathbf{M}_A|\mathbf{D_r}\mathbf{M}_{A-G})) = l_0 + \operatorname{rank} \mathbf{S}^{(A)}(\mathbf{r}).$$

Thus it suffices to show that $\operatorname{rank} \mathbf{S}^{(A)}(\mathbf{r}) \leq t$.

Now suppose that $\mathbf{r} = \mathbf{c} + \mathbf{e}$, where $\mathbf{c} \in C_L(D, G)$ and $\mathbf{e}$ an error-vector of Hamming weight $\operatorname{wt}(\mathbf{e}) = t$. From Remark 2.7 we see that $\mathbf{S}^{(A)}(\mathbf{r}) = \mathbf{S}^{(A)}(\mathbf{e})$ and therefore we have that

$$\operatorname{rank} \mathbf{S}^{(A)}(\mathbf{r}) = \operatorname{rank} \mathbf{S}^{(A)}(\mathbf{e}) \leq \operatorname{rank}(\mathbf{H}\mathbf{D_e}\mathbf{M}_{A-G}) \leq \operatorname{rank} \mathbf{D_e} = \operatorname{wt}(\mathbf{e}) = t.$$

□

This corollary enables one to analyse the performance of the basic algorithm in more detail than before.

Proposition 2.10. *Let* $c = \operatorname{Ev}_D(f) \in C_L(D, G)$ *be a codeword and* e *an error-vector of weight* $t < (n - \deg G - g)/2$. *Given the received word* $\mathbf{r} = \mathbf{c} + \mathbf{e}$, *there exists an interpolation polynomial* $Q(y) = Q_0 + Q_1 y$ *and a divisor* A *such that*

(1) $Q_0 \in L(A)$ *and* $Q_1 \in L(A - G)$,
(2) $\deg A < n - t$,
(3) $l(A - G) > t$,
(4) $f = -Q_0/Q_1$.

Proof. The above corollary implies that the number of linearly independent equations in system 2.4 is at most $l_0 + t$. Therefore if $l(A-G) > t$ and $\deg A < n-t$, an interpolation polynomial $Q(y) = Q_0 + Q_1 y$ with the desired properties exists. If $\deg A \geq \deg G + t + g$, then $l(A-G) > t$. It is therefore enough to assume that $\deg A < n - t$ and $\deg A \geq \deg G + t + g$. A divisor A satisfying these two conditions exists as long as $t < (n - \deg G - g)/2$.□

Example 2.11. In this example $\mathbb{F} = \mathbb{F}_{q^2}$, where q is a power of a prime number p. We state some general facts about the Hermitian curve χ defined over $\mathbb{F}$ by the equation

$$x_2^q + x_2 = x_1^{q+1}. \tag{2.9}$$

We actually consider its projective closure, but for convenience we usually work with equation (2.9). First we fix some notation. Given $\alpha, \beta \in \mathbb{F}$ and a point P with $x_1(P) = \alpha$ and $x_2(P) = \beta$, we write $P = P_{\alpha\beta}$. Let $\beta_1, \ldots, \beta_q$ be all solutions to the equation $t^q + t = 0$. Then we define $T_i := P_{0\beta_i}$ for $1 \leq i \leq q$. The projective point $(0 : 1 : 0)$ we denote by T_∞. Note that the points $T_1, \ldots, T_q, T_\infty$ are exactly those points on the Hermitian curve that also lie on the line $x_1 = 0$. All these points are rational.

It is well known that the genus of H is $g = q(q-1)/2$ and that it has $q^3 + 1$ rational points. We denote the $q^3 - q$ rational points different from $T_1, \ldots, T_q, T_\infty$ by $P_1, \ldots, P_{q^3-q}$ and define

$$D := P_1 + \cdots + P_{q^3-q}.$$

Also for any $(q+1)$-tuple $k_\infty, k_1, \ldots, k_q$ of integers we define

$$G(k_\infty, k_1, \ldots, k_q) := k_\infty T_\infty + \sum_{i=1}^{q} k_i T_i.$$

A basis of the space $L(G(k_\infty, k_1, \ldots, k_q))$ can be described as follows: first of all, a generating set for $L(G(k_\infty, k_1, \ldots, k_q))$ is given by the set of all functions $x_1^i \prod_{j=1}^{q} (x_2 - \beta_j)^{e(i,j)}$ satisfying:

- $0 \leq i \leq q$,
- $i + (q+1)e(i,j) \geq -k_j$ for all j with $1 \leq j \leq q$,
- $iq + \sum_{j=1}^{q} e(i,j)(q+1) \leq k_\infty$.

The resulting functions are not linearly independent in general, but this can be achieved in the following way: for each i between 0 and q and each number $d(i)$ between $-\sum_{j=1}^{q} \lfloor (k_j + i)/(q+1) \rfloor$ and $(k_\infty - iq)/(q+1)$, choose (if it exists) exactly one q-tuple $(\, e(i,1), \ldots, e(i,q)\,)$ satisfying the above conditions such that $e(i,1) + \cdots + e(i,q) = d(i)$. The corresponding functions constitute a basis.

For future reference we also note that the differential dx_1 has divisor

$$(dx_1) = (q^2 - q - 2)T_\infty. \tag{2.10}$$

Let $S \subset \mathbb{F}_{q^2}$ and suppose that

$$D = \sum_{\alpha \in S} \sum_{\beta:\, \beta^q + \beta = \alpha^{q+1}} P_{\alpha\beta}.$$

Then we have that

$$\left(\frac{dx_1}{\prod_{\alpha \in S}(x_1 - \alpha)} \right) = -D + (n + 2g - 2)T_\infty.$$

One can use this differential to show that for D as above, we obtain an isomorphism between $\Omega(-D+A)$ and $L(-A+(n+2g-2)T_\infty)$.

Example 2.12. In this example consider the Hermitian curve for $q=4$ and choose the divisor $G=T_1+2T_2+3T_3+4T_4+13T_\infty$. We write $\mathbb{F}_{16}$ as $\mathbb{F}_2[\gamma]$, where $\gamma^4=\gamma+1$. All solutions of $t^4+t=0$ are then given by $\beta_1=0$, $\beta_2=1$, $\beta_3=\gamma^5$, and $\beta_4=\gamma^{10}$. A basis for $L(G)$ is given by

- x_2^α, with $0\leq\alpha\leq 2$,
- $x_1x_2^\alpha/(x_2+\gamma^{10})$, with $0\leq\alpha\leq 2$,
- $x_1^2x_2^\alpha/(x_2^2+x_2+1)$, with $0\leq\alpha\leq 3$,
- $x_1^3x_2^\alpha/(x_2^3+1)$, with $0\leq\alpha\leq 3$, and
- $x_1^4x_2^\alpha/(x_2^4+x_2)$, with $0\leq\alpha\leq 3$.

Now let D be the sum of all 60 rational points not in $\operatorname{Supp}G$. We order the points by writing their coordinates as a power of γ and then ordering theses two exponents lexicographically. In this way we get $P_1=(1,\gamma),\ldots,P_{60}=(\gamma^{14},\gamma^{14})$.

The code $C_L(D,G)$ is an $[60,18,\geq 37]$ code and the basic algorithm can correct $t=15$ errors. Now we choose $A=G+21T_\infty$, since then $\deg A=44<60-15$ and $l(A-G)=l(21T_\infty)=16>15$. To write down system (2.6), we need, according to Proposition 2.8, to calculate a basis for the space $L(A-G)$ and differentials $\omega_1,\ldots,\omega_{21}$ such that their images under the residue map form a basis of the code $C_\Omega(D,A)$. In this case the last part amounts to calculating a basis for $\Omega(-D+A)$. Using the differential form $\omega:=(x_1^{15}+1)^{-1}dx_1$, we see that the spaces $L(-A+70T_\infty)$ and $\Omega(-D+A)$ are isomorphic via $f\mapsto f\omega$. A basis for $L(A-G)$ is given by:

- x_2^α, with $0\leq\alpha\leq 4$,
- $x_1x_2^\alpha$, with $0\leq\alpha\leq 3$,
- $x_1^2x_2^\alpha$, with $0\leq\alpha\leq 2$,
- $x_1^3x_2^\alpha$, with $0\leq\alpha\leq 1$, and
- $x_1^4x_2^\alpha$, with $0\leq\alpha\leq 1$.

For future convenience, we order this basis with respect to the pole-order in T_∞, so that $h_1=1,h_2=x_1,h_3=x_2,\ldots,h_{15}=x_2^4,h_{16}=x_1^4x_2$. A basis for $\Omega(-D+A)$ is given by:

- $(x_2^4 + x_2)x_2^\alpha\omega$, with $0 \le \alpha \le 3$,
- $x_1(x_2^3 + 1)x_2^\alpha\omega$, with $0 \le \alpha \le 3$,
- $x_1^2(x_2^2 + x_2 + 1)x_2^\alpha\omega$, with $0 \le \alpha \le 3$,
- $x_1^3(x_2 + \gamma^{10})x_2^\alpha\omega$, with $0 \le \alpha \le 3$, and
- $x_1^4x_2^\alpha\omega$, with $0 \le \alpha \le 4$.

Again we order this basis with respect to the pole-order in T_∞. We then get $\omega_1 = x_1^4\omega$, $\omega_2 = x_1^3(x_2 - \gamma^{10})\omega, \ldots,$ $\omega_{20} = (x_2^4 + x_2)x_2^3\omega$, $\omega_{21} = x_1^4x_2^4\omega$.

Now we will show an example of error-correction using the basic algorithm. Suppose that the sent codeword is $\mathbf{c} = \mathrm{Ev}_D\,(x_2^2 + x_1^4x_2^3/(x_2^4 + x_2))$ and that the error-vector $\mathbf{e} = (e_1, \ldots, e_{60})$ is given by $e_4 = 1$, $e_8 = \gamma$, $e_9 = \gamma^3$, $e_{16} = \gamma^7$, $e_{18} = \gamma^{11}$, $e_{25} = 1$, $e_{31} = \gamma$, $e_{37} = \gamma^6$, $e_{39} = \gamma^{10}$, $e_{42} = \gamma$, $e_{47} = 1$, $e_{52} = \gamma^{12}$, $e_{55} = \gamma^8$, $e_{58} = 1$, $e_{60} = \gamma^3$, and $e_i = 0$ for all other values of i. Then the matrix $\mathbf{S}^{(A)}(\mathbf{c} + \mathbf{e})$, which is independent of the sent codeword $\mathbf{c}$, is the following:

$$\begin{pmatrix}
\gamma^6 & \gamma^5 & \gamma^{14} & \gamma^{11} & \gamma^6 & \gamma & \gamma^{13} & \gamma & \gamma^6 & \gamma^2 & 0 & \gamma^{12} & 0 & \gamma^2 & \gamma^6 & \gamma^9 \\
\gamma^9 & \gamma^7 & \gamma^5 & \gamma^{13} & \gamma^3 & \gamma^{11} & \gamma^{11} & \gamma^{11} & \gamma^9 & \gamma^7 & \gamma^9 & \gamma^{11} & \gamma^5 & \gamma^4 & \gamma^6 & \gamma^5 \\
\gamma^3 & \gamma^{12} & \gamma^{10} & \gamma^{10} & \gamma^{14} & \gamma^9 & \gamma^5 & \gamma^{12} & \gamma^{14} & \gamma^8 & \gamma^6 & \gamma^2 & \gamma^9 & \gamma^4 & \gamma^3 & \gamma \\
1 & \gamma^{12} & \gamma^{11} & \gamma^5 & \gamma^{13} & \gamma & \gamma^3 & 0 & \gamma^{12} & 0 & \gamma & \gamma^8 & \gamma^9 & \gamma^{13} & 0 & 1 \\
\gamma^5 & \gamma^{11} & \gamma^6 & \gamma^{13} & \gamma & \gamma^6 & 0 & \gamma^{12} & 0 & \gamma^2 & \gamma^8 & \gamma^9 & \gamma^{13} & 0 & \gamma^{13} & \gamma^{10} \\
\gamma^{14} & \gamma^6 & \gamma & \gamma & \gamma^6 & \gamma^2 & \gamma^{12} & 0 & \gamma^2 & \gamma^6 & \gamma^9 & \gamma^{13} & 0 & \gamma^{13} & \gamma^8 & \gamma^7 \\
\gamma^5 & \gamma^3 & \gamma^{11} & \gamma^{11} & \gamma^9 & \gamma^7 & \gamma^{11} & \gamma^5 & \gamma^4 & \gamma^6 & \gamma^5 & 0 & \gamma & \gamma^{10} & \gamma^5 & \gamma^7 \\
\gamma^{10} & \gamma^{14} & \gamma^9 & \gamma^{12} & \gamma^{14} & \gamma^8 & \gamma^2 & \gamma^9 & \gamma^4 & \gamma^3 & \gamma & \gamma^8 & \gamma^{13} & \gamma^3 & \gamma^5 & \gamma^5 \\
\gamma^{11} & \gamma^{13} & \gamma & 0 & \gamma^{12} & 0 & \gamma^8 & \gamma^9 & \gamma^{13} & 0 & 1 & \gamma^{10} & \gamma^7 & \gamma^{11} & \gamma^5 & \gamma^{11} \\
\gamma^6 & \gamma & \gamma^6 & \gamma^{12} & 0 & \gamma^2 & \gamma^9 & \gamma^{13} & 0 & \gamma^{13} & \gamma^{10} & \gamma^7 & \gamma^{11} & \gamma^5 & \gamma & \gamma^{12} \\
\gamma & \gamma^6 & \gamma^2 & 0 & \gamma^2 & \gamma^6 & \gamma^{13} & 0 & \gamma^{13} & \gamma^8 & \gamma^7 & \gamma^{11} & \gamma^5 & \gamma & \gamma^7 & \gamma^4 \\
\gamma^{11} & \gamma^9 & \gamma^7 & \gamma^5 & \gamma^4 & \gamma^6 & 0 & \gamma & \gamma^{10} & \gamma^5 & \gamma^7 & \gamma^5 & \gamma^{10} & \gamma^4 & \gamma^6 & \gamma^{14} \\
\gamma^9 & \gamma^{14} & \gamma^8 & \gamma^9 & \gamma^4 & \gamma^3 & \gamma^8 & \gamma^{13} & \gamma^3 & \gamma^5 & \gamma^5 & \gamma^7 & \gamma & \gamma^7 & \gamma^4 & \gamma^{13} \\
\gamma & \gamma^{12} & 0 & \gamma^9 & \gamma^{13} & 0 & \gamma^{10} & \gamma^7 & \gamma^{11} & \gamma^5 & \gamma^{11} & \gamma^{12} & \gamma^4 & \gamma^8 & \gamma^7 & \gamma^{11} \\
\gamma^6 & 0 & \gamma^2 & \gamma^{13} & 0 & \gamma^{13} & \gamma^7 & \gamma^{11} & \gamma^5 & \gamma & \gamma^{12} & \gamma^4 & \gamma^8 & \gamma^7 & \gamma^9 & \gamma^{13} \\
\gamma^2 & \gamma^2 & \gamma^6 & 0 & \gamma^{13} & \gamma^8 & \gamma^{11} & \gamma^5 & \gamma & \gamma^7 & \gamma^4 & \gamma^8 & \gamma^7 & \gamma^9 & 1 & \gamma^8 \\
\gamma^7 & \gamma^4 & \gamma^6 & \gamma & \gamma^{10} & \gamma^5 & \gamma^5 & \gamma^{10} & \gamma^4 & \gamma^6 & \gamma^{14} & \gamma^9 & \gamma^2 & \gamma^8 & 0 & \gamma^2 \\
\gamma^8 & \gamma^4 & \gamma^3 & \gamma^{13} & \gamma^3 & \gamma^5 & \gamma^7 & \gamma & \gamma^7 & \gamma^4 & \gamma^{13} & \gamma^8 & \gamma^{12} & \gamma^{11} & \gamma^{12} & \gamma^{14} \\
0 & \gamma^{13} & 0 & \gamma^7 & \gamma^{11} & \gamma^5 & \gamma^{12} & \gamma^4 & \gamma^8 & \gamma^7 & \gamma^{11} & \gamma^{13} & \gamma^8 & \gamma^6 & \gamma^2 & \gamma^6 \\
\gamma^2 & 0 & \gamma^{13} & \gamma^{11} & \gamma^5 & \gamma & \gamma^4 & \gamma^8 & \gamma^7 & \gamma^9 & \gamma^{13} & \gamma^8 & \gamma^6 & \gamma^2 & 1 & 1 \\
\gamma^6 & \gamma^{13} & \gamma^8 & \gamma^5 & \gamma & \gamma^7 & \gamma^8 & \gamma^7 & \gamma^9 & 1 & \gamma^8 & \gamma^6 & \gamma^2 & 1 & \gamma^2 & \gamma^6
\end{pmatrix}.$$

One can check that the kernel of this matrix is one-dimensional. A corresponding error-locator is:

$$Q_1 = \gamma^{12}h_2 + h_3 + \gamma^2 h_4 + \gamma^2 h_5 + \gamma^4 h_6 + \gamma^{13}h_7 + \gamma^6 h_8 + \gamma^7 h_9 + \gamma^4 h_{10} + \gamma^3 h_{11} + \gamma^7 h_{12} + \gamma^6 h_{13} + \gamma^{11} h_{14} + \gamma^8 h_{15}.$$

The error-positions i can be found by computing the zeroes P_i of this polynomial. In this case we find that the 15 error-positions are contained in the set $\{4, 8, 9, 12, 16, 18, 19, 21, 25, 31, 37, 39, 42, 47, 48, 52, 55, 58, 60\}$.

Now that the variables $\mathbf{q}_1 = (q_{11}, \ldots, q_{1l_1})$ are known, we can substitute their values into system (2.7). In that way we obtain a system of 39 equations in the 39 variables $\mathbf{q}_0 = (q_{01}, \ldots, q_{0l_0})$. To find these equations we need to choose, as in Lemma 2.5, differentials $\omega_1, \ldots, \omega_{60}$ such that their images under the map Res_D form a basis of $\mathbb{F}_{16}^{60}$. The first 21 are simply the differentials defined above as the basis for $\Omega(-D + A)$. The remaining 39 we choose from $\Omega(-D + A - 45T_\infty)$. We can actually choose them in the following way

- $(x_2^4 + x_2)x_2^\alpha\omega$, with $4 \leq \alpha \leq 11$,
- $x_1(x_2^3 + 1)x_2^\alpha\omega$, with $4 \leq \alpha \leq 11$,
- $x_1^2(x_2^2 + x_2 + 1)x_2^\alpha\omega$, with $4 \leq \alpha \leq 11$,
- $x_1^3(x_2 + \gamma^{10})x_2^\alpha\omega$, with $4 \leq \alpha \leq 11$, and
- $x_1^4 x_2^\alpha\omega$, with $5 \leq \alpha \leq 11$.

Like for the given basis for $\Omega(-D+A)$, we order this basis by increasing pole order at T_∞. Then we get $\omega_{22} = x_1^3(x_2 + \gamma^{10})x_2^4\omega, \ldots, \omega_{60} = (x_2^4 + x_2)x_2^{11}$. We can now calculate the 60×60 matrix $\mathbf{H}$ as well as the vector $\mathbf{v} := \mathbf{H}\mathbf{D}_\mathbf{r}\mathbf{M}_{A-G}\mathbf{q}_1$. The first 21 coordinates of $\mathbf{v}$ are 0, since $\mathbf{q}_1$ is in the kernel of $\mathbf{S}^{(A)}(\mathbf{r})$. The remaining 39 coordinates of this vector are given by:

$$\begin{array}{llllll}
v_{22} = 0, & v_{23} = 0, & v_{24} = 0, & v_{25} = \gamma^8, & v_{26} = \gamma^7, & v_{27} = \gamma, \\
v_{28} = \gamma^{10}, & v_{29} = \gamma^4, & v_{30} = \gamma^7, & v_{31} = \gamma^3, & v_{32} = \gamma^{14}, & v_{33} = \gamma^5, \\
v_{34} = \gamma^{13}, & v_{35} = \gamma^4, & v_{36} = \gamma^{10}, & v_{37} = \gamma^5, & v_{38} = \gamma, & v_{39} = 0, \\
v_{40} = \gamma^2, & v_{41} = \gamma^8, & v_{42} = \gamma^{13}, & v_{43} = \gamma, & v_{44} = 0, & v_{45} = \gamma^4, \\
v_{46} = \gamma^3, & v_{47} = \gamma, & v_{48} = \gamma, & v_{49} = 0, & v_{50} = \gamma^4, & v_{51} = \gamma^{10}, \\
v_{52} = \gamma^5, & v_{53} = \gamma, & v_{54} = 0, & v_{55} = 1, & v_{56} = \gamma^{11}, & v_{57} = \gamma^{12}, \\
v_{58} = \gamma^8, & v_{59} = \gamma^4, & v_{60} = \gamma^3. & & &
\end{array}$$

We now choose the following basis for $L(A)$:

- x_2^α, with $0 \leq \alpha \leq 6$,
- $x_1 x_2^\alpha/(x_2 + \gamma^{10})$, with $0 \leq \alpha \leq 7$,
- $x_1^2 x_2^\alpha/(x_2^2 + x_2 + 1)$, with $0 \leq \alpha \leq 7$,
- $x_1^3 x_2^\alpha/(x_2^3 + 1)$, with $0 \leq \alpha \leq 7$, and
- $x_1^4 x_2^\alpha/(x_2^4 + x_2)$, with $0 \leq \alpha \leq 7$,

and order it with increasing pole order in T_∞. Then $g_1 = x_1^4/(x_2^4 + x_2)$, $g_2 = x_1^3/(x_2^3 + 1), \ldots, g_{39} = x_1 x_2^7/(x_2 + \gamma^{10})$. We can then calculate the matrix $\mathbf{B}$ from the proof of Proposition 2.8. By the way we have chosen and ordered the differentials and functions, we obtain more structure than was indicated in Proposition 2.8. In this case we obtain that

$$\mathbf{B}_{ij} = \begin{cases} 1 & \text{if } i + j = 40 \text{ or } i + j = 55, \\ 0 & \text{otherwise.} \end{cases}$$

This means that is straightforward to calculate Q_0 now and we obtain

$$\begin{aligned} Q_0 = \ & \gamma^{13} g_{12} + \gamma^2 g_{13} + \gamma^7 g_{14} + \gamma^3 g_{16} + \gamma^4 g_{17} + \gamma^8 g_{18} + \gamma^{12} g_{19} + \gamma^{11} g_{20} + \\ & g_{21} + \gamma g_{23} + \gamma^5 g_{24} + \gamma^{10} g_{25} + \gamma^4 g_{26} + \gamma^{13} g_{27} + \gamma^5 g_{28} + \gamma^{14} g_{29} + \gamma^3 g_{30} + \\ & \gamma^7 g_{31} + \gamma^4 g_{32} + \gamma^{10} g_{33} + \gamma g_{34} + \gamma^7 g_{35} + \gamma^8 g_{36}. \end{aligned}$$

Note that $Q_0/Q_1 = x_2^2 + x_1^4 x_2^3/(x_2^4 + x_2)$.

2.4. The generalized order bound

The advantage of the codes $C_L(D, G)$ and $C_\Omega(D, G)$ is an a priori lower bound on the minimum distance d. In case of $C_L(D, G)$ we know that $d \geq n - \deg G$, while for $C_\Omega(D, G)$ we know that $d \geq \deg G - 2g + 2$. These bounds are known as the Goppa-bounds. Though good in general, it is clear that if $\deg G \leq 2g - 2$ the bound $d \geq \deg G - 2g + 2$ is trivial, while if $\deg G \geq n$, the bound $d \geq n - \deg G$ lower bound is trivial. We will see that there exist a bound that improves the Goppa-bounds in the mentioned cases, but sometimes also if $2g - 2 < \deg G < n$. In this section we will show how to obtain this lower bound, called the generalized order bound.

Let $T \notin \operatorname{Supp} D$ be a rational point. We then define the ring

$$R(T) := \bigcup_{i \geq 0} L(iT). \tag{2.11}$$

There is a natural mapping ρ_T from $R(T)\backslash\{0\}$ to $\mathbb{N} = \{0, 1, 2, \dots\}$, namely

$$\begin{aligned} \rho_T : R(T)\backslash\{0\} &\to \mathbb{N} \\ f &\mapsto -v_T(f). \end{aligned} \tag{2.12}$$

The image $H(T)$ of this map is the so-called Weierstrass semigroup of T:

$$H(T) := \rho_T(R(T)\backslash\{0\}) \tag{2.13}$$

We will now define a certain type of $R(T)$-modules called order modules that will be useful when computing lowwer bounds on the minimum distance of algebraic geometry codes.

Definition 2.13. An *order module* $\mathcal{M}$ for $R(T)$ is a pair (M, φ), where M is an $R(T)$-module and φ a surjective $\mathbb{F}$-linear map $\varphi : M \to \mathbb{F}^n$ such that:

(1) $M = \bigcup_{i\in\mathbb{Z}} M_i$, with $M_i \subset M$ vector spaces such that for all integers $i \leq j$ we have that $M_i \subset M_j$,
(2) There exists an integer a such that $M_i = \{0\}$ for all $i < a$,
(3) For any integers i and j, we have that $L(iT)M_j \subset M_{i+j}$,
(4) For $f \in R(T)$ and $m \in M$ we have that $\varphi(fm) = \mathrm{Ev}_D(f) * \varphi(m)$, where $*$ denotes the coordinate-wise product on $\mathbb{F}^n$,
(5) For $m \in M_i\backslash M_{i-1}$ and $f \in R(T)$ satisfying $\rho_T(f) = j$, we have that $fm \in M_{i+j}\backslash M_{i+j-1}$,
(6) For all i, we have that $M_i = M_{i-1}$ or $\dim M_i = \dim M_{i-1} + 1$.

Remark 2.14. An analogue of the map ρ_T can be defined on $\mathcal{M}$ as follows:

$$\begin{aligned} \rho_{T,\mathcal{M}} : M\backslash\{0\} &\to \mathbb{Z} \\ m &\mapsto \min\{i \,|\, m \in M_i\}. \end{aligned} \tag{2.14}$$

Item 5 of Definition 2.13 is then equivalent to:

(5a) For any $f \in R(T)\backslash\{0\}$ and $m \in M\backslash\{0\}$ we have that

$$\rho_{T,M}(fm) = \rho_T(f) + \rho_{T,M}(m).$$

The linear subspaces $\varphi(M_i) \subset \mathbb{F}^n$ are interpreted as codes. Examples of order modules are

$$\mathcal{M}_L(D, G, T) := (\cup_{i\in\mathbb{Z}} L(G + iT), \mathrm{Ev}_D) \tag{2.15}$$

and

$$\mathcal{M}_\Omega(D,G,T) := (\cup_{i\in\mathbb{Z}}\Omega(-D+G-iT), \mathrm{Res}_D\,). \tag{2.16}$$

In the first case, we have that $\rho_{T,\mathcal{M}}(m) = -v_T(m) - v_T(G)$, while the corresponding codes are the codes $C_L(D, G+iT)$. In the second example we have that $\rho_{T,\mathcal{M}}(m) = -v_T(m) + v_T(G)$, while we now obtain the codes $C_\Omega(D, G-iT)$.

Remark 2.15. The codes coming from $\mathcal{M}_\Omega(D,G,T)$ are the same as those from $\mathcal{M}_L(D, K+D-G, T)$, where $K = (\omega)$ is the divisor of a differential ω that has poles of order one and residues equal to one in all points of $\mathrm{Supp}\, D$. If one wishes, we can therefore reduce computations in the module $\mathcal{M}_\Omega(D,G,T)$ to ones in $\mathcal{M}_L(D, K+D-G, T)$.

The analogue of the set $H(T)$ for an order module $\mathcal{M} = (M, \varphi)$ is:

$$H(T,\mathcal{M}) := \rho_{T,\mathcal{M}}(M\backslash\{0\}). \tag{2.17}$$

Note that this set is not a semigroup in general, but it does have the property that $i \in H(T,\mathcal{M})$ implies that $i + H(T) \subset H(T,\mathcal{M})$. An element from $\mathbb{N}\backslash H(T)$ is called a gap of the semigroup $H(T)$. It is well known that the number of gaps equals the genus g of the curve. We will define the analogue concepts for $H(T,\mathcal{M})$.

Definition 2.16. Let $a = \min H(T,\mathcal{M})$. The set $\mathbb{Z}_{\geq a}\backslash H(T,\mathcal{M})$ is called the set of *gaps* of $H(T,\mathcal{M})$. We denote the number of gaps by $g(\mathcal{M})$.

Remark 2.17. Since $a + H(T) \subset H(T,\mathcal{M})$, we always have $g(\mathcal{M}) \leq g$. Using Riemann-Roch's theorem, we find that in case $\mathcal{M} = \mathcal{M}_L(D,G,T)$, then $a = -\mathrm{deg}G + g - g(\mathcal{M})$. Similarly, if $\mathcal{M} = \mathcal{M}_\Omega(D,G,T)$, then we get that $a = -n + \mathrm{deg}G - g - g(\mathcal{M}) + 2$.

We now define the set

$$N(T,\mathcal{M},i) := \{(i_1,i_2)\,|\,i_1 \in H(T); i_2 \in H(T,\mathcal{M}); i_1 + i_2 = i+1\} \tag{2.18}$$

and its cardinality

$$\nu(T,\mathcal{M},i) := \#N(T,\mathcal{M},i). \tag{2.19}$$

Lemma 2.18. *Let* $p_T(t) := \sum_{i_1\in H(T)} t^{i_1}$ *and* $p_{T,\mathcal{M}}(t) := \sum_{i_2\in H(T,\mathcal{M})} t^{i_2}$. *Then* $\nu(T,\mathcal{M},i)$ *is the coefficient of* t^{i+1} *in* $p_T(t)p_{T,\mathcal{M}}(t)$.

Proof. This follows directly from the definition of $\nu(T,\mathcal{M},i)$. □

We can use this interpretation to give a lower bound on $\nu(T, \mathcal{M}, i)$. We do this in the following lemma.

Lemma 2.19. *Let $\mathcal{M}$ be an order module and let $a = \min H(T, \mathcal{M})$. Then $\nu(T, \mathcal{M}, i) \geq i - a + 2 - g - g(\mathcal{M})$.*

Proof. We can choose polynomials $q_T(t)$ and $q_{T,\mathcal{M}}(t)$ such that the following identities of Laurent series hold:

$$p_T(t) + q_T(t) = \frac{1}{1-t}$$

and

$$p_{T,\mathcal{M}}(t) + q_{T,\mathcal{M}}(t) = \frac{t^a}{1-t}.$$

Moreover, $q_T(t)$ is the sum of precisely g monomials, and $q_{T,\mathcal{M}}(t)$ of $g(\mathcal{M})$ monomials. These monomials all have coefficient 1. The above implies that

$$p_T(t)p_{T,\mathcal{M}}(t) = t^a \frac{1}{(1-t)^2} - \frac{t^a q_T(t) + q_{T,\mathcal{M}}(t)}{1-t} + q_T(t)q_{T,\mathcal{M}}(t).$$

Considering this as a Laurent series in t, we can compute the coefficient of t^{i+1}. The term $t^a/(1-t)^2$ contributes exactly with $i - a + 2$ to this coefficient, the term $-(t^a q_T(t) + q_{T,\mathcal{M}}(t))/(1-t)$ with at least $-g - g(\mathcal{M})$ and the term $q_T(t)q_{T,\mathcal{M}}(t)$ with a nonnegative number. All in all we get that the coefficient of t^{i+1} in $p_T(t)p_{T,\mathcal{M}}(t)$ is at least $i - a + 2 - g - g(\mathcal{M})$. The lemma now follows from Lemma 2.18. □

Given an order module $\mathcal{M} = (\cup_i M_i, \varphi)$, we can shift the order module by s as follows: $\mathcal{M}_{+s} = (\cup_i M_{i+s}, \varphi)$. Then $\nu(T, \mathcal{M}_{+s}, i) = \nu(T, \mathcal{M}, i + s)$ implying that $\nu(T, \mathcal{M}, s) = \nu(T, \mathcal{M}_{+s}, 0)$. Therefore it will be practical to simplify our notation when $i = 0$ by defining:

$$N(T, \mathcal{M}) := N(T, \mathcal{M}, 0)$$

and

$$\nu(T, \mathcal{M}) := \nu(T, \mathcal{M}, 0).$$

We now have the necessary notation to formulate the following proposition that is essential in order to obtain lower bounds on the minimum distance of codes coming from order modules.

Proposition 2.20. *Let $\mathcal{M} = (M, \varphi)$ be an order module for $R(T)$ and let $\mathbf{c} \in \varphi(M_i)^{\perp} \backslash \varphi(M_{i+1})^{\perp}$. Then $\mathrm{wt}(\mathbf{c}) \geq \nu(T, \mathcal{M}, i)$, with $\mathrm{wt}(\mathbf{c})$ the Hamming weight of $\mathbf{c}$.*

Proof. Let $\mathbf{c} = (c_1, \ldots, c_n) \in \varphi(M_i)^\perp \backslash \varphi(M_{i+1})^\perp$. We denote by $\mathbf{D_c}$ the diagonal matrix with $c_1, \ldots, c_n$ on its diagonal.

Let $H(T) = \{\rho_1, \rho_2, \ldots\}$, such that $\rho_k < \rho_l$ if $k < l$. For every $\rho_k \in H(T)$ we choose a function $f_k \in R(T)$ such that $\rho_T(f_k) = \rho_k$. Further we define $v_k := \mathrm{Ev}_D(f_k)$. Let N be a natural number such that $\mathrm{Ev}_D(L(NT)) = \mathbb{F}^n$ and $N > \max\{k \,|\, (\rho_k, l) \in N(T, \mathcal{M}, i)\}$. Then we let $\mathbf{H}_1$ be the $N \times n$ matrix whose k-th row is $\mathrm{Ev}_D(f_k)$ for $1 \leq k \leq N$. By choice of N, we have that $\operatorname{rank} \mathbf{H}_1 = n$.

By item 2 in Definition 2.13, there exists an integer N_1 such that $M_{N_1} = 0$. On the other hand, since φ is assumed to be a surjective linear map to $\mathbb{F}^n$, there exists an N_2 such that $\varphi(M_{N_2}) = \mathbb{F}^n$ and $N_2 > \max\{l \,|\, (\rho_k, l) \in N(T, \mathcal{M}, i)\}$. The set $H(T, \mathcal{M}) \cap [N_1, N_2]$ consists of finitely many integers, say $s_1, \ldots, s_L$. Then we can choose $m_k \in M_{s_k} \backslash M_{s_k - 1}$. By the choice of the m_k we see that $\rho_{T,\mathcal{M}}(m_k) < \rho_{T,\mathcal{M}}(m_l)$ if $k < l$. Now we define $h_k := \varphi(m_k)$ and $\mathbf{H}_2$ the $L \times n$ matrix with h_k as k-th row. By our choice of N_1, N_2 and by item 5 in Definition 2.13, we have that $\operatorname{rank} \mathbf{H}_2 = n$.

Consider the matrix $\mathbf{S}(\mathbf{c}) := \mathbf{H}_1 \mathbf{D_c} \mathbf{H}_2^t$. Since $\mathbf{H}_1$ and $\mathbf{H}_2$ have full rank, we see that $\operatorname{rank} \mathbf{S}(\mathbf{c}) = \mathrm{wt}(\mathbf{c})$. On the other hand we will show that $\operatorname{rank} \mathbf{S}(\mathbf{c}) \geq \nu(T, \mathcal{M}, i)$. Note that

$$\mathbf{S}(\mathbf{c})_{ij} = \sum_{\lambda=1}^{n} f_i(P_\lambda) c_\lambda \varphi(m_j)_\lambda = \sum_{\lambda=1}^{n} c_\lambda \varphi(f_i m_j)_\lambda = \langle \mathbf{c}, \varphi(f_i m_j) \rangle. \quad (2.20)$$

Let $(\rho_i, j) \in N(T, \mathcal{M}, i)$. By our choice of N we have that $i \leq N$ and therefore v_i occurs as a row in H_1. Similarly h_j occurs as a row in H_2. Now let $t := \nu(T, \mathcal{M}, i)$ and suppose that

$$N(T, \mathcal{M}, i) = \{(\rho_{i_1}, j_t), (\rho_{i_2}, j_{t-1}), \ldots, (\rho_{i_t}, j_1)\}.$$

For convenience, we define $\sigma_k := \rho_{i_k}$. Without loss of generality we can assume that $i_1 < i_2 < \cdots < i_t$. This implies that $j_1 < j_2 < \cdots < j_t$, since if both $k < l$ and $j_k > j_l$, then

$$i + 1 = \sigma_{t+1-l} + j_l < \sigma_{t+1-k} + j_l < \sigma_{t+1-k} + j_k = i + 1.$$

Let $\mathbf{H}$ be the $t \times t$ matrix obtained from $\mathbf{S}(\mathbf{c})$ by choosing all those entries $\mathbf{S}(\mathbf{c})_{ij}$ with $i \in \{i_1, \ldots, i_t\}$ and $j \in \{j_1, \ldots, j_t\}$. Clearly $\operatorname{rank} \mathbf{S}(\mathbf{c}) \geq \operatorname{rank} \mathbf{H}$, so the proposition follows if we show that $\mathbf{H}$ has full rank. Suppose that $k+l < t+1$. Then $\varphi(f_{i_k} m_{j_l}) \in \varphi(M_i)$, since $\rho_{T,\mathcal{M}}(f_{i_k} m_{j_l}) = \rho_T(f_{i_k}) + \rho_{T,\mathcal{M}}(m_{j_l}) = \sigma_k + j_l < \sigma_k + j_{t+1-k} = i + 1$. But this implies by equation (2.20) that

$$\mathbf{S}(\mathbf{c})_{i_k j_l} = \langle \mathbf{c}, \varphi(f_{i_k} m_{j_l}) \rangle = 0.$$

On the other hand, if $k+l=t+1$, then a similar computation shows that $\varphi(f_{i_k}m_{j_l}) \in \varphi(M_{i+1})$ and that $\mathbf{S}(\mathbf{c})_{i_k j_l} \neq 0$. This means that $\mathbf{H}$ is of the form

$$\mathbf{H} = \begin{pmatrix} 0 & & * \\ & \iddots & \\ * & & \end{pmatrix},$$

where a $*$ denotes a nonzero element of $\mathbb{F}$. Thus $\operatorname{rank}\mathbf{H} = t$. □

When using the above proposition, one needs to choose an order module. For example for the code $C_L(D,G)$ we could choose the module $\mathcal{M}_\Omega(D,G,T)$ and for the code $C_\Omega(D,G)$, we can use the module $\mathcal{M}_L(D,G,T)$.

Now we describe the generalized order bound. Let $D = P_1 + \cdots + P_n$ as usual and G a divisor such that $\operatorname{Supp} G \cap \operatorname{Supp} D = \varnothing$. Suppose that the set $\{T_1, T_2, \dots, \}$ consists of rational points that do not occur in $\operatorname{Supp} D$. Now let $S = (S_1, S_2, \dots)$ be a sequence of points, each of which is contained in $\{T_1, T_2, \dots, \}$. We also recursively define the divisors $G_0 := G$, $G_{i+1} := G_i + S_{i+1}$, $H_0 := G$, $H_{i+1} := H_i - S_{i+1}$ and modules

$$\mathcal{M}_S(i) := \mathcal{M}_\Omega(D, H_i, S_{i+1})$$

and

$$\mathcal{M}_S^\perp(i) := \mathcal{M}_L(D, G_i, S_{i+1}).$$

With this notation in mind, we can then define:

Definition 2.21.

$$d_S(G) := \min_i\{\nu(S_{i+1}, \mathcal{M}_S(i))\}$$

and

$$d_S^\perp(G) \geq \min_i\{\nu(S_{i+1}, \mathcal{M}_S^\perp(i))\}.$$

In the first (respectively second) case the minimum is taken over all i such that $i \geq 0$ and $C_L(D,H_i) \neq C_L(D,H_{i+1})$ (respectively $C_\Omega(D,G_i) \neq C_\Omega(D,G_{i+1})$).

With this notation we get the following theorem:

Theorem 2.22. *Let $\{T_1, T_2, \dots\}$ be a set of rational points not occurring in $\operatorname{Supp} D$ and let $S = (S_1, S_2, \dots)$ be a sequence of points, each of which*

is contained in the set $\{T_1, T_2, \ldots, \}$. *Then the minimum distance* d *of the code* $C_L(D, G)$ *satisfies*

$$d \geq d_S(G),$$

while the minimum distance $d^\perp$ *of the code* $C_\Omega(D, G)$ *satisfies*

$$d^\perp \geq d_S^\perp(G).$$

Proof. We will prove the statements about the code $C_L(D, G)$. The results for the code $C_\Omega(D, G)$ can be proved similarly.

Recall that $\nu(T, \mathcal{M}) := \nu(T, \mathcal{M}, 0)$. We can write $C_L(D, G)$ as the disjoint union $\cup_{i \geq 0} C_L(D, H_i) \backslash C_L(D, H_{i+1})$. If $C_L(D, H_i) \neq C_L(D, H_{i+1})$ and $\mathbf{c} \in C_L(D, H_i) \backslash C_L(D, H_{i+1})$, then from Proposition 2.20 we see that $\mathrm{wt}(\mathbf{c}) \geq \nu(S_{i+1}, \mathcal{M}_S(i))$. Then it follows that $d \geq \min_i \{\nu(S_{i+1}, \mathcal{M}_S(i))\}$, if we take the minimum over all nonnegative i such that $C_L(D, H_i) \neq C_L(D, H_{i+1})$. □

The original Goppa bounds now follow as a corollary, showing that the generalized order bound is always at least as good.

Corollary 2.23. *The minimum distance* d *of the code* $C_L(D, G)$ *satisfies*

$$d \geq n - \deg G,$$

while the minimum distance $d^\perp$ *of the code* $C_\Omega(D, G)$ *satisfies*

$$d^\perp \geq \deg G - 2g + 2.$$

Proof. Recall that $\mathcal{M}_S(i) = \mathcal{M}_\Omega(D, H_i, S_{i+1})$ and $H_i = G - S_0 - \cdots - S_i$. Remark 2.17 and Lemma 2.19 imply that $\nu(S_{i+1}, \mathcal{M}_S(i)\,) \geq n - \deg G + i \geq n - \deg G$. Therefore $d \geq d_S(G) \geq n - \deg G$. Similarly, we have that $\nu(S_{i+1}, \mathcal{M}_S^\perp(i)\,) \geq \deg G + i - 2g + 2 \geq \deg G - 2g + 2$, which implies that $d^\perp \geq d_S^\perp(G) \geq \deg G - 2g + 2$. □

Example 2.24. In this example we will study a code coming from the Hermitian curve defined over $\mathbb{F}_{64}$ by the equation $x_2^8 + x_2 = x_1^9$. This curve has 513 rational points, exactly one of which has a pole in x_1 and x_2. As usual, we denote this point by T_∞. We denote by T_0 the unique point having a zero in both x_1 and x_2. Further, we denote by D the sum of the 504 rational points P satisfying $x1(P) \neq 0$.

In this example we will consider the code $C_L(D, -T_0 + 490T_\infty)$. This is a $[504, 462, \geq 15]$ code, since $l(-T_0 + 490T_\infty) = 462$ and the Goppa bound gives that the minimum distance is at least $504 - 489 = 15$. We will show

that the Goppa bound is not sharp in this case and show that the minimum distance is at least 21.

We wish to use Theorem 2.22 to get a lower bound on the minimum distance of the code $C_L(D, -T_0 + 490T_\infty)$. First we need to choose a sequence S, which we take to be $S := (T_\infty, T_0, T_0, T_0, \dots)$ in this example. We will compute the quantity $d_S(-T_0 + 490T_\infty)$. In order to do so we will work in the modules ${\mathcal{M}^{(i)}}_\Omega(S)$. The first module we need to work in is therefore $\mathcal{M}_S(0) = \mathcal{M}_\Omega(D, -T_0 + 490T_\infty, T_\infty)$. We start by calculating $H(T_\infty, \mathcal{M}_S(0))$.

We will need to know what $\rho_{T_\infty}(\Omega(-D - T_0 + 490T_\infty))$ is. The Weierstrass semigroup $H(T_\infty)$ is as a semigroup generated by 8 and 9. Explicitly, we have that $H(T_\infty) = \langle 8, 9\rangle = \{0, 8, 9, 16, 17, 18, 24, \dots\}$. Moreover it holds that $H(T) = H(T_\infty)$ for any rational point T. This means that the Laurent series

$$p(t) := \sum_{i \in \langle 8,9\rangle} t^i \tag{2.21}$$

will play a central role in the evaluation of the generalized order bound.

For any order modules and for any $m \in M_i \backslash M_{i-1}$ we have $\rho_{T,\mathcal{M}}(m) = i$. We see that for $m \in \Omega(-D - T_0 + (490 - i)T_\infty) \backslash \Omega(-D - T_0 + (491 - i)T_\infty)$ we have $\rho_{T_\infty, \mathcal{M}_S(0)}(m) = \rho_{T_\infty}(m) + 490$. Further, using the differential $\omega = (x_1^{63} + 1)^{-1} dx_1$, we see that

$$\rho_{T_\infty}(\Omega(-D - T_0 + (490 - i)T_\infty)) = \{-558 + s \,|\, s \in \rho_{T_\infty}(L(T_0 + (68 + i)T_\infty)).$$

Using the description of L-spaces in Example 2.11, we see that

$$\bigcup_{i \in \mathbb{Z}} \rho_{T_\infty}(L(T_0 + (68 + i)T_\infty)) = H(T_\infty) \cup \{55\}.$$

Putting everything together, we find that

$$H(T_\infty, \mathcal{M}_S(0)) = \{s - 68 \,|\, s \in H(T_\infty)\} \cup \{-13\}.$$

Therefore

$$p_{T_\infty, \mathcal{M}_S(0)}(t) = t^{-13} + t^{-68} p(t)$$

Using equation (2.21), we can now calculate that

$$p(t) p_{T_\infty, \mathcal{M}_S(0)}(t) = \cdots + 24t + 21t^2 + 17t^3 + \cdots,$$

and therefore (see Lemma 2.18):

$$\nu(T_\infty, \mathcal{M}_S(0)) = 24.$$

For the next step we need to know the set $H(T_0, \mathcal{M}_S(1))$. Note that $H(T_0) = H(T_\infty)$. We will calculate $\rho_{T_0}(L((1+i)T_0+69T_\infty))$. Using the fact that $(x_2) = 9(T_0 - T_\infty)$, we see that $\rho_{T_0}(L((1+i)T_0+69T_\infty)) = \{s-63 \,|\, s \in \rho_{T_0}(L((64+i)T_0+6T_\infty))$. The automorphism τ defined by $\tau(x) = x_1/x_2$ and $\tau(y) = 1/x_2$, interchanges the points T_0 and T_∞. Using this automorphism, we can conclude that $\rho_{T_0}(L((64+i)T_0+6T_\infty)) = \rho_{T_\infty}(L((64+i)T_\infty+6T_0))$. Similarly as above we now find that

$$H(T_0, \mathcal{M}_S(1)) = \{s - 64 \,|\, s \in H(T_0)\} \cup \{-49, -41, -33, -25, -17, -9\}.$$

This implies that

$$p_{T_0, \mathcal{M}_S(1)}(t) = t^{-49} + t^{-41} + t^{-33} + t^{-25} + t^{-17} + t^{-9} + t^{-64}p(t),$$

enabling us to calculate that

$$p(t)p_{T_0, \mathcal{M}_S(1)}(t) = \cdots + 21t + 25t^2 + 27t^3 + 27t^4 + 25t^5 + \cdots . \qquad (2.22)$$

Hence $\nu(T_0, \mathcal{M}_S(1)) = 21$. Since the sequence S only contains T_0 apart from the very first point in the sequence, we now can remain working with the module $\mathcal{M}_S(1)$. For $i \geq 0$, we can see the module $\mathcal{M}_S(i+1)$ as the i-th shift of $\mathcal{M}_S(1)$. More precisely, we have that $\nu(T_0, \mathcal{M}_S(i+1)) = \nu(T_0, \mathcal{M}_S(1), i)$. This means that with the above computation of $H(T_0, \mathcal{M}_S(1))$, we have all information we need to calculate $d_S(-T_0 + 490T_\infty)$. More specifically, we see from equation (2.22) that $\nu(T_0, \mathcal{M}_S(2)) = \nu(T_0, \mathcal{M}_S(5)) = 25$ and $\nu(T_0, \mathcal{M}_S(3)) = \nu(T_0, \mathcal{M}_S(4)) = 27$. For $i \geq 6$, we can use Lemma 2.19 to show that $\nu(T_0, \mathcal{M}_S(i)) \geq 15 + i \geq 21$.

All in all, we have shown that $d_S(-T_0 + 490T_\infty) = 21$.

2.5. Majority voting

Given a code $C_L(D, G)$, we have seen that the basic algorithm is able to correct $\lfloor (n - \deg G - 1 - g)/2 \rfloor$ errors. This means that the full potential of the code has not been used yet. In this section we will describe an algorithm that can correct $\lfloor (d_S(G) - 1)/2 \rfloor$ errors, where $d_S(G)$ denotes the generalized order bound from Section 2.4. This main technique is that of majority voting for so-called unknown syndromes. Loosely speaking this technique enables one to obtain more information about the error-vector enabling one to correct more errors than with the basic algorithm.

Let $\mathbf{r} = \mathbf{c} + \mathbf{e}$. The $(n - l_0) \times l_1$ matrix $\mathbf{S}^{(A)}(\mathbf{r})$ plays a central role in the proof of the fact that the basic algorithm can correct $\lfloor (n - \deg G - 1 - g)/2 \rfloor$ errors. The reason is that if $\mathbf{c} \in C_L(D, G)$, then $\mathbf{S}^{(A)}(\mathbf{c}) = \mathbf{0}$ which implies

that $\mathbf{S}^{(A)}(\mathbf{c}) = \mathbf{S}^{(A)}(\mathbf{e})$. The matrix $\mathbf{S}^{(A)}(\mathbf{r})$ therefore gives information about the error-vector $\mathbf{e}$. More precisely, we have seen in Proposition 2.8 that its kernel determines the error-locator Q_1.

Definition 2.25. If ω and h are such that $h\omega \notin \Omega(-D+G)$, then the syndrome $s_{\omega,h}(\mathbf{r})$ will in general depend both on $\mathbf{c}$ and $\mathbf{e}$. Such a syndrome it said to be *unknown*.

Also we define the following syndrome:

Definition 2.26. Let ω be a differential form. Then we define

$$s_\omega(\mathbf{r}) := s_{\omega,1}(\mathbf{r}).$$

Let $T \notin \operatorname{Supp} G$ be a rational point. For now let us assume that $A = G + aT$. We can do this, since the only restrictions on A were that $\deg A < n - t$ and $l(A-G) > t$. If $t + g - 1 < a < n - t - \deg G$ both conditions are guaranteed to hold. It will be convenient to extend the matrix $\mathbf{S}^{(A)}(\mathbf{r})$ in this setup. The matrix $\mathbf{S}^{(A)}(\mathbf{r})$ itself depends on the choice of functions and differentials from $L(A-G)$ and $\Omega(A-D)$ (see Proposition 2.8). We now specify a more precise choice: let $H(T) = \{\rho_1, \rho_2, \dots\}$ and $h_1, h_2, \dots \in R(T)$ such that $\rho_T(h_i) = \rho_i$. Similarly, let $\mathcal{M} := \mathcal{M}_\Omega(D, G, T)$ and $H(T, \mathcal{M}) = \{\sigma_1, \sigma_2, \dots\}$. We can then choose differential forms $\omega_1, \omega_2, \dots \in \cup_i \Omega(-D + G - iT)$ such that $\rho_{T,\mathcal{M}}(\omega_j) = \sigma_j$. We then define the following matrices:

Definition 2.27.

$$\mathbf{S}_T^{tot}(\mathbf{r}) := \begin{pmatrix} s_{\omega_1,h_1}(\mathbf{r}) & s_{\omega_1,h_2}(\mathbf{r}) & \dots \\ s_{\omega_2,h_1}(\mathbf{r}) & s_{\omega_2,h_2}(\mathbf{r}) & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

and

$$\mathbf{S}_T^{tot}(\mathbf{r})|_{i,j} := \begin{pmatrix} s_{\omega_1,h_1}(\mathbf{r}) & \dots & s_{\omega_1,h_i}(\mathbf{r}) \\ \vdots & & \vdots \\ s_{\omega_j,h_1}(\mathbf{r}) & \dots & s_{\omega_j,h_i}(\mathbf{r}) \end{pmatrix}.$$

The matrix $\mathbf{S}_T^{tot}(\mathbf{r})$ extends the matrix $\mathbf{S}^{(A)}(\mathbf{r})$ in equation (2.8) in the case that $A = G + aT$. Note that $h_i\omega_j \in \Omega(-D+G-(\rho_i+\sigma_j)T)$. Therefore we have that all elements $s_{\omega_j,h_i}(\mathbf{r})$ of $\mathbf{S}_T^{tot}(\mathbf{r})$ such that $\rho_i + \sigma_j \leq 0$, are known syndromes, i.e. equal to $s_{\omega_j,h_i}(\mathbf{e})$.

Before proceeding, we need some terminology:

Definition 2.28. A position (i,j) in the matrix $\mathbf{S}_T^{tot}(\mathbf{e})$ is said to be a *candidate*, if the matrices $\mathbf{S}_T^{tot}(\mathbf{e})|_{i-1,j-1}$, $\mathbf{S}_T^{tot}(\mathbf{e})|_{i-1,j}$, and $\mathbf{S}_T^{tot}(\mathbf{e})|_{i,j-1}$ all have the same rank. If furthermore the matrices $\mathbf{S}_T^{tot}(\mathbf{e})|_{i-1,j-1}$ and $\mathbf{S}_T^{tot}(\mathbf{e})|_{i,j}$ do not have equal rank, then the position (i,j) is called a *discrepancy*.

Now suppose that $\mathbf{r} = \mathbf{c} + \mathbf{e}$, with $\mathbf{c} \in C_L(D,G)$ and that we are given a candidate (i,j) with $\rho_i + \sigma_j = 1$. We can determine these candidates, since the part of the matrix $\mathbf{S}_T^{tot}(\mathbf{e})$ that we need to determine them only involves known syndromes and therefore can be copied from $\mathbf{S}_T^{tot}(\mathbf{r})$.

Furthermore, suppose that $\omega_l \in \Omega(-D+G-T)\backslash\Omega(-D+G)$. Then there exists constants $\mu \in \mathbb{F}\backslash\{0\}$ and $\mu_k \in \mathbb{F}$ (only depending on (i,j)) such that

$$\omega_l = \mu h_i \omega_j + \sum_{k=0}^{l-1} \mu_k \omega_k. \tag{2.23}$$

Also there exists a unique element $\alpha \in \mathbb{F}$ such that the matrix $\mathbf{M}$ obtained from $\mathbf{S}_T^{tot}(\mathbf{r})|_{i,j}$ by replacing its $(i,j)-th$ element by α, has the same rank as the matrix $\mathbf{S}_T^{tot}(\mathbf{r})|_{i-1,j-1}$. We say that the candidate (i,j) votes for α concerning the syndrome $s_{\omega_j,h_i}(\mathbf{e})$. Using equation (2.23) we then also get a value for $s_{\omega_l}(\mathbf{e})$. If this value is correct, we say that the candidate votes correctly, otherwise we say that the candidate votes incorrectly. We now show that this voting procedure gives the right value for $s_{\omega_j,h_i}(\mathbf{e})$ in the majority of cases, if we assume that not too many errors have occurred.

Theorem 2.29. *Let $\mathbf{c} \in C_L(D,G)$ be a codeword and $\mathbf{r} = \mathbf{c} + \mathbf{e}$ a received word. Let $\omega_l \in \Omega(-D+G-T)\backslash\Omega(-D+G)$ and assume that $C_L(D,G) \neq C_L(D,G-T)$ as well as that $2\,\mathrm{wt}(\mathbf{e}) < \nu(T,\mathcal{M}_\Omega(D,G,T))$.*
Then the majority of candidates in $N(T,\mathcal{M}_\Omega(D,G,T))$ vote for the correct value of $s_{\omega_l}(\mathbf{e})$.

Proof. We consider the following sets:

$$\mathtt{K} := \{(i,j) \,|\, (i,j) \text{ a discrepancy}, \rho_i + \sigma_j < 1\},$$

$$\mathtt{F} := \{(i,j) \in N(T,\mathcal{M},0) \,|\, (i,j) \text{ a candidate voting incorrectly for } s_{\omega_l}(\mathbf{e})\}$$

and

$$\mathtt{T} := \{(i,j) \in N(T,\mathcal{M},0) \,|\, (i,j) \text{ a candidate voting correctly for } s_{\omega_l}(\mathbf{e})\}.$$

Let ρ_{N_1} (resp. σ_{N_2}) be the largest first (resp. second) coordinate occurring in $N(T,\mathcal{M},0)$. We consider the matrix $\mathbf{S}_T^{tot}(\mathbf{e})|_{N_1,N_2}$. Its rank

equals $\mathrm{wt}(\mathbf{e})$, but on the other hand it is at least $\#\mathtt{K}+\#\mathtt{F}$, since discrepancies are exactly pivot positions in the matrix $\mathbf{S}_T^{tot}(\mathbf{e})|_{N_1,N_2}$. Therefore we have that

$$2\#\mathtt{K} + 2\#\mathtt{F} \leq 2\,\mathrm{wt}(\mathbf{e}) < \nu(T,\mathcal{M}).$$

On the other hand, if an element $(i,j) \in N(T,\mathcal{M},0)$ is not a candidate, then there exists an element of $\mathtt{K}$ with first coordinate i or second coordinate j. Therefore, the number of non-candidates in $N(T,\mathcal{M},0)$ is at most $2\#\mathtt{K}$. The number of candidates in $N(T,\mathcal{M},0)$ is of course equal to $\#\mathtt{F}+\#\mathtt{T}$. All in all we find that

$$\nu(T,\mathcal{M}) \leq 2\#\mathtt{K} + \#\mathtt{F} + \#\mathtt{T}.$$

Combining this with the above, we see that $\#\mathtt{T} > \#\mathtt{F}$. □

Note that if $C_L(D,G) = C_L(D,G-T)$, but $\Omega(-D+G-T) \neq \Omega(-D+G)$ it is not hard to determine $s_{\omega_l}(\mathbf{e})$ for $\omega_l \in \Omega(-D+G-T)$, since then there exists $\omega \in \Omega(-D+G)$ such that $\mathrm{Res}_D(\omega) = \mathrm{Res}_D(\omega_l)$. This implies that $s_{\omega_l}(\mathbf{e}) = s_\omega(\mathbf{e})$, but the latter is a known syndrome. Combined with the above theorem, we see that we can always determine the value of $s_{\omega_l}(\mathbf{e})$ as long as $2\,\mathrm{wt}(\mathbf{e}) < \nu(T,\mathcal{M})$.

Suppose as before that the set $\{T_1, T_2, \dots\}$ consists of rational points that do not occur in $\mathrm{Supp}\, D$. Now let again $S = (S_1, S_2, \dots)$ be a sequence of points, each of which is contained in $\{T_1, T_2, \dots, \}$. Further we defined divisors $H_0 := G$, $H_{i+1} := H_i - S_{i+1}$ and modules $\mathcal{M}_S(i) := \mathcal{M}_\Omega(D, H_i, S_{i+1})$. Recall that by Theorem 2.22 the minimum distance d of the code $C_L(D,G)$ satisfies $d \geq d_S(G) := \min_i\{\nu(S_{i+1}, \mathcal{M}_S(i))\}$, where the minimum is taken over all i such that $C_L(D,H_i) \neq C_L(D,H_{i+1})$. We can decode the code $C_L(D,G)$ up to half this bound, since we can determine all unknown syndromes by using Theorem 2.29 iteratively on the sequence of codes $C_L(D,G) \supset \cdots \supset C_L(D,H_i) \supset C_L(D,H_{i+1}) \supset \cdots$. Eventually, we then know all syndromes, after which we can determine the error-vector $\mathbf{e}$.

One does not need to calculate all unknown syndromes, but one can stop the recursive computations when a code $C_L(D,H_i)$ is reached such that $n - \deg H_i - g \geq d_S(G)$. We prove this in the following proposition.

Proposition 2.30. *Let* $\mathbf{c} \in C_L(D,G)$ *and* $S = (S_1, S_2, \dots)$ *a sequence of points not occurring in* $\mathrm{Supp}\, D$. *Suppose that* $\mathbf{e} \in \mathbb{F}^n$ *of weight at most* $(d_S(G)-1)/2$. *Let* $\delta = d_S(G) - n + \deg G + g$. *Suppose that we know* $s_\omega(\mathbf{e})$

for all $\omega \in \Omega(-D+G-S_1-\cdots-S_\delta)$. Then we can find c using the basic algorithm on the code $C_L(D, G-S_1-\cdots-S_\delta)$.

Proof. We write $T = S_1$ and suppose that $\mathbf{c} = \mathrm{Ev}_D(f)$ with $f \in L(G)$. Let $f_1, \ldots, f_k$ be a basis of $L(G)$ such that $\rho_T(f_1) < \cdots < \rho_T(f_k)$ and ω_l an element of $\Omega(-D+G-T)$ of maximal pole order at T. We then have that any $\omega \in \Omega(-D+G-T)$ can be written as $\alpha\omega_l + \omega_r$ for certain $\omega_r \in \Omega(-D+G)$ and constant α. Also we can write

$$f = \sum_{i=1}^{k} \alpha_i f_i$$

and by assumption $s_{\omega_l}(\mathbf{c}) = s_{\omega_l}(\mathbf{r}) - s_{\omega_l}(\mathbf{e})$ is a known expression. Further, since $\rho_T(f_i) < \rho_T(f_k)$ for $1 \le i < k$ and $\mathbf{c} = \mathrm{Ev}_D(f)$, we have that

$$s_{\omega_l}(\mathbf{c}) = \sum_{i=1}^{k} \alpha_i s_{\omega_l}(\mathrm{Ev}_D(f_i)) = \alpha_k s_{\omega_l}(\mathrm{Ev}_D(f_k)).$$

We claim that we can always determine α_k. Indeed if $s_{\omega_m}(\mathrm{Ev}_D(f_k)) = 0$, then $s_{\omega_l}(\mathbf{c}) = 0$ implying that $\mathbf{c} \in C_L(D, G-T)$. But then $\alpha_k = 0$. If $s_{\omega_m}(\mathrm{Ev}_D(f_k)) \neq 0$, then

$$\alpha_k = \frac{s_{\omega_l}(\mathbf{c})}{s_{\omega_l}(\mathrm{Ev}_D(f_k))} = \frac{s_{\omega_l}(\mathbf{r}) - s_{\omega_l}(\mathbf{e})}{s_{\omega_l}(\mathrm{Ev}_D(f_k))}. \tag{2.24}$$

We can now repeat the above proces treating $r - \alpha_k \,\mathrm{ev}(f_k)$ as the received vector, taking $C_L(D, G-S_1)$ as the code we work with and defining $T = S_2$. Iterating this procedure δ times, we obtain as output a vector $r - \mathrm{Ev}_D(g)$ for an explicitly known function g such that $f - g \in L(G - S_1 - \cdots - S_\delta)$. The vector $r - \mathrm{Ev}_D(g)$ differs in $\mathrm{wt}(\mathbf{e}) < (n - \deg G + \delta - g)/2$ positions from $\mathrm{Ev}_D(f-g)$, so we can use the basic algorithm to find the function $f - g$ completing the decoding. □

Example 2.31. In this example we consider the curve χ defined over $\mathbb{F}_{64}$ given by the equation $x_2^2 + x_2 = x_1^9$. It is a hyperelliptic curve of genus 4 with 129 rational points. We denote by T_∞ the unique point that has a pole at x_1, by T_0 the point that has a zero at x_2 and by T_1 the point that has a zero at $x_2 + 1$. Let $G = -T_0 + 121T_\infty$ and D be the sum of the 126 rational points different from T_0, T_1 and T_∞. The code $C_L(D, G)$ is a $[126, 117, \ge 6]$ code. We first calculate the generalized order bound for this code using the sequence $S = (T_\infty, T_\infty, \ldots)$. We have that $H(T_\infty) = \langle 2, 9\rangle$. The differential $\omega = (x_1^{63} + 1)^{-1} dx_1$ has divisor $-D + 132T_\infty$ and can be

used to show that $H(T_\infty, \mathcal{M}_S(0)) = \{i - 11 \mid i \in H(T_\infty)\} \cup \{-4\}$. We find that

$$p_{T_\infty}(t)p_{T_\infty,\mathcal{M}_S(0)}(t) = \cdots + 7t + 7t^2 + 8t^3 + 9t^4 + 10t^5 + \cdots . \qquad (2.25)$$

This means that $d_S(G) = 7$ implying that the code we are studying is in fact a $[126, 117, \geq 7]$ code.

We represent $\mathbb{F}_{64}$ as $\mathbb{F}_2[\gamma]$, with γ a primitive element satisfying $\gamma^6 + \gamma + 1 = 0$. The points in $\operatorname{Supp} D$ have nonzero coordinates, and therefore we write them as powers of γ with exponents between 0 and 62. Then we can order these points lexicographically after these exponents. In this way we get $P_1 = (1, \gamma^{21}), \ldots, P_{126} = (\gamma^{62}, \gamma^{45})$. As in the proof of Proposition 2.30 we will need a basis $f_1, \ldots, f_{117}$ of $L(G)$ of increasing pole order in T_∞. We can take

$$f_i = \begin{cases} x_1^i & \text{if } 1 \leq i \leq 3, \\ x_1^{(i-5)/2}x_2 & \text{if } i \geq 5 \text{ and } i \text{ odd}, \\ x_1^{i/2} & \text{if } i \geq 4 \text{ and } i \text{ even}. \end{cases}$$

Following the notation in and just before Definition 2.27, we have in our case

i	1	2	3	4	5	6	7	8	9	10	11	12
ρ_i	0	2	4	6	8	9	10	11	12	13	14	15
h_i	1	x_1	x_1^2	x_1^3	x_1^4	x_2	x_1^5	x_1x_2	x_1^6	$x_1^2x_2$	x_1^7	$x_1^3x_2$

and (still using $\omega = (x_1^{63} + 1)^{-1}dx_1$)

j	1	2	3	4	5	6	7	8	9	10	11	12
σ_j	-11	-9	-7	-5	-4	-3	-2	-1	0	1	2	3
$\frac{\omega_j}{\omega}$	1	x_1	x_1^2	x_1^3	$\frac{x_1^8}{x_2}$	x_1^4	x_2	x_1^5	x_1x_2	x_1^6	$x_1^2x_2$	x_1^7

Now we define an error-vector $\mathbf{e}$ in the following way: $e_1 = 1$, $e_2 = \gamma^{42}$, $e_{93} = \gamma^{13}$, and $e_i = 0$ otherwise. Since $d_S(G) = 7$, we can correct this error-pattern with the majority voting algorithm. Goppa's bound for the minimum distance of the code $C_L(D, G)$ equals 6, so we need to determine $g + (7 - 6) = 5$ unknown syndromes in this case. We now assume that the sent codeword was $\mathbf{c} = \mathrm{Ev}_D(\gamma x_1^{60} + x_1^{56}x_2)$, so that the received word is

$\mathbf{r} = \mathbf{c} + \mathbf{e}$. Then we have that $\mathbf{S}^{tot}_{T_\infty}(\mathbf{c})|_{14,14}$ equals

$$\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \gamma & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & \gamma & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & \gamma & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & \gamma & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & \gamma & 0 & 0 & 0 & 0 & 0 & 1 & \gamma & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \gamma & 0 & 0 & 0 & 0 & 1 & \gamma & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\gamma & 0 & 0 & 0 & 1 & \gamma & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix},$$

while $\mathbf{S}^{tot}_{T_\infty}(\mathbf{e})|_{14,14}$ is equal to the matrix

$$\begin{pmatrix}
\gamma^7 & 1 & \gamma^{45} & \gamma^{43} & \gamma^{37} & \gamma^7 & \gamma^{54} & \gamma^{53} & \gamma^{26} & \gamma^{36} & \gamma^{30} & \gamma^{19} & \gamma^{62} & \gamma^2 \\
1 & \gamma^{45} & \gamma^{43} & \gamma^{37} & \gamma^{54} & \gamma^{53} & \gamma^{26} & \gamma^{36} & \gamma^{30} & \gamma^{19} & \gamma^{62} & \gamma^2 & \gamma^{46} & \gamma^{48} \\
\gamma^{45} & \gamma^{43} & \gamma^{37} & \gamma^{54} & \gamma^{26} & \gamma^{36} & \gamma^{30} & \gamma^{19} & \gamma^{62} & \gamma^2 & \gamma^{46} & \gamma^{48} & \gamma^{16} & \gamma^{31} \\
\gamma^{43} & \gamma^{37} & \gamma^{54} & \gamma^{26} & \gamma^{30} & \gamma^{19} & \gamma^{62} & \gamma^2 & \gamma^{46} & \gamma^{48} & \gamma^{16} & \gamma^{31} & \gamma^{13} & \gamma^{14} \\
\gamma^5 & 0 & \gamma^{51} & \gamma^{18} & \gamma^{23} & \gamma^{62} & \gamma^{15} & \gamma^{46} & \gamma^{49} & \gamma^{16} & \gamma^{25} & \gamma^{13} & \gamma^{47} & \gamma^{36} \\
\gamma^{37} & \gamma^{54} & \gamma^{26} & \gamma^{30} & \gamma^{62} & \gamma^2 & \gamma^{46} & \gamma^{48} & \gamma^{16} & \gamma^{31} & \gamma^{13} & \gamma^{14} & \gamma^{36} & \gamma^{60} \\
\gamma^7 & \gamma^{53} & \gamma^{36} & \gamma^{19} & \gamma^2 & \gamma^{50} & \gamma^{48} & \gamma^{60} & \gamma^{31} & \gamma^{28} & \gamma^{14} & \gamma^3 & \gamma^{60} & \gamma^{61} \\
\gamma^{54} & \gamma^{26} & \gamma^{30} & \gamma^{62} & \gamma^{46} & \gamma^{48} & \gamma^{16} & \gamma^{31} & \gamma^{13} & \gamma^{14} & \gamma^{36} & \gamma^{60} & \gamma^{22} & \gamma^{43} \\
\gamma^{53} & \gamma^{36} & \gamma^{19} & \gamma^2 & \gamma^{48} & \gamma^{60} & \gamma^{31} & \gamma^{28} & \gamma^{14} & \gamma^3 & \gamma^{60} & \gamma^{61} & \gamma^{43} & \gamma^7 \\
\gamma^{26} & \gamma^{30} & \gamma^{62} & \gamma^{46} & \gamma^{16} & \gamma^{31} & \gamma^{13} & \gamma^{14} & \gamma^{36} & \gamma^{60} & \gamma^{22} & \gamma^{43} & \gamma^{35} & \gamma^{26} \\
\gamma^{36} & \gamma^{19} & \gamma^2 & \gamma^{48} & \gamma^{31} & \gamma^{28} & \gamma^{14} & \gamma^3 & \gamma^{60} & \gamma^{61} & \gamma^{43} & \gamma^7 & \gamma^{26} & 1 \\
\gamma^{30} & \gamma^{62} & \gamma^{46} & \gamma^{16} & \gamma^{13} & \gamma^{14} & \gamma^{36} & \gamma^{60} & \gamma^{22} & \gamma^{43} & \gamma^{35} & \gamma^{26} & \gamma^{34} & \gamma^9 \\
\gamma^{19} & \gamma^2 & \gamma^{48} & \gamma^{31} & \gamma^{14} & \gamma^3 & \gamma^{60} & \gamma^{61} & \gamma^{43} & \gamma^7 & \gamma^{26} & 1 & \gamma^9 & \gamma^{45} \\
\gamma^{62} & \gamma^{46} & \gamma^{16} & \gamma^{13} & \gamma^{36} & \gamma^{60} & \gamma^{22} & \gamma^{43} & \gamma^{35} & \gamma^{26} & \gamma^{34} & \gamma^9 & \gamma^{48} & \gamma^{55}
\end{pmatrix}.$$

In the decoding algorithm, we know the matrix $\mathbf{S}^{tot}_{T\infty}(\mathbf{r})|_{14,14}$, which is the sum of the above two matrices, which are unknown to the receiver. However, note that $\mathbf{S}^{tot}_{T\infty}(\mathbf{r})$ and $\mathbf{S}^{tot}_{T\infty}(\mathbf{e})$ are guaranteed to be the same in all those positions (i,j) satisfying $\sigma_i + \rho_j \leq 0$, since these positions contain the known syndromes.

We now try to calculate $f = \gamma x_1^{60} + x_1^{56} x_2$. Since we know that $f \in L(G)$, we can write $f = \sum_{i=1}^{117} \alpha_i f_i$. We will determine α_{113} up till α_{117} using majority voting. In the first step of the algorithm we need to determine which positions (i,j) satisfying $\sigma_i + \rho_j = 1$, are candidates as well. From equation (2.25) we can deduce that there are at most 7 such positions (i,j). By row reduction of the matrix $\mathbf{S}^{tot}_{T\infty}(\mathbf{r})$ we can compute that in this case the positions $(1,1)$ and $(2,2)$ are the only discrepancies in the known part $\mathbf{S}^{tot}_{T\infty}(\mathbf{e})$. The candidates in the first and following steps can therefore not contain a 1 or a 2 in any of their coordinates. The votes can be calculated directly once the candidates are known. The results of the first step of the algorithm are given in the following table:

candidate	$(6,3)$	$(4,4)$	$(3,5)$
vote	γ^{26}	γ^{26}	γ^{26}

We conclude that $s_{\omega_{10}}(\mathbf{e}) = \gamma^{26}$. Using equation (2.24), we find that $\alpha_{117} = 1$. We can then update the matrix $\mathbf{S}^{tot}_{T\infty}(\mathbf{r})$ by replacing it with the matrix $\mathbf{S}^{tot}_{T\infty}(\mathbf{r} - \mathrm{Ev}_D(f_{117}))$. Since the voting is unanimous, there are no new discrepancies. In the second step of the algorithm, we find the following:

candidate	$(7,3)$	$(5,4)$	$(3,6)$
vote	γ^{36}	γ^{36}	γ^{36}

Therefore $s_{\omega_{10}}(\mathbf{e}) = \gamma^{36}$ and $\alpha_{116} = \gamma$. In this particular example the updated syndrome matrix now becomes $\mathbf{S}^{tot}_{T\infty}(\mathbf{e})$, because of our choice of the sent codeword $\mathbf{c}$. Continuing to the third step, we find:

candidate	$(8,3)$	$(6,4)$	$(4,5)$	$(3,7)$
vote	γ^{30}	γ^{30}	γ^{30}	γ^{30}

Thus $s_{\omega_{11}}(\mathbf{e}) = \gamma^{30}$ and $\alpha_{115} = 0$. The fourth step yields the following results:

candidate	$(9,3)$	$(7,4)$	$(5,5)$	$(4,6)$	$(3,8)$
vote	γ^{19}	γ^{19}	γ^{19}	γ^{19}	γ^{19}

This implies that $s_{\omega_{12}}(\mathbf{e}) = \gamma^{19}$ and $\alpha_{114} = 0$. The fifth and last step gives:

candidate	$(10,3)$	$(8,4)$	$(6,5)$	$(5,6)$	$(4,7)$	$(3,9)$
vote	γ^{62}	γ^{62}	γ^{62}	γ^{49}	γ^{62}	γ^{62}

In this case the voting is not unanimous and we find $s_{\omega_{13}}(\mathbf{e}) = \gamma^{62}$ and $\alpha_{113} = 0$. The reason the voting is not unanimous in this case, is that the $(5,6)$-th position is a discrepancy in the matrix of syndromes.

2.6. List decoding of algebraic geometry codes

In this section we will describe a list decoding algorithm for algebraic geometry codes, which is an extension of the basic algorithm presented in Section 2.2.

Suppose that we wish to use the code $C_L(D,G)$ and that we have received the word $(r_1,\ldots,r_n)$ containing at most τ errors. The algorithm works with a divisor A with $\operatorname{Supp} A \cap \operatorname{Supp} D = \varnothing$ satisfying certain conditions that we describe below and a natural number s.

The idea of the algorithm is to find a nonzero polynomial $Q(y) \in \mathscr{F}[y]$ such that:

(i) $Q(y) = Q_0 + Q_1 y + \cdots + Q_\lambda y^\lambda$ where $Q_i \in L(A - iG), i = 0,\ldots,\lambda$
(ii) $Q(y)$ has a zero of multiplicity s in (P_j, r_j), $j = 1,\ldots,n$

The meaning of (ii) is the following: Let t be a local parameter at P_j then $Q(y) = \sum \mu_{a,b} t^a (y - r_j)^b$. That $Q(y)$ has a zero of multiplicity s in (P_j, r_j) then means that $\mu_{a,b} = 0$ for $a + b < s$

This algorithm is an extension of the basic algorithm in two ways. The y-degree of the interpolation polynomial Q is allowed to be larger than one and the zeroes now shall be of multiplicity s. In this way, as we shall see, we are able to correct a larger number of errors if we accept a list of possible codewords.

The conditions on the divisor A are as follows.

(1) $\deg A < s(n - \tau)$
(2) $\deg A > \frac{ns(s+1)}{2(\lambda+1)} + \frac{\lambda \deg G}{2} + g - 1$

It can be seen that if $\tau < n - \frac{n(s+1)}{2(\lambda+1)} - \frac{\lambda \deg G}{2s} - \frac{g}{s}$ then such a divisor A exists.

Lemma 2.32. *Suppose the transmitted word is generated by $f \in L(G)$ and $Q(y)$ satisfies (i) and (ii) then $Q(f) = 0$*

Proof. Since $f \in L(G)$ and $Q_i \in L(A - iG)$ we have $f^i Q_i \in L(A)$ and therefore $Q(f) \in L(A)$. We also have that $Q(f(P_j))$ has a zero of multiplicity s in P_j for at least $n - \tau$ j's $\in \{1, 2, \ldots, n\}$ so that $Q(f) \in L(A - sP_{i_1} - \cdots - sP_{i_r})$ with $r \geq n - \tau$. But $\deg(A - sP_{i_1} - \cdots - sP_{i_r}) < 0$ and therefore $Q(f) = 0$. This implies that if the divisor A satisfies condition (1) above then the function f that generated the sent codeword gives a factor $y - f$ in $Q(y)$. □

Later we will discuss how such factors are actually found.

Lemma 2.33. *If* $\deg A$ *satisfies (2) above then a nonzero $Q(y) \in \mathscr{F}[y]$ exists satisfying (i) and (ii).*

Proof. By selecting bases for the spaces $L(A - iG), i = 0, 1, \ldots, \lambda$ the condition (ii) translates into a system of homogeneous linear equations in $\sum_{i=0}^{\lambda} l(A - iG)$ unknowns. The number of equations is $\frac{n(s+1)s}{2}$ which by (2) is smaller than the number of unknowns, so there is a nonzero solution to the system. □

This leads to the following algorithm:

Input: A received word $r = (r_1, r_2, \ldots, r_n)$.
Find a polynomial $Q(y)$ satisfying (i) and (ii).
Find factors of $Q(y)$ of the form $y - f$ with $f \in L(G)$.
If no such factors exist **Output**: Failure.
Else **Output** : $\mathrm{Ev}_D(f)$ for those f's where $d(\mathrm{Ev}_D(f), r) \leq \tau$.

It can be seen that this list decoding algorithm only improves on $\frac{n - \deg G}{2}$ if $\lambda \geq s$ and

$$n\left(1 - \tfrac{s+1}{\lambda+1}\right) > \left(\tfrac{\lambda}{s} - 1\right)\deg G + \tfrac{2g}{s} + 1$$

and also that for fixed λ the optimal s is

$$\left\lfloor \left[\tfrac{2(\lambda+1)}{n} \left(\tfrac{\lambda}{2} \deg G + g \right) \right]^{\frac{1}{2}} \right\rfloor$$

Example 2.34. This is a continuation of Example 2.12 where we considered the $[60, 18, \geq 37]$ code over $\mathbb{F}_{16}$. With $\lambda = 6$ and $s = 4$ we can correct 19 errors, with $\lambda = 10$ and $s = 7$, 20, and with $\lambda = 50$ and $s = 32$, 22 errors can be corrected with the list decoder.

As we have seen, and we will discuss this further in the next section, the polynomial $Q(y)$ can be found by solving a system of homogenous linear equations.

We will address the question of finding the relevant factors of the polynomial $Q(y)$ and present two different methods for doing that. The first one transforms the problem to that of finding factors of an univariate polynomial over a large finite field and the second one uses Hensel lifting.

The first algorithm reduces the problem of finding factors of $Q(y)$ of the form $y - f$ with $f \in L(G)$ to the problem of finding roots of a polynomial $\widehat{Q}(y)$ in $\mathbb{F}_{q^m}$ obtained by "reducing" the coefficients of $Q(y)$ modulo a point R of sufficiently large degree m where $R \notin \operatorname{Supp} A$ and $R \notin \operatorname{Supp} G$. It can be seen that such a point exists. The reduction is performed by evaluating the functions Q_i in R. One then finds zeroes of $\widehat{Q}(y)$ using a root-finding algorithm for finite fields and for those zeroes that lie in $\mathrm{Ev}_R(L(G))$ one finds the corresponding f's $\in L(G)$. For this to be possible the map $\mathrm{Ev}_R : L(G) \to \mathbb{F}_{q^m}$ shall be injective and this is the case if $\deg R > \deg G$. To turn these remarks into a proper algorithm one needs a way of evaluating functions from $L(G)$ and $L(A - iG)$ in R, and also a method for reconstructing an f from an element in $\mathrm{Ev}_R(L(G)) \subseteq \mathbb{F}_{q^m}$. We shall now assume w.l.o.g that the divisor G is effective and also that $A \geq G$. This implies that $L(G) \subseteq L(A)$ and also that $L(A - iG) \subseteq L(A)$.

Let $\phi_1, \phi_2, \ldots, \phi_k$ be a basis of $L(G)$ as a vector space over $\mathbb{F}_q$, and let $\phi_1, \ldots, \phi_k, \phi_{k+1}, \ldots, \phi_a$ be a basis of $L(A)$. We then "represent" R by the values $\phi_1(R), \phi_2(R), \ldots, \phi_a(R)$ i.e. an element of $\mathbb{F}_{q^m}^a$. Let $Q_i = \sum_{j=1}^{a} \gamma_{i,j}\phi_j$ then $Q(y) = \sum_{i=0}^{\lambda} \sum_{j=1}^{a} \gamma_{i,j}\phi_j y^i$ and $\widehat{Q}(y) = \sum_{i=0}^{\lambda} \sum_{j=1}^{a} \gamma_{i,j}\phi_j(R) y^i$. If $\beta \in \mathbb{F}_{q^m}$ is a zero of $\widehat{Q}(y)$ we shall then find $(f_1, f_2, \ldots, f_k) \in \mathbb{F}_q$ such that $\sum_{l=1}^{k} f_l \phi_l(R) = \beta$. Using a basis of $\mathbb{F}_{q^m}$ over $\mathbb{F}_q$ this gives m linear equations in k unknowns and we know that either this has no solution or a unique solution. In the latter case we have found an f and if $d(\mathrm{Ev}_D(f), r) \leq \tau$ we put $\mathrm{Ev}_D(f)$ on the list.

In the second algorithm the idea is the following: Let P be a point, $P \notin \operatorname{Supp} A$ and $P \notin \operatorname{Supp} G$ and let t be a local parameter at P. Then a function in $L(G)$ can be developed as a power series in t, $f = \sum_{i=0}^{\infty} a_i t^i$. The polynomial $Q(y)$ can also be considered as element of $\mathbb{F}_q[[t]][y]$, $Q(y) = Q_0(t, y) = \sum_{i=0,j=0}^{\infty,\lambda} \alpha_{i,j} t^i y^j$, so if $Q(f) = 0$ we get

$$Q_0(t, \sum_{i=0}^{\infty} a_i t^i) = 0 \tag{2.26}$$

If we consider this equation modulo increasing powers of t it is possible to determine the a_i's recursively.

In the first step we look at equation (2.26) mod t which is the same as $Q_0(0, a_0) = 0$ and this is

$$\sum_{j=0}^{\lambda} \alpha_{0,j} a_0{}^j = 0 \tag{2.27}$$

Here we can suppose that $\alpha_{0,j} \neq 0$ for some j since if not $Q_0(t, y) = tR(t, y)$ and we would get $R(t, f) = 0$. This means that we can determine a_0 as a zero in $\mathbb{F}_q$ of the polynomial $Q_0(0, T)$. In order to give an easy description of the updating procedure we let for $i \geq 0$, $\psi_i(t) = \sum_{s=i}^{\infty} a_s t^{s-i}$, $M_i(t, y) = t^{-r_i} Q_i(t, y)$ where r_i is the largest integer such that t^{r_i} divides $Q_i(t, ty + a_i)$ and

$$Q_{i+1}(t, y) = M_i(t, ty + a_i).$$

Note that $Q_{i+1}(t, y)$ as well as r_i may depend on the value found for a_i in the previous step of the algorithm, but for simplicity we suppress this in our notation. We then have

Lemma 2.35. *$Q_i(t, \psi_i(t)) = 0$, $M_i(0, a_i) = 0$ and $M_i(0, y) \neq 0$.*

Proof. Observe that the y degrees of $Q_i(t, y)$ are the same for all i and that $Q_i(t, y) \neq 0$ so r_i is well-defined. Also since t does not divide $M_i(t, y)$ we have $M_i(0, y) \neq 0$. We can now prove that $Q_i(t, \psi_i(t)) = 0$ by induction on i, where the basis $i = 0$ is obvious from (2.26). For the induction step if $Q_i(t, \psi_i(t)) = 0$ then $\psi_{i+1}(t) = (\psi_i(t) - a_i)/t$ is a y-root of $Q_i(t, ty + a_i)$ and hence of $Q_{i+1}(t, y) = t^{-r_i} Q_i(t, ty + a_i)$. By substituting $t = 0$ in $M_i(t, \psi_i(t)) = t^{-r_i} Q_i(t, \psi_i(t)) = 0$ we obtain $M_i(0, a_i) = 0$. □

This means that the coefficients a_i can be found by solving an equation of degree λ, but as we shall see the total number of solutions f is at most λ. This can be concluded from the following

Lemma 2.36. *Let $M_1(t, y) = \sum_{j=0}^{\lambda} M^{(j)}(t) y^j$ be a nonzero polynomial in $\mathbb{F}_q[[t]][y]$ and let β be zero of $M_1(0, y)$ of multiplicity m_β. Define*

$$M_2(t, y) = t^{-r} M_1(t, ty + \beta),$$

where r is the largest integer such that t^r divides $M_1(t, ty + \beta)$ then $\deg_y M_2(0, y) \leq m_\beta$.

Proof. Let $\widehat{M}(t,y) = M_1(t, y+\beta) = \sum_{j=0}^{\lambda} q_j(t)y^j$ then $q_j(0) = 0$ for $0 \leq j < m_\beta$ and $q_{m_\beta}(0) \neq 0$, or equivalently t divides $q_j(t)$ for $0 \leq j < m_\beta$ but it does not divide $q_{m_\beta}(0)$. This means that t divides $\widehat{M}(t, ty)$ but $t^{m_\beta+1}$ does not, so $r \leq m_\beta$. Since $M_2(t,y) = t^{-r}M_1(t, ty+\beta) = \sum_{j=m_\beta}^{\lambda} q_j(t)t^{j-r}y^j$ we get $M_2(0,y) = \sum_{j=m_\beta}^{\lambda}(q_j(t)t^{j-r})|_{t=0}y^j$ so $\deg_y M_2(0,y) \leq r \leq m_\beta$. □

Corollary 2.37. *The number of different f's is at most λ.*

Proof. Denote by A_i the set of all solutions $\mathbf{a} = (a_0, \ldots, a_i)$ the algorithm finds after i steps. We will show by induction that

$$\sum_{\mathbf{a} \in A_i} m_{a_i} \leq \lambda. \tag{2.28}$$

This will imply the corollary, since then $\#A_i \leq \lambda$ for all i. For $i = 0$ equation (2.28) is true, since all found a_0's in the start of the algorithm are roots of $Q_0(0,y)$ and $\deg_y Q_0(0,y) = \lambda$. Now suppose the result is true for i. Given a fixed $(a_0, \ldots, a_i)$ at this stage of the algorithm, the a_{i+1}'s the algorithm finds in the next step are, according to Lemma 2.36, roots of a polynomial of degree at most m_{a_i} so the sum of their multiplicities is at most m_{a_i}. This implies that $\sum_{\mathbf{a} \in A_{i+1}} m_{a_{i+1}} \leq \sum_{\mathbf{a} \in A_i} m_{a_i} \leq \lambda$. □

The only remaining issue is to bound the number of a_i's we have to determine in order to reconstruct the function $f \in L(G)$. To this end let $k = \dim L(G)$and let $b_1, b_2, \ldots, b_k$ be a basis of $L(G)$ such that $j_i = v_P(b_i) < v_P(b_{i+1}) = j_{i+1}$, $i = 1, \ldots, k-1$. This means that f is determined if we know the a_i's up to $i = j_k$. Since $b_k \in L(G - j_k P)$ we have $j_k \leq \deg G$.

Example 2.38. In this example we consider the Hermitian curve over $\mathbb{F}_4$ defined by $x_2^2 + x_2 = x_1^3$. We write $\mathbb{F}_4 = \mathbb{F}_2[\alpha]$ with $\alpha^2 = \alpha + 1$. Also we write $P_1 = (0,0)$, $P_2 = (0,1)$, $P_3 = (1,\alpha)$, $P_4 = (1,\alpha^2)$, $P_5 = (\alpha,\alpha)$, $P_6 = (\alpha,\alpha^2)$, $P_7 = (\alpha^2,\alpha)$, $P_8 = (\alpha^2,\alpha^2)$, and denote as before by T_∞ the unique pole of x_1. We now take $D = P_1 + \cdots + P_8$, $G = 4T_\infty$, and $A = 35T_\infty$. If we choose $s = 6$ and $\lambda = 8$, we can correct 2 errors using the list decoder. In order to describe the list-decoding procedure, we need to choose bases for the spaces $L(A - iG)$, whose dimension we denote by l_i. In this case we can for $0 \leq i \leq \lambda$ and $1 \leq j \leq l_i$ choose

$$g_{ij} = \begin{cases} 1 & \text{if } j = 1, \\ x_1 x_2^{(j-2)/3} & \text{if } j \equiv 2 \bmod 3, \\ x_2^{j/3} & \text{if } j \equiv 0 \bmod 3, \\ x_1^2 x_2^{(j-4)/3} & \text{if } j > 1 \text{ and } j \equiv 1 \bmod 3. \end{cases}$$

Suppose that we transmit the all zero word and receive.

$$(\alpha^2, 0, 0, \alpha^2, 0, 0, 0, 0).$$

The decoder from Section 2.5 fails to decode this word, but we can use list decoding if we choose $s = 6$ and $\lambda = 8$. To find an interpolation polynomial we could solve the linear system occurring in the proof of Lemma 2.33. This system has 168 equations and 171 variables. However, we will see in Section 2.7 that this approach is not optimal. Using the results from Section 2.7 (more specifically, Example 2.45), we get that the interpolation polynomial is equal to:

$$\begin{aligned} Q(y) = {} & (1 + x_2 + \alpha x_2^2 + \alpha x_1^2 x_2 + \alpha^2 x_1 x_2^2 + \alpha x_2^3 + \alpha^2 x_1^2 x_2^2 + \alpha x_1 x_2^3 + x_2^4 + \\ & \alpha x_1^2 x_2^3 + \alpha^2 x_1 x_2^4 + x_1^2 x_2^4 + \alpha x_1 x_2^5 + \alpha^2 x_1^2 x_2^5 + \alpha x_1 x_2^6 + x_2^7 + \alpha x_1^2 x_2^6 + \\ & x_1 x_2^7 + x_2^8 + x_1^2 x_2^7 + \alpha x_1 x_2^8 + \alpha x_2^9 + \alpha^2 x_1^2 x_2^8 + x_1 x_2^9 + \alpha^2 x_2^{10} + x_1^2 x_2^9) y + \\ & (\alpha^2 + \alpha x_1 + \alpha x_1^2 + x_2^2 + \alpha^2 x_1^2 x_2 + \alpha^2 x_2^3 + x_1^2 x_2^2 + \alpha^2 x_1 x_2^3 + \alpha^2 x_2^5 + x_1^2 x_2^4 + \\ & x_1^2 x_2^4 + \alpha^2 x_2^6 + \alpha x_1^2 x_2^5 + \alpha x_2^7 + \alpha^2 x_1^2 x_2^6 + \alpha x_1 x_2^7 + x_2^8 + \alpha^2 x_1 x_2^8 + \alpha x_2^9) y^2 + \\ & (\alpha^2 + \alpha x_2 + x_1 x_2 + \alpha^2 x_1^2 x_2 + x_1 x_2^2 + \alpha x_2^3 + x_1^2 x_2^2 + \alpha^2 x_2^4 + \alpha^2 x_1^2 x_2^3 + \\ & \alpha x_2^5 + \alpha x_1^2 x_2^4 + \alpha^2 x_1 x_2^5 + \alpha x_2^6 + \alpha^2 x_1^2 x_2^5 + \alpha^2 x_1 x_2^6) y^3 + (\alpha + x_1 + \alpha^2 x_2 + \\ & x_1 x_2 + \alpha x_2^2 + \alpha^2 x_1^2 x_2 + \alpha x_1 x_2^2 + x_2^3 + \alpha x_1 x_2^3 + \alpha x_2^4 + x_1^2 x_2^3) y^4 + (\alpha + \\ & \alpha^2 x_2 + \alpha^2 x_1 x_2 + x_2^2 + x_1^2 x_2 + x_1 x_2^2 + \alpha^2 x_1^2 x_2^2 + \alpha x_1 x_2^3) y^5 + (1 + \alpha^2 x_1 + \\ & \alpha x_2 + \alpha^2 x_1^2 + \alpha^2 x_1 x_2 + x_2^2 + \alpha^2 x_1^2 x_2) y^6 + y^7 + (\alpha^2 + \alpha x_1) y^8. \end{aligned}$$

In order to factorize this using the first method described above, we let

$$\mathbb{F}_{4^3} = \mathbb{F}_4[X_2]/\langle {X_2}^3 + \alpha X_2 + 1 \rangle \text{ and } \mathbb{F}_{4^{3\times 3}} = \mathbb{F}_{4^3}[X_1]/\langle {X_1}^3 + {X_2}^2 + X_2 \rangle.$$

This makes sense since the polynomial ${X_2}^3 + \alpha X_2 + 1$ is irreducible over $\mathbb{F}_4$ and for any root X_2 of it, the polynomial ${X_1}^3 + {X_2}^2 + X_2$ is irreducible over $\mathbb{F}_{4^3}$. If we let R be a point (x_1, x_2) on the curve in $\mathbb{F}_{4^{3\times 3}}$ corresponding to the description above we get

$$\begin{aligned}\widehat{Q}(y) = &((\alpha + \alpha x_2) + (\alpha x_2 + \alpha^2 x_2^2)\, x_1 + (\alpha x_2 + x_2^2)\, x_1^2)\, y + ((\alpha + \alpha^2 x_2) + \\ &(\alpha+\alpha^2 x_2)x_1+(1+\alpha x_2^2)x_1^2)y^2+((\alpha^2 x_2+\alpha^2 x_2^2)+(\alpha+\alpha x_2+\alpha^2 x_2^2)x_1+ \\ &(\alpha^2 + \alpha x_2 + \alpha^2 x_2^2)x_1^2)y^3+((\alpha^2 + x_2 + \alpha x_2^2)+(\alpha^2 + \alpha^2 x_2 + \alpha x_2^2)x_1+ \\ &(\alpha x_2 + x_2^2)\, x_1^2)\, y^4 + ((\alpha + x_2) + (\alpha + \alpha\, x_2^2)\, x_1 + (1 + \alpha\, x_2)\, x_1^2)\, y^5 + \\ &((x_2 + \alpha x_2^2) + (1 + \alpha x_2 + x_2^2)x_1 + (\alpha + \alpha^2 x_2^2)x_2^2)y^6 + ((1 + \alpha^2 x_2^2)+ \\ &(\alpha^2 + \alpha x_2^2)x_1 + (\alpha + x_2^2)x_1^2)y^7 + y^8.\end{aligned}$$

This polynomial has three factors of degree one namely

$$\begin{gathered} y \\ (\alpha^2 + \alpha^2 x_1 + \alpha^2 x_1^2) + y \\ \text{and } ((\alpha^2 + \alpha x_2 + x_2^2) + (\alpha x_2 + \alpha^2 x_2^2)x_1 + (1 + \alpha^2 x_2 + \alpha x_2^2)x_1^2) + y \end{gathered}$$

The last of these factors does not correspond to a codeword since it is not in $L(G)$ but the first two factors correspond to the codewords

$$\begin{gathered}(\alpha^2, \alpha^2, \alpha^2, \alpha^2, 0, 0, 0, 0) \\ (0, 0, 0, 0, 0, 0, 0, 0)\end{gathered}$$

which both have distance two to the received word.

Now we shall describe the Hensel-lifting approach to find y-roots of $Q(y)$. As the point in which we develop, we choose $P = P_{00}$ and as local parameter for P we pick $t = x_1$. Then we write $Q(y)$ explicitly as an element of $\mathbb{F}_4[[t]][y]$. Since $x_1 = t$, we find from the defining equation of the curve that $x_2 = t^3 + t^6 + t^{12} + \mathcal{O}(t^{24})$. Substituting this in $Q(y)$ we see that

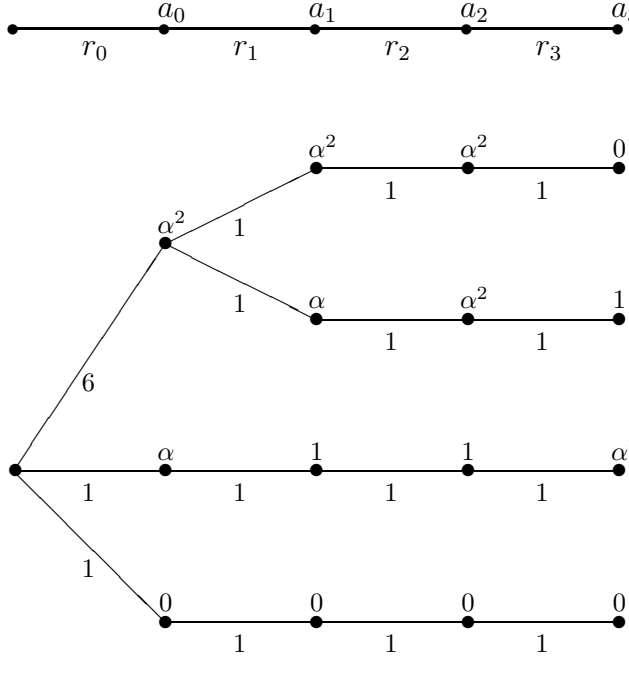
$$\begin{aligned}Q(y) = &(1 + t^3 + \alpha t^5 + \alpha^2 t^6 + \alpha^2 t^7 + t^8 + \alpha t^9)y+ \\ &(\alpha^2 + \alpha t + \alpha t^2 + \alpha^2 t^5 + t^6 + \alpha t^8 + \alpha^2 t^9)y^2+ \\ &(\alpha^2 + \alpha t^3 + t^4 + \alpha^2 t^5 + \alpha t^6 + \alpha t^8 + \alpha t^9)y^3+ \\ &(\alpha + t + \alpha^2 t^3 + t^4 + \alpha^2 t^5 + t^6 + \alpha^2 t^7 + \alpha^2 t^8 + t^9)y^4+ \\ &(\alpha + \alpha^2 t^3 + \alpha^2 t^4 + t^5 + \alpha t^6 + \alpha t^7 + \alpha t^8)y^5+ \\ &(1 + \alpha^2 t + \alpha^2 t^2 + \alpha t^3 + \alpha^2 t^4 + \alpha^2 t^5 + \alpha^2 t^6 + \alpha^2 t^7 + \alpha^2 t^8)y^6+ \\ &y^7 + (\alpha^2 + \alpha t)y^8 + \mathcal{O}(t^{10}).\end{aligned}$$

Now we can determine all possibilities for a_0, since from the above it has

to be a root of $Q_0(0,y) = \alpha^2 y(y-\alpha)(y-\alpha^2)^6$. Therefore there are three possibilities for a_0, namely 0, α and α^2. For each of them separately we can calculate the updated polynomial $Q_1(t,y)$. If a_0 equals 0 or α, it has multiplicity 1, implying by Lemma 2.36 that the next coefficient is the root of a polynomial of degree at most one, i.e. a_1 is uniquely determined if it exists. Since $a_0 = \alpha^2$ has multiplicity 6 this need not be true in that case. In fact, if $a_0 = \alpha^2$ then $Q_1(t,y) = t^{-6}Q_0(t, ty + \alpha^2)$ and we find

$$\begin{aligned} Q_1(t,y) = {} & 1 + t^3 + (t + \alpha t^2 + \alpha^2 t^3)y + (1 + \alpha^2 t + \alpha t^2 + \alpha t^3)y^2 + \\ & (\alpha + t + \alpha^2 t^2 + \alpha t^3)y^3 + (1 + \alpha t + \alpha t^2 + t^3)y^4 + (\alpha^2 t^2 + \alpha^2 t^3)y^5 + \\ & (\alpha + \alpha^2 t + \alpha^2 t^2 + \alpha t^3)y^6 + ty^7 + (\alpha^2 t^2 + \alpha t^3)y^8 + \mathcal{O}(t^4) \end{aligned}$$

and therefore $Q_1(0,y) = (y-\alpha)(y-\alpha^2)(\alpha y^4 + \alpha y^3 + y^2 + y + 1)$. We see that if $a_0 = \alpha^2$, then $a_1 = \alpha$ or $a_1 = \alpha^2$ both having multiplicity one. The degree 4 factor of $Q_1(0,y)$ does not give rise to $\mathbb{F}_4$-rational solutions and is therefore ignored. The outcome of the entire Hensel-lifting procedure including multiplicities and found values for the a_i's can be described in a tree structure. Below the edges we state the multiplicities r_i, while above the vertices we give the values for the a_i's.



From the Hensel-lifting procedure we get four outputs for (a_0, a_1, a_2, a_3), namely $(\alpha^2, \alpha^2, \alpha^2, 0)$, $(\alpha^2, \alpha, \alpha^2, 1)$, $(\alpha, 1, 1, \alpha^2)$, and $(0, 0, 0, 0)$. The corresponding functions are $\alpha^2 + \alpha^2 x + \alpha^2 x^2$, $\alpha^2 + \alpha x + \alpha^2 x^2 + y$, $\alpha + x + x^2 + \alpha^2$, and 0. The first and the last function give rise to solutions of the equation $Q(f) = 0$ and thus to two codewords, while the remaining two are not solutions.

2.7. Syndrome formulation of list decoding

In this section we will formulate the list decoding algorithm using syndromes. As in the case of the basic algorithm, the advantage is that one can eliminate variables from the system of linear equations used to determine the interpolation polynomial.

As discussed in Section 2.6 in order to list-decode we need a polynomial $Q(y) = \sum_{i=0}^{\lambda} Q_i y^i$ such that $Q_i \in L(A - iG)$ and such that (P_l, r_l) is a zero of $Q(y)$ of multiplicity s for all i between 1 and n. We denote by $g_{i1}, \ldots, g_{il_i}$ a basis of $L(A - iG)$ and write $Q_i = \sum_{j=1}^{l_i} q_{ij} g_{ij}$. The condition that (P_l, r_l) is a zero of $Q(y)$ of multiplicity s gives rise to $\binom{s+1}{2}$ linear equations in the coefficients q_{ij}. More explicitly, we can do the following: first for any $P_l \in \operatorname{Supp} D$ we choose a function $t_l \in \mathscr{F}$ such that $v_{P_l}(t_l) = 1$. Given such a t_l, we can write a function g that is regular at P_l as a power series in t_l, say $g = \alpha_0 + \alpha_1 t + \cdots + \alpha_a t^a + \cdots$. We have that $\alpha_0 = g(P_l)$. The α_a depend in general on P_l and the choice of $t_l \in \mathscr{F}$. Denoting by $D_{t_l}^{(a)}$ the a-th Hasse-derivative with respect to t_l, we then have that $D_{t_l}^{(a)}(g)(P) = \alpha_a$, so we can describe the power series purely in terms of Hasse-derivatives. We extend the Hasse-derivative to $\mathscr{F}[y]$ by first defining

$$D_y^{(b)} D_{t_l}^{(a)}(gy^j) := \binom{j}{b} D_{t_l}^{(a)}(g) y^{j-b}$$

and then by extending it linearly to $\mathscr{F}[y]$. This definition ensures that if we develop the polynomial $Q(y)$ in a power series in the variables t_l and $y - r_l$, then the coefficient of $t_l^a (y - r_l)^b$ is given exactly by $D_y^{(b)} D_{t_l}^{(a)}(Q(y))(P_l, r_l)$.

By the approximation theorem there exists $t \in \mathscr{F}$ such that $v_P(t) = 1$ for all $P \in \operatorname{Supp} D$. We will therefore for convenience assume from now on that $t_l = t$ does not depend on l. The $\binom{s+1}{2}$ equations coming from the condition that (P_l, r_l) is a zero of $Q(y)$ of multiplicity s can now be described as follows:

$$D_y^{(b)} D_t^{(a)}(Q(y))(P_l, r_l) = 0, \text{ for all } a, b \text{ with } a + b < s,$$

or equivalently

$$\sum_{i=b}^{\lambda} \binom{i}{b} r_l^{i-b} \sum_{j=1}^{l_i} q_{ij} D_t^{(a)}(g_{ij})(P_l) = 0, \tag{2.29}$$

for all $\binom{s+1}{2}$ pairs of nonnegative integers (a, b) such that $a + b < s$.

We would like to write these equations in matrix form

$$\mathbf{M} \begin{pmatrix} \mathbf{q}_0 \\ \vdots \\ \mathbf{q}_\lambda \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{2.30}$$

where $\mathbf{M}$ plays a similar role as the matrix $(\mathbf{M}_A|\mathbf{D_r}\mathbf{M}_{A-G})$ for equation (2.4). For $0 \le b \le s-1$ and $b \le i \le \lambda$, we therefore introduce the following $(s-b)n \times l_i$ matrix:

$$\mathbf{M}_i^{(i-b)} := \begin{pmatrix} g_{i1}(P_1) & \dots & g_{il_i}(P_1) \\ \vdots & & \vdots \\ D_t^{(s-1-b)}(g_{i1})(P_1) & \dots & D_t^{(s-1-b)}(g_{il_i})(P_1) \\ \vdots & & \vdots \\ g_{i1}(P_n) & \dots & g_{il_i}(P_n) \\ \vdots & & \vdots \\ D_t^{(s-1-b)}(g_{i1})(P_n) & \dots & D_t^{(s-1-b)}(g_{il_i})(P_n) \end{pmatrix} \tag{2.31}$$

and the $(s-b)n \times (s-b)n$ matrix

$$\mathbf{D}_i^{(b)} := \binom{i+b}{b} \begin{pmatrix} r_1^i & & & & & & & \\ & \ddots & & & & & & \\ & & r_1^i & & & & & \\ & & & \ddots & & & & \\ & & & & r_n^i & & & \\ & & & & & \ddots & & \\ & & & & & & r_n^i & \end{pmatrix}, \tag{2.32}$$

where every element r_l^i is repeated $s-b$ times on the diagonal. Using these definitions, we can then find the matrix $\mathbf{M}$ we are looking for. In other words, if we define $\mathbf{M}$ to be the matrix:

$$\left(\begin{array}{c|c|c|c|c|c} \mathbf{M}_0^{(0)} & \mathbf{D}_1^{(0)}\mathbf{M}_1^{(1)} & \cdots & \mathbf{D}_{s-1}^{(0)}\mathbf{M}_{s-1}^{(s-1)} & \cdots & \mathbf{D}_{\lambda}^{(0)}\mathbf{M}_{\lambda}^{(\lambda)} \\ \hline \mathbf{0} & \mathbf{M}_1^{(0)} & \cdots & \mathbf{D}_{s-2}^{(1)}\mathbf{M}_{s-1}^{(s-2)} & \cdots & \mathbf{D}_{\lambda-1}^{(1)}\mathbf{M}_{\lambda}^{(\lambda-1)} \\ \hline \vdots & \ddots & \ddots & \vdots & & \vdots \\ \hline \mathbf{0} & \cdots & \mathbf{0} & \mathbf{M}_{s-1}^{(0)} & \cdots & \mathbf{D}_{\lambda-s+1}^{(s-1)}\mathbf{M}_{\lambda}^{(\lambda-s+1)} \end{array}\right), \tag{2.33}$$

then we can reformulate equation (2.29) as matrix equation (2.30).

Example 2.39. In this example we show how to calculate the above equations in case of the Hermitian curve given by the equation $x_2^q + x_2 = x_1^{q+1}$ defined over $\mathbb{F}_{q^2}$. The function $t = x^{q^2} - x$ is a local parameter for all points on the curve different from T_∞. We will describe how to compute $D_t^{(a)}(f)$ for any function $f \in \mathscr{F}$. In the first place, Hasse derivatives satisfy the Leibniz rule:

$$D_t^{(a)}(fg) = \sum_{i=0}^{a} D_t^{(i)}(f) D_t^{(a-i)}(g)$$

and more general

$$D_t^{(a)}(f_1 \cdots f_m) = \sum_{i_1+\cdots+i_m=a} D_t^{(i_1)}(f_1) \cdots D_t^{(i_m)}(f_m).$$

Using this and the linearity of Hasse derivatives, we see that in order to describe them explicitly, it is enough to be able to calculate $D_t^{(a)}(x_1)$ and $D_t^{(a)}(x_2)$ for all natural numbers a.

We will now show how to calculate $D_t^{(a)}(x_1)$ recursively. We have that $D_t^{(0)}(x_1) = x_1$. Now suppose that $a > 0$ and that we know $D_t^{(\alpha)}(x_1)$ for all α between 0 and $a-1$. Using the equation $t = x_1^{q^2} + x_1$, we find that $D_t^{(a)}(x_1) = D_t^{(a)}(t) - D_t^{(a)}(x_1^{q^2})$. We have that $D_t^{(0)}(t) = t$, $D_t^{(1)}(t) = 1$ and $D_t^{(a)}(t) = 0$ if $a > 1$. Further using the general Leibniz rule, we find that $D_t^{(a)}(x_1^{q^2}) = \sum_{i_1+\cdots i_{q^2}=a} D_t^{i_1}(x_1) \cdots D_t^{(i_{q^2})}(x_1)$. If $i_j = a$ for some j, then remaining indices are zero implying that for this choice of indices we find the term $x_1^{a-1} D_t^{(a)}(x_1)$. By varying j between 1 and q^2, we see that there are exactly q^2 such terms. Thus these terms do not contribute to the sum. This means that $D_t^{(a)}(x_1) = D_t^{(a)}(t - x_1^{q^2})$ can be expressed as polynomial in $D_t^{(\alpha)}(x_1)$ for α varying between 0 and $a-1$.

It remains to show how to calculate $D_t^{(a)}(x_2)$ recursively. In the first place $D_t^{(0)}(x_2) = x_2$ and since $x_2^q + x_2 = x_1^{q+1}$, we also have that $D_t^{(a)}(x_2) = D_t^{(a)}(x_1^{q+1}) - D_t^{(a)}(x_2^q)$. We already know how to calculate $D_t^{(a)}(x_1^{q+1})$ recursively and similarly as above we can express $D_t^{(a)}(x_2^q)$ as a polynomial in $D_t^{(\alpha)}(x_2)$ with α between 0 and $a-1$. For future use, we state some explicit results for $q = 2$:

a	0	1	2	3	4	5
$D_t^{(a)}(x_1)$	x_1	1	0	0	1	0
$D_t^{(a)}(x_2)$	x_2	x_1^2	$x_1 + x_1^4$	1	x_1^8	0

Before continuing our discussion of equation (2.29), we will establish some facts on the matrices $\mathbf{M}_i^{(0)}$. We will think about them as generator matrices of certain codes that we will define now.

Definition 2.40. Let s be a natural number, $D = P_1 + \cdots + P_n$ as before and A be a divisor with support disjoint from $\operatorname{Supp} D$, but of arbitrary degree. Further, let $t \in \mathscr{F}$ be a local parameter for all $P \in \operatorname{Supp} D$ simultaneously. We define

$$\operatorname{Ev}_P^{(s)} : L(A) \to \mathbb{F}^s$$

$$f \mapsto (f(P), D_t^{(1)}(f)(P), \ldots, D_t^{(s-1)}(f)(P))$$

$$\operatorname{Ev}_D^{(s)} : L(A) \to \mathbb{F}^{sn}$$

$$f \mapsto (\operatorname{Ev}_{P_1}^{(s)}(f), \ldots, \operatorname{Ev}_{P_n}^{(s)}(f))$$

and

$$C_L^{(s)}(D, A) := \operatorname{Ev}_D^{(s)}(L(A)).$$

Note that if $s > 1$, the map $\operatorname{Ev}_P^{(s)}$ depends on the choice of the local parameter t. The point of the above definition is that the columns occurring in the matrix $\mathbf{M}_i^{(0)}$ are codewords in the code $C_L^{(s-i)}(D, A - iG)$. Moreover, we have that

$$\operatorname{rank} \mathbf{M}_i^{(0)} = \dim C_L^{(s-i)}(A - iG). \tag{2.34}$$

In order to define the analogue of the code $C_\Omega(D, A)$, we consider a differential $\omega \in \Omega(-sD + A)$. Locally at a point $P \in \operatorname{Supp} D$, one can then write $\omega = (\beta_s t^{-s} + \cdots + \beta_1 t^{-1} + \cdots)\, dt$. We can calculate β_i using residues, since $\beta_i = \operatorname{res}_P(t^{i-1}\omega)$. This motivates the following definition:

Definition 2.41. Let s, D, A and t be as in Definition 2.40. We define

$$\mathrm{Res}_P^{(s)} : \Omega(-sD + A) \to \mathbb{F}^s$$

$$\omega \mapsto (\mathrm{res}_P(\omega), \mathrm{res}_P(t\omega), \ldots, \mathrm{res}_P(t^{s-1}\omega)),$$

$$\mathrm{Res}_D^{(s)} : \Omega(-sD + A) \to \mathbb{F}^{sn}$$

$$\omega \mapsto (\mathrm{Res}_{P_1}^{(s)}(\omega), \ldots, \mathrm{Res}_{P_n}^{(s)}(\omega))$$

and

$$C_\Omega^{(s)}(D, A) := \mathrm{Res}_D^{(s)}(\Omega(-sD + A)).$$

If $s = 1$ it is well known that $C_L^{(s)}(D, A)$ and $C_\Omega^{(s)}(D, A)$ are dual to each other. We will now show that this also holds for arbitrary s. It is important that the choice of local parameter t is fixed when defining these codes.

Proposition 2.42. *We have that*

(1) $\dim C_L^{(s)}(D, A) = l(A) - l(-sD + A)$,
(2) $C_\Omega^{(s)}(D, A) = C_L^{(s)}(D, A)^\perp$.

Proof. Let $g \in L(A)$. We have that $\mathrm{Ev}_D^{(s)}(g) = (0, \ldots, 0)$ if and only if g has a zero of order at least s in every $P \in \mathrm{Supp}\, D$. This implies that the kernel of $\mathrm{Ev}_D^{(s)}$ is $L(-sD + A)$. This proves the first statement.

Now we prove the second statement. Let $\omega \in \Omega(-sD+A)$ and $g \in L(A)$. Locally at a $P \in \mathrm{Supp}\, D$, we can write $\omega = (\beta_s t^{-s} + \cdots + \beta_1 t^{-1} + \cdots)\, dt$ and $g = \alpha_0 + \alpha_1 t + \cdots + \alpha_{s-1} t^{s-1} + \cdots$. Then $\mathrm{Res}_P^{(s)}(\omega) = (\beta_1, \ldots, \beta_s)$ and $\mathrm{Ev}_P^{(s)}(g) = (\alpha_0, \ldots, \alpha_{s-1})$. The inner product $\langle \mathrm{Res}_P^{(s)}(\omega), \mathrm{Ev}_P^{(s)}(g) \rangle$ is exactly the coefficient of t^{-1} in the product $g\omega$. Therefore we have that

$$\langle \mathrm{Res}_P^{(s)}(\omega), \mathrm{Ev}_P^{(s)}(g) \rangle = \mathrm{res}_P(g\omega).$$

Also note that $g\omega \in \Omega(-sD)$. All in all, we can deduce that

$$\langle \mathrm{Res}_D^{(s)}(\omega), \mathrm{Ev}_D^{(s)}(g) \rangle = \sum_{i=0}^{n} \mathrm{res}_{P_i}(g\omega) = 0.$$

In the last equality, we used the residue theorem. This implies that $C_\Omega^{(s)}(D, A) \subset C_L^{(s)}(D, A)^\perp$.

The proposition now follows once we prove that

$$\dim C_\Omega^{(s)}(D, A) + \dim C_L^{(s)}(D, A) = sn.$$

However, similarly to the first statement, one can prove that $\dim C_{\Omega}^{(s)}(D,A) = \dim\Omega(-sD+A) - \dim\Omega(A)$. Therefore we have that

$$\begin{aligned}\dim C_{L}^{(s)}(D,A) + \dim C_{\Omega}^{(s)}(D,A) &= l(A) - l(-sD+A) + \\ \dim\Omega(-sD+A) - \dim\Omega(A) &= \deg(A) - \deg(-sD+A) = sn.\end{aligned}$$

We used Riemann-Roch's theorem to obtain the second equality. □

Recall that $l_i = l(A - iG)$. For convenience we also define

$$m_i := l(-(s-i)D + A - iG).$$

Combining the above proposition with equation (2.34), we find that

$$\operatorname{rank}\mathbf{M}_i^{(0)} = l_i - m_i. \tag{2.35}$$

Note that this implies that $\dim C_{L}^{(s)}(D,A) = l(A)$ if $\deg A < sn$. This is always the case in the setup of the list decoding algorithm.

We now have the right machinery to describe the analogue of the matrix $\mathbf{H}$ in Proposition 2.8 for the list-decoding case. Therefore we now give the following definition:

Definition 2.43. Let A and G be divisors as in the list decoding setup. Let b be an integer between 0 and $s-1$ and $\omega_1, \ldots, \omega_{(s-b)n}$ differential forms such that the vectors $\operatorname{Res}_D^{(s-b)}(\omega_i)$ with $1 \leq i \leq \dim C_{\Omega}^{(s-b)}(D, A - bG)$, form a basis of $C_{\Omega}^{(s-b)}(D, A - bG)$ and the vecors $\operatorname{Res}_D^{(s-b)}(\omega_1)$, …, $\operatorname{Res}_D^{(s-b)}(\omega_{(s-b)n})$ form a basis of $\mathbb{F}^{(s-b)n}$. Then we define the $(s-b)n \times (s-b)n$ matrix.

$$\mathbf{H}_b := \begin{pmatrix} \operatorname{Res}_D^{(s-b)}(\omega_1) \\ \vdots \\ \operatorname{Res}_D^{(s-b)}(\omega_{(s-b)n}) \end{pmatrix}$$

and for $0 \leq b \leq s-1$ and $b \leq i \leq \lambda$, the $(s-b)n \times l_i$ matrix

$$\mathbf{S}_i^{(i-b)} := \mathbf{H}_b \, \mathbf{D}_{i-b}^{(b)} \, \mathbf{M}_i^{(i-b)}.$$

Note that the matrices $\mathbf{H}_b$ are regular, since its rows form by choice of differentials a basis of $\mathbb{F}^{(s-b)n}$. We now obtain the following proposition.

Proposition 2.44. *The set of equations in (2.29) is row equivalent to the system*

$$\left(\begin{array}{c|c|c|c|c|c} \mathbf{S}_0^{(0)} & \mathbf{S}_1^{(1)} & \cdots & \mathbf{S}_{s-1}^{(s-1)} & \cdots & \mathbf{S}_\lambda^{(\lambda)} \\ \hline \mathbf{0} & \mathbf{S}_1^{(0)} & \cdots & \mathbf{S}_{s-1}^{(s-2)} & \cdots & \mathbf{S}_\lambda^{(\lambda-1)} \\ \hline \vdots & \ddots & \ddots & \vdots & & \vdots \\ \hline \mathbf{0} & \cdots & \mathbf{0} & \mathbf{S}_{s-1}^{(0)} & \cdots & \mathbf{S}_\lambda^{(\lambda-s+1)} \end{array}\right) \left(\begin{array}{c} \mathbf{q}_0 \\ \mathbf{q}_1 \\ \vdots \\ \mathbf{q}_\lambda \end{array}\right) = \left(\begin{array}{c} 0 \\ 0 \\ \vdots \\ 0 \end{array}\right). \tag{2.36}$$

Proof. The proposition follows after multiplying the b-th row of matrices in equation (2.33) with $\mathbf{H}_b$. □

The matrices $\mathbf{S}_0^{(0)}, \ldots, \mathbf{S}_{s-1}^{(0)}$ are independent of the received word $\mathbf{r}$ and by equation (2.35) we have

$$\operatorname{rank} \mathbf{S}_i^{(0)} = l_i - m_i. \tag{2.37}$$

If $l_i < (s-i)n$, this reduces to $\operatorname{rank} \mathbf{S}_i^{(0)} = l_i$. Similarly as in the proof of Corollary 2.9, we then have that if $l_i < (s-i)n$, then $\mathbf{S}_i^{(0)}$ can be written in the form

$$\mathbf{S}_i^{(0)} = \left(\begin{array}{c} \mathbf{0} \\ \hline \mathbf{B}_i^{(0)} \end{array}\right),$$

where $\mathbf{0}$ denotes the $(s-i)n - l_i \times l_i$ zero matrix. The $l_i \times l_i$ matrix $\mathbf{B}_i^{(0)}$ is regular, meaning that with Gaussian elimination, we can eliminate the variables $q_{i1}, \ldots, q_{il_i}$ in all rows different from those of $\mathbf{B}_i^{(0)}$. For $i = 0$ the situation is very simple, since the only rows in system (2.36) in which the variables $q_{01}, \ldots, q_{0l_0}$ occur, are the rows coming from $\mathbf{B}_0^{(0)}$. If $l_i \geq (s-i)n$, then we can eliminate $\operatorname{rank} \mathbf{S}_i^{(0)} = l_i - m_i$ variables among $q_{i1}, \ldots, q_{il_i}$.

All in all, we can simplify system (2.36) by eliminating $\sum_{i=0}^{s}(l_i - m_i)$ variables. This means that the remaining

$$\sum_{i=0}^{s} m_i + \sum_{i=s+1}^{\lambda} l_i$$

variables can be found by solving

$$\sum_{i=0}^{s}((s-i)n - l_i + m_i)$$

linear equations. This gives an in general significant reduction of the size of the original system.

Example 2.45. This example is a continuation of Example 2.38. In the formulation from Section 2.6, we needed to solve a linear system of 168 equations 171 variables in order to find an interpolation polynomial $Q(y)$. We have just seen however that we can reduce the size of the system. First we calculate the rank of the matrices $\mathbf{S}_i^{(0)}$ and find:

i	0	1	2	3	4	5
$\operatorname{rank}\mathbf{S}_i^{(0)}$	35	31	27	23	16	8

This means that we can eliminate 140 variables and equations thereby reducing the original system to a system of 28 equations in 31 variables. We can eliminate all 116 variables q_{ij} with $0 \leq i \leq 3$ and $1 \leq j \leq l_i$, since for $i \leq 3$ we have that $l_i < (s-i)n$. For $i = 4$ and $i = 5$, the situation is more complicated, but all we need to do is to compute the matrices $\mathbf{S}_4^{(0)}$ and $\mathbf{S}_5^{(0)}$ explicitly. In order to do this, we need to choose differentials as in Definition 2.43. Given a b between 0 and s, we can choose a basis for $\Omega(-(s-b)D + A - bG)$ with the desired properties as follows (recall $t = x_1 + x_1^4$):

$$\omega_i = \begin{cases} f_i\, dt/t^{s-b} & \text{if } 1 \leq i < (s-b)n, \\ f_{(s-b)n+1}\, dt/t^{s-b} & \text{if } i = (s-b)n. \end{cases}$$

Using this choice of differential, we can compute all matrices $\mathbf{S}_i^{(0)}$ explicitly. By our choice of bases, they have more structure than we indicated before. In the first place we find that $(\mathbf{B}_i^{(0)})_{pq} = 0$ if $p+q < l_i+1$ and $(\mathbf{B}_i^{(0)})_{pq} = 1$ if $p + q = l_i + 1$. This means that the Gaussian elimination steps needed to eliminate the q_{ij} (with $0 \leq i \leq 3$ and $1 \leq j \leq l_i$) are straightforward to do. We also find that the matrix $\mathbf{S}_4^{(0)}$ is equal to

$$\begin{pmatrix}
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&1&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&1&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&1\\
0&0&0&0&0&0&0&0&0&0&0&1&0&0&1&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&1&0&0&1&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&1&0&0&0&0&0&1&0&0&0\\
0&0&0&0&0&0&0&0&1&0&0&1&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&1&0&0&1&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&1&0&0&0&0&0&1&0&0&0&0&0&0\\
0&0&0&0&0&1&0&0&1&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&1&0&0&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&1&0&0&0&0&0&1&0&0&0&0&0&0&0&0&0\\
0&0&1&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&1&0&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
1&0&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&0
\end{pmatrix}.$$

Therefore we can eliminate the 16 variables q_{4j} with $1 \leq j \leq 15$ and $j = 17$. We also find that

$$\mathbf{S}_5^{(0)} = \begin{pmatrix}
0&0&0&0&0&0&0&0&1&0&0&0&0&0&0\\
0&0&0&0&0&0&1&0&0&1&0&0&0&0&0\\
0&0&0&0&0&1&0&0&0&0&0&0&0&0&1\\
0&0&0&0&1&0&0&1&0&0&0&0&0&1&0\\
0&0&0&1&0&0&1&0&0&0&0&0&1&0&0\\
0&0&1&0&0&0&0&0&0&0&0&1&0&0&0\\
0&1&0&0&1&0&0&0&0&0&1&0&0&1&0\\
1&0&0&0&0&0&0&0&1&0&0&0&0&0&0
\end{pmatrix},$$

enabling us to eliminate the 8 variables q_{5j} with $1 \leq j \leq 7$ and $j = 9$. What remains is to calculate the remaining 31 variables. Doing the elimination explicitly, we find that the vector consisting of these remaining 31 variables has to be in the kernel of the 28×31 matrix:

$$\left(\begin{array}{c|c} \mathbf{A}_1 & \mathbf{A}_2 \\ \hline \mathbf{A}_3 & \mathbf{A}_4 \end{array}\right),$$

with

$$\mathbf{A}_1 = \begin{pmatrix}
0 & 0 & 0 & 1 & \alpha & \alpha & \alpha & 0 & 0 & \alpha^2 & 1 & 0 & 1 & \alpha^2 & 1 \\
0 & 0 & \alpha^2 & \alpha & 1 & \alpha^2 & \alpha^2 & 0 & \alpha & \alpha^2 & 0 & \alpha^2 & \alpha^2 & 0 & \alpha \\
0 & \alpha^2 & 0 & \alpha^2 & 0 & 1 & 1 & \alpha & 1 & \alpha & 0 & \alpha^2 & 0 & 1 & \alpha^2 \\
0 & 0 & 0 & \alpha^2 & \alpha & 1 & \alpha & 1 & 1 & \alpha^2 & \alpha^2 & 0 & \alpha^2 & \alpha & \alpha \\
\alpha^2 & 0 & \alpha^2 & 1 & 0 & \alpha & \alpha & 1 & \alpha^2 & \alpha & \alpha^2 & 1 & \alpha^2 & 0 & 0 \\
0 & \alpha^2 & \alpha & 0 & \alpha^2 & 0 & \alpha^2 & 1 & \alpha^2 & 1 & 0 & 0 & 0 & 1 & \alpha^2 \\
0 & \alpha & 0 & \alpha & 1 & 1 & 0 & \alpha^2 & \alpha & \alpha^2 & 0 & 0 & 0 & \alpha^2 & 0 \\
\alpha^2 & 0 & 0 & 1 & 0 & \alpha^2 & 0 & 1 & \alpha & 0 & 1 & 0 & \alpha & 1 & 1 \\
\alpha & 0 & 0 & \alpha & \alpha^2 & \alpha^2 & 1 & 0 & 1 & \alpha & 0 & 0 & 0 & 0 & 0 \\
0 & \alpha & \alpha^2 & \alpha^2 & \alpha^2 & \alpha^2 & \alpha & \alpha & 1 & \alpha^2 & 0 & \alpha & 0 & \alpha^2 & 0 \\
0 & \alpha^2 & 0 & 0 & 0 & 0 & \alpha^2 & \alpha^2 & 0 & \alpha^2 & \alpha & 0 & \alpha & 1 & 1 \\
0 & 0 & \alpha & \alpha^2 & 1 & \alpha^2 & 0 & 1 & \alpha^2 & 0 & 0 & \alpha & 0 & 0 & \alpha \\
\alpha^2 & \alpha & 0 & 0 & \alpha^2 & 1 & \alpha^2 & 1 & \alpha & 1 & 0 & \alpha^2 & 0 & \alpha & \alpha^2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix},$$

$$\mathbf{A}_2 = \begin{pmatrix}
0 & 1 & 0 & \alpha^2 & \alpha^2 & 0 & \alpha & \alpha^2 & \alpha & \alpha^2 & \alpha^2 & \alpha & \alpha & 0 & 0 & 0 \\
1 & 0 & 0 & \alpha & \alpha^2 & 1 & \alpha^2 & \alpha^2 & 1 & \alpha & 1 & 1 & \alpha^2 & 0 & 0 & 0 \\
0 & 0 & \alpha & 0 & 1 & \alpha & \alpha & 0 & 1 & 1 & 0 & \alpha^2 & \alpha & 0 & 0 & 0 \\
0 & \alpha & 1 & \alpha^2 & \alpha & \alpha & \alpha^2 & 1 & \alpha & \alpha & 1 & 0 & 1 & 0 & 0 & 0 \\
\alpha & 0 & 1 & \alpha & 1 & 0 & 1 & \alpha & \alpha^2 & 0 & 0 & \alpha & \alpha^2 & 0 & 0 & 0 \\
\alpha & 0 & \alpha & 0 & \alpha & 0 & 1 & 0 & \alpha & 0 & 1 & 1 & 1 & 0 & 0 & \alpha \\
0 & 0 & 0 & 0 & \alpha & 0 & \alpha & 0 & \alpha & \alpha & \alpha & \alpha^2 & 0 & 0 & \alpha & 0 \\
\alpha & \alpha^2 & \alpha & \alpha & 0 & \alpha^2 & \alpha & \alpha^2 & \alpha^2 & \alpha & \alpha^2 & \alpha & \alpha & 0 & 0 & \alpha \\
\alpha^2 & 0 & \alpha^2 & 0 & \alpha^2 & \alpha^2 & 0 & \alpha^2 & 1 & \alpha & \alpha^2 & \alpha & 1 & \alpha & \alpha & 0 \\
1 & 0 & \alpha^2 & 0 & 1 & 1 & 1 & \alpha & \alpha & \alpha & 1 & 1 & \alpha^2 & 0 & \alpha & \alpha^2 \\
0 & \alpha^2 & 0 & \alpha & \alpha & 1 & \alpha^2 & \alpha & 1 & \alpha^2 & 1 & \alpha^2 & 1 & \alpha & \alpha^2 & \alpha \\
0 & 0 & \alpha & 0 & 0 & 1 & \alpha & \alpha & \alpha^2 & \alpha & \alpha & \alpha & 0 & 0 & \alpha & \alpha \\
0 & 0 & \alpha & 0 & 0 & \alpha & 0 & \alpha^2 & \alpha^2 & 0 & \alpha^2 & \alpha & \alpha^2 & \alpha^2 & \alpha & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \alpha & 0 & 0 & \alpha^2 & 1 & 0 & 0 & 0
\end{pmatrix},$$

$$\mathbf{A}_3 = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & \alpha^2 & 0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \alpha^2 & 0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 & \alpha^2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \alpha^2 & 0 & 0 & \alpha & 0 & \alpha^2 & \alpha & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \alpha^2 & \alpha & 0 & \alpha^2 & \alpha & \alpha^2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \alpha & 0 & \alpha^2 & 0 & \alpha^2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \alpha^2 & 0 & \alpha^2 & \alpha & \alpha^2 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \alpha & 0 & \alpha & \alpha^2 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha^2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha^2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha^2 & 0 & 0 & \alpha^2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha^2 & 0 & \alpha^2 & \alpha & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix},$$

and

$$\mathbf{A}_4 = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \alpha^2 & \alpha & \alpha^2 & \alpha & \alpha & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \alpha & \alpha^2 & 0 & \alpha^2 & \alpha & \alpha & \alpha & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha^2 & 1 & 1 & \alpha & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \alpha^2 & \alpha & \alpha^2 & 1 & \alpha^2 & 0 & \alpha^2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & \alpha & 0 & 1 & \alpha^2 & 1 & \alpha^2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \alpha & 0 & 0 & \alpha^2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \alpha & \alpha & 0 & \alpha^2 & \alpha & 1 & \alpha & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & \alpha^2 & 1 & \alpha^2 & \alpha & 0 & \alpha & 0 & 0 & 0 \\
0 & 0 & \alpha^2 & \alpha & 0 & 0 & \alpha & \alpha^2 & \alpha & \alpha^2 & 1 & 0 & \alpha^2 & 0 & 0 & 0 \\
\alpha^2 & \alpha & 0 & 0 & \alpha & \alpha^2 & \alpha^2 & \alpha^2 & \alpha & 1 & 1 & \alpha^2 & \alpha & 0 & 0 & 0 \\
\alpha & 0 & 0 & 0 & \alpha^2 & 0 & \alpha & \alpha & 1 & \alpha^2 & \alpha & \alpha^2 & \alpha & 0 & 0 & 0 \\
0 & \alpha^2 & \alpha & \alpha^2 & 0 & \alpha & \alpha^2 & \alpha^2 & \alpha & 1 & \alpha^2 & 1 & \alpha^2 & 0 & 0 & 0 \\
0 & \alpha & \alpha^2 & 0 & \alpha & 0 & \alpha & \alpha & 1 & \alpha^2 & \alpha & \alpha^2 & \alpha & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha^2 & \alpha^2 & \alpha & 0 & 0 & 0
\end{pmatrix}.$$

This matrix is much easier to handle than the original 168×171 matrix. Its kernel is 5-dimensional and one of the solutions is given by (only nonzero values are given, the rest of the 31 variables are zero):

q_{58}	q_{510}	q_{511}	q_{61}	q_{62}	q_{63}	q_{64}	q_{65}	q_{66}	q_{67}	q_{71}	q_{81}	q_{82}
1	α^2	α	1	α^2	α	α^2	α^2	1	α^2	1	α^2	α

Setting in these 31 values in system (2.36), we can then calculate the remaining 140 variables immediately and find the interpolation polynomial $Q(y)$ mentioned in Example 2.38.

2.8. Literature

In this section we give the reader a guide to the literature as well as give the source of some of the results used in this chapter. The goal of these references is not to be extensive and many more people have worked on these subjects than we refer to. The reader is asked to check the references in the material we refer to to obtain a more complete reference list.

All necessary algebraic geometry is covered by the books [7, 16]. The references [10, 11] survey the decoding of algebraic geometry codes up till 1995. Especially many references for and history about the basic algorithm, the order bound, and majority voting can be found there. Again we would like to stress that for a complete record the reader should consult the references in [10, 11].

We now give more detailed references for each section. Again we stress that for a complete record the reader should check the references in the mentioned literature. The basic algorithm presented in Sections 2.2 and 2.3 has been investigated in many papers. Usually the algorithm is presented for the code $C_\Omega(D, G)$, but we have chosen to adapt it to the code $C_L(D, G)$. The differences are small and the performance is the same, which can be expected since there exists a divisor H such that $C_L(D, G) = C_\Omega(D, H)$. The given basis for $L(G(k_\infty, k_1, \ldots, k_q))$ in Example 2.11 was calculated in [13]. The generalized order bound in Section 2.4 was described in [2]. It is an extension of the order bound. Independently, an extension of the order bound was discussed in [4]. The majority voting algorithm presented in Section 2.5 is an extension of the existing majority voting algorithm as presented in [10, 11] and is close to the algorithm described in [5].

The list-decoding algorithm from Section 2.6 was first described in [17] for Reed-Solomon codes and subsequently extended to algebraic geometry codes in [15]. In both articles the multiplicity parameter s is equal to 1. The concept of multiplicity occurs for the first time in [8], thus extending the original algorithm to all rates. For Hermitian codes, the root finding part was done using points of high degree in [12], for general algebraic geometry codes in [9]. The existence of a point (or equivalently: "place" in the language of function fields) of high enough degree follows from Corollary V.2.10 in [16]. Root finding using Hensel lifting was done for Reed-Solomon

codes in [14] and for algebraic geometry codes in [1]. The Hensel lifting algorithm we present is an extension to cover general algebraic geometry codes with multiplicity parameter s possibly larger than one. The syndrome reformulation of the list-decoding algorithm in this generality (i.e. $s > 1$) is new. In [3] it was done for Reed-Solomon and Hermitian codes. For $s = 1$ it was described in [14] in the case of Reed-Solomon codes. For some stated facts about Hasse-derivatives and Taylor series see Sections 1.3 and 2.5 in [7].

References

[1] D. Augot and L. Pecquet, A Hensel lifting to replace factorization in list-decoding of algebraic-geometric and Reed-Solomon codes, IEEE Transactions on Information Theory, vol. 46, pp. 2605–2614, 2000.

[2] P. Beelen, The order bound for general algebraic geometric codes, Finite Fields and their Applications 13, pp. 665–680, 2007.

[3] P. Beelen and T. Høholdt, List decoding using syndromes, to appear in proceedings of SAGA, Tahiti, 2007.

[4] C. Carvalho, C. Munuera, E. Silva, and F. Torres, Near orders and codes, IEEE Trans. Inf. Theory, vol. 53, pp. 1919–1924, May 2007.

[5] I. Duursma, Majority coset decoding, IEEE Trans. Inform. Theory, vol. 39, pp. 1067–1071, May 1993.

[6] P. Elias, List decoding for noisy channels, In: 1957-IRE WESCON Convention record (now IEEE), Pt. 2, pp. 94–104, 1957.

[7] D.M. Goldschmidt, Algebraic functions and projective curves, Springer, Berlin, 2003.

[8] V. Guruswami and M. Sudan, Improved decoding of Reed-Solomon and algebraic-geometric codes, IEEE Trans. Inform. Theory, vol. 45, pp. 1757–1767, Sept. 1999.

[9] V. Guruswami and M. Sudan, On representation of algebraic-geometric codes, IEEE Trans. Inform. Theory, vol. 47, pp. 1610–1613, May 2001.

[10] T. Høholdt, J.H. van Lint and R. Pellikaan, Volume I, Chapter 10 of Handbook of coding theory, V.S. Pless and W.C. Huffman (eds.), Elsevier, Amsterdam, 1998.

[11] T. Høholdt and R. Pellikaan, On the decoding of algebraic-geometric codes, IEEE Trans. Inf. Theory, vol. 41, pp. 1589–1614, November 1995.

[12] T.Høholdt and R.R. Nielsen, Decoding Hermitian codes with Sudan's algorithm, In: Proc. AAECC-13 (Lecture Notes on Computer Science), Berlin, Germany: Springer-Verlag, pp. 260–270, 1999.

[13] H. Maharaj, G.L. Matthews and G. Pirsic, Riemann-Roch spaces of the Hermitian function field with applications to algebraic geometry codes and low-discrepancy sequences, Journal of Pure and Applied Algebra, vol. 195, pp. 261–280, 2005.

[14] R.M. Roth and G. Ruckenstein, Efficient decoding of Reed-Solomon codes

beyond half the minimum distance, IEEE Trans. Inform. Theory, vol. 46, pp. 246–257, Jan. 2000.
[15] M.A. Shokrollahi and H. Wasserman, List decoding of algebraic-geometric codes, IEEE Trans. Inform. Theory, vol. 45, pp. 432–437, March. 1999.
[16] H. Stichtenoth, Algebraic function fields and codes, Springer, Berlin, 1993.
[17] M. Sudan, Decoding of Reed-Solomon codes beyond the error-correcting bound, J. Compl., vol. 13, pp. 180–193, 1997.

Chapter 3

The Key Equation for One-Point Codes

Michael E. O'Sullivan and Maria Bras-Amorós*

Department of Mathematics and Statistics
San Diego State University
5500 Campanile Drive San Diego, CA 92182-7720, USA
mosulliv@sciences.sdsu.edu
Departament dEnginyeria Informàtica i Matemàtiques
Universitat Rovira i Virgili
Avinguda Països Catalans, 26
43007 Tarragona, Catalonia, Spain
maria.bras@urv.cat

For Reed-Solomon codes, the key equation relates the syndrome polynomial—computed from the parity check matrix and the received vector—to two unknown polynomials, the locator and the evaluator. The roots of the locator polynomial identify the error positions. The evaluator polynomial, along with the derivative of the locator polynomial, gives the error values via the Forney formula. The Berlekamp-Massey algorithm efficiently computes the two unknown polynomials.

This chapter shows how the key equation, the Berlekamp-Massey algorithm, the Forney formula, and another formula for error evaluation due to Horiguchi all generalize in a natural way to one-point codes. The algorithm presented here is based on Kötter's adaptation of Sakata's algorithm.

Contents

*This work was partly supported by the Spanish Ministry of Education through projects TSI2007-65406-C03-01 "E-AEGIS" and CONSOLIDER CSD2007-00004 "ARES", and by the Government of Catalonia under grant 2005 SGR 00446.

3.1. Introduction

For Reed-Solomon codes, the key equation relates the syndrome polynomial—computed from the parity check matrix and the received vector—to two unknown polynomials, the locator and the evaluator. The exact formulation of the key equation has evolved since Berlekamp's introduction of the term [2]. There are also key equations for other algorithms, such as Sugiyama et al [42], and Berlekamp-Welch [44]. The goal of this chapter is to show that the key equation, the Berlekamp-Massey algorithm and the error evaluation formulas of Forney and Horiguchi [18] all generalize to one-point codes. An important aspect of the generalization is to treat the ideal of error locator polynomials as a module over a polynomial ring in one variable, which is essentially the approach Kötter used in his version of the Berlekamp-Massey-Sakata algorithm [21]. The chapter is divided into three main sections, Reed-Solomon codes, Hermitian codes, and one-point codes. We have attempted to make each section as self-contained as possible, and to minimize the mathematical background required.

The section on Reed-Solomon codes gives a concise treatment of the key-equation, the Berlekamp-Massey algorithm, and the error evaluation formulas in a manner that will generalize easily to one-point codes. Two aspects of our approach are atypical, though certainly not new. First, the locator polynomial vanishes at the error positions—as opposed to the usual definition which uses the reciprocals of the positions—because this is more

natural in the context of algebraic geometry codes. Second, the syndrome is a rational polynomial—rather than a polynomial—because this is in accord with the duality of codes on algebraic curves. Theorem 3.7 gives a very formal statement of the properties satisfied by the intermediate polynomials computed in the Berlekamp-Massey algorithm. Analogous results are established in the later sections for Hermitian and one-point codes. At the end of the section on Reed-Solomon codes we briefly discuss the usual formulation of the key equation—see for example [3, 37]—and the connections with the Euclidean algorithm and the algorithm of Sugiyama et al. There are also interesting connections to the Berlekamp-Welch algorithm and to the list decoding algorithm of Lee and O'Sullivan [24], but these are not developed here.

The section on Hermitian codes requires little if any background in algebraic geometry, and only minimal familiarity with the algebra of polynomial rings and Gröbner bases. The presentation of this section closely parallels that of the section on Reed-Solomon codes, so that overall similarity between the two as well as the new complexities are as clear as possible. The locator polynomial is replaced with the ideal of polynomials vanishing at the error locations, and the problem is to find several locator polynomials of minimal degree, one for each congruence class modulo q, where the field size is q^2. The syndrome is again a rational polynomial, and the property of a locator is that its product with the syndrome eliminates the denominator, giving a polynomial. The product of the locator and the syndrome also may be used for error evaluation. Kötter's algorithm is essentially q Berlekamp-Massey algorithms operating in parallel, and the only place in which the algebra of the curve is used is in the computation of recursions of candidate locator polynomials with the syndrome. The Forney formula and Horiguchi formula for error evaluation are simple, but not obvious, generalizations of those for Reed-Solomon codes.

The section on one-point codes shows that the decoding algorithms and formulas for Hermitian codes need only minor modification to apply to general one-point codes. The focus of this section is not reproving the decoding results in the more general setting; instead, it is to establish the algebraic structure that makes the algorithms work. In particular, we will need to use differentials, residues of differentials, and duality with respect to the residue map. This section does require the theory of curves and algebraic function fields, but we have tried to build the exposition using a small number of key results as a base. The treatment is based on O'Sullivan [32, 33], with, we hope, improvements in exposition.

3.2. The key equation for Reed-Solomon codes

In this section, we briefly discuss Reed-Solomon codes, set up the decoding problem and introduce the locator and evaluator polynomials. The syndrome is defined as a rational polynomial, but it may also be seen as a power series. We then present the key equation and the Berlekamp-Massey algorithm in a form that we will generalize to codes from algebraic curves. We derive Horiguchi's formula for error evaluation, which removes the need to compute the error evaluator polynomial. Finally, we explore connections with the Euclidean algorithm.

3.2.1. *Reed-Solomon codes*

Let $\mathbb{F}_q$ be the finite field of q elements. Given n different elements $\alpha_1, \ldots, \alpha_n$ of $\mathbb{F}_q$, define the map $\mathrm{ev} : \mathbb{F}_q[x] \to \mathbb{F}_q^n$, $f \mapsto (f(\alpha_1), \ldots, f(\alpha_n))$. The *generalized Reed-Solomon code* $GRS(\bar{\alpha}, k)$, where $\bar{\alpha} = (\alpha_1, \ldots, \alpha_n)$, is defined as the image by ev of the polynomials in $\mathbb{F}_q[x]$ with degree at most $k-1$. It has generator matrix

$$\begin{pmatrix} 1 & 1 & \ldots & 1 \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \ldots & \alpha_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_1^{k-1} & \alpha_2^{k-1} & \ldots & \alpha_n^{k-1} \end{pmatrix}$$

It is well known (see for instance [37, §5.1]) that the parity check matrix of $GRS(\bar{\alpha}, k)$ is then

$$\begin{pmatrix} 1 & 1 & \ldots & 1 \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \ldots & \alpha_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_1^{n-k-1} & \alpha_2^{n-k-1} & \ldots & \alpha_n^{n-k-1} \end{pmatrix} \begin{pmatrix} \beta_1 & 0 & \ldots\ldots & 0 \\ 0 & \beta_2 & 0 & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0 & \ddots & 0 \\ 0 & \ldots\ldots & 0 & \beta_n \end{pmatrix}$$

for some $\beta_1, \beta_2, \ldots, \beta_n \in \mathbb{F}_q$. That is, $(c_1, c_2, \ldots, c_n)$ is in $GRS(\bar{\alpha}, k)$ if and only if $(c_1\beta_1, c_2\beta_2, \ldots, c_n\beta_n)$ is in $GRS^{\perp}(\bar{\alpha}, n-k)$.

If the field size is q and $n = q-1$ then it is said to be a *conventional Reed-Solomon code* or just Reed-Solomon code and we denote it by $RS(k)$.

In this case it can be proven that $\beta_i = \alpha_i$. So the parity check matrix is

$$\begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \alpha_1^2 & \alpha_2^2 & \dots & \alpha_n^2 \\ \alpha_1^3 & \alpha_2^3 & \dots & \alpha_n^3 \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_1^{n-k} & \alpha_2^{n-k} & \dots & \alpha_n^{n-k} \end{pmatrix}.$$

3.2.2. *Polynomials for decoding*

Suppose that a word $c \in GRS^{\perp}(\bar{\alpha}, n-k)$ is transmitted and that the vector u is received. The vector $e = u - c$ is the error vector. We assume that e has $t \leq \frac{n-k}{2}$ non-zero positions. We will use c, u, e and t throughout this section. The decoding task is to recover e from u and thereby get $c = u - e$.

We define the *error locator polynomial* associated to e as

$$f^e = \prod_{j:e_j \neq 0} (x - \alpha_j)$$

and the *error evaluator polynomial* as

$$\varphi^e = \sum_{j:e_j \neq 0} e_j \prod_{\substack{k:e_k \neq 0 \\ k \neq j}} (x - \alpha_k).$$

The utility of the error locator polynomial and the error evaluator polynomial is that the error positions can be identified as the indices j such that $f^e(\alpha_j) = 0$ and the error values can be computed by the so-called Forney formula given in the next lemma, whose verification is straightforward.

Lemma 3.1. *If $e_j \neq 0$ then $e_j = \frac{\varphi^e(\alpha_j)}{f^{e\prime}(\alpha_j)}$.*

Another useful fact about f^e and φ^e is that from the received vector we know the first coefficients of the power series in $\frac{1}{x}$ obtained when dividing φ^e by f^e. This is shown in the next lemma.

Lemma 3.2. *$\frac{\varphi^e}{f^e} = \frac{1}{x}\left(s_0 + \frac{s_1}{x} + \frac{s_2}{x^2} + \cdots\right)$, where $s_a = \sum_{j=1}^{n} e_j \alpha_j^a$. In particular, for $a \leq n-k-1$, $s_a = \sum_{j=1}^{n} u_j \alpha_j^a$.*

Proof.

$$\begin{aligned}\frac{\varphi^e}{f^e} = \sum_{j=1}^{n} \frac{e_j}{x-\alpha_j} &= \frac{1}{x}\sum_{j=1}^{n} \frac{e_j}{1-\frac{\alpha_j}{x}} \\ &= \frac{1}{x}\sum_{j=1}^{n} e_j \sum_{a=0}^{\infty} \left(\frac{\alpha_j}{x}\right)^a \\ &= \frac{1}{x}\sum_{a=0}^{\infty} \frac{1}{x^a} \sum_{j=1}^{n} e_j \alpha_j^a \\ &= \frac{1}{x}\sum_{a=0}^{\infty} \frac{s_a}{x^a}\end{aligned}$$

Looking at the parity check matrix of $GRS(\bar{\alpha}, n-k)^{\perp}$, it can be deduced that for $a \leq n-k-1$, $\sum_{j=1}^{n} c_j \alpha_j^a = 0$. Hence, $s_a = \sum_{j=1}^{n} e_j \alpha_j^a = \sum_{j=1}^{n} (u_j - c_j)\alpha_j^a = \sum_{j=1}^{n} u_j \alpha_j^a$. □

Definition 3.3. For a vector e, the *syndrome* of e is $S = \frac{\varphi^e}{f^e}$. The *syndrome of order* a is $s_a = \sum_{j=1}^{n} u_j \alpha_j^a$.

Just as any element of the field $\mathbb{F}_q(x)$ may be written as a Laurent series in x, any $h \in \mathbb{F}_q(x)$ also may be written as a Laurent series in $1/x$, $h = \sum_{a \leq d} h_a x^a$ for some $d \in \mathbb{Z}$. If h_d is nonzero in this expression, we say the *degree* of h is d, and if $h_d = 1$ we say that h is *monic*. Notice that $h \in \mathbb{F}_q[x]$ if and only if $h_a = 0$ for all $a < 0$ and that our definition of degree coincides with the usual one on $\mathbb{F}_q[x]$. Henceforth, we will not use the form for h given above. Instead we will write Laurent series in $1/x$ in the form $h = \frac{1}{x}\sum_a h_a x^{-a}$. It is understood that the sum is over all integers $a \geq -d-1$ where d is the degree of h. In this form, h is a polynomial when $h_a = 0$ for all $a \geq 0$. As an example, the syndrome is $S = \frac{1}{x}\sum_{a \geq 0} s_a x^{-a}$. Its degree is -1, unless $s_0 = 0$.

Lemma 3.4. *Let f be a polynomial and let $\alpha \in \mathbb{F}_q$. If the Laurent series in $\frac{1}{x}$ given by $\frac{f}{x-\alpha}$ has no term of degree -1 then $f(\alpha) = 0$.*

Proof. There exists $g \in \mathbb{F}_q[x]$ such that $f(x) = f(\alpha) + (x-\alpha)g(x)$. Then

$$\begin{aligned}\frac{f(x)}{x-\alpha} &= \frac{f(\alpha)}{x-\alpha} + g(x) \\ &= g(x) + \frac{f(\alpha)}{x}\left(1 + \frac{\alpha}{x} + \left(\frac{\alpha}{x}\right)^2 + \cdots\right) \\ &= g(x) + \frac{f(\alpha)}{x} + \frac{\alpha f(\alpha)}{x^2} + \frac{\alpha^2 f(\alpha)}{x^3} + \cdots\end{aligned}$$

If the term of degree -1 is zero, then $f(\alpha) = 0$. □

Proposition 3.5. *If fS has no terms of degrees $-1, -2, \ldots, -t$ then f is a multiple of f^e. In particular, if fS is a polynomial then f is a multiple of f^e.*

Proof. Suppose fS has no terms of degrees $-1, -2, \ldots, -t$. Suppose $e_j \neq 0$ and let

$$g(x) = \prod_{\substack{k:e_k \neq 0 \\ k \neq j}} (x - \alpha_k).$$

Note that $\deg g = t - 1$ and so fgS has no term of degree -1. Now,

$$\begin{aligned}fgS &= \sum_{k:e_k \neq 0} \frac{e_k f g}{x - \alpha_k} \\ &= e_j \frac{fg}{x - \alpha_j} + \sum_{\substack{k:e_k \neq 0 \\ k \neq j}} e_k f \frac{g}{x - \alpha_k}.\end{aligned}$$

Since fgS has no term of degree -1 and the right term in the previous sum is a polynomial, we deduce that $\frac{fg}{x-\alpha_j}$ has no term of degree -1. By the previous lemma, $x - \alpha_j$ must divide f. Since j was chosen arbitrarily such that $e_j \neq 0$, we conclude that f^e must divide f. □

3.2.3. *The key equation and the Berlekamp-Massey algorithm*

We now present the version of the Berlekamp-Massey algorithm that will be our model for generalization to codes from algebraic curves. The Berlekamp-Massey algorithm finds the minimal solution to the key equation.

Definition 3.6. We will say that polynomials f, φ satisfy the *key equation* for syndrome S when $fS = \varphi$.

The Berlekamp-Massey Algorithm

Initialize: $\begin{pmatrix} f^{(0)} & \varphi^{(0)} \\ g^{(0)} & \psi^{(0)} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

Algorithm: For $m = 0$ to $n - k - 1$,

$d = \deg f^{(m)}$

$\mu = \sum_{a=0}^{d} f_a^{(m)} s_{a+(m-d)}$

$p = 2d - m - 1$

$$U^{(m)} = \begin{cases} \begin{pmatrix} 1 & -\mu x^p \\ 0 & 1 \end{pmatrix} & \text{if } \mu = 0 \text{ or } p \geq 0 \\ \begin{pmatrix} x^{-p} & -\mu \\ 1/\mu & 0 \end{pmatrix} & \text{otherwise.} \end{cases}$$

$$\begin{pmatrix} f^{(m+1)} & \varphi^{(m+1)} \\ g^{(m+1)} & \psi^{(m+1)} \end{pmatrix} = U^{(m)} \begin{pmatrix} f^{(m)} & \varphi^{(m)} \\ g^{(m)} & \psi^{(m)} \end{pmatrix}$$

Output: $f^{(n-k)}$, $\varphi^{(n-k)}$.

Notice that this algorithm uses only the syndromes of order up to $n - k - 1$ and these are exactly the syndromes that can be computed from the received vector. We may think of $f^{(m)}, \varphi^{(m)}$ and also $g^{(m)}, \psi^{(m)}$ as approximate solutions of the key equation. The algorithm takes a linear combination of two approximate solutions to create a better approximation.

Theorem 3.7. *For all $m \geq 0$,*

(1) $f^{(m)}$ is monic of degree at most m.
(2) $\deg(f^{(m)}S - \varphi^{(m)}) \leq -m + \deg f^{(m)} - 1$. In particular, $f^{(m)}S$ has no terms in degrees $-1, -2, \ldots, -m + \deg f^{(m)}$.
(3) $g^{(m)}S - \psi^{(m)}$ is monic of degree $-\deg f^{(m)}$.
(4) $\deg(g^{(m)}) \leq m - \deg f^{(m)}$.

Proof. We will proceed by induction on m. It is easy to verify the case $m = 0$. Assume the statements are satisfied at step m. Let $d = \deg f^{(m)}$. Notice that $d \leq m$ by item (1), and μ is the coefficient of x^{d-m-1} in $f^{(m)}S$. Furthermore, since $d - m - 1 < 0$, and $\varphi^{(m)}$ is a polynomial, μ is the coefficient of x^{d-m-1} in $f^{(m)}S - \varphi^{(m)}$.

If $\mu = 0$, then the algorithm retains the polynomials from the mth iteration, *e.g.* $f^{(m+1)} = f^{(m)}$. The induction hypothesis immediately gives items (1), (3), and (4) of the theorem, and item (2) follows from $\mu = 0$.

Consider the case when $p = 2d - m - 1 \geq 0$ and $\mu \neq 0$. The algorithm sets $f^{(m+1)} = f^{(m)} - \mu x^p g^{(m)}$. By the induction hypothesis,

$$\deg(x^p g^{(m)}) \leq 2d - m - 1 + m - d = d - 1,$$

so $\deg(f^{(m+1)}) = \deg(f^{(m)}) = d < m$ and $f^{(m+1)}$ is monic, so item (1) holds. Now,

$$f^{(m+1)}S - \varphi^{(m+1)} = (f^{(m)}S - \varphi^{(m)}) - \mu x^p (g^{(m)}S - \psi^{(m)}).$$

The degree of each term is $d-m-1$ and the coefficients of x^{d-m-1} cancel. Thus $\deg(f^{(m+1)}S - \varphi^{(m+1)}) \leq -(m+1) + \deg(f^{(m+1)}) - 1$, as required. This proves item (2). Items (3) and (4) are trivial in this case, since $g^{(m+1)} = g^{(m)}$ and $\varphi^{(m+1)} = \varphi^{(m)}$.

Finally, consider the case when $p = 2d - m - 1 < 0$, in which $f^{(m+1)} = x^{-p} f^{(m)} - \mu g^{(m)}$. By computing the degrees of each summand, one can see that $f^{(m+1)}$ is monic of degree $m + 1 - d \leq m + 1$ as claimed in item (1). We have

$$f^{(m+1)}S - \varphi^{(m+1)} = x^{-p}(f^{(m)}S - \varphi^{(m)}) - \mu(g^{(m)}S - \psi^{(m)}).$$

The degree of each term is $-d$ and the coefficients cancel. Thus $\deg(f^{(m+1)}S - \varphi^{(m+1)}) < -d$. We can see that item (2) holds since $-(m+1) + \deg(f^{(m+1)}) - 1 = -d - 1$. The algorithm sets $g^{(m+1)} = \mu^{-1} f^{(m)}$ and $\psi^{(m+1)} = \mu^{-1}\varphi^{(m)}$. Item (4) holds since $(m+1) - \deg(f^{(m+1)}) = d = \deg(g^{(m)})$. Item (3) holds since $g^{(m+1)}S - \psi^{(m+1)} = \mu^{-1}(f^{(m)}S - \varphi^{(m)})$, which has degree exactly $d - m - 1 = -\deg(f^{(m+1)})$ and it is monic. □

The next few results show that the algorithm produces the minimal solution to the key equation, f^e and $f^e S$.

Lemma 3.8. *For all m, $\deg f^{(m)} \leq t$.*

Proof. Consider $f^e g^{(m)} S - f^e \psi^{(m)}$. This is a polynomial since $f^e S$, $g^{(m)}$, $\psi^{(m)}$ and f^e are. Since the degree of f^e is t, we have $\deg(f^e g^{(m)} S - f^e \psi^{(m)}) = t - \deg f^{(m)}$, using item (3) in Theorem 3.7. Thus $t - \deg f^{(m)} \geq 0$. □

Lemma 3.9. *When $m \geq 2t$, $f^{(m)} = f^e$ and $\varphi^{(m)} = \varphi^e$.*

Proof. Theorem 3.7 tells us that $f^{(m)}S$ has no terms of degree $-1, \ldots, -m + \deg(f^{(m)})$. From the previous lemma, if $m \geq 2t$ then $-m + \deg(f^{(m)}) \leq -2t + t = -t$. Thus, $f^{(m)}S$ has no terms of degree $-1, \ldots, -t$. By Proposition 3.5, $f^{(m)}$ must be a multiple of f^e; by Theorem 3.7 it is monic; and, by the preceding lemma, its degree is at most t. Thus, it must be equal to f^e.

On the other hand, $\deg(f^e S - \varphi^{(m)}) \leq -m + t - 1 \leq -t - 1 < 0$. Since both $f^e S$ and $\varphi^{(m)}$ are polynomials, this means $\varphi^{(m)} = f^e S = \varphi^e$. □

Proposition 3.10. *If $t \leq \frac{d-1}{2}$ then the previous algorithm outputs f^e and φ^e.*

Proof. If $t \leq \frac{d-1}{2}$ then $n - k \geq d - 1 \geq 2t$ and the result follows from Lemma 3.9. □

3.2.4. *Error evaluation without the evaluator polynomial*

We now derive a formula for error evaluation that does not use the error evaluator polynomial, and thereby removes the need for computing it. It is called the Horiguchi-Kötter algorithm in [3] and appears in [18, 20].

From the algorithm it is clear that

$$\begin{pmatrix} f^{(m)} & \varphi^{(m)} \\ g^{(m)} & \psi^{(m)} \end{pmatrix} = U^{(m-1)}U^{(m-2)} \ldots U^{(1)}U^{(0)} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Taking determinants, since each $U^{(m)}$ has determinant 1, we get

$$f^{(m)}\psi^{(m)} - g^{(m)}\varphi^{(m)} = -1. \tag{3.1}$$

In particular, when $m \geq 2t$,

$$f^e\psi^{(m)} - g^{(m)}\varphi^e = -1.$$

Let j be such that $e_j \neq 0$. Evaluating at α_j we get $g^{(m)}(\alpha_j)\varphi^e(\alpha_j) = 1$ and so $\varphi^e(\alpha_j) = (g^{(m)}(\alpha_j))^{-1}$. Using Lemma 3.1 we can establish the following proposition.

Proposition 3.11. *For $m \geq 2t$ and $g^{(m)}$ as in the Berlekamp-Massey algorithm, if $e_j \neq 0$ then*

$$e_j = (f^{e\prime}(\alpha_j)g^{(m)}(\alpha_j))^{-1}.$$

The last proposition tells us that in the Berlekamp-Massey algorithm we do not need to multiply $U^{(m)}$ by all the matrix

$$\begin{pmatrix} f^{(m)} & \varphi^{(m)} \\ g^{(m)} & \psi^{(m)} \end{pmatrix}$$

but by the vector

$$\begin{pmatrix} f^{(m)} \\ g^{(m)} \end{pmatrix}.$$

Then the initialization step will be

$$\begin{pmatrix} f^{(0)} \\ g^{(0)} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and the updating step will be

$$\begin{pmatrix} f^{(m+1)} \\ g^{(m+1)} \end{pmatrix} = U^{(m)} \begin{pmatrix} f^{(m)} \\ g^{(m)} \end{pmatrix}.$$

3.2.5. *Connections to the Euclidean algorithm*

Suppose that all the α_i defining the GRS code are nth roots of unity. In particular, we could demand that none of the α_i are zero and take $n = q-1$. From the definition of S it is easy to see that $s_a = s_{n+a}$ for all $a \geq 0$, and consequently,

$$S(x^n - 1) = s_0 x^{n-1} + s_1 x^{n-2} + \cdots + s_{n-2}x + s_{n-1}.$$

Call this polynomial $\overline{S}$. We might alter our definition of the key equation to say that f, and φ are solutions when $f\overline{S} = \varphi(x^n - 1)$. That is

$$f(s_0 x^{n-1} + s_1 x^{n-2} + \cdots + s_{n-2}x + s_{n-1}) = \varphi(x^n - 1). \tag{3.2}$$

Of course, the solution set is the same as for our original equation, and f^e, φ^e are the minimal degree solutions such that f^e is monic. The result analogous to Theorem 3.7 states that $\deg(f^{(m)}\bar{S} - \varphi^{(m)}) \leq n - 1 - m + \deg f^{(m)}$ and $g^{(m)}S - \psi^{(m)}$ is monic of degree $n - \deg f^{(m)}$. When the weight of e is t, $f^{(2t)} = f^e$ and $\varphi^{(2t)} = \varphi^e$ give the least common multiple of $\overline{S}$ and $x^n - 1$; the lcm is $f^e S = \varphi^e(x^n - 1)$.

For a linear algebra perspective, write $f = f_0 + f_1 x + \cdots + f_t x^t$. The key equation requires that $\sum_{i=0}^{t} f_i s_{i+a} = 0$ for all $0 \leq a \leq n - t - 1$. Setting $f_t = 1$, there are t unknowns, $f_0, \ldots, f_{t-1}$, so the t equations where $a = 0, \ldots, t-1$, are enough to determine the coefficients of f. Thus we need to know the syndromes s_0 to s_{2t-1} to compute f^e. This verifies that

the Berlekamp-Massey algorithm used with a code of redundancy $2t$ can correct t errors.

Equation (3.2) leads to a relationship between the Berlekamp-Massey algorithm and the Euclidean algorithm. Let $m_0, m_1, \ldots, m_r$ be the iterations of the algorithm in which $p^{(m)} < 0$ and let $m_{r+1} = 2t$ where t is the weight of e. One can check that $p^{(m)} < -p^{(m_\ell)}$ for all $m_\ell < m \leq m_{\ell+1}$. For each $\ell = 0, \ldots, r$, let

$$V^{(\ell)} = U^{(m_{\ell+1}-1)} \cdots U^{(m_\ell+1)} U^{(m_\ell)}.$$

Then

$$V^{(\ell)} = \begin{pmatrix} q_\ell & -\mu^{(m_\ell)} \\ (\mu^{(m_\ell)})^{-1} & 0 \end{pmatrix}$$

where q_ℓ is a monic polynomial of degree $p^{(m_\ell)}$. Define recursively,

$$\begin{pmatrix} A_0 \\ B_0 \end{pmatrix} = \begin{pmatrix} \overline{S} \\ x^n - 1 \end{pmatrix} \qquad \text{and}$$

$$\begin{pmatrix} A_{\ell+1} \\ B_{\ell+1} \end{pmatrix} = V^{(\ell)} \begin{pmatrix} A_\ell \\ B_\ell \end{pmatrix}$$

so that $A_{\ell+1} = -\mu^{(m_\ell)} B_\ell + q_\ell A_\ell$ and $B_{\ell+1} = (\mu^{(m_\ell)})^{-1} A_\ell$. Rearranging, we get $B_{\ell+1} = \frac{-\mu^{(m_{\ell-1})}}{\mu^{(m_\ell)}} (B_{\ell-1} - q_{\ell-1} B_\ell)$. This is a variant of the classical Euclidean algorithm for computing the greatest common divisor with the modification that the remainders are all monic. We will sketch the main points and leave verification of the details to the reader.

Notice that $A_\ell = f^{(m_\ell-1)}\overline{S} - \varphi^{(m_\ell-1)}(x^n - 1)$ and similarly $B_\ell = g^{(m_\ell-1)}\overline{S} - \psi^{(m_\ell-1)}(x^n - 1)$. From the discussion after (3.2), $\deg B_\ell = n - \deg f^{(m_\ell-1)}$. Referring to the Berlekamp-Massey algorithm, $\deg f^{(m_{l+1}-1)} = \deg f^{(m_\ell)} < \deg f^{(m_\ell-1)}$ so we have $\deg B_{\ell+1} < \deg B_\ell$ and the sequence of B_ℓ does indeed satisfy the requirements of the Euclidean algorithm with monic quotients.

At the final iteration, $m_{r+1} = 2t$, $f^{(m_{r+1})} = f^e$ and $\varphi^{(m_{r+1})} = \varphi^e$ so that $A_{r+1} = f^e\overline{S} - \varphi^e(x^n - 1) = 0$. As noted earlier, $f^e\overline{S}$ is a constant multiple of the lcm of $\overline{S}$ and $x^n - 1$. We also have $B_{r+1} = g^{(2t)}\overline{S} - \psi^{(2t)}(x^n - 1)$ is the monic greatest common divisor of $\overline{S}$ and $(x^n - 1)$, namely $\prod_{i:e_i=0}(x - \alpha_i)$.

Thus we see that the Berlekamp-Massey algorithm breaks each division of this version of the Euclidean algorithm into several steps, one for each subtraction of a monomial multiple of the divisor. The Berlekamp-Massey algorithm is also more efficient than the Euclidean algorithm, because it

never computes the B_ℓ. It takes advantage of the fact that $B_0 = x^n - 1$ is very sparse, and just computes the critical coefficients $\mu^{(m)}$ via the polynomials $f^{(m)}$ and $\overline{S}$.

Berlekamp's formulation of the key equation was different from the one presented here. To obtain his formulation, let

$$\sigma^e = x^t f^e\left(\frac{1}{x}\right) = \prod_{k:e_k\neq 0}(1-\alpha_k x)$$

$$\omega^e = x^{t-1}\varphi\left(\frac{1}{x}\right) = \sum_{j:e_j\neq 0} e_j \prod_{\substack{k:e_k\neq 0\\ k\neq j}}(1-\alpha_k x).$$

These polynomials are $\Lambda(x)$ and $\Gamma(x)$ respectively in [3, 37]. Then

$$x^{n+t-1} f^e\left(\frac{1}{x}\right)\overline{S}\left(\frac{1}{x}\right) = x^{n+t-1}\varphi^e\left(\frac{1}{x}\right)\left(\left(\frac{1}{x}\right)^n - 1\right)$$

$$\sigma^e\left(s_0 + s_1 x + \cdots + s_{n-1}x^{n-1}\right) = \omega^e(1-x^n)$$

$$\sigma^e\left(s_0 + s_1 x + \cdots + s_{2t-1}x^{2t-1}\right) \equiv \omega^e \bmod x^{2t}$$

This is essentially the key equation in [2, 3, 37], modulo minor changes due to different choices of parity check matrix.

The algorithm of Sugiyama et al [42] is based on the equation

$$\sigma^e\left(s_0 + s_1 x + \cdots + s_{2t-1}x^{2t-1}\right) + x^{2t}T = \omega^e$$

One can run the Euclidean algorithm on $R_0 = x^{2t}$ and $R_1 = s_0 + \cdots + s_{2t-1}x^{2t-1}$ until the remainder has degree less than t. Sugiyama et al showed that the resulting combination of $(s_0 + s_1 x + \cdots + s_{2t-1}x^{2t-1})$ and x^{2t} obtained is ω^e and that the coefficient of $(s_0 + s_1 x + \cdots + s_{2t-1}x^{2t-1})$ is σ^e. The article [42] actually treats the more general situation of Goppa codes and error-erasure decoding.

3.3. The key equation for Hermitian codes

The most widely studied algebraic geometry codes are those from Hermitian curves. One reason for the interest in Hermitian curves is that they are maximal curves, meeting the Weil bound on the number of points for a given genus. They also have a very simple formula, and a great deal of symmetry, which leads to lots of structure that makes them useful in coding. The short articles of Stichtenoth [40] and Tiersma [43], and Stichtenoth's book [41] are good references for information on Hermitian curves and codes.

In this section we will derive the key equation and the algorithm for solving it in a manner that parallels the section on Reed-Solomon codes. We will not discuss one very important issue: The syndromes computed from the received vector are insufficient for exploiting the full error correction capability of the Berlekamp-Massey-Sakata decoding algorithm. The majority voting algorithm of Feng-Rao [10] and Duursma [7] is required to compute more syndrome values. We will not discuss majority voting. Instead, we simply deal with the problem solved by the BMS algorithm, computing the error locator ideal from the syndrome of the error vector. A detailed treatment of majority voting may be found in the chapter on algebraic geometry codes by Høholdt et al [17]. The conditions ensuring success in the majority voting algorithm are best understood in terms of the "footprint" of the error vector, which is discussed below, and can lead to decoding beyond the minimum distance [4].

3.3.1. *The Hermitian curve*

Let q be a prime power. We will use the following equation for the Hermitian curve over $\mathbb{F}_{q^2}$,

$$X^{q+1} = Y^q + Y.$$

For each $\alpha \in \mathbb{F}_{q^2}$, α^{q+1} is the norm of α with respect to the extension $\mathbb{F}_{q^2}/\mathbb{F}_q$, so α^{q+1} belongs to $\mathbb{F}_q$. On the other hand, $\beta^q + \beta$ is the trace of β with respect to $\mathbb{F}_{q^2}/\mathbb{F}_q$, so $\beta^q + \beta$ also belongs to $\mathbb{F}_q$. Each element $\gamma \in \mathbb{F}_q$ has q preimages under the trace map, and γ has $q+1$ preimages under the norm map (unless $\gamma = 0$ when there is one). Thus there are $n = q + (q-1)q(q+1) = q^3$ points on the curve. We label them $P_1 = (\alpha_1, \beta_1), P_2 = (\alpha_2, \beta_2), \ldots, P_n = (\alpha_n, \beta_n)$.

Let $\mathbb{F}_{q^2}[X,Y]/(X^{q+1} - Y^q - Y) = \mathbb{F}_{q^2}[x,y]$, where x is the image of X in the quotient and y is the image of Y. Since $y^q = x^{q+1} - y$, each element f in $\mathbb{F}_{q^2}[x,y]$ can be expressed in a unique way as a sum $f_0(x) + f_1(x)y + f_2(x)y^2 + \cdots + f_{q-1}(x)y^{q-1}$. That is, $\{1, y, \ldots, y^{q-1}\}$ is a basis of $\mathbb{F}_{q^2}[x,y]$ as an $\mathbb{F}_{q^2}[x]$-module. Also, $\mathcal{M} = \{x^a y^b : 0 \leq a, 0 \leq b < q\}$ is a basis of $\mathbb{F}_{q^2}[x,y]$ as a $\mathbb{F}_{q^2}$-vector space.

We wish to introduce a function on $\mathbb{F}_{q^2}[x,y]$ akin to the degree function on $\mathbb{F}_q[x]$. Notice that any weighted degree in $\mathbb{F}_{q^2}[X,Y]$ such that X^{q+1} and $Y^q + Y$ have equal weights is obtained by assigning to X a weight kq and to Y a weight $k(q+1)$ for some non-negative integer k. Letting $k = 1$, we define the *order function* ρ by $\rho(x^a y^b) = \deg_{(q,q+1)}(X^a Y^b) = aq + b(q+1)$

and for $f = \sum_{a,b\geq 0} f_{a,b}x^a y^b$ we define $\rho(f) = \max_{f_{a,b}\neq 0} \rho(x^a y^b)$.

One can see that $x^a y^b$ and $x^{a'} y^{a'}$ in $\mathcal{M}$ satisfy $\rho(x^a y^b) = \rho(x^{a'} y^{a'})$ if and only if $a = a'$ and $b = b'$ and that $\rho(\mathbb{F}_{q^2}[x,y]) = \rho(\mathcal{M})$. Define $\Lambda = \rho(\mathbb{F}_{q^2}[x,y]) = q\mathbb{N}_0 + (q+1)\mathbb{N}_0$. The map $\rho : \mathbb{F}_{q^2}[x,y] \to \Lambda$ satisfies $\rho(fg) = \rho(f) + \rho(g)$. This suggests extending it to the quotient field of $\mathbb{F}_{q^2}[x,y]$, which we will write $\mathbb{F}_{q^2}(x,y)$, by defining $\rho(f/g) = \rho(f) - \rho(g)$. Now the image of ρ is all of $\mathbb{Z}$.

3.3.2. *Hermitian codes*

We define the evaluation map

$$\begin{aligned} \mathrm{ev} : \mathbb{F}_{q^2}[x,y] &\longrightarrow \mathbb{F}^n \\ f &\longmapsto (f(\alpha_1,\beta_1), f(\alpha_2,\beta_2), \ldots, f(\alpha_n,\beta_n)). \end{aligned}$$

The Hermitian code $H(m)$ over $\mathbb{F}_{q^2}$ is the linear code generated by $\{(f(P_1),\ldots,f(P_n)) : f \in \mathcal{M}, \rho(f) \leq m\}$. It is shown in [40] (see also [19]) that $H(m) = \mathbb{F}_{q^2}^n$ when $m \geq q^3+q^2-q-1$ and that for $m < q^3+q^2-q-1$ the dual of $H(m)$ is $H(q^3+q^2-q-2-m)$. Clearly, the monomials $x^a y^b$ such that $0 \leq b < q$ and $aq + b(q+1) \leq m$ are a basis for the space $\{f \in \mathbb{F}_{q^2} : \rho(f) \leq m\}$, so they may be used to create a generating matrix for $H(m)$. Since $x^{q^2} - x$ vanishes on all points P_k, we should not use monomials $x^a y^b$ with $a \geq q^2$ in the generating matrix. This is only an issue when $m \geq q^3$. Thus for $m \in \Lambda$ and $m = aq + b(q+1)$, with $b < q$, a generator matrix of $H(m)$ is obtained by evaluating monomials $x^{a'} y^{b'}$ whose weighted degree is at most m and such that $a' < q^2$.

$$\begin{pmatrix} 1 & 1 & \ldots & 1 \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \\ \beta_1 & \beta_2 & \ldots & \beta_n \\ \alpha_1^2 & \alpha_2^2 & \ldots & \alpha_n^2 \\ \alpha_1\beta_1 & \alpha_2\beta_2 & \ldots & \alpha_n\beta_n \\ \beta_1^2 & \beta_2^2 & \ldots & \beta_n^2 \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_1^a\beta_1^b & \alpha_2^a\beta_2^b & \ldots & \alpha_n^a\beta_n^b \end{pmatrix}.$$

3.3.3. *Polynomials for decoding*

Suppose that a word $c \in H(m)^\perp$ was transmitted and that a vector u is received. The vector $e = u - c$ is the error vector. Let t be the weight of

e. Define the *error locator ideal* of e as

$$I^e = \{f \in \mathbb{F}_{q^2}[x,y] : f(\alpha_k, \beta_k) = 0 \text{ for all } k \text{ with } e_k \neq 0\}$$

and the *syndrome* for e as

$$S = \sum_{k=1}^{n} e_k \frac{x^{q+1} - \alpha_k^{q+1}}{(x-\alpha_k)(y-\beta_k)} = \sum_{k=1}^{n} e_k \frac{y^q + y - \beta_k^q - \beta_k}{(x-\alpha_k)(y-\beta_k)}. \tag{3.3}$$

Notice that the order of each term in the summand is $q^2 - q - 1$.

We will give three justifications for this definition of the syndrome in the lemmas below. We note first that for any $(\alpha, \beta) \in \mathbb{F}_{q^2}^2$ on the Hermitian curve,

$$\begin{aligned}
\frac{x^{q+1} - \alpha^{q+1}}{x - \alpha} &= \alpha^q \frac{\left(\frac{x}{\alpha}\right)^{q+1} - 1}{\frac{x}{\alpha} - 1} \\
&= \alpha^q \left(\left(\frac{x}{\alpha}\right)^q + \left(\frac{x}{\alpha}\right)^{q-1} + \cdots + \frac{x}{\alpha} + 1 \right) \\
&= x^q + \alpha x^{q-1} + \cdots + \alpha^{q-1} x + \alpha^q
\end{aligned}$$

and

$$\begin{aligned}
\frac{y^q + y - \beta^q - \beta}{y - \beta} &= 1 + \frac{y^q - \beta^q}{y - \beta} \\
&= 1 + y^{q-1} + \beta y^{q-2} + \cdots + \beta^{q-2} y + \beta^{q-1}
\end{aligned}$$

We will use these identities several times during this presentation.

The first lemma gives a nice relationship between I^e and S. We will show later that the converse also holds.

Lemma 3.12. *If $f \in I^e$ then $fS \in \mathbb{F}_{q^2}[x,y]$.*

Proof. If $f \in I^e$ and $P_{k_1}, \ldots, P_{k_t}$ are the error positions then there exist $g_{k_1}, \ldots, g_{k_t}$ and $h_{k_1}, \ldots, h_{k_t}$ in $\mathbb{F}_{q^2}[x,y]$ such that

$$\begin{aligned}
f &= g_{k_1}(x - \alpha_{k_1}) + h_{k_1}(y - \beta_{k_1}) \\
&= g_{k_2}(x - \alpha_{k_2}) + h_{k_2}(y - \beta_{k_2}) \\
&\vdots \\
&= g_{k_t}(x - \alpha_{k_t}) + h_{k_t}(y - \beta_{k_t})
\end{aligned}$$

Hence

$$\begin{aligned} fS &= \sum_{k_i : e_{k_i} \neq 0} e_{k_i} \left(g_{k_i} \frac{y^q + y - \beta_{k_i}^q - \beta_{k_i}}{y - \beta_{k_i}} + h_{k_i} \frac{x^{q+1} - \alpha_{k_i}^{q+1}}{x - \alpha_{k_i}} \right) \\ &= \sum_{k_i : e_{k_i} \neq 0} e_{k_i} g_{k_i} \left(1 + y^{q-1} + \beta_{k_i} y^{q-2} + \cdots + \beta_{k_i}^{q-2} y + \beta_{k_i}^{q-1} \right) \\ &\quad + \sum_{k_i : e_{k_i} \neq 0} e_{k_i} h_{k_i} \left(x^q + \alpha_{k_i} x^{q-1} + \cdots + \alpha_{k_i}^{q-1} x + \alpha_{k_i}^q \right) \end{aligned}$$

which belongs to $\mathbb{F}_{q^2}[x, y]$. □

The next lemma shows that for $f \in I^e$, the product fS may be used for error evaluation. We will need the derivative of y with respect to x. Since $q = 0$ in $\mathbb{F}_{q^2}$, and $d(y^q + y)/dx = d(x^{q+1})/dx$, we deduce that $dy/dx = x^q$. We say that f has a *simple zero* at a point P when $f(P) = 0$ but $f'(P) \neq 0$.

Lemma 3.13. *If $f \in I^e$ and P_k is an error position then $e_k f'(P_k) = fS(P_k)$. If f has a simple zero at P_k then*

$$e_k = \frac{fS(P_k)}{f'(P_k)}.$$

Proof. The rational function $\frac{x^{q+1} - \alpha_j^{q+1}}{(x-\alpha_j)(y-\beta_j)} = \frac{y^q + y - \beta_j^q - \beta_j}{(x-\alpha_j)(y-\beta_j)}$ gives a well defined value at any point different from (α_j, β_j), so when $j \neq k$, $\left(f \frac{x^{q+1} - \alpha_j^{q+1}}{(x-\alpha_j)(y-\beta_j)} \right)(P_k) = 0$. Consequently,

$$fS(P_k) = e_k \left(f \frac{x^{q+1} - \alpha_k^{q+1}}{(x - \alpha_k)(y - \beta_k)} \right)(P_k).$$

Since $f(P_k) = 0$, there are $g, h \in \mathbb{F}_{q^2}[x, y]$ such that $f = (x - \alpha_k)g + (y - \beta_k)h$. Hence,

$$\begin{aligned} fS(P_k) &= e_k \left((1 + y^{q-1} + \beta_k y^{q-2} + \cdots + \beta_k^{q-2} y + \beta_k^{q-1})(P_k) \right) g(P_k). \\ &\quad + e_k \left((x^q + \alpha_k x^{q-1} + \cdots + \alpha_k^{q-1} x + \alpha_k^q)(P_k) \right) h(P_k) \\ &= e_k \left((1 + q\beta_k^{q-1}) g(P_k) + (q+1)\alpha_k^q h(P_k) \right) \\ &= e_k \left(g(P_k) + \alpha_k^q h(P_k) \right). \end{aligned}$$

On the other hand, $f' = g + x^q h + (x - \alpha_k)g' + (y - \beta_k)h'$. Evaluating f' at P_k,

$$\begin{aligned} f'(P_k) &= g(P_k) + \alpha_k^q h(P_k), \text{ so} \\ e_k f'(P_k) &= fS(P_k). \end{aligned}$$

When f has a simple zero at P_k

$$e_k = fS(P_k)/f'(P_k).$$

□

As our final justification for our definition of S, we show that the syndrome values for the vector e, that is the products $\mathrm{ev}(x^a y^b) \cdot e$, appear as coefficients in a particular expansion of S.

Lemma 3.14. *Let* $s_{a,b} = \sum_{k=1}^{n} e_k \alpha_k^a \beta_k^b$ *and let* δ_b *be 1 when* $b = 0$ *and* 0 *otherwise.*

$$S = \frac{1}{x} \sum_{b=0}^{q-1} \sum_{a=0}^{\infty} s_{a,b} x^{-a} (y^{q-1-b} + \delta_b)$$

Proof.

$$\begin{aligned} \frac{y^q + y - \beta_k^q - \beta_k}{(x - \alpha_k)(y - \beta_k)} &= \left(1 + \frac{y^q - \beta_k^q}{y - \beta_k}\right) \frac{1}{x} \left(\frac{1}{1 - \frac{\alpha_k}{x}}\right) \\ &= (1 + y^{q-1} + \beta_k y^{q-2} + \cdots + \beta_k^{q-2} y + \beta_k^{q-1})(\frac{1}{x} + \frac{\alpha_k}{x^2} + \frac{\alpha_k^2}{x^3} + \dots) \\ &= \sum_{0 \le a} \sum_{0 \le b < q} \alpha_k^a \beta_k^b x^{-a-1} (y^{q-1-b} + \delta_b). \end{aligned}$$

Hence,

$$\begin{aligned} S &= \sum_{k=1}^{n} e_k \sum_{0 \le a} \sum_{0 \le b < q} \alpha_k^a \beta_k^b x^{-a-1} (y^{q-1-b} + \delta_b) \\ &= \sum_{0 \le a} \sum_{0 \le b < q} \left(\sum_{k=1}^{n} e_k \alpha_k^a \beta_k^b \right) x^{-a-1} (y^{q-1-b} + \delta_b) \\ &= \frac{1}{x} \sum_{0 \le a} \sum_{0 \le b < q} s_{a,b} x^{-a} (y^{q-1-b} + \delta_b). \end{aligned}$$

□

3.3.4. *Another basis for* $\mathbb{F}_{q^2}(x, y)$

The final result of the previous section suggests that we introduce a new basis for $\mathbb{F}_{q^2}(x, y)$ over $\mathbb{F}_{q^2}(x)$. For $0 \le b < q$, let

$$z_b^* = \begin{cases} y^{q-1} + 1 & \text{if } b = 0 \\ y^{q-1-b} & \text{otherwise.} \end{cases} \tag{3.4}$$

Notice that $\rho(z_b^*) = (q+1)(q-1-b) = q^2 - 1 - b(q+1)$. We will call $\{z_b^* : b = 0, \ldots, q-1\}$ the $*$-basis. We will write the syndrome, and products of the syndrome with polynomials in x, y, using the $*$-basis. An element $f \in \mathbb{F}_{q^2}[x, y]$ is *monic* in the $*$-basis when its leading term, say $f_{a,b}x^a z_b^*$, has $f_{a,b} = 1$. The following two lemmas show how this basis is useful for decoding.

Lemma 3.15. *The coefficient of z_0^* in $y^b z_c^*$ is 1 if $b = c$ and is 0 otherwise.*

Proof. One can prove by a straightforward computation that for $0 \le b, c < q$,

$$y^b z_c^* = \begin{cases} z_0^* & \text{if } b = c = 0 \\ z_0^* - z_{q-1}^* & \text{if } b = c \neq 0 \\ x^{q+1} z_{q-b}^* & \text{if } b > c = 0 \\ x^{q+1} z_{q+c-b}^* - z_{q-1+c-b}^* & \text{if } b > c > 0 \\ z_{c-b}^* & \text{if } c > b \end{cases}$$

For example, if $b > c > 0$ then

$$\begin{aligned} y^b z_c^* &= y^{q-1+b-c} \\ &= y^{b-c-1}(x^{q+1} - y) \\ &= x^{q+1} z_{q+c-b}^* - z_{q-1+c-b}^* \end{aligned}$$

Notice that $2 \le q + c - b \le q - 1$, so that each of the indices in this case is between 1 and $q-1$. Thus the coefficient of z_0^* in $y^b z_c^*$ is 0 when $0 \le b \le c$. Similar arguments apply to the other cases. □

Any element of $\mathbb{F}_{q^2}(x, y)$ may be expressed uniquely as $\sum_{b=0}^{q-1} h_b z_b^*$ for $h_b \in \mathbb{F}_{q^2}(x)$. We will write h_b in the form used for the syndrome in Section 3.2.1, $h_b = \frac{1}{x}\sum_a h_{a,b}x^{-a}$, where it is understood that a varies over all integers larger than some unspecified bound. For example, we will write the syndrome as $S = \frac{1}{x}\sum_{b=0}^{q-1}\sum_a s_{a,b}x^{-a}z_b^*$, where it is understood that $s_{a,b} = 0$ for $a < 0$.

Lemma 3.16. *Let $f \in \mathbb{F}_{q^2}[x,y]$, let $a \in \mathbb{Z}$ and let b satisfy $0 \le b < q$. The coefficient of $x^{-a-1}z_b^*$ in fS equals the coefficient of z_0^*/x in $x^a y^b fS$.*

More precisely, expand $\tilde{f} = y^b f$, S, and fS as follows.

$$\begin{aligned}
\tilde{f} &= \sum_{c=0}^{q-1} \tilde{f}_c y^c = \sum_{c=0}^{q-1}\sum_a \tilde{f}_{a,c}\, x^a y^c \\
S &= \frac{1}{x}\sum_{c=0}^{q-1} s_c z_c^* = \frac{1}{x}\sum_{c=0}^{q-1}\sum_a s_{a,c}\, x^{-a} z_c^* \\
fS &= \frac{1}{x}\sum_{c=0}^{q-1} t_c z_c^* = \frac{1}{x}\sum_{c=0}^{q-1}\sum_a t_{a,c}\, x^{-a} z_c^*
\end{aligned}$$

Here $s_c = \sum_a s_{a,c}x^{-a}$ and similar definitions hold for t_c and $\tilde{f}_c$. Then

$$t_b = \sum_{c=0}^{q-1} \tilde{f}_c s_c \qquad \text{and} \qquad t_{a,b} = \sum_{c=0}^{q-1}\sum_i \tilde{f}_{i,c}\, s_{i+a,c}.$$

Proof. From the previous lemma, the coefficient of z_0^* in

$$y^b(fS) = \frac{1}{x}\sum_{c=0}^{q-1} t_c y^b z_c^*$$

is $(1/x)t_b$. On the other hand, $y^b f = \tilde{f}$, so

$$\begin{aligned}
(y^b f)S &= \left(\sum_{c=0}^{q-1} \tilde{f}_c y^c\right)\left(\frac{1}{x}\sum_{d=0}^{q-1} s_d z_d^*\right) \\
&= \frac{1}{x}\sum_{c=0}^{q-1}\sum_{d=0}^{q-1} \tilde{f}_c s_d y^c z_d^*.
\end{aligned}$$

Applying the previous lemma, the coefficient of z_0^* is $(1/x)\sum_{c=0}^{q-1}\tilde{f}_c s_c$. We conclude that $t_b = \sum_{c=0}^{q-1}\tilde{f}_c s_c$. Writing $\tilde{f}_c = \sum_i \tilde{f}_{i,c}\, x^i$ and $s_c = \sum_j s_{j,c}\, x^{-j}$ we have

$$\begin{aligned}
\tilde{f}_c s_c &= \sum_i \tilde{f}_{i,c}\, x^i \sum_j s_{j,c}\, x^{-j} \\
&= \sum_a x^{-a} \sum_i \tilde{f}_{i,c}\, s_{i+a,c}.
\end{aligned}$$

This sum is finite since $\tilde{f}_c$ has finite support, and it gives the formula for $t_{a,b}$. □

This lemma tells us that to identify the coefficient of $x^{-a-1}z_b^*$ in fS, we write $\tilde{f} = y^b f$ in the standard basis and then compute the recursion $t_{a,b} = \sum_{c=0}^{q-1} \sum_i \tilde{f}_{i,c}\, s_{i+a,c}$.

3.3.5. *The key equation*

We now define the key equation and approximate solutions to the key equation. We establish some simple lemmas that show basic properties of approximate solutions and how two approximate solutions can be combined to get a better approximation.

Definition 3.17. We say that $f, \varphi \in \mathbb{F}_{q^2}[x, y]$ solve the key equation for syndrome S when $fS = \varphi$.

For a nonzero $f \in \mathbb{F}_{q^2}[x, y]$, writing $fS = \frac{1}{x} \sum_{b=0}^{q-1} \sum_a t_{a,b}\, x^{-a} z_b^*$, we see that f, φ satisfy the key equation when $t_{a,b} = 0$ for $a \geq 0$ and $\varphi = \frac{1}{x} \sum_{b=0}^{q-1} \sum_{a<0} t_{a,b}\, x^{-a} z_b^*$.

Definition 3.18. We say that f and φ in $\mathbb{F}_{q^2}[x, y]$, with f nonzero, solve the Kth approximation of the key equation for syndrome S (or the Kth key equation, for short) when the following two conditions hold.

(1) $\rho(fS - \varphi) \leq q^2 - q - 1 - K$,
(2) φ, written in the $*$-basis, is a sum of terms whose order is at least $q^2 - q - K$.

We will also say that 0 and $x^{-a-1}z_b^*$, for $a < 0$, solve the $aq + b(q+1)$ key equation.

Notice that $\rho(x^{-a-1}z_b^*) = q^2 - q - 1 - (aq + b(q+1))$, and when $a < 0$, we have $x^{-a-1}z_b^* \in \mathbb{F}_{q^2}[x, y]$. Thus, for $0, x^{-a-1}z_b^*$, condition (1) holds with $K = aq + b(q+1)$, but condition (2) is not satisfied. It is convenient to make this pair a solution to the key equation, so we have included the special case in the definition.

For $f \neq 0$, (1) means that $fS - \varphi$ has only terms $x^{-a-1}z_b^*$ with $aq + b(q+1) \geq K$, while (2) means that φ has only terms $x^{-a-1}z_b^*$ with $aq + b(q+1) < K$ and with $a < 0$ because φ is a polynomial. Consequently, using the

expression for fS above, f, φ solve the Kth key equation if and only if

$$t_{a,b} = 0 \text{ whenever } a \geq 0 \text{ and } aq + b(q+1) < K, \text{ and}$$
$$\varphi = \frac{1}{x} \sum_{b=0}^{q-1} \sum_{\substack{a<0 \\ aq+b(q+1)<K}} t_{a,b} x^{-a} z_b^*.$$

Example 3.19. The pair $y^b, 0$ satisfies the $-b(q+1)$ key equation. We have

$$\rho(y^b S) \leq b(q+1) + q^2 - q - 1 = q^2 - q - 1 - (-b(q+1)).$$

The pair $0, z_b^*$ satisfies the $b(q+1) - q$ key equation.

$$\rho(z_b^*) = (q - 1 - b)(q+1) = q^2 - q - 1 - (b(q+1) - q)$$

The following technical lemmas will be used to simplify the proof of Theorem 3.26, which establishes the properties of the decoding algorithm.

Lemma 3.20. *Suppose that $f \neq 0$ and that f, φ satisfy the Kth key equation for syndrome S. Let $g \in \mathbb{F}_{q^2}[x, y]$ with $\rho(g) < K$. Then, in the $\star$-basis expansion of gfS the coefficient of z_0^*/x is 0. Consequently, if g and h are both monic of order K then the coefficients of z_0^*/x in gfS and hfS are equal.*

Proof. It is sufficient to establish this result for a monomial, $g = x^a y^b$, with $aq + b(q+1) < K$. By Lemma 3.16, the coefficient of z_0^*/x in $x^a y^b fS$ is equal to the coefficient of $x^{-a-1} z_b^*$ in fS. Expanding fS as in Lemma 3.16, this coefficient is $t_{a,b}$. The discussion after the definition of approximate solutions to the key equation shows that $t_{a,b} = 0$ for $a \geq 0$ and $aq + b(q+1) < K$. The final statement of the lemma follows from $\rho(g - h) < K$. □

Lemma 3.21. *Suppose that f, φ satisfy the Kth key equation. For any nonnegative integer i, the $K - iq$ key equation is satisfied by $x^i f, x^i \varphi$.*

Proof. It is trivial to check the lemma for the case when $f = 0$ and $\varphi = x^{-a-1} z_b^*$. For $f \neq 0$, we certainly have $x^i f, x^i \varphi \in \mathbb{F}_{q^2}[x, y]$. The terms in φ have order at least $q^2 - q - K$, so the terms in $x^i \varphi$ have order at least $q^2 - q - K + iq = q^2 - q - (K - iq)$. We also assume $\rho(fS - \varphi) \leq q^2 - q - 1 - K$, so $\rho(x^i(fS - \varphi)) \leq q^2 - q - 1 - (K - iq)$. □

Notice that an analogous result does *not* hold for multiplication by y. The example above shows that $0, z_1^*$ solves the $K = 1$ key equation. Yet, 0,

yz_1^* does not solve the key equation of order $K-(q+1) = -q$. Indeed, $yz_1^* = z_0^* - z_{q-1}^*$ and the $-z_{q-1}^*$ term violates the requirements of the definition.

Lemma 3.22. *Suppose that f, φ and g, ψ satisfy the Kth key equation where $K = aq + b(q+1)$. Suppose in addition that $f \neq 0$ and $gS - \psi$ is monic of order $q^2 - q - 1 - K$. Let the coefficient of $x^{-a-1}z_b^*$ in fS be μ. Then $f - \mu g$, $\varphi - \mu\psi$ satisfy the $(K+1)$th key equation.*

Proof. By assumption, $\rho(fS - \varphi) \leq q^2 - q - 1 - K$ and φ has terms of order $q^2 - q - K$ or larger. Furthermore, μ is the coefficient of $x^{-a-1}z_b^*$ in fS. Since $\rho(x^{-a-1}z_b^*) = q^2 - q - 1 - K$, when $\mu = 0$ the inequality above is strict, and f, φ solve the $(K+1)$th key equation. Suppose $\mu \neq 0$. Then both $fS - \varphi$ and $gS - \psi$ have order $q^2 - q - 1 - K$. Since $gS - \psi$ is monic, $\rho((fS - \varphi) - \mu(gS - \psi)) < q^2 - q - 1 - K$. Furthermore, ψ has terms of order at least $q^2 - q - 1 - K$ (allowing for the case in which $g = 0$) so $\varphi - \mu\psi$ has terms of order at least $q^2 - q - (K+1)$, as required for the $(K+1)$th key equation. □

Proposition 3.24 below is a generalization of Proposition 3.5 to Hermitian curves. It also gives the converse of Lemma 3.12. First, we need a lemma.

Lemma 3.23. *Let $f \in \mathbb{F}_{q^2}[x, y]$ and let $(\alpha, \beta) \in \mathbb{F}_{q^2}$ be a point on the Hermitian curve. Then $f(\alpha, \beta)$ is the coefficient of z_0^*/x in the $*$-basis expansion of $f\frac{x^{q+1} - \alpha^{q+1}}{(x-\alpha)(y-\beta)}$.*

Proof. We know that $f(x, y) = f(\alpha, \beta) + (x - \alpha)g + (y - \beta)h$ for some $g, h \in \mathbb{F}_{q^2}[x, y]$. Thus $f\frac{x^{q+1} - \alpha^{q+1}}{(x-\alpha)(y-\beta)}$ has a polynomial part plus

$$f(\alpha, \beta)\frac{x^{q+1} - \alpha^{q+1}}{(x-\alpha)(y-\beta)} = \frac{f(\alpha, \beta)}{x - \alpha}\left(y^{q-1} + 1 + \beta y^{q-2} + \cdots + \beta^{q-2}y + \beta^{q-1}\right) \tag{3.5}$$

$$= f(\alpha, \beta)\frac{1}{x}\sum_{b=0}^{q-1}\sum_{a}\alpha^a\beta^b x^{-a}z_b^* \tag{3.6}$$

The coefficient of z_0^*/x is $f(\alpha, \beta)$ as claimed. □

We need some facts about generators for I^e. We summarize here material that is treated in depth in [17]. Recall that $\Lambda = \{\rho(f) : f \in \mathbb{F}_{q^2}[x, y]\}$. Define the *footprint* of e as $\Delta^e = \Lambda - \rho(I^e)$. The quotient ring $\mathbb{F}_{q^2}[x, y]/I^e$ is a t-dimensional $\mathbb{F}_{q^2}$-vector space. A basis for this space is obtained by taking the classes of $x^a y^b$ for $\rho(x^a y^b) \in \Delta^e$, so $|\Delta^e| = t$.

Since $\mathbb{F}_{q^2}[x]$ is a principal ideal domain, for any ideal I with $\mathbb{F}_{q^2}[x,y]/I$ finite dimensional over $\mathbb{F}_{q^2}$, the ideal I is a free module over $\mathbb{F}_{q^2}[x]$ of rank q. For each i with $0 \leq i \leq q-1$ let f_i be such that $\rho(f_i)$ is minimal among $\{f \in I : \rho(f) \equiv i \bmod q\}$. Then $\{f_i : 0 \leq i \leq q-1\}$ is a Gröbner basis for I. By reducing f_i by multiples of f_j for $j \neq i$ we may assume that all nonzero terms of f_i, except the leading term, have order in $\Delta = \Lambda - \rho(I)$.

Proposition 3.24. *If the expansion of fS in the $*$-basis has zero coefficients for all $x^{-a-1}z_b^*$ such that $aq + b(q+1) \in \Delta^e$ then $f \in I^e$. In particular, let K be the maximal element of Δ^e. If f, φ satisfy the $(K+1)$th key equation then $f \in I^e$.*

Proof. Let f satisfy the hypotheses of the proposition. Let $e_k \neq 0$ and let $P_k = (\alpha_k, \beta_k)$. We will prove that $f(P_k) = 0$. Consider the ideal

$$I' = \{h \in \mathbb{F}_{q^2}[x,y] : h(P_j) = 0 \text{ for all } j \text{ with } e_j \neq 0 \text{ and } j \neq k\}$$

and let $\Delta' = \Lambda \setminus \{\rho(f) : f \in I'\}$. Notice that $|\Delta'| = t - 1$, so there exists some $g \in I'$ such that $\rho(g) \in \Delta^e \setminus \Delta'$. Reducing g modulo a reduced Gröbner basis for I', we can ensure that every monomial in g has order in Δ^e.

As in Lemma 3.16, write $fS = \frac{1}{x}\sum_{b=0}^{q-1}\sum_a t_{a,b}x^{-a}z_b^*$. Lemma 3.16 shows that $t_{a,b}$ is the coefficient of z_0^*/x in $x^a y^b fS$. For $aq + b(q+1) \in \Delta^e$, the hypothesis of this lemma is that $t_{a,b} = 0$, so the coefficient of z_0^*/x in $x^a y^b fS$ is 0. Since g is a linear combination of monomials with order in Δ^e, the coefficient of z_0^*/x in gfS is 0.

On the other hand, Lemma 3.23 and the definition of the syndrome, (3.3), shows that the coefficient of z_0^*/x in gfS is

$$\sum_{j=1}^{n} e_j g(P_j) f(P_j) = e_k g(P_k) f(P_k).$$

Here we have used $g(P_j) = 0$ for $j \neq k$ since $g \in I'$. Since $g \notin I^e$ we must have $g(P_k) \neq 0$, so we conclude that $f(P_k) = 0$.

The final statement of the proposition follows immediately from the observation following the definition of approximate solutions to the key equation. If f, φ satisfy the Kth key equation for $K = 1 + \max \Delta^e$, then the coefficient of $x^{-a-1}z_b^*$ is 0 for any $a \geq 0$ and $aq + b(q+1) < K$. □

3.3.6. *Solving the key equation*

As noted in the introduction, this algorithm is based on Kötter's version of Sakata's generalization of the Berlekamp-Massey algorithm. The

algorithm uses the algebra of $\mathbb{F}_{q^2}[x,y]$ in only one place, the computation of $\tilde{f}$, otherwise all the computations involve polynomials in x, which are easily implementable using shift-registers.

The value M determining the final iteration of the algorithm is given in Proposition 3.28 below.

Decoding algorithm for Hermitian codes

Initialize: For $i = 0$ to $q-1$, set $\begin{pmatrix} f_i^{(0)} & \varphi_i^{(0)} \\ g_i^{(0)} & \psi_i^{(0)} \end{pmatrix} = \begin{pmatrix} y^i & 0 \\ 0 & -z_i^* \end{pmatrix}$

Algorithm: For $m = 0$ to M, and for each pair i, j such that $m \equiv i + j \bmod q$, set

$$d_i = \rho(f_i^{(m)}) \qquad d_j = \rho(f_j^{(m)})$$
$$r_i = \tfrac{m-d_i-j(q+1)}{q} \qquad r_j = \tfrac{m-d_j-i(q+1)}{q}$$
$$\tilde{f}_i = y^j f_i \qquad \tilde{f}_j = y^i f_j$$
$$\mu_i = \textstyle\sum_{c=0}^{q-1} \sum_a (\tilde{f}_i)_{a,c} s_{a+r_i,c} \qquad \mu_j = \sum_{c=0}^{q-1} \sum_a (\tilde{f}_j)_{a,c} s_{a+r_j,c}$$
$$p = \tfrac{d_i+d_j-m}{q} - 1$$

The update for j is analogous to the one for i given below.

$$U_i^{(m)} = \begin{cases} \begin{pmatrix} 1 & -\mu_i x^p \\ 0 & 1 \end{pmatrix} & \text{if } \mu_i = 0 \text{ or } p \geq 0 \\ \begin{pmatrix} x^{-p} & -\mu_i \\ 1/\mu_i & 0 \end{pmatrix} & \text{otherwise.} \end{cases}$$

$$\begin{pmatrix} f_i^{(m+1)} & \varphi_i^{(m+1)} \\ g_j^{(m+1)} & \psi_j^{(m+1)} \end{pmatrix} = U_i^{(m)} \begin{pmatrix} f_i^{(m)} & \varphi_i^{(m)} \\ g_j^{(m)} & \psi_j^{(m)} \end{pmatrix}$$

Output: $f_i^{(M+1)}, \varphi_i^{(M+1)}$ for $0 \leq i < q$.

Remark 3.25. The monomial $x^{r_i} y^j$ used to define $\widetilde{f_i}$ is the *shift* necessary so that $x^{r_i} y^j f_i^{(m)}$ has leading term of order m. Indeed, $\rho(x^{r_i} y^j f_i^{(m)}) = r_i q + j(q+1) + d_i = m$. Lemma 3.16 says that μ_i is the coefficient of $x^{-r_i-1} z_j^*$ in $f_i^{(m)} S$.

Theorem 3.26. *For $m \geq 0$,*

(1) $f_i^{(m)}$ is monic and $\rho(f_i^{(m)}) \equiv i \bmod q$.

(2) $f_i^{(m)}, \varphi_i^{(m)}$ *satisfy the* $m - \rho(f_i^{(m)})$ *approximation of the key equation.*
(3) $g_i^{(m)}, \psi_i^{(m)}$ *satisfy the* $\rho(f_i^{(m)}) - q$ *approximation of the key equation and* $g_i^{(m)} S - \psi_i^{(m)}$ *is monic of order* $q^2 - 1 - \rho(f_i^{(m)})$.
(4) $\rho(g_i^{(m)}) < m - \rho(f_i^{(m)}) + q$.

Proof. We will proceed by induction on m. Example 3.19 establishes the base step, $m = 0$.

Assume that the statements of the theorem are true for m, we will prove them for $m+1$. It is sufficient to consider a pair i, j with $0 \le i, j < q-1$ satisfying $i + j \equiv m \bmod q$. Let d_i, r_i, μ_i, p be as defined in the algorithm.

The induction hypothesis says that $f_i^{(m)}, \varphi_i^{(m)}$ satisfy the $m - d_i$ key equation. By Lemma 3.16, μ_i is the coefficient of $x^{-r_i-1} z_j^*$ in $f_i^{(m)} S$. A simple computation shows $\rho(x^{-r_i-1} z_j^*) = q^2 - q - 1 - (m - d_i)$. Consequently, if $\mu_i = 0$, then f_i^m, φ_i^m solve the $(m + 1 - d_i)$ key equation. In this case, the algorithm retains the data from the iteration m, *e.g.* $f^{(m+1)} = f^{(m)}$. It is easy to verify that the properties of the theorem hold.

If $\mu_i \neq 0$, we consider two cases. First, suppose $p \ge 0$. The algorithm sets $f_i^{(m+1)} = f_i^{(m)} - \mu_i x^p g_i^{(m)}$. Notice that $\rho(\mu_i x^p g_i^{(m)}) < (d_i + d_j - m - q) + (m - d_j + q) = d_i$. This shows that $\rho(f_i^{(m+1)}) = d_i$ and $f_i^{(m+1)}$ is monic, as claimed in item (1). By the induction hypothesis and Lemma 3.21, $x^p g_j^{(m)}, x^p \psi_j^{(m)}$ satisfy the $d_j - q - pq = m - d_i$ key equation and $x^p (g_j^{(m)} S - \psi_j^{(m)})$ is monic of order $q^2 - q - 1 - (m - d_i)$. Lemma 3.22 shows that $f_i^{(m)} - \mu_i x^p g_j^{(m)}, \varphi_i^{(m)} - \mu_i x^p \psi_j^{(m)}$ solves the $m + 1 - d_i$ key equation. Since $d_i = \rho(f_i^{(m+1)})$, we have established item (2) of the theorem. Items (3) and (4) follow because $g_i^{(m+1)} = g_i^{(m)}$ and $\psi_i^{(m+1)} = \psi_i^{(m)}$.

Suppose now that $p < 0$ and $\mu_i \neq 0$. In this case, $f_i^{(m+1)} = x^{-p} f_i^{(m)} - \mu_i g_j^{(m)}$. A simple computation shows $\rho(x^{-p} f_i^{(m)}) = m - d_j + q$ while $\rho(g_j^{(m)}) < m - d_j + q$. Thus, $f_i^{(m+1)}$ is monic, and

$$\rho(f_i^{(m+1)}) = \rho(x^{-p} f_i^{(m)}) = m - d_j + q \equiv i \bmod q.$$

From Lemma 3.21, $x^{-p} f_i^{(m)}, x^{-p} \varphi_i^{(m)}$ satisfy the key equation of order $m - d_i + pq = d_j - q$. By the induction hypothesis, $g_j^{(m)}$, $\psi_j^{(m)}$ satisfy the key equation of the same order. Furthermore, μ_i is the coefficient of $x^{-p-r_i-1} z_j^*$ in $x^{-p} f_i^{(m)} S$. Noting that $q(p + r_i) + j(q + 1) = d_j - q$ we may apply Lemma 3.22 to obtain that $f_i^{(m+1)}$, $\varphi_i^{(m+1)}$ satisfy the key equation of order $d_j - q + 1 = m + 1 - \rho(f_i^{(m+1)})$. This proves item (2).

To prove items (4) and (3), we first establish that $\mu_j = \mu_i$. We claim that each is the coefficient of z_0^*/x in $x^{-p-1} f_j^{(m)} f_i^{(m)} S$. We know μ_i is the coefficient of $x^{-r_i-1} z_j^*$ in $f_i^{(m)} S$, which by Lemma 3.16 is the coefficient of z_0^*/x in $x^{r_i} y^j f_i^{(m)} S$. Since $x^{-p-1} f_j^{(m)}$ and $x^{r_i} y^j$ are both monic of order $m - d_i$, Lemma 3.20 says the coefficients of z_0^*/x in $x^{r_i} y^j f_i^{(m)} S$ and $x^{-p-1} f_j^{(m)} f_i^{(m)} S$ are equal. A similar argument works for j, which establishes the claim.

Since $\mu_j = \mu_i \neq 0$, the algorithm sets $g_i^{(m+1)} = \mu_i^{-1} f_j^{(m)}$ and $\psi_i^{(m)} = \mu_i^{-1} \varphi_j^{(m)}$. We can verify item (4),

$$\begin{aligned}(m+1) - \rho(f_i^{(m+1)}) + q &= m + 1 - (m - d_j + q) + q \\ &= d_j + 1 \\ &> \rho(g_i^{(m+1)})\end{aligned}$$

Item (3) also follows since $g_i^{(m+1)}, \psi_i^{(m+1)}$ satisfy the $m - d_j$ key equation and $m - d_j = \rho(f_i^{(m+1)}) - q$. Furthermore, $g_i^{(m+1)} S - \psi_i^{(m+1)} = \mu_j^{-1}(f_j^{(m)} S - \varphi_j^{(m)})$ is monic. □

The next results establish the iteration number M at which the algorithm may be terminated. This depends on the footprint of e, Δ^e, introduced earlier as well as the orders of the Gröbner basis for I^e. For $i = 0$ up to $q - 1$ define $\sigma_i = \min\{\rho(f) : f \in I^e \text{ and } \rho(f) \equiv i \bmod q\}$.

Lemma 3.27. *For all m and for all i, $\rho(f_i^{(m)}) \leq \sigma_i$.*

Proof. Let $f_i^e \in I^e$ have pole order σ_i and consider $f_i^e g_i^{(m)} S - f_i^e \psi_i^{(m)}$. By Theorem 3.26 (3), we have $\rho(f_i^e g^{(m)} S - f_i^e \psi^{(m)}) = \sigma_i + q^2 - 1 - \rho(f_i^{(m)})$. This must be an element of Λ because $f_i^e S$, $g^{(m)}$ and $\psi^{(m)}$ are all in $\mathbb{F}_{q^2}[x, y]$. Since $\sigma_i - \rho(f_i^{(m)})$ is a multiple of q, and $q^2 - q - 1 \notin \Lambda$, we must have $\sigma_i - \rho(f_i^{(m)}) \geq 0$. □

Proposition 3.28. *Let $\sigma_{\max} = \max\{\sigma_i : 0 \leq i \leq q-1\}$ and let $\delta_{\max} = \max\{c \in \Delta^e\}$. For $m > \sigma_{\max} + \delta_{\max}$, each of the polynomials $f_i^{(m)}$ belongs to I^e. Let $M = \sigma_{\max} + \max\{\delta_{\max}, q^2 - q - 1\}$. Each of the pairs $f_i^{(M+1)}, \varphi_i^{(M+1)}$ satisfies the key equation.*

Proof. By Theorem 3.26, $f_i^{(m)}, \varphi_i^{(m)}$ satisfy the $m - \rho(f_i^{(m)})$ key equation. If $m > \sigma_{\max} + \delta_{\max}$, then, $m - \rho(f_i^{(m)}) > \delta_{\max}$, so the result follows from Lemma 3.24. For $M = \sigma_{\max} + \max\{\delta_{\max}, q^2 - q - 1\}$, we have

$$\rho(f_i^{(M+1)} S - \varphi_i^{(M+1)}) \leq q^2 - q - 1 - (M + 1 - \rho(f_i^{(M+1)})) < 0.$$

Since $f_i^{(M+1)}$ is a locator, $\varphi_i^{(M+1)}$ must equal $f_i^{(M+1)}S$. □

3.3.7. *Error evaluation without the error evaluator polynomials*

In this section we generalize the error evaluation formula in Proposition 3.11 that uses just the error locator polynomial f and the update polynomial g to determine error values. The main result is Theorem 3.30, which is readily derived from Proposition 3.29. Unfortunately, the proposition requires a result that takes some work to establish: In the algorithm, when $i + j \equiv m \bmod q$, $\mu_i = \mu_j$. This was shown for $p < 0$ in the proof of Theorem 3.26, but in order to show it for $p \geq 0$ we need a rather technical result, Proposition 3.48. Since the result is easier to state using the language of residues—instead of referring to the coefficient of z_0^*/x—we have deferred it to the section on general one-point codes.

Proposition 3.29. *Let* $B_i^{(M)} = \begin{pmatrix} f_i^{(m)} & \varphi_i^{(m)} \\ g_i^{(m)} & \psi_i^{(m)} \end{pmatrix}$. *Then for all* m,

$$\sum_{i=0}^{q-1} \det B_i^{(m)} = -\sum_{i=0}^{q-1} y^i z_i^* = -1 \tag{3.7}$$

Proof. We proceed by induction. The case $m = 0$ is a simple calculation. Assume that the statement of the theorem is true for m; we will prove it for $m + 1$. It is sufficient to show that $\det B_i^{(m+1)} = \det B_i^{(m)}$ if $2i \equiv m \bmod q$ and $\det B_i^{(m+1)} + \det B_j^{(m+1)} = \det B_i^{(m)} + \det B_j^{(m)}$ if $i + j \equiv m \bmod q$ and $i \neq j$.

If $2i \equiv m \bmod q$, then $B_i^{(m+1)} = U_i^{(m)} B_i^{(m)}$, where

$$U_i^{(m)} = \begin{cases} \begin{pmatrix} 1 & -\mu_i x^p \\ 0 & 1 \end{pmatrix} & \text{if } \mu_i = 0 \text{ or } p \geq 0 \\ \begin{pmatrix} x^{-p} & -\mu_i \\ 1/\mu_i & 0 \end{pmatrix} & \text{otherwise.} \end{cases}$$

Since $\det U_i^{(m)} = 1$ in either case, we have $\det B_i^{(m+1)} = \det B_i^{(m)}$.

Assume now that $i + j \equiv m \bmod q$ and that $i \neq j$. Proposition 3.48 shows that $\mu_i = \mu_j$, so, from the algorithm

$$B_i^{(m+1)} = \begin{cases} \begin{pmatrix} f_i^{(m)} - \mu_i x^p g_j^{(m)} & \varphi_i^{(m)} - \mu_i x^p \psi_j^{(m)} \\ g_i^{(m)} & \psi_i^{(m)} \end{pmatrix} & \text{if } \mu_i = 0 \text{ or } p \geq 0 \\ \begin{pmatrix} x^{-p} f_i^{(m)} - \mu_i g_j^{(m)} & x^{-p}\varphi_i^{(m)} - \mu_i \psi_j^{(m)} \\ \mu_i^{-1} f_j^{(m)} & \mu_i^{-1}\varphi_j^{(m)} \end{pmatrix} & \text{otherwise.} \end{cases}$$

The two cases lead respectively to

$$\det B_i^{(m+1)} = \begin{cases} f_i^{(m)}\psi_i^{(m)} - g_i^{(m)}\varphi_i^{(m)} - \mu_i x^p (g_j^{(m)}\psi_i^{(m)} - g_i^{(m)}\psi_j^{(m)}) & \text{or} \\ f_j^{(m)}\psi_j^{(m)} - g_j^{(m)}\varphi_j^{(m)} - \mu_i x^p (g_i^{(m)}\psi_j^{(m)} - g_j^{(m)}\psi_i^{(m)}) \end{cases}$$

To obtain $\det B_j^{(m+1)}$ one simply switches i and j in these formulas. When we take the sum of $\det B_i^{(m+1)}$ and $\det B_j^{(m+1)}$, the final terms cancel, so $\det B_i^{(m+1)} + \det B_j^{(m+1)} = \det B_i^{(m)} + \det B_j^{(m)}$. □

We now take M as in Proposition 3.28, so that the algorithm of the previous section has produced solutions to the key equation. Let

$$f_i = f_i^{(M+1)} \quad \varphi_i = \varphi_i^{(M+1)}$$
$$g_i = g_i^{(M+1)} \quad \psi_i = \psi_i^{(M+1)}$$

Then the f_i are a basis for I^e as a module over $\mathbb{F}_{q^2}[x, y]$ and f_i, φ_i satisfy the key equation.

Theorem 3.30. *If P_k is an error position.*

$$e_k = \left(\sum_{i=0}^{q-1} f_i'(P_k) g_i(P_k) \right)^{-1} \tag{3.8}$$

Proof. From the preceding lemma,

$$\sum_{i=0}^{q-1} (f_i \psi_i - g_i \varphi_i) = -1$$

Evaluating at an error position P_k we have $-\sum_{i=0}^{q-1} g_i \varphi_i(P_k) = -1$. Apply Lemma 3.13, to get $\sum_{i=0}^{q-1} g_i(P_k) e_k f_i'(P_k) = 1$. Solving for e_k gives the formula. □

3.3.8. *An example*

Consider the Hermitian curve associated to the field extension $\mathbb{F}_9 = \mathbb{F}_3[\alpha]$ where $\alpha^2 = \alpha + 1$. Let x and y be the classes of X and Y in the quotient $\mathbb{F}_9[X,Y]/(X^4 - Y^3 - Y)$. The basis monomials are $x^a y^b$ for $a \geq 0$ and $0 \leq b \leq 2$. The order of x is 3 and the order of y is 4. So,

$$\Lambda = \{0, 3, 4, 6, 7, 8, 9, 10, \dots\}.$$

The Hermitian curve in this case has 27 points, which we take in the following order: $(1,\alpha)$, $(1,\alpha^3)$, $(1,2)$, $(\alpha,1)$, (α,α^5), (α,α^7), (α^2,α), (α^2,α^3), $(\alpha^2,2)$, $(\alpha^3,1)$, (α^3,α^5), (α^3,α^7), $(2,\alpha)$, $(2,\alpha^3)$, $(2,2)$, $(\alpha^5,1)$, (α^5,α^5), (α^5,α^7), (α^6,α), (α^6,α^3), $(\alpha^6,2)$, $(\alpha^7,1)$, (α^7,α^5), (α^7,α^7), $(0,\alpha^2)$, $(0,\alpha^6)$, $(0,0)$.

Let us consider correction of two errors. There are two choices for Δ^e when the weight of e is two, $\{0,4\}$ when the points are on a vertical line, and $\{0,3\}$ when they are not. Following [4], we will call the latter case "generic" and former "non-generic." In either case $\sigma_{\max} = 8$. From Proposition 3.28, the computation of all error locators and evaluators is complete after iteration number $M = \sigma_{\max} + \max\{\delta_{\max}, q^2 - q - 1\}$. Thus, for an algorithm to correct either of the two errors we terminate the algorithm with iteration $M = 8+5 = 13$, and take the data for superscript 14. We will explain in detail the first steps in the generic case. All the computations are summarized in Table 3.5. The computations in the non-generic case are summarized in Table 3.5.

For the generic error vector we take error values α^2 at the point $(\alpha, 1)$ and α^7 at the point (α^6, α^3), so the error vector is

$$e = (000\alpha^2 0000000000000000\alpha^7 000000).$$

The associated syndromes are

	$s_{0,b}$	$s_{1,b}$	$s_{2,b}$	$s_{3,b}$	$s_{4,b}$	$s_{5,b}$	$s_{6,b}$	$s_{7,b}$	$s_{8,b}$
$s_{a,0}$	α^5	α^2	α^5	0	1	2	α^6	α^5	α^5
$s_{a,1}$	2	1	α^2	α	α	α^6	α	0	2
$s_{a,2}$	α^5	α^2	α^5	0	1	2	α^6	α^5	α^5

To initialize f, g, φ, ψ we take

$$\begin{array}{cccc} f_0 = 1 & g_0 = 0 & \varphi_0 = 0 & \psi_0 = 2y^2 + 2 \\ f_1 = y & g_1 = 0 & \varphi_1 = 0 & \psi_1 = 2y \\ g_2 = 0 & f_2 = y^2 & \varphi_2 = 0 & \psi_2 = 2 \end{array}$$

We start with $m = 0$. The pairs i, j with $i + j \equiv m \bmod 3$ are $0,0$ and $1,2$. The data computed in the algorithm is,

$$\begin{array}{lll} r_0 = 0 & \tilde{f}_0 = 1 & \mu_0 = s_{0,0} = \alpha^5, \\ r_1 = -4 & \tilde{f}_1 = x^4 + 2y & \mu_1 = s_{0,0} + 2s_{-4,1} = \alpha^5, \\ r_2 = -4 & \tilde{f}_2 = x^4 + 2y & \mu_2 = s_{0,0} + 2s_{-4,1} = \alpha^5. \end{array}$$

For the pair $0, 0$, $p = -1$, and for the pair $1, 2$, $p = 3$, so

$$U_0^{(0)} = \begin{pmatrix} x & \alpha \\ \alpha^3 & 0 \end{pmatrix} \qquad \text{and} \qquad U_1^{(0)} = U_2^{(0)} = \begin{pmatrix} 1 & \alpha x^3 \\ 0 & 1 \end{pmatrix}.$$

As a result,

$$\begin{array}{llll} f_0^{(1)} = x & \varphi_0^{(1)} = \alpha^5 y^2 + \alpha^5 & g_0^{(1)} = \alpha^3 & \psi_0^{(1)} = 0, \\ f_1^{(1)} = y & \varphi_1^{(1)} = \alpha^5 x^3 & g_1^{(1)} = 0 & \psi_1^{(1)} = 2y \\ f_2^{(1)} = y^2 & \varphi_2^{(1)} = \alpha^5 x^3 y & g_2^{(1)} = 0 & \psi_2^{(1)} = 2 \end{array}$$

For $m = 1$, the pairs i, j with $i + j \equiv m \bmod 3$ are $0, 1$ and $2, 2$, and,

$$\begin{array}{lll} r_0 = -2 & \tilde{f}_0 = xy & \mu_0 = s_{-1,1} = 0, \\ r_1 = -1 & \tilde{f}_1 = y & \mu_1 = s_{-1,1} = 0, \\ r_2 = -5 & \tilde{f}_2 = x^4 y + 2y^2 & \mu_2 = s_{-1,1} + 2s_{-5,2} = 0. \end{array}$$

This means that

$$U_0^{(1)} = U_1^{(1)} = U_2^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and $f_i, \varphi_i, g_i, \psi_i$ remain unchanged.

For $m = 2$, the pairs i, j with $i + j \equiv m \bmod 3$ are $0, 2$ and $1, 1$, and,

$$\begin{array}{lll} r_0 = -3 & \tilde{f}_0 = xy^2 & \mu_0 = s_{-2,2} = 0 \\ r_1 = -2 & \tilde{f}_1 = y^2 & \mu_1 = s_{-2,2} = 0 \\ r_2 = -2 & \tilde{f}_2 = y^2 & \mu_2 = s_{-2,2} + 2s_{-2,2} = 0 \end{array}$$

Again, this means $U_0^{(2)} = U_1^{(2)} = U_2^{(2)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, and $f_i, \varphi_i, g_i, \psi_i$ remain unchanged.

For $m = 3$, the pairs i, j with $i + j \equiv m \bmod 3$ are $0, 0$ and $1, 2$. Now,

$$\begin{array}{lll} r_0 = 0 & \tilde{f}_0 = x & \mu_0 = s_{1,0} = \alpha^2 \\ r_1 = -3 & \tilde{f}_1 = x^4 + 2y & \mu_1 = s_{1,0} + 2s_{-3,1} = \alpha^2 \\ r_2 = -3 & \tilde{f}_2 = x^4 + 2y & \mu_2 = s_{1,0} + 2s_{-3,1} = \alpha^2 \end{array}$$

For the pair $0, 0$ we have $p = 0$ and for the pair $1, 2$ we have $p = 2$ so,

$$U_0^{(3)} = \begin{pmatrix} 1 & \alpha^6 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad U_1^{(3)} = U_2^{(3)} = \begin{pmatrix} 1 & \alpha^6 x^2 \\ 0 & 1 \end{pmatrix}$$

Consequently,

$$\begin{array}{llll} f_0^{(4)} = x + \alpha & \varphi_0^{(4)} = \alpha^5 y^2 + \alpha^5 & g_0^{(4)} = \alpha^3 & \psi_0^{(4)} = 0 \\ f_1^{(4)} = y & \varphi_1^{(4)} = \alpha^5 x^3 + \alpha^2 x^2 & g_1^{(4)} = 0 & \psi_1^{(4)} = 2y \\ f_2^{(4)} = y^2 & \varphi_2^{(4)} = \alpha^5 x^3 y + \alpha^2 x^2 y & g_2^{(4)} = 0 & \psi_2^{(4)} = 2 \end{array}$$

The subsequent steps are summarized in Table 3.5. Let $f_i = f_i^{(14)}$ and similarly for the other data. The locators and associated derivatives are (using $\frac{dy}{dx} = x^3$),

$$\begin{aligned} f_0 &= x^2 + x + \alpha^7 & (f_0)' &= 2x + 1 \\ f_1 &= y + \alpha^5 x + \alpha & (f_1)' &= x^3 + \alpha^5 \\ f_2 &= y^2 + \alpha^7 x^2 + \alpha^7 x + \alpha^3 & (f_2)' &= 2x^3 y + \alpha^3 x + \alpha^7 \end{aligned}$$

The points where f_0, f_1, and f_2 vanish are exactly $P_4 = (\alpha, 1)$ and $P_{21} = (\alpha^6, 2)$, coinciding with the error positions. The polynomials φ are:

$$\begin{aligned} \varphi_0 &= \alpha^5 xy^2 + y^2 + 2xy + \alpha x + 2, \\ \varphi_1 &= \alpha^5 x^3 + \alpha^2 y^2 + \alpha^2 x^2 + \alpha y + \alpha^5 x + a^6, \\ \varphi_2 &= \alpha^5 x^3 y + 2xy^2 + \alpha^2 x^2 y + 2x^3 + \alpha^7 y^2 + \alpha^2 xy + x^2 + \alpha^7 x + 1 \end{aligned}$$

The error values at these positions can be computed using the formula in Lemma 3.13. For example,

$$e_4 = \frac{\varphi_1(P_4)}{(f_1)'(P_4)} = \frac{\alpha^5}{\alpha^3} = \alpha^2 \qquad e_{21} = \frac{\varphi_1(P_{21})}{(f_1)'(P_{21})} = \frac{\alpha^6}{\alpha^7} = \alpha^7.$$

The same error values could have been obtained using f_0 and φ_0 instead of f_1 and φ_1. However, we could not have used f_2 and φ_2, because the zero of f_2 at P_{21} is not simple.

By Theorem 3.30, the error values can also be obtained using g_0, g_1, g_2 instead of $\varphi_0, \varphi_1, \varphi_2$. Since $g_0 = \alpha^6 x + \alpha^7$, and $g_1 = g_2 = 0$ there is only one term to compute.

$$\begin{aligned} e_4 &= (f_0'(P_4) g_0(P_4))^{-1} & e_{21} &= \left((f_0)'(P_{21}) g_0(P_{21})\right)^{-1} \\ &= \left(\alpha^3 \cdot \alpha^3\right)^{-1} & &= \left(\alpha^7 \cdot \alpha^2\right)^{-1} \\ &= \alpha^2 & &= \alpha^7. \end{aligned}$$

An example of a non-generic error vector is

$$(000000\alpha^2 0 \alpha^7 000000000000000000).$$

The error positions correspond to the points $P_7 = (\alpha^2, \alpha)$ and $P_9(\alpha^2, 2)$ which lie on the line $x = \alpha^2$. The associated syndromes are

	$s_{0,c}$	$s_{1,c}$	$s_{2,c}$	$s_{3,c}$	$s_{4,c}$	$s_{5,c}$	$s_{6,c}$	$s_{7,c}$	$s_{8,c}$
$s_{a,0}$	α^5	α^7	α	α^3	α^5	α^7	α	α^3	α^5
$s_{a,1}$	α^7	α	α^3	α^5	α^7	α	α^3	α^5	α^7
$s_{a,2}$	α^2	2	α^6	1	α^2	2	α^6	1	α^2

The steps of the algorithm are summarized in Table 3.5. Notice that after step $m = 4$, $f_0 = x - \alpha^2 = x + \alpha^6$ is already a locator.

3.4. The key equation for one-point codes

Sakata's generalization of the Berlekamp-Massey algorithm was originally designed for a monomial ordering on a polynomial ring in several variables [38]. It has been adapted to the more general setting of a ring with an order function [17], which corresponds to an algebraic variety (curve, surface or higher dimensional object) and a choice of valuation on the variety [31]. In the case of a curve $\mathcal{C}$, one takes the ring R of functions having poles only at a single point Q on $\mathcal{C}$, and the pole order function. The one-point codes defined by $\mathcal{C}$ and Q are obtained by evaluating functions in R at rational points $P_1, P_2, \ldots, P_n$ that are distinct from Q.

In this section we show that the results in the Hermitian codes section, with very minor modifications, apply to one-point codes. The main challenge is to establish the dual bases in which we write the locator polynomial and the evaluator, which is now a differential. Once this foundation is set, the decoding material falls in place via the same arguments as were used for Hermitian codes. We simply state the results here and leave verification to the reader. The section starts with a quick tour of the main properties of uniformizing parameters, differentials, residues, and other topics that are needed to establish the algorithms and formulas for decoding. Our primary reference for this section is Stichtenoth's book [41], but another valuable resource is Pretzel's book [36].

3.4.1. *Curves, function fields and differentials*

Let K be a function field of transcendence degree one over $\mathbb{F}_q$. Let $\mathcal{C}$ be the smooth curve over $\mathbb{F}_q$ defined by K. We assume that $\mathbb{F}_q$ is algebraically closed in K, which is equivalent to $\mathcal{C}$ being absolutely irreducible. Let Q be a rational point of $\mathcal{C}$ and let ν_Q be the associated valuation of K. Let $L(mQ)$ be the space of functions on $\mathcal{C}$ having poles only at Q and of order at most m there. Each $L(mQ)$ contains $L((m-1)Q)$, and is either equal to it, when we say m is a *gap*, or of dimension one larger, when m is a *nongap*. Let Λ be the set of nongaps and let Λ^c be its complement in $\mathbb{Z}$. Λ is called the Weierstrass semigroup of $\mathcal{C}$ at Q. The union of the $L(mQ)$ is a ring,

$$R = \bigcup_{m=0}^{\infty} L(mQ)$$

For $f \in R$, we define $\rho(f) = -\nu_Q(f)$ to be the pole order of f at Q. Formally, we set $\rho(0) = -\infty$.

Let κ be the smallest positive element of Λ. For each $b = 0, \ldots, \kappa - 1$, let λ_b be the smallest element of Λ congruent to b modulo κ. Any integer may be written in a unique way as $\lambda_b + a\kappa$ for some $b \in \{0, \ldots, \kappa - 1\}$ and $a \in \mathbb{Z}$. Elements of Λ have $a \geq 0$ and elements of Λ^c have $a < 0$. The set $\lambda_1, \ldots, \lambda_{\kappa-1}$ is usually known as the *Apéry* set of Λ (named so after [1]). Let $x \in R$ have pole order κ, and for each b, let z_b have pole order λ_b. We also assume that some uniformizing parameter u_Q at Q has been selected, and that x, and z_b are *monic* with respect to u_Q. That is, when either x or z_b is written as a power series in u_Q the initial term has coefficient 1. In particular, $z_0 = 1$.

Proposition 3.31. *With the notation above, R is a free module over $\mathbb{F}_q[x]$ with basis $\{z_b\}_{b=0}^{\kappa-1}$. This is also a basis for K over $\mathbb{F}_q(x)$.*

Proof. Let $y \in R$ satisfy $\rho(y) \equiv b \bmod \kappa$. Since λ_b is the smallest element of Λ congruent to b, there is some nonnegative a such that $\rho(y) = \lambda_b + a\kappa$. Now $\rho(y) = \rho(x^a z_b)$ so there is some $\beta \in \mathbb{F}_q$ such that $\rho(y - \beta x^a z_b) < \rho(y)$. Continuing in this manner, we find that for some $g_j \in \mathbb{F}_q[x]$, the pole order of $y - \sum_j g_j z_j$ is negative. Since $y - \sum_j g_j z_j \in R$, the pole order must be $-\infty$; that is $y - \sum_j g_j z_j = 0$.

On the other hand, no nontrivial combination $\sum_j g_j z_j$ can equal 0. If $g_j \neq 0$ then $\rho(g_j z_j) \equiv j \bmod \kappa$. Thus $\rho(\sum_j g_j z_j) = \max_{j:g_j \neq 0}\{\rho(g_j z_j)\}$ which is not $-\infty$. Thus, R is free over $\mathbb{F}_q[x]$ with basis $\{z_j\}_{j=0}^{\kappa-1}$. The argument for linear independence holds for $g_j \in \mathbb{F}_q(x)$ as well. Since x has only one pole, and that of order κ, the dimension of K over $F_q(x)$ is κ, [41, I.4.11]. Thus $\{z_j\}_{j=0}^{\kappa-1}$ is a basis for K over $\mathbb{F}_q(x)$. □

There are parallel constructions for differentials. The module of differentials of K over $\mathbb{F}_q$, which we denote Ω, is a one-dimensional vector space over K. For any separating element $u \in K$, in particular for a uniformizing parameter, du is a basis for Ω. If u_P is a uniformizing parameter at a point P, then any $\omega \in \Omega$ may be written in the form $\sum_{i=r}^{\infty} c_i u_P^i du_P$ with $c_i \in \mathbb{F}_q$ and $c_r \neq 0$. One defines $\nu_P(\omega) = r$ and $\mathrm{res}_P(\omega) = c_{-1}$ (or $\mathrm{res}_P(\omega) = 0$ if $r > -1$). These definitions are independent of the choice of uniformizing parameter. We will say that ω is *monic*, relative to u_P, when $c_r = 1$. The divisor of ω is $(\omega) = \sum_P \nu_P(\omega)$, where the sum is over all points of $\mathcal{C}$. For any divisor D, $\Omega(D)$ is the space of differentials such that $(\omega) \geq D$. Thus, $\Omega(mQ)$ is the space of differentials which have valuation *at least* m at Q

and which have nonnegative valuation elsewhere. Let

$$\Omega(-\infty Q) = \bigcup_{m=0}^{\infty} \Omega(-mQ)$$

It is evident that $\Omega(-\infty Q)$ is a module over R.

The most fundamental invariant of the curve $\mathcal{C}$ is its genus, g. We will use the following fundamental results about divisors and the genus.

- The degree of any differential is $2g-2$.
- For the point Q, the number of positive gaps, $|\mathbb{N} \setminus \Lambda|$, is g.
- $\Omega(-\infty Q)$ is isomorphic to R when $(2g-2)Q$ is a canonical divisor.
- The Riemann-Roch theorem: For any divisor D,

$$\dim L(D) - \dim \Omega(D) = m + 1 - g.$$

- The residue theorem: For any differential ω, $\sum_P \mathrm{res}_P(\omega) = 0$, where the sum is over all points of $\mathcal{C}$.

3.4.2. *One-point codes and their duals*

Let $P_1, P_2, \ldots, P_n$ be distinct rational points on $\mathcal{C}$, each different from Q, and let $D = P_1 + P_2 + \cdots + P_n$. We define the evaluation map ev as follows.

$$\begin{aligned} \mathrm{ev} : R &\longrightarrow \mathbb{F}_q^n \\ f &\longmapsto (f(P_1), f(P_2), \ldots, f(P_n)) \end{aligned}$$

Similarly, we have the residue map

$$\begin{aligned} \mathrm{res} : \Omega(-\infty Q - D) &\longrightarrow \mathbb{F}_q^n \\ \omega &\longmapsto (\mathrm{res}_{P_1}(\omega), \mathrm{res}_{P_2}(\omega), \ldots, \mathrm{res}_{P_n}(\omega)) \end{aligned}$$

Restricting the evaluation map to $L(mQ)$ and the residue map to $\Omega(mQ - D)$ we get exact sequences.

$$\begin{array}{ccccccc} 0 & \longrightarrow & L(mQ-D) & \longrightarrow & L(mQ) & \longrightarrow & \mathbb{F}_q^n \\ 0 & \longrightarrow & \Omega(mQ) & \longrightarrow & \Omega(mQ-D) & \longrightarrow & \mathbb{F}_q^n \end{array}$$

The image codes are $C_L(D, mQ) = \mathrm{ev}(L(mQ))$ and $C_\Omega(D, mQ) = \mathrm{res}(\Omega(mQ - D))$.

Proposition 3.32. *The codes $C_L(D, mQ)$ and $C_\Omega(D, mQ)$ are dual.*

Proof. For $f \in L(mQ)$ and $\omega \in \Omega(mQ - D)$, the poles of $f\omega$ are supported on D. From the residue theorem,

$$\mathrm{ev}(f) \cdot \mathrm{res}(\omega) = \sum_{k=1}^{n} \mathrm{res}_{P_k}(f\omega) = -\mathrm{res}_Q(f\omega) = 0$$

The Riemann-Roch theorem says

$$\begin{aligned} \dim L(mQ) - \dim \Omega(mQ) &= m + 1 - g \\ \dim L(mQ - D) - \dim \Omega(mQ - D) &= m - n + 1 - g \end{aligned}$$

Taking the difference,

$$(\dim L(mQ) - \dim L(mQ - D)) + (\dim \Omega(mQ - D) - \dim \Omega(mQ)) = n$$

Thus, the codes are of complementary dimension and are orthogonal, so they are dual codes. □

One consequence of the proposition is that the code $C_\Omega(D, -Q)$ is the whole space $\mathbb{F}_q^n$. In a later section we will identify a differential, $h_{P_k}dx \in \Omega(-Q - D)$, whose image under res is 1 in position k and 0 elsewhere. The syndrome of an error vector e will be $\sum_{k=1}^{n} e_k h_{P_k} dx$.

We will consider the family of codes $C_\Omega(D, mQ)$. The check matrix is constructed by taking rows of the form $\mathrm{ev}(x^a z_b)$ for $a\kappa + \lambda_b \leq m$, arranged by increasing pole order. As in earlier sections, we assume $c \in C_\Omega(D, mQ)$ is sent, the vector $u \in \mathbb{F}_q^n$ is received, and $e = u - c$, the error vector, has weight t.

3.4.3. *The trace and a dual basis*

We have identified a basis for K over $\mathbb{F}_q(x)$; we now seek a dual basis for Ω. The dual basis is constructed using the intimate relationship between differentials and the trace map of an extension of function fields (see [41, II.4, IV.3], or [36, 13.12-13]). Let Tr be the trace map from K to $\mathbb{F}_q(x)$. Recall that the dual basis to $\{z_b\}_{b=0}^{\kappa-1}$ is the unique set of elements of K, $z_0^*, \ldots, z_{\kappa-1}^*$ such that $\mathrm{Tr}(z_b z_j^*)$ is 1 if $b = j$ and 0 otherwise.

We will use a result that appears as Proposition 8 in Ch. X of [22]: Let F be a separable finite extension of $k(x)$ and let $Q_1, \ldots, Q_r$ be the distinct points over a point P of $k(x)$. Let y be an element of F. Then

$$\sum_{i=1}^{r} \mathrm{res}_{Q_i}(y dx) = \mathrm{res}_P(\mathrm{Tr}(y) dx)$$

The theorem assumes k is an algebraically closed field. It is also true if k is not algebraically closed provided P, Q_i are rational points since the residues are defined for rational points and unchanged when one passes to the algebraic closure. In our case, let ∞ be the point on the projective line where x has a pole. On $\mathcal{C}$, x will also have a pole at any point mapping to ∞. Since the only pole of x is Q, the formula says $\mathrm{res}_Q(ydx) = \mathrm{res}_\infty(\mathrm{Tr}(y)dx)$ for any $y \in K$.

Proposition 3.33. *For each $b \in \{0, \ldots, \kappa - 1\}$, $z_b^* dx$ is an element of $\Omega(-\infty Q)$, $-z_b^* dx$ is monic, relative to u_Q, and $\nu_Q(z_b^* dx) = \lambda_b - \kappa - 1$. Additionally,*

$$\mathrm{res}_Q(z_j z_b^* x^a dx) = \begin{cases} -1 & \text{when } a = -1 \text{ and } j = b \\ 0 & \text{otherwise} \end{cases}$$

Proof. We will prove the residue formula first. Using the formula for the residue at Q and the property of the dual basis,

$$\begin{aligned} \mathrm{res}_Q(z_j z_b^* x^a dx) &= \mathrm{res}_\infty(x^a \,\mathrm{Tr}(z_j z_b^*) dx) \\ &= \begin{cases} \mathrm{res}_\infty(x^a dx) & \text{when } j = b \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

For the case $j = b$, note that $u = 1/x$ is a uniformizing parameter at ∞, and $x^a dx = u^{-a}(-u^{-2}du) = -u^{-a-2}du$. The residue is -1 when $a = -1$ and is zero otherwise.

Now let $j \equiv \nu_Q(z_b^* dx) + 1 \bmod \kappa$ and let a be such that

$$\nu_Q(z_b^* dx) = \lambda_j - 1 - (a+1)\kappa$$

Then

$$\nu_Q(z_j z_b^* x^{-a-1} dx) = -\lambda_j + (\lambda_j - 1 - (a+1)\kappa) + (a+1)\kappa = -1$$

Therefore, $\mathrm{res}_Q(z_j z_b^* x^{-a-1} dx) \neq 0$. By what we proved earlier, this can only be true when $j = b$ and $a = 0$. Therefore, $\nu_Q(z_b^* dx) = \lambda_b - 1 - \kappa$. Furthermore, since $\mathrm{res}(z_b z_b^* x^{-1} dx) = -1$, and z_b is monic, $-z_b^* dx$ is also monic (relative to u_Q).

Finally, we show $z_b^* dx \in \Omega(-\infty Q)$. From the residue formula we can see that for each z_j and any $h_j \in \mathbb{F}_q[x]$, $\mathrm{res}_Q(h_j z_j z_b^*) = 0$. Since any element of R can be expressed in the form $\sum_{j=0}^{\kappa-1} h_j z_j$, we conclude that $\mathrm{res}_Q(f z_b^* dx) = 0$ for any $f \in R$. Now suppose that $z_b^* dx$ has a pole at some point $P \neq Q$. By the strong approximation theorem, we may choose $f \in R$ to eliminate any other poles of $z_b^* dx$ away from P and Q and we may also

ensure that $\nu_P(fz_b^*dx) = -1$. Then $\mathrm{res}_Q(fz_b^*dx) = -\mathrm{res}_P(fz_b^*dx) \neq 0$, which contradicts what was shown above. Thus z_b^*dx can have a pole only at Q. □

Proposition 3.34. *With the notation above, $\Omega(-\infty Q)$ is a free module over $\mathbb{F}_q[x]$ with basis $\{z_b^*dx\}_{b=0}^{\kappa-1}$. This is also a basis for Ω over $\mathbb{F}_q(x)$.*

Proof. Let $l(mQ) = \dim L(mQ)$ and $i(mQ) = \dim\Omega(mQ)$. From the Riemann-Roch theorem one can show

$$l((m-1)Q) - l(mQ) = i((m-1)Q) - i(mQ) - 1$$

If $m \in \Lambda$, the left hand side is -1, so $i((m-1)Q) = i(mQ)$. Conversely, if $m \in \Lambda^c$ then the left hand side is 0, so $i((m-1)Q) = i(mQ) + 1$ and there is some $\omega \in \Omega(-\infty Q)$ such that $\nu_Q(\omega) = m - 1$. Thus

$$\{\nu_Q(\omega) + 1 : \omega \in \Omega(-\infty Q)\} = \Lambda^c = \bigcup_{b=0}^{\kappa-1} \{\lambda_b - a\kappa : a > 0\}$$

We now proceed as in Proposition 3.31. Let $\omega \in \Omega(-\infty Q)$ and let i and $a > 0$ be such that $\nu_Q(\omega) = \lambda_b - a\kappa - 1$. There is some $\alpha \in \mathbb{F}_q$ such that $\nu_Q(\omega - \alpha x^{a-1} z_b^* dx) > \lambda_b - a\kappa - 1$. Continuing in this manner, there exist $g_b \in \mathbb{F}_q[x]$ such that $\omega - \sum_{b=0}^{\kappa-1} g_b z_b^* dx$ has valuation at Q larger than $(2g-2)$. It is also in $\Omega(-\infty Q)$, so it has no poles away from Q. Thus $\omega - \sum_{b=0}^{\kappa-1} g_b z_b^* dx = 0$, for otherwise it would have degree greater than $(2g-2)$. This shows any $\omega \in \Omega(-\infty Q)$ is a combination of z_b^*dx with coefficients in $\mathbb{F}_q[x]$.

Uniqueness and the extension to Ω are shown as in Proposition 3.31.□

The next result is required to derive the error evaluation formula that is analogous to Theorem 3.30.

Proposition 3.35. *Let M/L be a finite separable field extension and let* Tr *be the trace map from M to L. Let $z_1, \ldots, z_n$ be a basis for M over L and let $z_1^*, \ldots, z_n^*$ be the dual basis. Then*

$$\sum_{i=1}^{n} z_i z_i^* = 1 \tag{3.9}$$

Proof. Since M is finite and separable over L there is some $y \in M$ such that $M = L(y)$. We will show the result first for the basis $1, y, \ldots, y^{n-1}$. Let $F(T) \in L[T]$ be the minimal polynomial of y and let $F'(T)$ be its formal

derivative. Let

$$\begin{aligned} C(T) &= \frac{F(T)}{T-y} \\ &= c_{n-1}T^{n-1} + c_{n-2}T^{n-2} + \cdots + c_1 T + c_0 \end{aligned}$$

where $c_i \in M$ and $c_{n-1} = 1$. The proof of [41, III.5.10] (or [23, VI.5.5]) shows that the dual basis to $1, y, y^2, \ldots, y^{n-1}$ is $c_0/F'(y), \ldots, c_{n-1}/F'(y)$. For this basis, the sum in (3.9) is

$$\sum_{i=1}^{\kappa-1} y^i \frac{c_i}{F'(y)} = \frac{1}{F'(y)} C(y) \tag{3.10}$$

In some algebraic closure of M, let $y_1, y_2, \ldots, y_{n-1}$ be the roots of F that are distinct from y and let $y_n = y$. Then $C(y) = \prod_{i=1}^{n-1}(y - y_i)$. Since $F'(T) = \sum_{i=1}^{n} \prod_{j\neq i}(T - y_i)$, $F'(y) = \prod_{i=1}^{n-1}(y - y_i) = C(y)$, so the sum in (3.10) is 1 as claimed.

Now suppose $\{z_i\}$ is another basis let $\{z_i^*\}$ be its dual basis, and let $\{y_i^*\}$ be the dual basis to $\{y^i\}$. Let M be the change of basis matrix from the z-basis to the y-basis: $z_a = \sum_{i=1}^{n} m_{a,i} y^i$. The change of basis matrix $\overline{M}$ from the z^* basis to the y^* basis is $(M^T)^{-1}$, as the following computation shows.

$$\begin{aligned} \delta_{a,b} = \operatorname{Tr}(z_a z_b^*) &= \operatorname{Tr}\left(\sum_{i=1}^{n} m_{a,i} y^i \sum_{j=1}^{n} \overline{m}_{b,j} y_j^* \right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} m_{a,i}\overline{m}_{b,j} \operatorname{Tr}(y^i y_j^*) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} m_{a,i}\overline{m}_{b,i} \end{aligned}$$

A similar computation shows $\sum_{a=1}^{n} z_a z_a^* = 1$,

$$\begin{aligned} \sum_{a=1}^{n} z_a z_a^* &= \sum_{a=1}^{n}\sum_{i=1}^{n}\sum_{j=1}^{n} m_{a,i} y^i \overline{m}_{a,j} y_j^* \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} y^i y_j^* \sum_{a=1}^{n} m_{a,i}\overline{m}_{a,j} \\ &= \sum_{i=1}^{n} y^i y_i^* = 1. \end{aligned}$$

□

Example 3.36. A natural generalization of Hermitian codes is the norm-trace codes, which were studied in [13]. Consider the field extension, $\mathbb{F}_{q^r}/\mathbb{F}_q$. Let N be the norm function and Tr the trace function for this extension. The norm-trace curve is $\mathrm{Tr}(y) = N(x)$, that is

$$\sum_{i=0}^{r-1} y^{q^i} = x^{\frac{q^r-1}{q-1}}$$

In the function field of this curve, y is a solution to the polynomial $F(T) \in \mathbb{F}_q(x)[T]$, $F(T) = \sum_{i=0}^{r-1} T^{q^i} - x^{\frac{q^r-1}{q-1}}$. Dividing by $T-y$ and substituting $\sum_{i=0}^{r-1} y^{q^i}$ for $x^{\frac{q^r-1}{q-1}}$ we get

$$\begin{aligned} C(t) &= \frac{1}{T-Y}\left(\sum_{i=0}^{r-1} T^{q^i} - \sum_{i=0}^{r-1} y^{q^i}\right) \\ &= \sum_{i=0}^{r-1} (T^{q^i} - y^{q^i})/(T-y) \\ &= \sum_{i=0}^{r-1} \sum_{j=0}^{q^i-1} T^j y^{q^i-1-j} \\ &= \sum_{j=0}^{q^{r-1}-1} T^j \sum_{i=\lceil \log_q(j+1) \rceil}^{r-1} y^{q^i-1-j} \end{aligned}$$

We also have $F'(T) = 1$. Thus the dual basis to $1, y, \ldots, y^{q^r-1}$ is $y_0^*, \ldots y_{q^r-1}^*$ where $y_j^* = \sum_{i=\lceil \log_q(j+1) \rceil}^{r-1} y^{q^i-1-j}$.

3.4.4. *Polynomials for decoding*

Define the *error locator ideal* of e to be

$$I^e = \{f \in R : f(P_k) = 0 \text{ for all } k \text{ with } e_k \neq 0\}$$

For a point P, let

$$h_P = \frac{1}{x - x(P)} \sum_{b=0}^{\kappa-1} z_b(P) z_b^*.$$

We define the *syndrome* of e to be

$$S = \sum_{k=1}^{n} e_k h_{P_k}.$$

As we did with Hermitian codes, we will give three justifications for this definition of the syndrome. The first is that the coefficients of S are the products $\mathrm{ev}(x^a z_b) \cdot e$.

Lemma 3.37. *Let* $s_{a,b} = \sum_{k=1}^{n} e_k (x(P_k))^a (z(P_k))^b$. *Then*

$$S = \frac{1}{x} \sum_{b=0}^{\kappa-1} \sum_{a=0}^{\infty} s_{a,b} x^{-a} z_b^*$$

Proof. Writing $(x - x(P))^{-1}$ as a series in $1/x$ we have

$$\begin{aligned} h_P &= \frac{1}{x} \left(\sum_{a=0}^{\infty} \left(\frac{x(P)}{x} \right)^a \right) \left(\sum_{b=0}^{\kappa-1} z_b(P) z_b^* \right) \qquad (3.11) \\ &= \sum_{b=0}^{\kappa-1} \sum_{a=0}^{\infty} (x(P))^a z_b(P) x^{-a} z_b^* \end{aligned}$$

Thus

$$\begin{aligned} S &= \frac{1}{x} \sum_{k=1}^{n} e_k \sum_{b=0}^{\kappa-1} \sum_{a=0}^{\infty} (x(P_k))^a z_b(P_k) x^{-a} z_b^* \\ &= \frac{1}{x} \sum_{b=0}^{\kappa-1} \sum_{a=0}^{\infty} x^{-a} z_b^* \sum_{k=1}^{n} e_k (x(P_k))^a z_b(P_k) \\ &= \frac{1}{x} \sum_{b=0}^{\kappa-1} \sum_{a=0}^{\infty} s_{a,b} x^{-a} z_b^* \qquad (3.12) \end{aligned}$$

□

For the next two properties of the syndrome, we first need the following lemma.

Lemma 3.38. *The differential* $h_P dx$ *has simple poles at* P *and* Q *and no other poles. Furthermore* $\mathrm{res}_Q\, h_P dx = -1$, *so* $-h_P dx$ *is monic with respect to* u_Q.

Proof. The valuation at Q of $\frac{1}{x-x(P)} z_b^* dx$ is $\lambda_b - 1$, and this is minimal for $b = 0$. Since $z_0 = 1$ and $-z_0^* dx$ is monic, $\nu_Q(h_p dx) = \nu_Q(\frac{1}{x-x(P)} z_0^* dx) = -1$ and the residue is -1.

Using the expansion for h_P in (3.11),

$$\begin{aligned} \mathrm{res}_Q(x^i z_j h_P dx) &= \sum_{a=0}^{\infty} \sum_{b=0}^{\kappa-1} (x(P))^a z_b(P) \,\mathrm{res}_Q(x^{i-a-1} z_j z_b^* dx) \\ &= -(x(P))^i z_j(P) \end{aligned}$$

Extending by the linearity of the residue map, for any $g \in R$, $\operatorname{res}_Q(gh_P dx) = -g(P)$.

We now show that $h_P dx$ has no pole at $P' \neq P, Q$. Suppose the contrary, $h_P dx$ has a pole at some $P' \neq P, Q$. By the strong approximation theorem, there is some $g \in R$ such that $gh_P dx$ has a zero at P, a simple pole at P' and no other poles, except at Q. Using the residue theorem we get a contradiction,

$$0 = g(P) = -\operatorname{res}_Q(gh_P dx) = \operatorname{res}_{P'}(gh_P dx) \neq 0$$

Similarly, we may show that the pole of $h_P dx$ at P is simple. If not, we could find a $g \in R$ with $\nu_P(g) = -\nu_P(h_P dx) - 1 > 0$. Again, we get a contradiction,

$$0 = g(P) = -\operatorname{res}_Q(gh_P dx) = \operatorname{res}_P(gh_p dx) \neq 0$$

□

The connection between the error locator ideal and the syndrome is now clear.

Lemma 3.39. *For $f \in R$, $f \in I^e$ if and only if $fSdx \in \Omega(-\infty Q)$.*

Proof. From the previous lemma, Sdx has a simple pole at each P_k where e_k is nonzero. Thus $fSdx \in \Omega(-\infty Q)$ if and only if $\nu_{P_k}(f) \geq 1$ whenever $e_k \neq 0$. This is just saying $f \in I^e$. □

Finally, we show that for $f \in I^e$, fS may be used for error evaluation.

Lemma 3.40. *Let P_k be an error position and let u_k be a uniformizing parameter at P_k. If f is an error locator and $\varphi = fSdx$, then*

$$e_k \frac{df}{du_k}(P_k) = \frac{\varphi}{du_k}(P_k) \tag{3.13}$$

Proof. Since f vanishes at P_k we can write $f = a_1 u_k + a_2 u_k^2 + \cdots$. Each h_{P_k} has a simple pole at P_k and no pole at P_j for $j \neq k$, so from the definition of S,

$$Sdx = \left(e_k u_k^{-1} + c_0 + c_1 u_k + \cdots\right) du_k$$

Thus

$$\frac{fSdx}{du_k} = e_k a_1 + \cdots$$

On the other hand,

$$\frac{df}{du_k} = a_1 + 2a_2u_k + \cdots$$

Evaluating the two at P_k amounts to setting $u_k = 0$, which gives the result.

□

To compute e_k using this formula, we need f to have a simple zero at P_k. The formula simplifies when $x - x(P_k)$ itself is a uniformizing parameter at P_k, $e_k \frac{df}{dx}(P_k) = fS(P_k)$.

3.4.5. *The key equation and its solution*

The key equation and the algorithm for solving it are little changed from those for Hermitian codes. We use κ instead of q in the indexing of z and z^*. The key equation uses differentials, not just polynomials. The key equation for Hermitian codes can be derived from the one in this section by dividing by dx, whose divisor is $(2g-2)Q$, and thereby shifting the pole order by $2g - 2 = q^2 - q - 2$. We will simply state the main results, and leave adaptations of the proofs in the previous section to the reader.

Definition 3.41. We say that $f \in R$ and $\varphi \in \Omega(-\infty Q)$ solve the key equation for syndrome S when $fSdx = \varphi$. We say that a nonzero $f \in R$ and $\varphi \in \Omega(-\infty Q)$ solve the K-th approximation of the key equation for syndrome S when the following two conditions hold.

(1) $\rho(fSdx - \varphi) \leq 1 - K$,
(2) φ, written in the $*$-basis, is a sum of terms whose order is at least $2-K$.

We will also say that 0 and $x^{-a-1}z_b^*dx$, for $a < 0$, solve the $a\kappa + \lambda_b$ key equation.

One could also express this definition in terms of the valuation ν_Q, f and φ solve the K-th key equation when $\nu_Q(fSdx - \varphi) \geq K - 1$. Since each $h_P dx$ has a simple pole at Q, $\nu_Q(Sdx) \geq -1$. Therefore, $\rho(z_b Sdx) \leq 1 - \lambda_b$, so the pair $z_b, 0$ satisfies the $-\lambda_b$ key equation. The pair $0, z_b^*$ solves the $\lambda_b - \kappa$ key equation.

Here are the three lemmas used in the proof that the decoding algorithm works.

Lemma 3.42. *Suppose that $f \neq 0$ and that f, φ satisfy the Kth key equation for syndrome S. If g and h are both monic of order K then* $\operatorname{res}_Q(gfS) = \operatorname{res}_Q(hfS)$.

Lemma 3.43. *Suppose that f, φ satisfy the Kth key equation. For any nonnegative integer i, $x^i f, x^i \varphi$ satisfy the $K - i\kappa$ key equation.*

Lemma 3.44. *Suppose that f, φ and g, ψ satisfy the Kth key equation where $K = a\kappa + \lambda_b$. Suppose in addition that $f \neq 0$ and $gSdx - \psi$ is monic of order $1 - K$. Let the coefficient of $x^{-a-1} z_b^*$ in $fSdx$ be μ. Then $f - \mu g$, $\varphi - \mu\psi$ satisfy the $(K+1)$th key equation.*

The decoding algorithm has only minor changes: κ replaces q, z_i replaces y^i and λ_i replaces $i(q+1)$.

Decoding algorithm for one-point codes

Initialize: For $i = 0$ to $\kappa - 1$, set $\begin{pmatrix} f_i^{(0)} & \varphi_i^{(0)} \\ g_i^{(0)} & \psi_i^{(0)} \end{pmatrix} = \begin{pmatrix} z_i & 0 \\ 0 & -z_i^* dx \end{pmatrix}$

Algorithm: For $m = 0$ to M, and for each pair i, j such that $m \equiv i + j \bmod \kappa$, set

$$d_i = \rho(f_i^{(m)}) \qquad d_j = \rho(f_j^{(m)})$$
$$r_i = \tfrac{m - d_i - \lambda_j}{\kappa} \qquad r_j = \tfrac{m - d_j - \lambda_i}{\kappa}$$
$$\tilde{f}_i = z_j f_i \qquad \tilde{f}_j = z_i f_j$$
$$\mu_i = \textstyle\sum_{c=0}^{\kappa-1} \sum_a (\tilde{f}_i)_{a,c} s_{a+r_i,c} \qquad \mu_j = \sum_{c=0}^{\kappa-1} \sum_a (\tilde{f}_j)_{a,c} s_{a+r_j,c}$$
$$p = \tfrac{d_i + d_j - m}{\kappa} - 1$$

The update for j is analogous to the one for i given below.

$$U_i^{(m)} = \begin{cases} \begin{pmatrix} 1 & -\mu_i x^p \\ 0 & 1 \end{pmatrix} & \text{if } \mu_i = 0 \text{ or } p \geq 0 \\ \begin{pmatrix} x^{-p} & -\mu_i \\ 1/\mu_i & 0 \end{pmatrix} & \text{otherwise.} \end{cases}$$

$$\begin{pmatrix} f_i^{(m+1)} & \varphi_i^{(m+1)} \\ g_j^{(m+1)} & \psi_j^{(m+1)} \end{pmatrix} = U_i^{(m)} \begin{pmatrix} f_i^{(m)} & \varphi_i^{(m)} \\ g_j^{(m)} & \psi_j^{(m)} \end{pmatrix}$$

Output: $f_i^{(M+1)}, \varphi_i^{(M+1)}$ for $0 \leq i < \kappa$.

One can check that at iteration m, $\rho(x^{r_i} z_j f_i^{(m)}) = m$. Using an argument analogous to the one in Lemma 3.16 one can show that $\mu_i = \operatorname{res}_Q(x^{r_i} z_j f_i^{(m)} Sdx)$ and that this is the coefficient of $x^{-r_i - 1} z_j^*$ in $f_i^{(m)} S$.

Theorem 3.45. *For $m \geq 0$,*

(1) $f_i^{(m)}$ *is monic and* $\rho(f_i^{(m)}) \equiv i \bmod \kappa$.
(2) $f_i^{(m)}, \varphi_i^{(m)}$ *satisfy the* $m - \rho(f_i^{(m)})$ *approximation of the key equation.*
(3) $g_i^{(m)}, \psi_i^{(m)}$ *satisfy the* $\rho(f_i^{(m)}) - \kappa$ *approximation of the key equation and* $g_i S dx - \psi_i^{(m)}$ *is monic of order* $1 + \kappa - \rho(f_i^{(m)})$.
(4) $\rho(g_i^{(m)}) < m - \rho(f_i^{(m)}) + \kappa$.

The iteration at which the algorithm can terminate depends on the set $\Delta^e = \Lambda - \rho(I^e)$ and the values $\sigma_i = \min\{\rho(f) : f \in I^e \text{ and } \rho(f) \equiv i \bmod \kappa\}$.

Proposition 3.46. *If* $\operatorname{res}_Q(x^a z_b^* f S dx) = 0$ *for all* a, b *such that that* $a\kappa + \lambda_b \in \Delta^e$ *then* $f \in I^e$. *In particular, if* f, φ *satisfy the* $\max \Delta^e$ *key equation, then* $f \in I^e$.

Proposition 3.47. *Let* $\sigma_{\max} = \max\{\sigma_i : 0 \le i \le \kappa - 1\}$ *and let* $\delta_{\max} = \max\{c \in \Delta^e\}$. *For* $m > \sigma_{\max} + \delta_{\max}$, *each of the polynomials* $f_i^{(m)}$ *belongs to* I^e. *Let* $M = \sigma_{\max} + \max\{\delta_{\max}, 2g - 1\}$. *Each of the pairs* $f_i^{(M+1)}, \varphi_i^{(M+1)}$ *satisfies the key equation.*

3.4.6. *Error evaluation without the error evaluator polynomials*

The error evaluation formula that we derived for Hermitian codes carries over to one-point codes. We have to stipulate that $x - x(P_k)$ has a simple zero at P_k, though it may be possible to remove this restriction. As was mentioned in the section on Hermitian codes, the derivation of the formula depends on the fact that at iteration m, and for $i + j \equiv m \bmod \kappa$, $\mu_i = \mu_j$ in the decoding algorithm. This is proven in Proposition 3.48 below.

In the proof of the proposition we will use the Cauchy-Binet Theorem. Let $\mathbf{B}, \mathbf{C}$ be $n \times 2$ matrices and let $\mathbf{T}$ be an $n \times n$ matrix such that $\mathbf{C} = \mathbf{TB}$. For I, J two-element subsets of $\{1, \ldots, n\}$, let $\mathbf{C}_I$ be the two rows of $\mathbf{C}$ indexed by I and let $\mathbf{T}_I^J$ be the 2×2 submatrix of $\mathbf{T}$ consisting of entries from the rows in I and the columns in J. The Cauchy-Binet theorem says that

$$\det \mathbf{C}_I = \sum_J \det \mathbf{T}_I^J \det \mathbf{B}_J$$

where the sum runs over all two-element subsets J of $\{1, \ldots, n\}$.

Proposition 3.48. *In the* m*th iteration of the algorithm,* $\mu_i = \mu_j$ *for* $i + j \equiv m \bmod \kappa$. *Furthermore for* $i \neq j$, *the coefficient of* z_0^* *in the* $\star$*-basis*

expansion of each of the following determinants is 0*:*

$$\det\begin{pmatrix} f_i^{(m)} & \varphi_i^{(m)} \\ f_j^{(m)} & \varphi_j^{(m)} \end{pmatrix}, \qquad \det\begin{pmatrix} g_i^{(m)} & \psi_i^{(m)} \\ g_j^{(m)} & \psi_j^{(m)} \end{pmatrix}, \qquad \det\begin{pmatrix} f_i^{(m)} & \varphi_i^{(m)} \\ g_j^{(m)} & \psi_j^{(m)} \end{pmatrix}.$$

The coefficient of z_0^* *is* $-dx$ *in*

$$\det\begin{pmatrix} f_i^{(m)} & \varphi_i^{(m)} \\ g_i^{(m)} & \psi_i^{(m)} \end{pmatrix}. \tag{3.14}$$

The formulas may also be expressed using residues via Proposition 3.33. The coefficient of z_0^* in D is 0 if and only if $\operatorname{res}_Q(x^a D) = 0$ for all a. Equation (3.14) is equivalent to saying that

$$\operatorname{res}_Q\left(x^a \det\begin{pmatrix} f_i^{(m)} & \varphi_i^{(m)} \\ g_i^{(m)} & \psi_i^{(m)} \end{pmatrix}\right) = \begin{cases} 1 & \text{if } a = -1 \\ 0 & \text{otherwise} \end{cases}.$$

Proof. The proof proceeds by induction. The determinental conditions are readily verified for $m = 0$. The inductive step has two parts. First, we show that if the determinental conditions hold for m, then $\mu_i = \mu_j$ for $i + j \equiv m \bmod \kappa$ in the mth iteration of the algorithm. Then we show that the determinental conditions hold for $m + 1$.

Assume the determinental conditions hold for m. Let $i + j \equiv m \bmod \kappa$ and let μ_i, μ_j, r_i, and r_j be as in the algorithm. We will suppress the superscript (m) on $f_i^{(m)}$ and the other data. We will show below that

$$\mu_i = \operatorname{res}_Q\left(x^{-p-1} f_j f_i S dx - x^{-p-1} f_j \varphi_i\right). \tag{3.15}$$

One of the hypotheses of the lemma is that the coefficient of z_0^* in $f_j\varphi_i - f_i\varphi_j$ is 0. Thus, we may substitute $f_i\varphi_j$ for $f_j\varphi_i$ in (3.15) to say that

$$\mu_i = \operatorname{res}_Q\left(x^{-p-1} f_j f_i S dx - x^{-p-1} f_j \varphi_i\right). \tag{3.16}$$

The right hand side of this formula is the analogue of (3.15) with j and i switched. This shows that $\mu_j = \mu_i$.

To establish (3.15), we apply item (2) of Theorem 3.45 to obtain

$$\rho(f_i S dx - \varphi_i) \leq 1 + \rho(f_i) - m.$$

As we noted before Theorem 3.45, μ_i is the coefficient of $x^{-r_i-1} z_j^*$ in $f_i S$. Thus,

$$\rho(f_i S dx - \varphi_i - \mu_i x^{-r_i-1} z_j^* dx) < 1 + \rho(f_i) - m.$$

Multiplying by $x^{-p-1} f_j$ we have

$$\rho(x^{-p-1} f_j f_i S dx - x^{-p-1} f_j \varphi_i - \mu_i x^{-p-1} f_j x^{-r_i-1} z_j^* dx) < 1.$$

Equivalently, the valuation of the expression is nonnegative. This shows that $\mathrm{res}_Q(x^{-p-1}f_jf_iSdx - x^{-p-1}f_j\varphi_i) = \mathrm{res}_Q(\mu_i x^{-p-r_i-2}f_jz_j^*dx)$. The expression on the right has valuation -1, and residue μ_i, which establishes (3.15).

We now prove that the determinental conditions of the lemma hold for $m+1$. Let

$$B_i^{(m)} = \begin{pmatrix} f_i^{(m)} & \varphi_i^{(m)} \\ g_i^{(m)} & \psi_i^{(m)} \end{pmatrix}, \qquad \text{and let} \quad \mathbf{B}^{(m)} = \begin{pmatrix} B_0^{(m)} \\ B_1^{(m)} \\ B_2^{(m)} \\ \dots \\ B_{\kappa-1}^{(m)} \end{pmatrix}$$

Let $\mathbf{T}$ be the update matrix for the mth iteration, so $\mathbf{B}^{(m+1)} = \mathbf{T}\mathbf{B}^{(m)}$. We want to show that for $I \subseteq \{1,\dots,2\kappa\}$ and $\mathbf{B}_I^{(m+1)}$ the appropriate 2×2 submatrix, the coefficient of z_0^* in $\det\mathbf{B}_I^{(m+1)}$ is 0 unless I is a consecutive pair of the form $\{2i+1, 2i+2\}$ for $i = 0,\dots,\kappa-1$. From the inductive hypotheses, the coefficient of z_0^* in $\det\mathbf{B}_I^{(m)}$ is only nonzero for these I. Consequently, from the Cauchy-Binet theorem

$$\mathrm{res}_Q\left(x^a \det\mathbf{B}_I^{(m+1)}\right) = \sum_J \mathrm{res}_Q\left(x^a \det\mathbf{T}_I^J \det\mathbf{B}_J^{(m)}\right) \tag{3.17}$$

where the sum runs over all J of the form $\{2j+1, 2j+2\}$.

From the algorithm, for $i + j \equiv m \bmod \kappa$ and $i \neq j$,

$$\begin{pmatrix} f_i^{(m+1)} & \varphi_i^{(m+1)} \\ g_i^{(m+1)} & \psi_i^{(m+1)} \\ f_j^{(m+1)} & \varphi_j^{(m+1)} \\ g_j^{(m+1)} & \psi_j^{(m+1)} \end{pmatrix} = \begin{cases} \begin{pmatrix} 1 & 0 & 0 & -\mu x^p \\ 0 & 1 & 0 & 0 \\ 0 & -\mu x^p & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f_i^{(m)} & \varphi_i^{(m)} \\ g_i^{(m)} & \psi_i^{(m)} \\ f_j^{(m)} & \varphi_j^{(m)} \\ g_j^{(m)} & \psi_j^{(m)} \end{pmatrix}, & \text{if } \mu = 0 \text{ or } p \geq 0 \\ \begin{pmatrix} x^{-p} & 0 & 0 & -\mu \\ 0 & 0 & 1/\mu & 0 \\ 0 & -\mu & x^{-p} & 0 \\ 1/\mu & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} f_i^{(m)} & \varphi_i^{(m)} \\ g_i^{(m)} & \psi_i^{(m)} \\ f_j^{(m)} & \varphi_j^{(m)} \\ g_j^{(m)} & \psi_j^{(m)} \end{pmatrix} & \text{otherwise.} \end{cases} \tag{3.18}$$

Notice that we have used $\mu = \mu_i = \mu_j$. Of course, if $i = j$, *i.e.* $2i = m \bmod \kappa$, then the formula is simpler, $B_i^{(m+1)} = U_i^{(m)}B_i^{(m)}$, with $U_i^{(m)}$ from the algorithm.

In the formula (3.17), we consider two cases for I. If there is no i, j with $i + j \equiv m \bmod \kappa$ such that $I \subseteq \{2j+1, 2j+2, 2i+1, 2i+2\}$ then for all J

of the form $\{2j+1, 2j+2\}$, $\mathbf{T}_I^J$ has a row that is all zeros, and $\det\mathbf{T}_I^J = 0$. For such I, we therefore have $\mathrm{res}_Q\left(x^a \det\mathbf{B}_I^{(m+1)}\right) = 0$.

Now consider $I \subseteq \{2j+1, 2j+2, 2i+1, 2i+2\}$ with $i + j \equiv m \bmod \kappa$. Similar reasoning shows that

$$\mathrm{res}_Q\left(x^a \det\mathbf{B}_I^{(m+1)}\right) = \mathrm{res}_Q\left(x^a \det\mathbf{T}_I^J \det\mathbf{B}_J^{(m)}\right) + \mathrm{res}_Q\left(x^a \det\mathbf{T}_I^{\bar{J}} \det\mathbf{B}_{\bar{J}}^{(m)}\right)$$

where $J = \{2j+1, 2j+2\}$ and $\bar{J} = \{2i+1, 2i+2\}$. There are $\binom{4}{2}$ choices of I to check for each of the two possible update matrices. If $I = J$ or $I = \bar{J}$, then either $\det\mathbf{T}_I^J = 1$ and $\det\mathbf{T}_I^J = 0$ or vice-versa depending on the matrix. Thus the induction hypothesis shows that the coefficient of z_0^* in $\det\mathbf{B}_I^{(m)}$ is $-dx$ as desired. For $I = \{2i+1, 2j+2\}$ or $\{2i+2, 2j+2\}$, and for either update matrix, $\det\mathbf{T}_I^J = \det\mathbf{T}_I^{\bar{J}} = 0$. Thus the coefficient of z_0^* in $\det\mathbf{B}_I^{(m)}$ is 0 as desired. Finally, for $I = \{2i+1, 2j+1\}$, and for either update matrix, $\det\mathbf{T}_I^J = -\det\mathbf{T}_I^{\bar{J}}$ and this is a monomial in x.

$$\mathrm{res}_Q\left(x^a \det\mathbf{B}_I^{(m+1)}\right) = \mathrm{res}_Q\left(x^a \det\mathbf{T}_I^J \left(\det\mathbf{B}_J^{(m)} - \det\mathbf{B}_{\bar{J}}^{(m)}\right)\right)$$

The induction hypothesis says that the coefficient of z_0^* is the same in $\det\mathbf{B}_J^{(m)}$ and $\det\mathbf{B}_{\bar{J}}^{(m)}$. Thus the coefficient of z_0^* in $\det\mathbf{B}_I^{(m+1)}$ is 0 as desired. □

Proposition 3.49. *Let* $B_i^{(M)} = \begin{pmatrix} f_i^{(m)} & \varphi_i^{(m)} \\ g_i^{(m)} & \psi_i^{(m)} \end{pmatrix}$. *Then for all* m,

$$\sum_{i=0}^{\kappa-1} \det B_i^{(m)} = -dx\left(\sum_{i=0}^{\kappa-1} z_i z_i^*\right) = -dx \tag{3.19}$$

Theorem 3.50. *Suppose that* $x - x(P_k)$ *is a uniformizing parameter at an error position* P_k. *Let* $f' = df/dx$. *Then*

$$e_k = \left(\sum_{i=0}^{\kappa-1} f_i'(P_k) g_i(P_k)\right)^{-1} \tag{3.20}$$

3.5. Bibliographical notes

The history of the key equation may be divided into three stages. In the first stage there is the key equation and iterative solution of it in Berlekamp's book [2], and a more implementation oriented approach in Massey's article [27]. These articles build on the Peterson-Gorenstein-Zierler decoding algorithm [14, 34] and Forney's improvements [12], which use matrices and are less efficient.

The second stage includes two new algorithms. Sugiyama et al [42] define a key equation and give an efficient solution to it using the Euclidean algorithm. The Welch-Berlekamp algorithm [44] is related to the rational interpolation problem and has its own key equation. A number of articles explore the algebraic formulation of these algorithms, efficient implementation, or the relationship between the different algorithms. Among these we mention Fitzpatrick's article on the key equation [11], comparisons of the Euclidean and Berlekamp-Massey algorithms by Dornstetter [6] and Heydtmann and Jensen [16], and comparisons of key equations in Moon and Gunther [28], Morii and Kasahara [29], and Yaghoobian and Blake [45]. A more extensive discussion and bibliography may be found in Roth's textbook [37, Ch. 6].

A third stage concerns the extension of the key equation and decoding algorithms to algebraic geometry codes. The key breakthrough was Sakata's algorithm for finding linear recurrence relations for higher dimensional arrays [38]. We are using Kötter's version of the algorithm for algebraic curves [21], in which the ring of functions is treated as a module over a polynomial ring. The Forney formula is generalized for one-point codes in Hansen et al [15] and in Leonard [25, 26]. Several generalizations of the key equation have appeared. Chabanne and Norton [5] work with a polynomial ring in several variables and express the syndrome as a power series. The key equation is generalized to arbitrary codes on curves by Ehrhard [8], Porter, Shen and Pellikaan [35], and by Farrán [9]. A later paper by Shen and Tzeng [39], deals with one-point codes. There are elements of all these approaches in this chapter, but we have maintained the focus on one-point codes, where the generalizations are particularly simple, and the treatment is based on the articles of O'Sullivan [30, 32, 33].

References

[1] Roger Apéry. Sur les branches superlinéaires des courbes algébriques. *C. R. Acad. Sci. Paris*, 222:1198–1200, 1946.

[2] Elwyn R. Berlekamp. *Algebraic coding theory.* McGraw-Hill Book Co., New York, 1968.

[3] Richard Blahut. *Algebraic codes for data transmission.* Cambridge University Press, Cambridge, UK, 2003.

[4] Maria Bras-Amorós and Michael E. O'Sullivan. The correction capability of the Berlekamp-Massey-Sakata algorithm with majority voting. *Appl. Algebra Engrg. Comm. Comput.*, 17(5):315–335, 2006.

[5] Hervé Chabanne and Graham H. Norton. The n-dimensional key equation

and a decoding application. *IEEE Trans. Inform. Theory*, 40(1):200–203, 1994.

[6] Jean-Louis Dornstetter. On the equivalence between Berlekamp's and Euclid's algorithms. *IEEE Trans. Inform. Theory*, 33(3):428–431, 1987.

[7] Ivan M. Duursma. Majority coset decoding. *IEEE Trans. Inform. Theory*, 39(3):1067–1070, 1993.

[8] Dirk Ehrhard. Decoding algebraic-geometric codes by solving a key equation. In *Coding theory and algebraic geometry (Luminy, 1991)*, volume 1518 of *Lecture Notes in Math.*, pages 18–25. Springer, Berlin, 1992.

[9] José-Ignacio Farrán. Decoding algebraic geometry codes by a key equation. *Finite Fields Appl.*, 6(3):207–217, 2000.

[10] Gui Liang Feng and Thammavarapu R. N. Rao. Decoding algebraic-geometric codes up to the designed minimum distance. *IEEE Trans. Inform. Theory*, 39(1):37–45, 1993.

[11] Patrick Fitzpatrick. On the key equation. *IEEE Trans. Inform. Theory*, 41(5):1290–1302, 1995.

[12] G. David Forney, Jr. On decoding BCH codes. *IEEE Trans. Inform. Theory*, IT-11:549–557, 1965.

[13] Olav Geil. On codes from norm-trace curves. *Finite Fields Appl.*, 9(3):351–371, 2003.

[14] Daniel Gorenstein and Neal Zierler. A class of error-correcting codes in p^m symbols. *J. Soc. Indust. Appl. Math.*, 9:207–214, 1961.

[15] Johan P. Hansen, Helge Elbrønd Jensen, and Ralf Kötter. Determination of error values for algebraic-geometry codes and the Forney formula. *IEEE Trans. Inform. Theory*, 44(5):1881–1886, 1998.

[16] Agnes E. Heydtmann and Jørn M. Jensen. On the equivalence of the Berlekamp-Massey and the Euclidean algorithms for decoding. *IEEE Trans. Inform. Theory*, 46(7):2614–2624, 2000.

[17] Tom Høholdt, Jacobus H. van Lint, and Ruud Pellikaan. Algebraic geometry of codes. In *Handbook of coding theory, Vol. I, II*, pages 871–961. North-Holland, Amsterdam, 1998. Vol. I.

[18] Toshio Horiguchi. High-speed decoding of BCH codes using a new error-evaluation algorithm. *Electronics and Communications in Japan*, 72(12):63–71, 1989.

[19] Jørn Justesen and Tom Høholdt. *A course in error-correcting codes*. EMS Textbooks in Mathematics. European Mathematical Society (EMS), Zürich, 2004.

[20] Ralf Kötter. On the determination of error values for codes from a class of maximal curves. In *Proc. 35-th Allerton Conference on Communication, Control, and Computing*, pages 44–53, 1997.

[21] Ralf Kötter. A fast parallel implementation of a Berlekamp-Massey algorithm for algebraic-geometric codes. *IEEE Trans. Inform. Theory*, 44(4):1353–1368, 1998.

[22] Serge Lang. *Introduction to algebraic geometry*. Addison-Wesley Publishing Co., Inc., Reading, Mass., 1972. Third printing, with corrections.

[23] Serge Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*.

Springer-Verlag, New York, third edition, 2002.

[24] Kwankyu Lee and Michael E. O'Sullivan. List decoding of Reed-Solomon codes from a Gröbner basis perspective. *Journal of Symbolic Computation*, to appear.

[25] Douglas A. Leonard. A generalized Forney formula for algebraic-geometric codes. *IEEE Trans. Inform. Theory*, 42(4):1263–1268, 1996.

[26] Douglas A. Leonard. Efficient Forney functions for decoding AG codes. *IEEE Trans. Inform. Theory*, 45(1):260–265, 1999.

[27] James L. Massey. Shift-register synthesis and BCH decoding. *IEEE Trans. Inform. Theory*, IT-15:122–127, 1969.

[28] Todd K. Moon and Jacob H. Gunther. On the equivalence of two Welch-Berlekamp key equations and their error evaluators. *IEEE Trans. Inform. Theory*, 51(1):399–401, 2005.

[29] Masakatu Morii and Masao Kasahara. Generalized key-equation of remainder decoding algorithm for Reed-Solomon codes. *IEEE Trans. Inform. Theory*, 38(6):1801–1807, 1992.

[30] Michael E. O'Sullivan. Decoding of Hermitian codes: the key equation and efficient error evaluation. *IEEE Trans. Inform. Theory*, 46(2):512–523, 2000.

[31] Michael E. O'Sullivan. New codes for the Berlekamp-Massey-Sakata algorithm. *Finite Fields Appl.*, 7(2):293–317, 2001.

[32] Michael E. O'Sullivan. The key equation for one-point codes and efficient error evaluation. *J. Pure Appl. Algebra*, 169(2-3):295–320, 2002.

[33] Michael E. O'Sullivan. On Koetter's algorithm and the computation of error values. *Des. Codes Cryptogr.*, 31(2):169–188, 2004.

[34] W. Wesley Peterson. Encoding and error-correction procedures for the Bose-Chaudri codes. *IRE Transactions on Information Theory*, 6:459–470, 1960.

[35] Sidney C. Porter, Ba-Zhong Shen, and Ruud Pellikaan. Decoding geometric Goppa codes using an extra place. *IEEE Trans. Inform. Theory*, 38(6):1663–1676, 1992.

[36] Oliver Pretzel. *Codes and algebraic curves*, volume 8 of *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press Oxford University Press, New York, 1998.

[37] Ron Roth. *Introduction to coding theory*. Cambridge University Press, Cambridge, 2006.

[38] Shojiro Sakata. Extension of Berlekamp-Massey algorithm to n dimensions. *Inform. and Comput.*, 84(2):207–239, 1990.

[39] Ba-Zhong Shen and Kenneth K. Tzeng. Decoding geometric Goppa codes up to designed minimum distance by solving a key equation in a ring. *IEEE Trans. Inform. Theory*, 41(6, part 1):1694–1702, 1995. Special issue on algebraic geometry codes.

[40] Henning Stichtenoth. A note on Hermitian codes over $\mathrm{GF}(q^2)$. *IEEE Trans. Inform. Theory*, 34(5, part 2):1345–1348, 1988.

[41] Henning Stichtenoth. *Algebraic function fields and codes*. Universitext. Springer-Verlag, Berlin, 1993.

[42] Yasuo Sugiyama, Masao Kasahara, Shigeichi Hirasawa, and Toshihiko Namekawa. A method for solving key equation for decoding Goppa codes.

Information and Control, 27:87–99, 1975.

[43] Herman J. Tiersma. Remarks on codes from Hermitian curves. *IEEE Trans. Inform. Theory*, 33(4):605–609, 1987.

[44] Lloyd R. Welch and Elwyn R. Berlekamp. Error correction for algebraic block codes, 1983. US Patent 4 633 470.

[45] Tomik Yaghoobian and Ian F. Blake. Two new decoding algorithms for Reed-Solomon codes. *Appl. Algebra Engrg. Comm. Comput.*, 5(1):23–43, 1994.

Table 3.1. Steps for correcting two errors in general position.

$\mathbf{m}$	$\mathbf{i}$	$\mathbf{j}$	$\mathbf{r_i}$	$\tilde{\mathbf{r}}_\mathbf{i}$	$\mu_\mathbf{i}$	$\mathbf{p}$	$\mathbf{U_i^{(m)}}$	$\mathbf{f_i^{(m+1)}}$	$\varphi_\mathbf{i}^{(\mathbf{m+1})}$	$\mathbf{g_i^{(m+1)}}$	$\psi_\mathbf{i}^{(\mathbf{m+1})}$
-1	0							1	0	0	$2y^2+2$
	1							y	0	0	$2y$
	2							y^2	0	0	2
0	0	0	0	1	α^5	-1	$[[x,\alpha],[\alpha^3,0]]$	x	$\alpha^5y^2+\alpha^5$	α^3	0
	1	2	-4	x^4+2y	α^5	3	$[[1,\alpha x^3],[0,1]]$	y	α^5x^3		
	2	1	-4	x^4+2y	α^5	3	$[[1,\alpha x^3],[0,1]]$	y^2	α^5x^3y		
1	0	1	-2	xy	0	1					
	1	0	-1	y	0	1					
	2	2	-5	x^4y+2y^2	0	4					
2	0	2	-3	xy^2	0	2					
	1	1	-2	y^2	0	1					
	2	0	-2	y^2	0	2					
3	0	0	0	x	α^2	0	$[[1,\alpha^6],[0,1]]$	$x+\alpha$	$\alpha^5y^2+\alpha^5$		
	1	2	-3	x^4+2y	α^2	2	$[[1,\alpha^6x^2],[0,1]]$	y	$\alpha^5x^3+\alpha^2x^2$		
	2	1	-3	x^4+2y	α^2	2	$[[1,\alpha^6x^2],[0,1]]$	y^2	$\alpha^5x^3y+\alpha^2x^2y$		
4	0	1	-1	$xy+\alpha y$	2	0	$[[1,1],[0,1]]$	$x+\alpha$	$\alpha^5y^2+2y+\alpha^5$		
	1	0	0	y	2	0	$[[1,1],[0,1]]$	$y+\alpha^3$	$\alpha^5x^3+\alpha^2x^2$		
	2	2	-4	x^4y+2y^2	2	3	$[[1,x^3],[0,1]]$	y^2	$\alpha^5x^3y+\alpha^2x^2y+2x^3$		
5	0	2	-2	$xy^2+\alpha y^2$	0	1					
	1	1	-1	$y^2+\alpha^3y$	0	0					
	2	0	-1	y^2	0	1					
6	0	0	1	$x+\alpha$	α^2	-1	$[[x,\alpha^6],[\alpha^6,0]]$	$x^2+\alpha x+\alpha$	$\alpha^5xy^2+2xy+\alpha^5x$	$\alpha^6x+\alpha^7$	$\alpha^3y^2+\alpha^2y+\alpha^3$
	1	2	-2	$x^4+\alpha^3y^2+2y$	α^5	1	$[[1,\alpha x],[0,1]]$	$y+\alpha^3$	$\alpha^5x^3+\alpha^2x^2+\alpha^5x$		
	2	1	-2	x^4+2y	α^5	1	$[[1,\alpha x],[0,1]]$	y^2	$\alpha^5x^3y+\alpha^2x^2y+2x^3+\alpha^5xy$		
7	0	1	-1	$x^2y+\alpha xy+\alpha y$	α^3	0	$[[1,\alpha^7],[0,1]]$	$x^2+\alpha x+\alpha$	$\alpha^5xy^2+2xy+\alpha^3y+\alpha^5x$		
	1	0	1	$y+\alpha^3$	α^3	0	$[[1,\alpha^7],[0,1]]$	$y+\alpha^5x+\alpha$	$\alpha^5x^3+\alpha^2y^2+\alpha^2x^2+\alpha y+\alpha^5x+\alpha^2$		
	2	2	-3	x^4y+2y^2	1	2	$[[1,2x^2],[0,1]]$	y^2	$\alpha^5x^3y+\alpha^2x^2y+2x^3+\alpha^5xy+x^2$		
8	0	2	-2	$x^2y^2+\alpha xy^2+\alpha y^2$	α^5	1	$[[1,\alpha x],[0,1]]$	$x^2+\alpha x+\alpha$	$\alpha^5xy^2+2xy+\alpha^3y+\alpha x$		
	1	1	0	$y^2+\alpha^5xy+\alpha y$	0	-1					
	2	0	0	y^2	α^5	1	$[[1,\alpha x],[0,1]]$	$y^2+\alpha^7x^2+x$	$\alpha^5x^3y+2xy^2+\alpha^2x^2y+2x^3+\alpha^2xy+x^2+2x$		
9	0	0	1	$x^2+\alpha x+\alpha$	α	0	$[[1,\alpha^5],[0,1]]$	$x^2+x+\alpha^7$	$\alpha^5xy^2+y^2+2xy+\alpha x+1$		
	1	2	-1	$x^4+\alpha^5xy^2+\alpha y^2+2y$	α^2	0	$[[1,\alpha^6],[0,1]]$	$y+\alpha^5x+\alpha$	$\alpha^5x^3+\alpha^2y^2+\alpha^2x^2+\alpha y+\alpha^5x+\alpha^6$		
	2	1	-1	$x^4+\alpha^7x^2y+xy+2y$	α^2	0	$[[1,\alpha^6],[0,1]]$	$y^2+\alpha^7x^2+x$	$\alpha^5x^3y+2xy^2+\alpha^2x^2y+2x^3+\alpha^2xy+x^2+\alpha^2y+2x$		
10	0	1	0	$x^2y+xy+\alpha^7y$	0	-1					
	1	0	2	$y+\alpha^5x+\alpha$	0	-1					
	2	2	-2	$x^4y+\alpha^7x^2y^2+xy^2+2y^2$	α	1	$[[1,\alpha^5x],[0,1]]$	$y^2+\alpha^7x^2+x$	$\alpha^5x^3y+2xy^2+\alpha^2x^2y+2x^3+\alpha^2xy+x^2+\alpha^2y+\alpha^7x$		
11	0	2	-1	$x^2y^2+xy^2+\alpha^7y^2$	1	0	$[[1,2],[0,1]]$	$x^2+x+\alpha^7$	$\alpha^5xy^2+y^2+2xy+\alpha x+2$		
	1	1	1	$y^2+\alpha^5xy+\alpha y$	0	-2					
	2	0	1	$y^2+\alpha^7x^2+x$	1	0	$[[1,2],[0,1]]$	$y^2+\alpha^7x^2+\alpha^7x+\alpha^3$	$\alpha^5x^3y+2xy^2+\alpha^2x^2y+2x^3+\alpha^7y^2+\alpha^2xy+x^2+\alpha^7x+\alpha^7$		
12	0	0	2	$x^2+x+\alpha^7$	0	-1					
	1	2	0	$x^4+\alpha^5xy^2+\alpha y^2+2y$	0	-1					
	2	1	0	$x^4+\alpha^7x^2y+\alpha^7xy+\alpha^5y$	0	-1					
13	0	1	1	$x^2y+xy+\alpha^7y$	0	-2					
	1	0	3	$y+\alpha^5x+\alpha$	0	-2					
	2	2	-1	$x^4y+\alpha^7x^2y^2+\alpha^7xy^2+\alpha^5y^2$	α^6	0	$[[1,\alpha^2],[0,1]]$	$y^2+\alpha^7x^2+\alpha^7x+\alpha^3$	$\alpha^5x^3y+2xy^2+\alpha^2x^2y+2x^3+\alpha^7y^2+\alpha^2xy+x^2+\alpha^7x+1$		

Table 3.2. Steps for correcting two errors in positions on a vertical line.

$\mathbf{m}$	$\mathbf{i}$ $\mathbf{j}$ $\mathbf{r_i}$ $\tilde{\mathbf{f}}_\mathbf{i}$	$\mu_\mathbf{i}$ $\mathbf{p}$	$\mathbf{U_i^{(m)}}$	$\mathbf{f_i^{(m+1)}}$	$\varphi_\mathbf{i}^{(\mathbf{m+1})}$	$\mathbf{g_i^{(m+1)}}$	$\psi_\mathbf{i}^{(\mathbf{m+1})}$
-1	0			1	0	0	$2y^2+2$
	1			y	0	0	$2y$
	2			y^2	0	0	2
0	0 0 0 1	α^5 -1	$[[x,\alpha],[\alpha^3,0]]$	x	$\alpha^5y^2+\alpha^5$	α^3	0
	1 2 -4 x^4+2y	α^5 3	$[[1,\alpha x^3],[0,1]]$	y	α^5x^3		
	2 1 -4 x^4+2y	α^5 3	$[[1,\alpha x^3],[0,1]]$	y^2	α^5x^3y		
1	0 1 -2 xy	0 1					
	1 0 -1 y	0 1					
	2 2 -5 x^4y+2y^2	0 4					
2	0 2 -3 xy^2	0 2					
	1 1 -2 y^2	0 1					
	2 0 -2 y^2	0 2					
3	0 0 0 x	α^7 0	$[[1,\alpha^3],[0,1]]$	$x+\alpha^6$	$\alpha^5y^2+\alpha^5$		
	1 2 -3 x^4+2y	α^7 2	$[[1,\alpha^3x^2],[0,1]]$	y	$\alpha^5x^3+\alpha^7x^2$		
	2 1 -3 x^4+2y	α^7 2	$[[1,\alpha^3x^2],[0,1]]$	y^2	$\alpha^5x^3y+\alpha^7x^2y$		
4	0 1 -1 $xy+\alpha^6y$	α^7 0	$[[1,\alpha^3],[0,1]]$	$x+\alpha^6$	$\alpha^5y^2+\alpha^7y+\alpha^5$		
	1 0 0 y	α^7 0	$[[1,\alpha^3],[0,1]]$	$y+\alpha^6$	$\alpha^5x^3+\alpha^7x^2$		
	2 2 -4 x^4y+2y^2	α^7 3	$[[1,\alpha^3x^3],[0,1]]$	y^2	$\alpha^5x^3y+\alpha^7x^2y+\alpha^7x^3$		
5	0 2 -2 $xy^2+\alpha^6y^2$	0 1					
	1 1 -1 $y^2+\alpha^6y$	0 0					
	2 0 -1 y^2	0 1					
6	0 0 1 $x+\alpha^6$	0 -1					
	1 2 -2 $x^4+\alpha^6y^2+2y$	α 1	$[[1,\alpha^5x],[0,1]]$	$y+\alpha^6$	$\alpha^5x^3+\alpha^7x^2+\alpha x$		
	2 1 -2 x^4+2y	α 1	$[[1,\alpha^5x],[0,1]]$	y^2	$\alpha^5x^3y+\alpha^7x^2y+\alpha^7x^3+\alpha xy$		
7	0 1 0 $xy+\alpha^6y$	0 -1					
	1 0 1 $y+\alpha^6$	0 -1					
	2 2 -3 x^4y+2y^2	α 2	$[[1,\alpha^5x^2],[0,1]]$	y^2	$\alpha^5x^3y+\alpha^7x^2y+\alpha^7x^3+\alpha xy+\alpha x^2$		
8	0 2 -1 $xy^2+\alpha^6y^2$	α^2 0	$[[1,\alpha^6],[0,1]]$	$x+\alpha^6$	$\alpha^5y^2+\alpha^7y+1$		
	1 1 0 $y^2+\alpha^6y$	1 -1	$[[x,2],[1,0]]$	$xy+\alpha^6x$	$\alpha^5x^4+\alpha^7x^3+\alpha x^2+y$	$y+\alpha^6$	$\alpha^5x^3+\alpha^7x^2+\alpha x$
	2 0 0 y^2	α^2 0	$[[1,\alpha^6],[0,1]]$	$y^2+\alpha$	$\alpha^5x^3y+\alpha^7x^2y+\alpha^7x^3+\alpha xy+\alpha x^2$		
9	0 0 2 $x+\alpha^6$	0 -2					
	1 2 -2 $x^5+\alpha^6xy^2+2xy$	α^3 1	$[[1,\alpha^7x],[0,1]]$	$xy+\alpha^6x$	$\alpha^5x^4+\alpha^7x^3+\alpha x^2+y+\alpha^3x$		
	2 1 -1 $x^4+\alpha^7y$	α^3 1	$[[1,\alpha^7x],[0,1]]$	$y^2+\alpha^7xy+\alpha^5x+\alpha$	$\alpha^5x^3y+2x^4+\alpha^7x^2y+x^3+\alpha xy+\alpha^2x^2$		
10	0 1 1 $xy+\alpha^6y$	0 -1					
	1 0 1 $xy+\alpha^6x$	0 -1					
	2 2 -2 $x^4y+\alpha^7x^5+\alpha^5xy^2+\alpha^7y^2+\alpha^3xy$	2 1	$[[1,x],[0,1]]$	$y^2+\alpha^7xy+\alpha^5x+\alpha$	$\alpha^5x^3y+2x^4+\alpha^7x^2y+x^3+\alpha xy+\alpha^2x^2+2x$		
11	0 2 0 $xy^2+\alpha^6y^2$	0 -1					
	1 1 0 $xy^2+\alpha^6xy$	α^2 0	$[[1,\alpha^6],[0,1]]$	$xy+\alpha^6y+\alpha^6x+2$	α^5x^4+y		
	2 0 1 $y^2+\alpha^7xy+\alpha^5x+\alpha$	0 -1					
12	0 0 3 $x+\alpha^6$	0 -3					
	1 2 -1 $x^5+\alpha^6x^4+\alpha^6xy^2+2y^2+2xy+\alpha^2y$	α^6 0	$[[1,\alpha^2],[0,1]]$	$xy+\alpha^6y+\alpha^6x+2$	$\alpha^5x^4+y+\alpha^6$		
	2 1 0 $x^4+\alpha^7xy^2+\alpha^5xy+\alpha^7y$	α^6 0	$[[1,\alpha^2],[0,1]]$	$y^2+\alpha^7xy+\alpha^2y+\alpha^5x+\alpha^2$	$\alpha^5x^3y+2x^4+\alpha^7x^2y+\alpha x^3+\alpha xy+\alpha^3x^2+\alpha^5x$		
13	0 1 2 $xy+\alpha^6y$	0 -2					
	1 0 2 $xy+\alpha^6y+\alpha^6x+2$	0 -2					
	2 2 -1 $x^4y+\alpha^7x^5+\alpha^2x^4+\alpha^5xy^2+\alpha y^2+\alpha^3xy+\alpha^6y$	2 0	$[[1,1],[0,1]]$	$y^2+\alpha^7xy+\alpha^2y+\alpha^5x+\alpha^2$	$\alpha^5x^3y+2x^4+\alpha^7x^2y+\alpha x^3+\alpha xy+\alpha^3x^2+\alpha^5x+2$		

Chapter 4

Evaluation Codes from an Affine Variety Code Perspective

Olav Geil

Department of Mathematical Sciences, Aalborg University,
Fr. Bajersvej 7G, 9220 Aalborg Ø, Denmark,
olav@math.aau.dk

Evaluation codes (also called order domain codes) are traditionally introduced as generalized one-point geometric Goppa codes. In the present chapter we will give a new point of view on evaluation codes by introducing them instead as particular nice examples of affine variety codes. Our study includes a reformulation of the usual methods to estimate the minimum distances of evaluation codes into the setting of affine variety codes. Finally we describe the connection to the theory of one-point geometric Goppa codes.

Contents

4.1. Introduction

Over the years the theory of geometric Goppa codes has produced many interesting results. The only drawback is that the codes are often described theoretically and that concrete generator matrices or parity check matrices

are often not rendered. As an attempt to simplify the description of one-point geometric Goppa codes and to support an easy generalization of such codes to higher dimensional objects than curves, Høholdt, van Lint, and Pellikaan founded the theory of order domains in [20]. You may say that order domains are manufactured to simplify the concrete code constructions. That is, generator matrices and parity check matrices are easily described. The codes defined from order domains are often called evaluation codes or order domain codes. The minimum distance and in larger generality the generalized Hamming weights of evaluation codes can be found by applying one of two bounds that rely only on some relatively simple theory. For a parity check matrix description one applies the order bound [20], [19] and [18]. This bound is an incidence of the Feng-Rao bound [11], [12], [29]. If instead a generator matrix description is given then one uses the bound in [2] which relies on the same notion as does the more well-known order bound.

Although evaluation codes have their origin in the study of geometric Goppa codes in the present chapter we will turn things upside down and introduce them as particular nice examples of affine variety codes. This adds a new perspective to the theory of evaluation codes as well as to the theory of affine variety codes. We reformulate the Feng-Rao bound and the bound from [2] into the setting of affine variety codes. Having done this we see that the affine variety codes for which we get maximal information from the above two bounds are the affine variety codes related to order domains. We conclude the chapter by describing the connection to the theory of one-point geometric Goppa codes.

4.2. Affine variety codes

Affine variety codes were introduced by Fitzgerald and Lax in [13]. The definition of the codes calls for an ideal $I \subseteq \mathbf{F}_q[X_1, \ldots, X_m]$ from which we start by defining

$$I_q = I + \langle X_1^q - X_1, \ldots, X_m^q - X_m \rangle \tag{4.1}$$

$$R_q = \mathbf{F}_q[X_1, \ldots, X_m]/I_q. \tag{4.2}$$

Let

$$V = \{P_1, \ldots, P_n\} = \mathcal{V}_{\mathbf{F}_q}(I_q) = \mathcal{V}_{\bar{\mathbf{F}}_q}(I_q)$$

be the variety of I_q. Here, $\bar{k}$ means the algebraic closure of the field k and

$P_i \neq P_j$ for $i \neq j$. Define an $\mathbf{F}_q$ linear map $\mathrm{ev} : R_q \to \mathbb{F}_q^n$ by

$$\mathrm{ev}(F + I_q) = (F(P_1), \ldots, F(P_n)).$$

We will call this map an evaluation map. Writing $P_j = (P_j^{(1)}, \ldots, P_j^{(m)})$ for $j = 1, \ldots, n$ we see that the i-th entry of

$$\mathrm{ev}\left(\left(\prod_{s=1,\ldots,m} \prod_{\substack{j=1,\ldots,n \\ P_j^{(s)} \neq P_i^{(s)}}} (X_s - P_j^{(s)})\right) + I_q\right)$$

is nonzero whereas all other entries equal zero. Therefore, the map ev is surjective. We next show that ev is also injective. To this end we first recall from [4, Pro. 8.14] that if J is an ideal in a polynomialring $k[X_1, \ldots, X_m]$ where k is perfect and if J contains a squarefree univariate polynomial in every variable then J is a radical ideal. This clearly makes I_q radical. Next we recall from The Strong Nullstellensatz [7, Th. 6, Sec. 4.2] that if an ideal $J \subseteq \bar{k}[X_1, \ldots, X_m]$ is radical then the vanishing ideal of the variety $\mathcal{V}_{\bar{k}}(J)$ is J itself. This implies that the vanishing ideal in $\mathbf{F}_q[X_1, \ldots, X_m]$ of V equals I_q and therefore the map ev is injective. We have shown that ev is a vector space isomorphism. We can now define the affine variety codes.

Definition 4.1. Let I_q and R_q be as in (4.1) and (4.2) and assume that L is an $\mathbf{F}_q$-vector subspace of R_q. Define the affine variety code $C(I, L) = \mathrm{ev}(L)$, and the affine variety code $C(I, L)^\perp$ to be the orthogonal complement of $C(I, L)$ with respect to the usual inner product on $\mathbf{F}_q^n$. That is,

$$C(I, L)^\perp = \{\vec{c} \mid \vec{c} \cdot \mathrm{ev}(F + I_q) = 0 \text{ for all } F + I_q \in L\}$$

where $\vec{f} \cdot \vec{h}$ denotes the inner product of $\vec{f}$ and $\vec{h}$.

4.3. Some Gröbner basis theoretical tools

In this section we present some Gröbner basis theoretical tools that will be very useful in the construction of affine variety codes. The tools will also help us to estimate the parameters of the codes. We start by recalling the concept of a footprint.

Definition 4.2. Let $J \subseteq k[X_1, \ldots, X_m]$ be an ideal and let $\prec$ be a fixed monomial ordering. Denote by $\mathcal{M}(X_1, \ldots, X_m)$ the monomials in the variables $X_1, \ldots, X_m$. The footprint of J with respect to $\prec$ is the set

$$\Delta_\prec(J) = \{M \in \mathcal{M}(X_1, \ldots, X_m) \mid M \text{ is not the leading monomial of any polynomial in } J\}.$$

Given a basis for the ideal J it may indeed not be obvious at a first glance what is the footprint. However, every polynomial ideal possesses a particular type of basis from which the footprint can be easily read off. These are the Gröbner bases.

Definition 4.3. Let $J \subseteq k[X_1, \ldots, X_m]$ be an ideal and $\prec$ a monomial ordering. A finite subset $\mathcal{G}$ of J is called a Gröbner basis (with respect to $\prec$) if for every polynomial $P(X_1, \ldots, X_m) \in J$ there exists a $G \in \mathcal{G}$ such that the leading monomial of G divides the leading monomial of P.

One of the main results in Gröbner basis theory is that a Gröbner basis $\mathcal{G}$ for J is indeed a basis for J. Given a basis for J we can extend it to a Gröbner basis by applying Buchberger's algorithm. Hence, there is a method to detect the footprint $\Delta_\prec(J)$.

The next couple of results explain our interest in the footprint. From [7, Pro. 4, Sec. 5.3] we have the following proposition.

Proposition 4.4. *Let the notation be as in Definition 4.2. The set*

$$\{M + J \mid M \in \Delta_\prec(J)\} \tag{4.3}$$

constitutes a basis for $k[X_1, \ldots, X_m]/J$ as a vector space over k.

Throughout this chapter we will make extensively use of the division algorithm for multivariate polynomials [7, Sec. 2.3] with which we will assume the reader to be familiar. Given a monomial ordering, a polynomial H and an ordered list of polynomials $(G_1, \ldots, G_r)$ the algorithm calculates the remainder of H modulo $(G_1, \ldots, G_r)$. This remainder is written H rem $(G_1, \ldots, G_r)$. When $\mathcal{G} = \{G_1, \ldots, G_s\}$ constitutes a Gröbner basis (for the ideal $\langle G_1, \ldots, G_r \rangle$) the remainder does not depend on how we order the elements in the list $(G_1, \ldots, G_r)$ and therefore in this case we will simply talk about the remainder modulo $\mathcal{G}$. We observe that to write an element $H + J \in k[X_1, \ldots, X_m]/J$ as a linear combination of the elements in (4.3) we need only find the remainder of H modulo the Gröbner basis $\mathcal{G}$. Moreover, as a consequence of Proposition 4.4 and the definition of a Gröbner basis, H rem $\mathcal{G}$ are the same no matter which Gröbner basis is chosen for J as long as $\prec$ is fixed.

Applying the above theory to the case $R_q = \mathbf{F}_q[X_1, \ldots, X_m]/I_q$ we see that for every fixed choice of $\prec$ Proposition 4.4 gives us a basis $\{M + I_q \mid M \in \Delta_\prec(I_q)\}$ for R_q. If $\{B_1 + I_q, \ldots, B_{\dim(L)} + I_q\}$ is a basis for a subspace $L \subseteq R_q$ we may therefore without loss of generality assume

that $\mathrm{Supp}(B_1), \ldots, \mathrm{Supp}(B_{\dim(L)}) \subseteq \Delta_\prec(I_q)$. Here, $\mathrm{Supp}(F)$ means the support of F. Once the variety $\mathcal{V}_{\mathbb{F}_q}(I_q)$ is found we can then easily specify the generator matrix for $C(I,L)$ as well as easily specify the parity check matrix for $C(I,L)^\perp$. The length of the codes clearly is

$$n = \#\mathcal{V}_{\mathbb{F}_q}(I_q) = \#\mathcal{V}_{\mathbb{F}_q}(I) = \#\Delta_\prec(I_q).$$

As ev is an isomorphism the dimension of $C(I,L)$ is $\dim(L)$ whereas the dimension of $C(I,L)^\perp$ equals $n - \dim(L)$. What remains is to estimate the minimum distances of the codes. This will be done in Section 4.4 and Section 4.5 below.
In Section 4.4 we will need the following corollary to Proposition 4.4. It is an incidence of the more general footprint bound [8, Cor. 2.5, Sec. 4.2].

Corollary 4.5. *Let $F_1, \ldots, F_s \in \mathbf{F}_q[X_1, \ldots, X_m]$. The number of common zeros of $F_1, \ldots, F_s$ over $\mathbf{F}_q$ is $\#\Delta_\prec(\langle F_1, \ldots, F_s, X_1^q - X_1, \ldots, X_m^q - X_m\rangle)$ (here $\prec$ is any monomial ordering).*

Proof. Let n be the number of common zeros. As explained in the previous section R_q is isomorphic to $\mathbf{F}_q^n$ as a vector space over $\mathbf{F}_q$ under the isomorphism ev. By Proposition 4.4 the dimension of R_q is $\#\Delta_\prec(I_q)$. The proof is complete. □

4.4. A bound on the minimum distance of $C(I,L)$

We now estimate the minimum distance of $C(I,L)$. The bound that we present can be viewed as an interpretation of the bound in [2, Th. 8]. Let $\prec$ and $I \subseteq \mathbf{F}_q[X_1, \ldots X_m]$ be fixed and consider a subspace $L \subseteq R_q$. By using Gaussian elimination any basis of L can be transformed into a basis of the following form.

Definition 4.6. A basis $\{B_1 + I_q, \ldots, B_{\dim(L)} + I_q\}$ for $L \subseteq R_q$ where $\mathrm{Supp}(B_i) \subseteq \Delta_\prec(I_q)$ for $i = 1, \ldots, \dim(L)$ and where $\mathrm{lm}(B_1) \prec \cdots \prec \mathrm{lm}(B_{\dim(L)})$ is said to be well-behaving with respect to $\prec$. Here, $\mathrm{lm}(F)$ means the leading monomial of F.

For fixed $\prec$ the sequence $\big(\mathrm{lm}(B_1), \ldots, \mathrm{lm}(B_{\dim(L)})\big)$ is the same for all choices of well-behaving bases of L. Therefore the following definition makes sense.

Definition 4.7. Let L be a subspace of R_q and define

$$\square_\prec(L) = \{\mathrm{lm}(B_1), \ldots, \mathrm{lm}(B_{\dim(L)})\}$$

where $\{B_1 + I_q, \ldots, B_{\dim(L)} + I_q\}$ is any well-behaving basis of L with respect to $\prec$.

Definition 4.8. Let $\mathcal{G}$ be a Gröbner basis for I_q with respect to $\prec$. An ordered pair of monomials (M_1, M_2), $M_1, M_2 \in \Delta_\prec(I_q)$ is said to be one-way well-behaving (OWB) if for all H with $\mathrm{Supp}(H) \subseteq \Delta_\prec(I_q)$ and $\mathrm{lm}(H) = M_1$

$$\mathrm{lm}(M_1 M_2 \text{ rem } \mathcal{G}) = \mathrm{lm}(H M_2 \text{ rem } \mathcal{G})$$

holds.

As already mentioned F rem $\mathcal{G} = F$ rem $\mathcal{G}'$ if $\mathcal{G}$ and $\mathcal{G}'$ are Gröbner bases for I_q with respect to identical ordering. Therefore the definition of OWB is independent of which Gröbner basis $\mathcal{G}$ we consider as long as $\prec$ is fixed.

Theorem 4.9. *Let $\prec$ be fixed. The minimum distance of $C(I, L)$ is at least*

$$\min\Big\{\#\{K \in \Delta_\prec(I_q) \mid \exists N \in \Delta_\prec(I_q) \text{ such that } (P, N) \text{ is OWB and } lm(PN \text{ rem } \mathcal{G}) = K\} \Big| P \in \square_\prec(L)\Big\}.$$

Proof. Let $\vec{c} \in C(I, L)$. Then there exists an F such that $\mathrm{Supp}(F) \subseteq \Delta_\prec(I_q)$, $\mathrm{lm}(F) = P \in \square_\prec(L)$ and $\mathrm{ev}(F + I_q) = \vec{c}$. By Corollary 4.5 the Hamming weight of $\vec{c}$ is equal to $n - \#\Delta_\prec(I_q + \langle F\rangle)$ and therefore we take a closer look at $\Delta_\prec(I_q + \langle F\rangle)$. If $N, K \in \Delta_\prec(I_q)$ satisfy that (P, N) is OWB and $\mathrm{lm}(PN \text{ rem } \mathcal{G}) = K$ then

$$K \in \Delta_\prec(I_q) \backslash \Delta_\prec(I_q + \langle F\rangle).$$

Hence,

$$\#\Delta_\prec(I_q + \langle F\rangle) \leq \#\Delta_\prec(I_q) - \#\big\{K \in \Delta_\prec(I_q) \mid \exists N \in \Delta_\prec(I_q) \text{ such that } (P, N) \text{ is OWB and } \mathrm{lm}(PN \text{ rem } \mathcal{G}) = K\big\}. \quad (4.4)$$

But $n = \#\Delta_\prec(I_q)$ and therefore the Hamming weight of $\vec{c}$ is at least

$$\#\big\{K \in \Delta_\prec(I_q) \mid \exists N \in \Delta_\prec(I_q) \text{ such that } (P, N) \text{ is OWB and } \mathrm{lm}(PN \text{ rem } \mathcal{G} = K\big\}. \square$$

It is of course possible to apply Theorem 4.9 for different choices of $\prec$ to see which one gives the sharpest estimate. To get the full advantage of Theorem 4.9 we need to have some information of the algebraic structure of R_q. The following Corollary, however, easily applies to any affine variety code. Also this bound could be applied for different choices of $\prec$ to get the sharpest estimate.

Corollary 4.10. *Let $\prec$ be fixed. The minimum distance of $C(I,L)$ is at least*

$$\min\{\#\{K \in \Delta_\prec(I_q) \mid P \text{ divides } K\} \big| P \in \Box_\prec(L)\}. \tag{4.5}$$

Proof. Let K, P be as in (4.5). Clearly $\frac{K}{P} \in \Delta_\prec(I_q)$. To see that $(P, \frac{K}{P})$ is OWB let H be a polynomial with $\mathrm{lm}(H) = P$ and $\mathrm{Supp}(H) \subseteq \Delta_\prec(I_q)$. Clearly, the leading monomial of $H\frac{K}{P}$ is equal to K. The division algorithm, when applied to $H\frac{K}{P}$ and $\mathcal{G}$, starts by moving K to the remainder. This is due to $K \in \Delta_\prec(I_q)$. When we run the division algorithm all other terms A are either moved to the remainder, are replaced with with polynomials S such that $\mathrm{lm}(S) \prec \mathrm{lm}(A)$ holds, or are replaced with 0. Therefore,

$$\mathrm{lm}\big(H\frac{K}{P} \text{ rem } \mathcal{G}\big) = K = \mathrm{lm}\big(P\frac{K}{P} \text{ rem } \mathcal{G}\big). \qquad \Box$$

Remark 4.11. It is possible to modify Theorem 4.9 and Corollary 4.10 to also deal with generalized Hamming weights. For the case of Theorem 4.9 this corresponds to interpreting the bound in [2, Th. 10].

Example 4.12. Let $I = \langle 0 \rangle \subseteq \mathbb{F}_q[X_1, \ldots, X_m]$. Then

$$\mathcal{G} = \{X_1^q - X_1, \ldots, X_m^q - X_m\}$$

is a Gröbner basis for I_q (regardless of the ordering $\prec$ chosen). Hence,

$$\Delta_\prec(I_q) = \{X_1^{i_1} \cdots X_m^{i_m} \mid 0 \le i_1 < q, \ldots, 0 \le i_m < q\}$$

holds and

$$\{X_1^{i_1} \cdots X_m^{i_m} + I_q \mid 0 \le i_1 < q, \ldots, 0 \le i_m < q\}$$

is a basis for $R_q = \mathbb{F}_q[X_1, \ldots, X_m]/I_q$ as a vectorspace over $\mathbb{F}_q$. It follows that the corresponding affine variety codes are of length $n = \#\Delta_\prec(I_q) = q^m$. Let s be an integer $0 \le s \le m(q-1)$. If we choose L to be the space generated by the basis elements $X_1^{i_1} \cdots X_m^{i_m} + I_q$ with $i_1 + \cdots + i_m \le s$ then we get

$$L = \{F(X_1, \ldots, X_m) + I_q \mid \deg(F) \le s\}. \tag{4.6}$$

Here, $\deg(F)$ means the total degree of F. Clearly,

$$\Box_\prec(I_q) = \{X_1^{i_1} \cdots X_m^{i_m} \mid 0 \le i_1 < q, \ldots, 0 \le i_m < q, i_1 + \cdots + i_m \le s\}.$$

The code $C(I,L)$ is known as the generalized Reed-Muller code $\mathrm{RM}_q(s,m)$, and Corollary 4.10 tells us that the minimum distance of $\mathrm{RM}_q(s,m)$ is at least

$$\min\{(q-i_1)\cdots(q-i_m) \mid 0 \le i_1 < q, \ldots, 0 \le i_m < q, i_1 + \cdots + i_m \le s\} \tag{4.7}$$

as

$$\#\{X_1^{j_1}\cdots X_m^{j_m} \in \Delta_\prec(I_q) \mid X_1^{i_1}\cdots X_m^{i_m} \text{ divides } X_1^{j_1}\cdots X_m^{j_m}\} \\ = (q-i_1)\cdots(q-i_m).$$

Writing $s = a(q-1)+b$ with $a, b \in \mathbb{N}_0$ and $0 \leq b < q-1$ the number in (4.7) can be shown to be equal to $(q-b)q^{m-a-1}$. Now letting $\mathbb{F}_q = \{\alpha_1, \ldots, \alpha_q\}$ and defining

$$F = (X_1^{q-1}-1)\cdots(X_a^{q-1}-1)(X_{a+1}-\alpha_1)\cdots(X_{a+1}-\alpha_b)$$

we see that $\mathrm{ev}(F+I_q) \in C(I,L)$ is of Hammingweight equal to $(q-b)q^{m-a-1}$. Hence, Corollary 4.10 produces the correct value of the minimum distance of the generalized Reed-Muller codes. It is interesting to observe that the minimum distance of the generalized Reed-Muller codes was originally established using quite different and more complicated methods [23].

If the goal is to produce codes with good parameters then there is better choice of L than (4.6) namely

$$L = \mathrm{Span}_{\mathbb{F}_q}\{X_1^{i_1}\cdots X_m^{i_m} \mid 0 \leq i_1 < q, \ldots, 0 \leq i_m, (q-i_1)\cdots(q-i_m) \geq \delta\}. \tag{4.8}$$

Corollary 4.10 tells us that the corresponding code $C(I,L)$ is of minimum distance at least δ and it is the largest code of prescribed minimum distance δ. If actually $i_1, \ldots, i_m$ exists with $(q-i_1)\cdots(q-i_m) = \delta$ then, as above, we can detect a codeword of Hammingweight δ and we conclude that Corollary 4.10 produces the actual minimum distance in this case. The codes $C(I,L)$ corresponding to (4.8) are called Massey-Costello-Justesen codes [26], [22] and are of course examples of improved generalized Reed-Muller codes.

4.5. The Feng-Rao bound for $C(I,L)^\perp$

In this section we reformulate the Feng-Rao bound into the setting of affine variety codes.

Theorem 4.13. *Let $\prec$ be fixed. The minimum distance of $C(I,L)^\perp$ is at least*

$$\min\big\{\#\{P \in \Delta_\prec(I_q) \mid \exists N \in \Delta_\prec(I_q) \text{ such that } (P,N) \text{ is OWB} \\ \text{and } lm(PN \text{ rem } \mathcal{G}) = K\}\big| K \in \Delta_\prec(I_q)\backslash\Box_\prec(L)\big\}. \tag{4.9}$$

Proof. Let $\{B_1 + I_q, \ldots, B_{\dim(L)} + I_q\}$ be a well-behaving basis for L. Consider $\vec{c} \in C(I, L)^{\perp} \backslash \{\vec{0}\}$. That is, $\vec{c}$ satisfies $\vec{c} \cdot \mathrm{ev}(B_i + I_q) = 0$ for $i = 1, \ldots, \dim(L)$ but

$$\vec{c} \cdot \mathrm{ev}(K + I_q) \neq 0 \tag{4.10}$$

holds for some $K \in \Delta_{\prec}(I_q)$. Let $K \in \Delta_{\prec}(I_q)$ be smallest possible with respect to $\prec$ such that (4.10) holds. By linearity of the inner product and the minimality of K we have $K \notin \Box_{\prec}(L)$. Consider OWB pairs $(P_1, N_1), \ldots, (P_\delta, N_\delta)$, where $P_1, N_1, \ldots, P_\delta, N_\delta \in \Delta_{\prec}(I_q)$, $P_1 \prec \cdots \prec P_\delta$ and $\mathrm{lm}(P_i N_i \text{ rem } \mathcal{G}) = K$ for $i = 1, \ldots, \delta$. The minimality of K and the OWB property of (P_i, N_i) ensure that

$$\vec{c} \cdot \mathrm{ev}\Bigg(\Big(\sum_{\substack{t=1,\ldots,i \\ a_i \neq 0}} a_t P_t \Big) N_i \text{ rem } \mathcal{G} + I_q \Bigg) \neq 0 \tag{4.11}$$

holds for any $i \in \{1, \ldots, \delta\}$. Let $*$ be the componentwise product on $\mathbb{F}_q^n$ given by

$$(a_1, \ldots, a_n) * (b_1, \ldots, b_n) = (a_1 b_1, \ldots, a_n b_n).$$

As

$$\Big(\sum_{\substack{t=1,\ldots,i \\ a_i \neq 0}} a_t P_t \Big) N_i \text{ rem } \mathcal{G} + I_q = \Big(\sum_{\substack{t=1,\ldots,i \\ a_i \neq 0}} a_t P_t \Big) N_i + I_q$$

we conclude from (4.11) that

$$\vec{c} * \mathrm{ev}\Bigg(\Big(\sum_{\substack{t=1,\ldots,i \\ a_i \neq 0}} a_t P_t \Big) + I_q \Bigg) \neq \vec{0}$$

for any $i \in \{1, \ldots, \delta\}$. Hence, $\vec{c} * \vec{e} \neq \vec{0}$ for all

$$\vec{e} \in \Big\{ \mathrm{ev}\Big(\Big(\sum_{t=1}^{\delta} a_t P_t \Big) + I_q \Big) \Big| a_1, \ldots, a_\delta \in \mathbb{F}_q, \text{ not all } a_i \text{ equal } 0 \Big\}. \tag{4.12}$$

The space consisting of (4.12) and $(0, \ldots, 0)$ is of dimension δ and therefore the Hamming weight of $\vec{c}$ needs to be at least δ. □

It is of course possible to apply Theorem 4.13 to different choices of $\prec$ to see which one gives the sharpest estimate. Theorem 4.13 requires

that we have some information about the algebraic structure of R_q. The following Corollary, however, easily applies to any affine variety code. Also this bound could be applied for different choices of $\prec$ to get the sharpest estimate.

Corollary 4.14. *Let the notation be as in Theorem 4.13. The minimum distance of* $C(I,L)^{\perp}$ *is at least*

$$min\big\{\#\{P \in \Delta_{\prec}(I_q) \mid P \text{ divides } K\} \big| K \in \Delta_{\prec}(I_q) \backslash \Box_{\prec}(L)\big\}.$$

Proof. See the proof of Corollary 4.10. □

Remark 4.15. It is possible to modify Theorem 4.13 and Corollary 4.14 to also deal with generalized Hamming weights. For the case of Theorem 4.13 this corresponds to interpreting the last part of [18, Th. 1].

Example 4.16. This is a continuation of Example 4.12. It is well-known that the dual code of a generalized Reed-Muller code is again a generalized Reed-Muller code. More precisely,

$$\mathrm{RM}_q(s,m) = \mathrm{RM}_q((q-1)m-1-s,m)^{\perp}$$

holds [9, Th. 2.2.1]. Applying Corollary 4.14 to $\mathrm{RM}((q-1)m-1-s,m)^{\perp}$ we see that the minimum distance of $\mathrm{RM}_q(s,m)$ is at least

$$\min\big\{(i_1+1)\cdots(i_m+1) \mid 0 \le i_1 < q, \ldots, 0 \le i_m < q, \\ i_1+\cdots+i_m \ge (q-1)m-s\big\}. \quad (4.13)$$

Writing again $s = a(q-1)+b$ with $0 \le b < q-1$ (4.13) becomes equal to $(q-b)q^{m-a-1}$ which we in Example 4.12 have seen to be equal to the true minimum distance of $\mathrm{RM}_q(s,m)$. Hence, also Corollary 4.14 produces the true value of the minimum distance of generalized Reed-Muller codes. If the goal is to produce codes $C(I,L)^{\perp}$ with good parameters then choosing L to be

$$L = \mathrm{Span}_{\mathbb{F}_q}\big\{X_1^{i_1}\cdots X_m^{i_m} \mid 0 \le i_1 < q, \ldots, 0 \le i_m < q, \\ (i_1+1)\cdots(i_m+1) < q^m - s\big\} \quad (4.14)$$

would be a better choice. The codes $C(I,L)^{\perp}$ corresponding to (4.14) are called hyperbolic codes and are denoted $\mathrm{Hyp}_q(s,m)$ [14, Def. 6]. By [14, Th. 3] $\mathrm{Hyp}_q(s,m)$ equals $C(I,L')$ where L' is the space in (4.8) with $r = q^m - s$. That is, hyperbolic codes are the same as Massey-Costello-Justesen codes. We showed in Example 4.12 that the minimum distance of $C(I,L')$ is at

least $q^m - s$. Applying Corollary 4.14 to $\mathrm{Hyp}_q(s,m)$ also gives the result that the minimum distance is at least $q^m - s$. Hence, Corollary 4.10 and Corollary 4.14 produce the same results for generalized Reed-Muller codes and for Hyperbolic codes.

4.6. Using weighted degree orderings

In this section we consider two examples where the monomial ordering is a weighted degree lexicographic ordering.

Definition 4.17. Let $w(X_1), \ldots, w(X_m) \in \mathbf{N}$ and define the weight of $X_1^{i_1} \cdots X_m^{i_m}$ to be the number $w(X_1^{i_1} \cdots X_m^{i_m}) = i_1 w(X_1) + \cdots + i_m w(X_m)$. The weighted degree lexicographic ordering on $\mathcal{M}(X_1, \ldots, X_m)$ is the ordering with $X_1^{i_1} \cdots X_m^{i_m} \prec X_1^{j_1} \cdots X_m^{j_m}$ if either $w(X_1^{i_1} \cdots X_m^{i_m}) < w(X_1^{j_1} \cdots X_m^{j_m})$ holds or $w(X_1^{i_1} \cdots X_m^{i_m}) = w(X_1^{j_1} \cdots X_m^{j_m})$ holds but $X_1^{i_1} \cdots X_m^{i_m} \prec_{lex} X_1^{j_1} \cdots X_m^{j_m}$. Here, $\prec_{lex}$ is the lexicographic ordering with $X_m \prec_{lex} \cdots \prec_{lex} X_1$.

One of the qualities of weighted degree lexicographic orderings is the following lemma. The proof of the lemma is left for the reader.

Lemma 4.18. *Let a weighted degree lexicographic ordering be given as in Definition 4.17. If H has got exactly one monomial of highest weight w' in its support and G has exactly two monomials of highest weight in its support then H rem (G) has exactly one monomial of highest weight in its support and this weight is w'.*

The codes $C(I,L)^{\perp}$ in the next example were originally treated in [24] whereas the codes $C(I,L)$ are treated for the first time in the present chapter.

Example 4.19. Consider the ideals

$$I = \langle X^3Y + Y^3 + X \rangle \subseteq \mathbf{F}_8[X,Y]$$

$$I_q = I + \langle X^8 + X, Y^8 + Y \rangle \subseteq \mathbf{F}_8[X,Y].$$

Let $\prec$ be the weighted degree lexicographic ordering defined by setting $w(X) = 2$, $w(Y) = 3$ and by interpreting X as X_1 and Y as X_2. Clearly, $\mathcal{B} = \{X^3Y + Y^3 + X\}$ is a Gröbner basis for I and

$$\Delta_{\prec}(I) = \{X^iY^j \mid \text{ if } i \geq 3 \text{ then } j = 0\}$$

holds. Using Buchberger's algorithm we find the following Gröbner basis for I_q

$$\mathcal{G} = \{X^3Y + Y^3 + X, X^8 + X, XY^5 + X^5 + X^2Y^2 + Y, Y^7 + X^7\}$$

and therefore

$$\begin{aligned}\Delta_{\prec}(I_q) = \big\{&1, X, Y, X^2, XY, Y^2, X^3, X^2Y, XY^2, X^4, Y^3, X^2Y^2,\\ &X^5, XY^3, Y^4, X^6, X^2Y^3, XY^4, X^7, Y^5, X^2Y^4, Y^6\big\} \quad (4.15)\end{aligned}$$

with corresponding weights

$$\{0, 2, 3, 4, 5, 6, 6, 7, 8, 8, 9, 10, 10, 11, 12, 12, 13, 14, 14, 15, 16, 18\}.$$

The elements in (4.15) are listed in increasing order with respect to $\prec$. Using Lemma 4.18 and some other results we can detect altogether 166 useful OWB pairs plus a few more that we will not use. We illustrate the method used to check for the OWB property by considering a few OWB pairs. First to see that (X^3, X) is OWB we must show that HX rem $\mathcal{G} = X^3X$ rem $\mathcal{G}$ for all H with $\mathrm{lm}(H) = X^3$. We have

$$\mathrm{lm}\big((a_1 + a_2X + a_3Y + a_4X^2 + a_5XY + a_6Y^2 + X^3)X \text{ rem } \mathcal{G}\big) = X^4 \quad (4.16)$$

no matter what are $a_1, \ldots, a_6$. This is because $X^3X = X^4 \in \Delta_{\prec}(I_q)$ and therefore X^4 is moved to the remainder upon division with $\mathcal{G}$. The proof that (X^3, X) is OWB is complete. To see that (XY, X^2) is OWB we cannot apply the same argument as above as $XYX^2 = X^3Y \notin \Delta_{\prec}(I_q)$. We have

$$w(1 \cdot X^2), w(X \cdot X^2), w(Y \cdot X^2), w(X^2 \cdot X^2) \;<\; w(XY \cdot X^2) \;=\; 9.$$

That is, there is only one monomial of highest weight in $(a_1 + a_2X + a_3Y + a_4X^2 + XY)X^2$ and this weight is 9. As $X^3Y + Y^3 + Y$ has exactly two monomials of highest weight in its support Lemma 4.18 tells us that the monomial

$$\mathrm{lm}\big((a_1 + a_2X + a_3Y + a_4X^2 + XY)X^2 \text{ rem } \mathcal{B}\big)$$

is also of weight 9. There is only one such monomial in $\Delta_{\prec}(I)$ namely Y^3. As Y^3 also belongs to $\Delta_{\prec}(I_q)$ we conclude

$$\mathrm{lm}\big((a_1 + a_2X + a_3Y + a_4X^2 + XY)X^2 \text{ rem } \mathcal{G}\big) = Y^3$$

no matter what are $a_1, \ldots, a_4$. Hence, (XY, X^2) is OWB. Finally, to see that (XY, X^2Y) is OWB we start by recognizing from Lemma 4.18 that the weight of

$$\mathrm{lm}\big((a_1 + a_2X + a_3Y + a_4X^2 + XY)X^2Y \text{ rem } \mathcal{B}\big)$$

equals $w(XY \cdot X^2Y) = 12$. However, now there are the two possibilities X^6 and Y^4 of leading monomials as both are of weight 12 and both belong to $\Delta_\prec(I)$. A closer analysis reveals that

$$\mathrm{lm}\big((a_1 + a_2X + a_3Y + a_4X^2 + XY)X^2Y \text{ rem } \mathcal{B}\big) = Y^4.$$

As Y^4 also belongs to $\Delta_\prec(I_q)$ we conclude that

$$\mathrm{lm}\big((a_1 + a_2X + a_3Y + a_4X^2 + XY)X^2Y \text{ rem } \mathcal{G}\big) = Y^4$$

and (XY, X^2Y) is OWB.
Observe that for fixed P and K there can exist more choices of N such that (P, N) is OWB and $\mathrm{lm}(PN \text{ rem } \mathcal{G}) = K$. As an example both (XY, Y^2) and (XY, X^3) are OWB and satisfy

$$\mathrm{lm}(XY \cdot Y^2 \text{ rem } \mathcal{G}) = \mathrm{lm}(XY \cdot X^3 \text{ rem } \mathcal{G}) = XY^3.$$

In table 4.1 we list some information about the OWB pairs. By $\bar{\sigma}(P)$ we denote the number of detected $K \in \Delta_\prec(I_q)$ such that an $N \in \Delta_\prec(I_q)$ exists with (P, N) OWB and $\mathrm{lm}(PN \text{ rem } \mathcal{G}) = K$. By $\bar{\mu}(K)$ we denote the number of detected $P \in \Delta_\prec(I_q)$ such that an $N \in \Delta_\prec(I_q)$ exists with (P, N) OWB and $\mathrm{lm}(PN \text{ rem } \mathcal{G}) = K$.

Table 4.1. Information about the OWB pairs

M	1	X	Y	X^2	XY	Y^2	X^3	X^2Y	XY^2	X^4	Y^3
$\bar{\sigma}(M)$	22	19	14	16	12	11	5	10	9	4	8
$\bar{\mu}(M)$	1	2	2	3	4	3	4	6	6	5	8

M	X^2Y^2	X^5	XY^3	Y^4	X^6	X^2Y^3	XY^4	X^7	Y^5	X^2Y^4	Y^6
$\bar{\sigma}(M)$	7	3	6	5	2	4	3	1	2	2	1
$\bar{\mu}(M)$	9	6	10	11	7	12	13	8	14	15	17

For the code construction $C(I, L)$ we choose L to be spanned by the $(M + I_q)$'s with $M \in \Delta_\prec(I_q)$ and $\bar{\sigma}(M) \geq \delta$. By Theorem 4.9 this gives us codes of highest possible dimension with prescribed minimum distance at least δ. For the code construction $C(I, L)^\perp$ we choose L to be spanned by the $(M + I_q)$'s with $M \in \Delta_\prec(I_q)$ and $\bar{\mu}(M) < \delta$. By Theorem 4.13 this gives codes of highest possible dimension with prescribed minimum distance at least δ. The length of the codes equals $n = \#\Delta_\prec(I_q)$. From (4.15) we

therefore have $n = 22$. In Table 4.2 we list the parameters $[k, \delta]$ that can be realized from Theorem 4.9 and Theorem 4.13. Here k is the dimension and δ is the prescribed minimum distance. We conclude that although

Table 4.2. Parameters of the codes

$C(I,L)$	[1,22]	[2,19]	[3,16]	[4,14]	[5,12]	[6,11]
	[7,10]	[8,9]	[9,8]	[10,7]	[11,6]	[13,5]
	[15,4]	[17,3]	[20,2]	[22,1]		
$C(I,L)^{\perp}$	[1,17]	[2,15]	[3,14]	[4,13]	[5,12]	[6,11]
	[7,10]	[8,9]	[10,8]	[11,7]	[14,6]	[15,5]
	[17,4]	[19,3]	[21,2]			

the bound in Theorem 4.9 relies on the same notion as does the bound in Theorem 4.13 the two bounds can sometimes produce completely different results.

In Example 4.19 it was quite involved to detect which pairs are OWB. This is due to the fact that in $\Delta_{\prec}(I)$ as well as in $\Delta_{\prec}(I_q)$ there were more monomials of the same weight. In the next example no two different monomials in $\Delta_{\prec}(I)$ will be of the same weight. As a consequence it becomes very easy to find OWB pairs.

Example 4.20. Consider the ideals

$$I = \langle X^4 - Y^3 - Y \rangle \subseteq \mathbf{F}_9[X, Y]$$

$$I_q = \langle X^4 - Y^3 - Y, X^9 - X, Y^9 - Y \rangle \subseteq \mathbf{F}_9[X, Y].$$

Let $\prec$ be the weighted degree lexicographic ordering given by $w(X) = 3$, $w(Y) = 4$ and by interpreting X as X_2 and Y as X_1. Clearly,

$$\mathcal{B} = \{X^4 - Y^3 - Y\}$$

is a Gröbner basis for I and applying Buchberger's algorithm we find that

$$\mathcal{G} = \{X^4 - Y^3 - Y, X^9 - X\}$$

is a Gröbner basis for I_q. Hence,

$$\begin{aligned} \Delta_{\prec}(I) &= \{X^i Y^j \mid 0 \leq i, 0 \leq j < 3\} \\ \Delta_{\prec}(I_q) &= \{X^i Y^j \mid 0 \leq i < 9, 0 \leq j < 3\}. \end{aligned} \tag{4.17}$$

The map $w : \Delta_{\prec}(I) \to \langle 3, 4 \rangle$ given by $w(X^i Y^j) = i3 + j4$ is a bijection. Here, $\langle 3, 4 \rangle$ means the semigroup generated by 3 and 4. Hence, we can identify any monomial $M \in \Delta_{\prec}(I)$ uniquely by its weight. Consider a

polynomial F with $\mathrm{Supp}(F) \subseteq \Delta_\prec(I_q)$ and write $P = \mathrm{lm}(F)$. Let $N \in \Delta_\prec(I_q)$ be arbitrary. By Lemma 4.18 the leading monomial of FN rem $\mathcal{B}$ is the unique monomial $K \in \Delta_\prec(I)$ of weight equal to $w(PN) = w(P)+w(N)$. If $K \in \Delta_\prec(I_q)$ holds then (P, N) is OWB. Hence, given $P, N \in \Delta_\prec(I_q)$ then (P, N) is OWB if $w(P) + w(N) \in w(\Delta_\prec(I_q))$. Next we show that if $K \in \Delta_\prec(I_q)$ and $P, N \in \Delta_\prec(I)$ satisfy $w(P) + w(N) = w(K)$ then also $P, N \in \Delta_\prec(I_q)$ holds. This in particular implies that (P, N) is OWB. Aiming for a contradiction assume that $P \notin \Delta_\prec(I_q)$. By the definition of the footprint there exists a polynomial $H \in I_q$ having P as leading monomial. As $P \in \Delta_\prec(I)$ we may without loss of generality assume that H is reduced modulo $\mathcal{B}$. That is, we may assume that $\mathrm{Supp}(H) \subseteq \Delta_\prec(I)$ holds. From $H \in I_q$ we conclude that

$$HN \text{ rem } \mathcal{B} \in I_q. \tag{4.18}$$

On the other hand the assumption $\mathrm{Supp}(H) \subseteq \Delta_\prec(I)$ in combination with Lemma 4.18 implies $\mathrm{lm}(HN \text{ rem } \mathcal{B}) = K$. Here we used the fact that no two monomials in $\Delta_\prec(I)$ are of the same weight. But K is assumed to be in $\Delta_\prec(I_q)$ and therefore (4.18) cannot be true. We have reached at a contradiction. Assuming $N \notin \Delta_\prec(I_q)$ would lead to a similar contradiction. The above observations imply that to detect OWB pairs it is enough to study the weights. For this purpose define

$$\Gamma = w(\Delta_\prec(I)) = \langle 3, 4 \rangle$$

and for $\lambda \in w(\Delta_\prec(I_q))$ let

$$\sigma(\lambda) = \#\{\eta \in w(\Delta_\prec(I_q)) \mid \eta - \lambda \in \Gamma\}$$

and for $\lambda \in \Gamma$ let

$$\mu(\lambda) = \#\{\alpha \in \Gamma \mid \lambda - \alpha \in \Gamma\}.$$

We have shown above that if $P \in \Delta_\prec(I_q)$ then there exist pairwise different elements $K_1, \ldots, K_{\sigma(w(P))} \in \Delta_\prec(I_q)$ and corresponding elements $N_1, \ldots, N_{\sigma(w(P))} \in \Delta_\prec(I_q)$ such that for $i = 1, \ldots, \sigma(w(P))$ (P, N_i) is OWB with $\mathrm{lm}(PN_i \text{ rem } \mathcal{G}) = K_i$. Similarly, if $K \in \Delta_\prec(I_q)$ then there exist pairwise different elements $P_1, \ldots, P_{\mu(w(K))} \in \Delta_\prec(I_q)$ and corresponding elements $N_1, \ldots, N_{\mu(w(K))} \in \Delta_\prec(I_q)$ such that (P_i, N_i) is OWB with $\mathrm{lm}(P_iN_i \text{ rem } \mathcal{G}) = K$. In Table 4.3 we list $\sigma(w)$ and $\mu(w)$ for all $w \in w(\Delta_\prec(I_q))$. For the purpose of the code constructions define the fol-

Table 4.3.

w	0	3	4	6	7	8	9	10	11
$\sigma(w)$	27	24	23	21	20	19	18	17	16
$\mu(w)$	1	2	2	3	4	3	4	6	6
w	12	13	14	15	16	17	18	19	20
$\sigma(w)$	15	14	13	12	11	10	9	8	7
$\mu(w)$	7	8	9	10	11	12	13	14	15
w	21	22	23	24	25	26	28	29	32
$\sigma(w)$	6	6	4	3	4	3	2	2	1
$\mu(w)$	16	17	18	19	20	21	23	24	27

lowing subspaces of $R_q = \mathbf{F}_9[X,Y]/I_q$

$$L_1 = \mathrm{Span}_{\mathbb{F}_9}\{M + I_q \mid M \in \Delta_\prec(I_q), w(M) \leq s\}$$
$$L_2 = \mathrm{Span}_{\mathbb{F}_9}\{M + I_q \mid M \in \Delta_\prec(I_q), \sigma(w(M)) \geq \delta\}$$
$$L_3 = \mathrm{Span}_{\mathbb{F}_9}\{M + I_q \mid M \in \Delta_\prec(I_q), \mu(w(M)) < \delta\}.$$

The corresponding affine variety codes are all of length $n = \#\Delta_\prec(I_q) = 27$. From Theorem 4.9 the minimum distance of $C(I, L_2)$ is at least δ and from Theorem 4.13 also the minimum distance of $C(I, L_3)^\perp$ is at least δ. The codes $C(I, L_2)$ and $C(I, L_3)^\perp$ respectively are so to speak the largest codes with designed minimum distance δ with respect to Theorem 4.9 and Theorem 4.13 respectively. Applying Theorem 4.9 and Theorem 4.13 respectively to the codes $C(I, L_1)$ and $C(I, L_1)^\perp$ respectively we get lower bounds on the minimum distances. As an example choosing $s = 23$ the code $C(I, L_1)$ is of dimension 21 and minimum distance at least 4. Choosing $\delta = 4$ the code $C(I, L_2)$ is of dimension 22 and minimum distance also at least 4. As another example choosing $s = 7$ the code $C(I, L_1)^\perp$ is of dimension 22 and of minimum distance at least 3. Choosing $\delta = 4$ the code $C(I, L_3)^\perp$ is also of dimension 22 but is of minimum distance at least 4.

4.7. The order domain conditions

In the previous section we demonstrated that the weighted degree lexicographic ordering can sometimes be very useful when we look for OWB pairs. In particular the task of finding OWB pairs were rather simple in Example 4.20 due to the fact that no two monomials in $\Delta_\prec(I)$ were of the same weight and due to the fact that the defining polynomial of I possessed exactly two monomials of highest weight in its support. In this section we

generalize the construction in Example 4.20. All proofs will be straightforward generalizations of the arguments from Example 4.20 and so they are mostly left out. We start by generalizing the concept of a weighted degree lexicographic ordering.

Definition 4.21. Let $w(X_1), \ldots, w(X_m) \in \mathbf{N}_0^r$ and assume $\prec_{\mathbf{N}_0^r}$ is a monomial ordering on $\mathbf{N}_0^r$. Extend w to a monomial function on $\mathcal{M}(X_1, \ldots, X_m)$ by

$$w(X_1^{i_1} \cdots X_m^{i_m}) = i_1 w(X_1) + \cdots + i_m w(X_m).$$

Let $\prec_{\mathcal{M}}$ be a monomial ordering on $\mathcal{M}(X_1, \ldots, X_m)$. The generalized weighted degree ordering defined from $w(X_1), \ldots, w(X_m)$, $\prec_{\mathbf{N}_0^r}$ and $\prec_{\mathcal{M}}$ is the ordering $\prec_w$ given by

$$X_1^{i_1} \cdots X_m^{i_m} \prec_w X_1^{j_1} \cdots X_m^{j_m}$$

if

$$w(X_1^{i_1} \cdots X_m^{i_m}) \prec_{\mathbf{N}_0^r} w(X_1^{j_1} \cdots X_m^{j_m})$$

holds or if

$$w(X_1^{i_1} \cdots X_m^{i_m}) = w(X_1^{j_1} \cdots X_m^{j_m})$$

holds but

$$X_1^{i_1} \cdots X_m^{i_m} \prec_{\mathcal{M}} X_1^{j_1} \cdots X_m^{j_m}.$$

The weighted degree of a polynomial F is $\mathrm{wdeg}(F) = w(\mathrm{lm}(F))$.

We now state the order domain conditions which play a central role in the present chapter.

Definition 4.22. Consider an ideal $I \subseteq k[X_1, \ldots, X_m]$ where k is a field. Let a generalized weighted degree ordering $\prec_w$ be given as in Definition 4.21. Assume I possesses a Gröbner basis $\mathcal{B}$ such that any $G \in \mathcal{B}$ has exactly two monomials of highest weight and such that no two monomials in $\Delta_{\prec}(I)$ is of the same weight. Then we say that I and $\prec_w$ satisfy the order domain conditions.

The following lemma is an immediate generalization of Lemma 4.18. Again we leave the proof for the reader.

Lemma 4.23. *Let I, $\prec_w$ and $\mathcal{B}$ be as in Definition 4.22. Let F be a polynomial with exactly one monomial of highest weight. Then $w(lm(F)) = w(lm(F \text{ rem } \mathcal{B}))$. In particular $w(lm(F)) = w(lm(F \text{ rem } \mathcal{B}))$ holds for all F with $Supp(F) \subseteq \Delta_{\prec_w}(I)$.*

Remark 4.24. If I and $\prec_w$ satisfy the order domain conditions then any polynomial G in any Gröbner basis $\mathcal{B}$ of I must contain exactly two monomials of highest weight. Hence, the choice of $\mathcal{B}$ is of no importance in Definition 4.22. This result is a consequence of Lemma 4.23 and the fact that the remainder is independent of the Gröbner basis chosen.

The following proposition is an immediate generalization of similar results in Example 4.20.

Proposition 4.25. *Assume $I \subseteq \mathbf{F}_q[X_1, \ldots, X_m]$ and $\prec_w$ satisfy the order domain conditions. Consider $I_q = I + \langle X_1^q - X_1, \ldots, X_m^q - X_m \rangle$. A pair (P, N) where $P, N \in \Delta_{\prec_w}(I_q)$ is OWB if $w(P) + w(N) \in w(\Delta_{\prec_w}(I_q))$. If $K \in \Delta_{\prec_w}(I_q)$ and $P, N \in \Delta_{\prec_w}(I)$ satisfy $w(P) + w(N) = w(K)$, then $P, N \in \Delta_{\prec_w}(I_q)$, and (P, N) is OWB.*

Definition 4.26. Assume I and $\prec_w$ satisfy the order domain conditions. Let $\Gamma = w(\Delta_{\prec_w}(I))$ and define for all $\lambda \in w(\Delta_{\prec_w}(I_q))$

$$\sigma(\lambda) = \#\{\eta \in w(\Delta_{\prec_w}(I_q)) \mid \eta - \lambda \in \Gamma\}$$

and for all $\lambda \in \Gamma$

$$\mu(\lambda) = \#\{\alpha \in \Gamma \mid \lambda - \alpha \in \Gamma\}.$$

Applying Theorem 4.9 and Theorem 4.13 in combination with Proposition 4.25 we get the following theorem.

Theorem 4.27. *Assume I and $\prec_w$ satisfy the order domain conditions. Let L be a subspace of $R_q = \mathbb{F}_q[X_1, \ldots, X_m]/I_q$ and assume*

$$\{B_1 + I_q, \ldots, B_{\dim(L)} + I_q\}$$

is a well-behaving basis (Definition 4.6). The minimum distance of $C(I, L)$ is at least

$$\min\{\sigma(w(lm(B_1))), \ldots, \sigma(w(lm(B_{\dim(L)})))\}.$$

The minimum distance of $C(I, L)^\perp$ is at least

$$\min\{\mu(w(M)) \mid M \in \Delta_{\prec_w}(I_q) \backslash \{lm(B_1), \ldots, lm(B_{\dim(L)})\} \geq \min\{\mu(\lambda) \mid \lambda \in \Gamma \backslash \{w(B_1), \ldots, w(B_{\dim(L)})\}\}.$$

Consider the following choices of L. Let $\vec{s} \in \mathbf{N}_0^r$ and $\delta \in \mathbf{N}$.

$$L_1 = \mathrm{Span}_{\mathbb{F}_q}\{M + I_q \mid M \in \Delta_{\prec_w}(I_q), w(M) \preceq_{\mathbf{N}_0^r} \vec{s}\} \tag{4.19}$$

$$L_2 = \mathrm{Span}_{\mathbb{F}_q}\{M + I_q \mid M \in \Delta_{\prec_w}(I_q), \sigma(w(M)) \geq \delta\} \tag{4.20}$$

$$L_3 = \mathrm{Span}_{\mathbb{F}_q}\{M + I_q \mid M \in \Delta_{\prec_w}(I_q), \mu(w(M)) < \delta\}. \tag{4.21}$$

Theorem 4.27 tells us that the minimum distance of $C(I, L_2)$ and $C(I, L_3)^\perp$ is at least δ. By construction $C(I, L_2)$ and $C(I, L_3)^\perp$ are the largest codes with prescribed minimum distance δ. We shall in Section 4.10 see that whenever the weights are numerical, that is whenever $\vec{s} = s$ is an integer, then the minimum distance of $C(I, L_1)$ is at least $n - s$. Here, $n = \#\Delta_{\prec_w}(I_q)$. Similarly we will derive in Section 4.10 a simple expression for a lower bound on the minimum distance of $C(I, L_1)^\perp$ whenever the weights are numerical.

Example 4.28. This is a continuation of Example 4.12 and Example 4.16. Choose the weights $w(X_1) = (1, 0, \ldots, 0)$, $w(X_2) = (0, 1, 0, \ldots, 0)$, $\ldots, w(X_m) = (0, \ldots, 0, 1) \in \mathbf{N}_0^m$. Let $\prec_{\mathbf{N}_0^m}$ be the graded ordering on $\mathbf{N}_0^m$ with $(1, 0, \ldots, 0) \prec_{\mathbf{N}_0^m} \cdots \prec_{\mathbf{N}_0^m} (0, \ldots, 0, 1)$. Let $\prec_{\mathcal{M}}$ be any monomial ordering on $\mathcal{M}(X_1, \ldots, X_m)$. Using the convention that the empty set is a Gröbner basis for the ideal $I = \langle 0 \rangle \subseteq \mathbb{F}_q[X_1, \ldots, X_m]$ we see that the order domain conditions are trivially satisfied. The code $C(I, L_1)$ with $\vec{s} = (0, \ldots, 0, s)$ is the generalized Reed-Muller code $\mathrm{RM}_q(s, m)$. Similarly, the codes $C(I, L_2)$ and $C(I, L_3)^\perp$ are the improved generalized Reed-Muller codes considered in Example 4.12 and Example 4.16. Applying Theorem 4.27 we count exactly the same OWB pairs that we count by applying Corollary 4.10 and Corollary 4.14.

Given I and $\prec_w$ such that the order domain conditions are satisfied we might want to construct codes by evaluating in a subset $U \subsetneq \mathcal{V}_{\mathbb{F}_q}(I)$ rather than in the entire variety $\mathcal{V}_{\mathbb{F}_q}(I)$. The following remark deals with this situation

Remark 4.29. Assume that the pair I and $\prec_w$ satisfies the order domain conditions. Let $U \subseteq \mathcal{V}_{\mathbb{F}_q}(I)$. Every finite set of points is a variety and therefore there exists polynomials $H_1, \ldots, H_r$ such that the vanishing ideal of U equals

$$I_U = I_q + \langle H_1, \ldots, H_r \rangle.$$

The estimates of the minimum distances of $C(I, L)$ and $C(I, L)^\perp$ still hold if these codes are made by evaluating in U rather than in the entire variety $\mathcal{V}_{\mathbb{F}_q}(I)$. All we need to do is to replace I_q with I_U in Definition 4.6, Definition 4.7, Proposition 4.25, Definition 4.26 and Theorem 4.27.

4.8. Weight functions and order domains

The concept of an order function was introduced by Høholdt et al. in [20] to simplify the treatment of one-point geometric Goppa codes and to provide a language for easy generalization of one-point geometric Goppa codes to objects of higher dimensions than curves. The concept was further developed in [33] and [17]. Here, we describe some terminology from [17].

Definition 4.30. Let R be a k-algebra and let Γ be a subsemigroup of $\mathbb{N}_0^r$ for some r. Let $\prec$ be a monomial ordering on $\mathbb{N}_0^r$. A surjective map $\rho : R \to \Gamma_{-\infty} = \Gamma \cup \{-\infty\}$ that satisfies the following six conditions is said to be a weight function

(W.0) $\rho(f) = -\infty$ if and only if $f = 0$
(W.1) $\rho(af) = \rho(f)$ for all nonzero $a \in \mathbb{F}_q$
(W.2) $\rho(f+g) \preceq \max\{\rho(f), \rho(g)\}$ and equality holds when $\rho(f) \prec \rho(g)$
(W.3) If $\rho(f) \prec \rho(g)$ and $h \neq 0$, then $\rho(fh) \prec \rho(gh)$
(W.4) If f and g are nonzero and $\rho(f) = \rho(g)$, then there exists a nonzero $a \in \mathbb{F}_q$ such that $\rho(f - ag) \prec \rho(g)$
(W.5) If f and g are nonzero then $\rho(fg) = \rho(f) + \rho(g)$.

A k-algebra with a weight function is called an order domain and Γ is called the value semigroup of ρ.

From [17][Th. 9.1 and Th. 10.4] we know that if the value semigroup Γ is finitely generated then it can be described in the language of Gröbner basis theory. We have the following result which connects Definition 4.30 to the theory of the previous section.

Theorem 4.31. *Let $\prec_w$ be a generalized weighted degree ordering on $\mathcal{M}(X_1, \ldots, X_m)$ and let $I \subset k[X_1, X_2, \ldots, X_m]$ be an ideal. If I and $\prec_w$ satisfy the order domain conditions (Definition 4.22) then $R = k[X_1, X_2, \ldots, X_m]/I$ is an order domain with a weight function defined as follows: Given a nonzero $f \in k[X_1, X_2, \ldots, X_m]/I$ write $f = F + I$ where $Supp(F) \subseteq \Delta_{\prec_w}(I)$. We have $\rho(f) = wdeg(F)$ and $\rho(0) = -\infty$.*
Any weight function with a finitely generated value semigroup Γ can be described as above.

Proof. We only show the first part of the theorem. Regarding the last part we refer to the proof in [17]. Assume I and $\prec_w$ satisfy the order domain conditions. The properties (W.0), (W.1), and (W.2) are obviously satisfied. Given $f = F_1 + I$ and $g = F_2 + I$ with $\mathrm{Supp}(F_1) \subseteq \Delta_{\prec_w}(I)$ and

$\mathrm{Supp}(F_2) \subseteq \Delta_{\prec_w}(I)$ let b be the leading coefficient of F_1 and let c be the leading coefficient of F_2. If we choose $a = b/c$ then the result in (*W*.4) holds. Property (*W*.5) follows immediately from Lemma 4.23. Finally, property (*W*.3) is a consequence of (*W*.5) (in fact (*W*.3) is not needed in the definition of a weight function). □

As mentioned earlier the ideals and the monomial orderings considered in Example 4.20 and Example 4.28 satisfy the order domain conditions. Therefore by Theorem 4.31 the corresponding factor rings are order domains and the weights correspond to weight functions following Theorem 4.31.

4.9. Codes form order domains

We now describe the codes related to order domains. We will need a couple of definitions.

Definition 4.32. Let R be an $\mathbb{F}_q$-algebra. A surjective map $\varphi : R \to \mathbb{F}_q^n$ is called a morphism of $\mathbb{F}_q$-algebras if φ is $\mathbb{F}_q$ linear and if

$$\varphi(fg) = \varphi(f) * \varphi(g)$$

for all $f, g \in R$ (here $*$ is the componentwise product described in Section 4.5).

Definition 4.33. Let $\rho : R \to \Gamma \cup \{-\infty\}$ be a weight function. A set

$$\{f_\lambda \mid \rho(f_\lambda) = \lambda, \lambda \in \Gamma\}$$

is called a well-behaving basis for R.

It is clear that all order domains possess well-behaving bases. Recall that we in Definition 4.6 introduced the concept of a well-behaving basis for $L \subseteq R_q$. The concept of a well-behaving basis for an order domain R as defined above is not the same. However, the two concepts are closely related.

Proposition 4.34. *Assume R is an order domain over k. If $\{f_\lambda \mid \rho(f_\lambda) = \lambda, \lambda \in \Gamma\}$ is a well-behaving basis for R then it is a basis for R as a vectorspace over k.*

Proof. For the case of weight functions with finitely generated value semigroup the result follows by combining the characterization in Theorem 4.31 with the result in Proposition 4.4. For the general case we refer to [17, Th. Pro. 3.2 and Def. 3.1]. □

Remark 4.35. Given two well-behaving bases $\{f_\lambda \mid \rho(f_\lambda) = \lambda, \lambda \in \Gamma\}$ and $\{g_\lambda \mid \rho(g_\lambda) = \lambda, \lambda \in \Gamma\}$ then for all $\eta \in \Gamma$, g_η is a linear combination of the elements in $\{f_\lambda \mid \lambda \preceq \eta\}$ and the coefficients of f_η in this expression is nonzero.

It follows from Remark 4.35 that it is of no importance in the next definition which well-behaving basis is considered.

Definition 4.36. Let R be an order domain over $\mathbb{F}_q$ with a weight function $\rho : R \to \Gamma \cup \{-\infty\}$ and let $\{f_\lambda \mid \rho(f_\lambda) = \lambda, \lambda \in \Gamma\}$ be a well-behaving basis. Let $\varphi : R \to \mathbb{F}_q^n$ be a morphism as in Definition 4.32. Define $\alpha(1) = 0$. For $i = 2, \ldots, n$ define recursively $\alpha(i)$ to be the smallest element in Γ that is greater than $\alpha(1), \ldots, \alpha(i-1)$ and satisfies

$$\varphi(f_{\alpha(i)}) \notin \mathrm{Span}_{\mathbb{F}_q}\{\varphi(f_\lambda) \mid \lambda \prec_{\mathbf{N}_0^r} \alpha(i)\}.$$

Write $\Delta(R, \rho, \varphi) = \{\alpha(1), \ldots, \alpha(n)\}$.

Definition 4.37. For $\lambda \in \Delta(R, \rho, \varphi)$ define

$$\sigma(\lambda) = \#\{\gamma \in \Delta(R, \rho, \varphi) \mid \gamma - \lambda \in \Gamma\}.$$

For $\lambda \in \Gamma$ define

$$\mu(\lambda) = \#\{\alpha \in \Gamma \mid \lambda - \alpha \in \Gamma\}.$$

We can now define the codes.

Definition 4.38. Let R be an order domain over $\mathbb{F}_q$ and let φ be a morphism as in Definition 4.32. Consider a fixed well-behaving basis $\{f_\lambda \mid \rho(f_\lambda) = \lambda, \lambda \in \Gamma\}$. For $\lambda \in \Gamma$ and $\delta \in \mathbf{N}$ consider the codes

$$E(\lambda) = \mathrm{Span}_{\mathbb{F}_q}\{\varphi(f_\eta) \mid \eta \preceq_{\mathbf{N}_0^r} \lambda\}$$
$$\tilde{E}(\delta) = \mathrm{Span}_{\mathbb{F}_q}\{\varphi(f_\eta) \mid \eta \in \Delta(R, \rho, \varphi) \text{ and } \sigma(\eta) \geq \delta\}$$
$$C(\lambda) = \{\vec{c} \in \mathbb{F}_q^n \mid \vec{c} \cdot \varphi(f_\eta) = 0 \text{ for all } \eta \text{ with } \eta \preceq_{\mathbf{N}_0^r} \lambda\}$$
$$\tilde{C}(\delta) = \{\vec{c} \in \mathbb{F}_q^n \mid \vec{c} \cdot \varphi(f_\eta) = 0 \text{ for all } \eta \in \Delta(R, \rho, \varphi) \text{ with } \mu(\eta) < \delta\}.$$

Remark 4.39. From Remark 4.35 we conclude that the choice of well-behaving basis is of no importance for the definition of the codes $E(\lambda)$ and $C(\lambda)$.

From [20, Th. 4.13 and Pro. 4.23] and [2, Th. 33] we have the following theorem. The result concerning $C(\lambda)$ and $\tilde{C}(\delta)$ is known as the order bound.

Theorem 4.40. *The minimum distance of $E(\lambda)$ is at least*

$$\min\{\sigma(\eta) \mid \eta \preceq_{\mathbf{N}_0^r} \lambda\} \tag{4.22}$$

and the minimum distance of $C(\lambda)$ is at least

$$\min\{\mu(\eta) \mid \lambda \prec_{\mathbf{N}_0^r} \eta \text{ and } \eta \in \Delta(R,\rho,\varphi)\} \geq \min\{\mu(\eta) \mid \lambda \prec_{\mathbf{N}_0^r} \eta\}. \tag{4.23}$$

The minimum distances of $\tilde{E}(\delta)$ and $\tilde{C}(\delta)$ are at least δ.

Recall from Theorem 4.31 that if Γ is a finitely generated value semigroup then the corresponding order domain R can be described as a factor ring. We now show that for such order domains Theorem 4.40 is a direct consequence of the theory developed in Section 4.7. We start with the following easy characterization of φ.

Proposition 4.41. *Let $\varphi : R = \mathbb{F}_q[X_1,\ldots,X_m]/I \to \mathbb{F}_q^n$ be a morphism as in Definition 4.32. There exists a set*

$$U = \{P_1,\ldots,P_n\} \subseteq \mathcal{V}_{\mathbb{F}_q}(I)$$

such that $\varphi(F+I) = (F(P_1),\ldots,F(P_n))$ for all $F+I \in R$. The P_i's are pairwise different.

Applying Proposition 4.41 to order domains with finitely generated value semigroup we see that the codes in Definition 4.38 are of the type covered by Remark 4.29 of Section 4.7. Rather than dealing with the general case $U \subseteq \mathcal{V}_{\mathbb{F}_q}(I)$ we will in the following concentrate on the situation $U = \mathcal{V}_{\mathbb{F}_q}(I)$. The reader can easily generalize our findings by replacing, as in Remark 4.29, any occurrence of I_q with I_U.
Our most important observation is that

$$\Delta(R,\rho,\varphi) = w(\Delta_{\prec_w}(I_q)). \tag{4.24}$$

To show (4.24) we start by noting that both sets are of size n. Hence, (4.24) must hold if we can show

$$\Delta(R,\rho,\varphi) \subseteq w(\Delta_{\prec_w}(I_q)).$$

Clearly, $\alpha(1) = 0$ is in $w(\Delta_{\prec_w}(I_q))$ as any non-empty footprint contains 1. Aiming for a contradiction assume $\alpha(i) \notin w(\Delta_{\prec_w}(I_q))$ for some $2 \leq i \leq n$. Let $f_{\alpha(i)} = F + I$, $w(\mathrm{lm}(F)) = \alpha(i)$. We have

$$\varphi(F+I) = \varphi(F \text{ rem } \mathcal{G} + I) \tag{4.25}$$

where $\mathcal{G}$ is a Gröbner basis for I_q. The very definition of a Gröbner basis ensures that $\mathrm{lm}(F \text{ rem } \mathcal{G}) \in \Delta_{\prec_w}(I_q)$. Hence, $\mathrm{lm}(F \text{ rem } \mathcal{G}) \prec_w \mathrm{lm}(F)$. But

then, by (4.25) and Definition 4.36, $\alpha(i) \notin \Delta(R, \rho, \varphi)$. We have reached at a contradiction and therefore (4.24) holds.
With (4.24) in hand we establish the following connections: $E(\lambda)$ and $C(\lambda)$ respectively equals $C(I, L_1)$ and $C(I, L_1)^{\perp}$ respectively where L_1 is as in (4.19). $\tilde{E}(\delta)$ equals $C(I, L_2)$ where L_2 is as in (4.20) and $\tilde{C}(\delta)$ equals $C(I, L_3)^{\perp}$ where L_3 is as in (4.21). We conclude that the bounds in Theorem 4.40 on the minimum distances of $E(\lambda)$, $\tilde{E}(\delta)$, $C(\lambda)$ and $\tilde{C}(\delta)$ are consequences of Theorem 4.27.

4.10. One-point geometric Goppa codes

One of the main reasons for introducing order domains in [20] was to have an easy description of one-point geometric Goppa codes and to have an easy way of generalizing the construction of one-point geometric Goppa codes to algebraic structures of higher transcendence degree. Presenting in the present chapter things in reverse order of what is normally done we now finally come to the one-point geometric Goppa codes.
Let $\mathcal{P}$ be a rational place in an algebraic function field $\mathbb{F}$ of one variable with constant field $\mathbb{F}_q$. Let $\nu_{\mathcal{P}}$ be the valuation corresponding to $\mathcal{P}$. Consider the algebraic structure

$$R = \cup_{m=0}^{\infty} \mathcal{L}(m\mathcal{P}). \tag{4.26}$$

Defining $\rho = -\nu_{\mathcal{P}}$ we have $\rho(R) = \Gamma \cup \{-\infty\}$ where $\Gamma \subseteq \mathbf{N}_0$ is known as the Weierstrass semigroup corresponding to $\mathcal{P}$. By inspection the map $\rho : R \to \Gamma \cup \{-\infty\}$ satisfies the six conditions in Definition 4.30 and therefore is a weight function.
Unfortunately it is not in general an easy task to determine the structure R above and therefore it is often difficult to find the factor ring description of R as guaranteed by Theorem 4.31. Observe, that one such description was given in Example 4.20 in the case of a Hermitian curve over $\mathbb{F}_9$.
The geometric Goppa codes coming from the structure in (4.26) are known as one-point geometric Goppa codes. We now explain the connection between these codes and the affine variety codes in Section 4.7. Let $\mathcal{Q}_1, \ldots, \mathcal{Q}_n$ be rational places, pairwise different, and all different from $\mathcal{P}$. The map $\varphi : R \to \mathbb{F}_q^n$, $\varphi(f) = (f(\mathcal{Q}_1), \ldots, f(\mathcal{Q}_n))$ is a morphism as in Definition 4.32. Therefore from Proposition 4.41 the rational places $\mathcal{Q}_1, \ldots, \mathcal{Q}_n$ correspond to n different affine points $P_1, \ldots, P_n$ in $\mathcal{V}(I_q)$ and $\varphi(F + I) = (F(P_1), \ldots, F(P_n))$ holds. We have

$$C_{\mathcal{L}}(\mathcal{Q}_1 + \cdots + \mathcal{Q}_n, \lambda\mathcal{P}) = C(I, L)$$

and

$$C_\Omega(\mathcal{Q}_1 + \cdots + \mathcal{Q}_n, \lambda\mathcal{P}) = C(I, L)^\perp$$

where

$$L = \{f \in R \mid \rho(f) \leq \lambda\}.$$

Let $\Gamma = \{\lambda_1, \lambda_2, \ldots\}$ where $\lambda_1 < \lambda_2 < \cdots$ holds. The Goppa bounds from algebraic geometry applied to the case of one-point geometric Goppa codes state.

Theorem 4.42. *Let $\mathcal{P}$ be a rational place as above and let R be the corresponding order domain as in (4.26). The minimum distance of $E(\lambda)$ is at least*

$$n - \lambda. \tag{4.27}$$

The minimum distance of $C(\lambda_t)$ is at least

$$t + 1 - g. \tag{4.28}$$

Now we show that the bounds in Theorem 4.42 can be viewed as being a consequence of Theorem 4.40. We will need the following technical lemma from [20, Lem. 5.15 and Th. 5.24].

Lemma 4.43. *Let $\Gamma = \{\lambda_1, \lambda_2, \ldots\}$ with $\lambda_1 < \lambda_2 < \cdots$ be a semigroup in $\mathbf{N}_0$ with finitely many gaps. Define*

$$g(i) = \#\{\lambda \in \mathbf{N}_0 \backslash \Gamma \mid \lambda < \lambda_i\}.$$

For any λ_i we have $\#(\Gamma \backslash (\lambda_i + \Gamma)) = \lambda_i$ and $\mu(\lambda_i) = i - g(i) + D(i)$ where

$$D(i) = \{(x, y) | x, y \in \mathbf{N}_0 \backslash \Gamma \text{ and } x + y = \lambda_i\}.$$

Here, $\lambda + \Gamma$ means $\{\lambda + \lambda_1, \lambda + \lambda_2, \ldots\}$.

Theorem 4.44. *For the case of one-point geometric Goppa codes the bound in (4.22) is always at least as good as (and sometimes better than) the bound in (4.27). Similarly, the bound in (4.23) is always at least as good as (and sometimes better than) the bound in (4.28).*

Proof. To prove the first claim we need only consider numbers $\lambda_i \in \Delta(R, \rho, \varphi)$, $\lambda_i \leq s$. We have $\sigma(\lambda_i) = \#(\Delta(R, \rho, \varphi) \cap (\lambda_i + \Gamma))$. From the first part of Lemma 4.43 we see that the number of elements in $\Delta(R, \rho, \varphi)$ that are not in $\lambda_i + \Gamma$ is at most λ_i. Therefore $\sigma(\lambda_i) \geq n - \lambda_i$ holds with

equality only when $\Gamma \backslash (\lambda_i + \Gamma) \subseteq \Delta(R, \rho, \varphi)$. We conclude $\min\{\sigma(\lambda_i) \mid \lambda_i \in \Delta(R, \rho, \varphi), \lambda_i \leq s\} \geq n - s$. Concerning the last claim we have

$$\min\{\mu(\eta) \mid \eta \in \Gamma \text{ and } \lambda_t < \eta\} = \min\{i - g(i) + \#D(i) \mid t < i\} \geq t + 1 - g$$

with equality if and only if $\lambda_{t+1} = \lambda_t + 1$, $g(t+1) = g$ and $\#D(t+1) = 0$ hold. □

Having shown that the bounds in Theorem 4.40 on the minimum distances of the codes $E(\lambda)$ and $C(\lambda)$ are at least as good as the Goppa bounds in the case of R being of the form (4.26) it is clear that we can consider the codes $\tilde{E}(\delta)$ and the codes $\tilde{C}(\delta)$ related to (4.26) as improved one-point geometric Goppa codes.
It was shown in [27, Th. 1] that all numerical weight functions (i. e. weight functions with weights in $\mathbb{N}_0$) are either of the form (4.26) or is a sub algebra of such a structure. Turning to semigroups that are not numerical the related structures are no longer curves but are higher dimensional [17, Sec. 11]. The related codes can be viewed as generalizations of one-point geometric Goppa codes.

4.11. Bibliographical Notes

The theory of evaluation codes has grown relatively large in its ten years' lifetime and therefore it is not possible to give a full list of references on the topic in the present chapter. Instead we will give just a few references on different aspects of the theory.
Examples of evaluation codes coming from higher dimensional objects than curves are given in [25] and [2]. Regarding generalized Hamming weights of evaluation codes more results can be found in [19], [3], [2], and [18]. The Feng-Rao bound as described in [11], [12], and [24] is more general than the order bound [20] in that it does not only deal with evaluation codes. The most general version of the Feng-Rao bound deals with linear codes [29], [18]. The Gröbner basis theoretical point of view on order domains are studied in [30], [31], [28], [33], [21], and [17]. Evaluation codes are described in a Gröbner basis theoretical setting in [30], [31], [1], and [2]. For the case of affine variety codes decoding algorithms can be found in [13], [10], [32]. Many papers deal with decoding of evaluation codes. Among these are [20], [6], and [16]. A study of the function μ on different families of semigroups Γ can be found in [5] and [34].

References

[1] H. E. Andersen, Codes from Order Domains, *PhD Report Series,* **12**, (2005.)

[2] H. E. Andersen and O. Geil, Evaluation codes from order domain theory, *Finite Fields and Their Applications,* **14**, (2008), pp. 92-123.

[3] A. I. Barbero and C. Munuera, The Weight Hierarchy of Hermitian Codes, *SIAM Journ. Discr. Math.,* **13**, (2000), pp. 79-104.

[4] T. Becker and V. Weispfenning, "Gröbner Bases - A Computational Approach to Commutative Algebra," Springer Verlag, Berlin, (1993).

[5] M. Bras-Amorós, Acute Semigroups, the Order Bound on the Minimum Distance, and the Feng-Rao Improvements, *IEEE Trans. Inform. Theory,* **50**, (2004), pp. 1282-1289.

[6] M. Bras-Amorós and M. E. O'Sullivan, The Correction Capability of the Berlekamp-Massey-Sakata Algorithm with Majority Voting, *Appl. Alpgebra Engrg. Comm. Comput.,* **17**, (2006). pp. 315-335.

[7] D. Cox, J. Little, and D. O'Shea, "Ideals, Varieties, and Algorithms, 2nd ed.," Springer, Berlin, (1997).

[8] D. Cox, J. Little, and D. O'Shea, "Using Algebraic Geometry," Springer, Berlin, (1998).

[9] P. Delsarte, J. M. Goethals, and F. J. Mac Williams, On generalized Reed-Muller codes and their relatives, *Information and Control*, **16**, (1970), 403-442.

[10] J. B. Farr and S. Gao, Gröbner bases Padé approximation and decoding of linear codes, *Proc. of Coding theory and quantum computing (Virginia 2003), Contemp. Math.,* **381**, Amer. Math. Soc. (2005), pp. 3-18.

[11] G.-L. Feng and T.R.N. Rao, Decoding of algebraic geometric codes up to the designed minimum distance, *IEEE Trans. Inf. Theory*, **39**, (1993), pp. 37-46.

[12] G.-L. Feng and T.R.N. Rao, Improved Geometric Goppa Codes, Part I:Basic theory, *IEEE Trans. Inf. Theory*, **41**, (1995), pp. 1678-1693.

[13] J. Fitzgerald and R. F. Lax, Decoding Affine Variety Codes Using Gröbner Bases, *Designs, Codes and Cryptography*, **13**, **2**, (1998), pp. 147-158.

[14] O. Geil and T. Høholdt, On Hyperbolic Codes, *Proc. of AAECC-14, Lecture Notes in Comput. Sci. 2227*, Springer, Berlin, (2001), pp. 159-171.

[15] O. Geil and T. Høholdt, On Hyperbolic Type Codes, Proc. of 2003 IEEE International Symposium on Inform. Theory, Yokohama, Japan, June 29-July 4, (2003), p. 331.

[16] O. Geil and R. Matsumoto, Generalized Sudan's List Decoding for Order Domain Codes, *Proc. of AAECC-17, Lecture Notes in Comput. Sci. 4851*, Springer, Berlin, (2007), pp. 50-59.

[17] O. Geil and R. Pellikaan, On the Structure of Order Domains, *Finite Fields and their Applications*, **8**, (2002), pp. 369-396.

[18] O. Geil and C. Thommesen, On the Feng-Rao Bound for Generalized Hamming Weights, *Proc. of AAECC-16, Lecture Notes in Comput. Sci. 3857*, Springer, Berlin, (2006), pp. 295-306.

[19] P. Heijnen and R. Pellikaan, Generalized Hamming weights of q-ary Reed-

Muller codes, *IEEE Trans. Inf. Theory*, **44**, (1998), pp. 181-196.

[20] T. Høholdt, J. van Lint, and R. Pellikaan, "Algebraic Geometry Codes," Chapter 10 in *Handbook of Coding Theory* (V.S. Pless and W.C. Huffman, eds.), vol. 1, Elsevier, Amsterdam, (1998), pp. 871-961.

[21] R. Pellikaan, On the existence of order functions, *Journal of Statistical Planning and Inference,* **94**, (2001), pp. 287-301.

[22] G. Kabatiansky, Two Generalizations of Product Codes, *Proc. of Academy of Science USSR, Cybernetics and Theory of Regulation*, **232**, vol. 6, (1977), pp. 1277-1280 (in Russian).

[23] T. Kasami, S. Lin, and W. Peterson, New generalizations of the Reed-Muller codes. I. Primitive codes, *IEEE Transactions on Information Theory*, **14**, (1968), pp. 189-199.

[24] M. S. Kolluru, G. L. Feng, and T. R. N. Rao, Construction of Improved Geometric Gopps Codes from Klein Curves and Klein-like Curves, *Applicable Algebra in Engineering, Communication and Computing*, **10**, (2000), pp. 433-464.

[25] J. B. Little, The Ubiquity of Order Domains for the Construction of Error Control Codes, *Advances in Mathematics of Communications*, **1**, (2007), pp. 151-171.

[26] J. Massey, D. J. Costello, and J. Justesen, Polynomial Weights and Code Constructions, *IEEE Trans. Inf. Theory*, **19** (1973), pp. 101-110.

[27] R. Matsumoto, Miura's Generalization of One-Point AG codes is Equivalent to Høholdt, van Lint and Pellikaan's Generalization, *IEICE Trans. Fundamentals,* **E82-A**, no. 10 (1999), 2007-2010.

[28] R. Matsumoto and S. Miura, On Construction and Generalization of Algebraic Geometry Codes, textitProc. of Algebraic Geometry, Number Thoery, Coding Theory and Cryptography, Univ. of Tokyo, January 19-20, 2000, (Ed. T. Katsura et al.), (2000) pp. 3-15.

[29] R. Matsumoto and S. Miura, On the Feng-Rao Bound for the $\mathcal{L}$-Construction of Algebraic Geometry Codes, *IEICE Trans. Fund.,* **E83-A**, no. 5 (2000), pp. 923-927.

[30] S. Miura, Linear Codes on Affine Algebraic Varieties, *Trans. IEICE*, **J81-A**, no. 10 (1998), pp. 1386-1397 (in Japanese).

[31] S. Miura, Linear Codes on Affine Algebraic Curves, *Trans. IEICE*, **J81-A**, no. 10 (1998), pp. 1398-1421 (in Japanese).

[32] E. Orsini and M. Sala, Improved Decoding of Affine-Variety Codes, *BCRI preprint*, www.bcri.ucc.ie, University College Cork, Boole Centre BCRI, UCC Cork, Ireland, (2007).

[33] M. E. O'Sullivan, New codes for the Berlekamp-Massey-Sakata algorithm, *Finite Fields and Their Applications*, **7**, 2001, pp. 293-317.

[34] D. Ruano, Computing the Feng-Rao Distances for Codes from Order Domains, *Journ. of Algebra,*, **309**, (2007), pp. 672-682.

Chapter 5

Asymptotically Good Codes

Harald Niederreiter

Department of Mathematics, National University of Singapore
2 Science Drive 2, Singapore 117543, Republic of Singapore
nied@math.nus.edu.sg

Ferruh Özbudak

Department of Mathematics and Institute of Applied Mathematics
Middle East Technical University, İnönü Bulvarı, 06531, Ankara, Turkey
ozbudak@metu.edu.tr

For a prime power q, let α_q be the standard function in the asymptotic theory of codes, that is, $\alpha_q(\delta)$ is the largest asymptotic information rate that can be achieved for a given asymptotic relative minimum distance δ of q-ary codes. In recent years the Tsfasman-Vlăduţ-Zink lower bound on $\alpha_q(\delta)$ was improved by Elkies, Xing, Niederreiter and Özbudak, Stichtenoth and Xing, and Maharaj. In this chapter we give an exposition of these results.

Contents

5.1. Introduction

For any prime power q, let $\mathbb{F}_q$ denote the finite field of order q. We write $|M|$ for the cardinality of a finite set M. If C is a code over $\mathbb{F}_q$ (also called a q-ary code), we always assume that $|C| \geq 2$. We write $n(C)$ for

the length of C and $d(C)$ for the minimum distance of C.

For any prime power q, let α_q and α_q^{lin} denote the important functions in the asymptotic theory of codes which are defined by

$$\alpha_q(\delta) = \sup\{R \in [0,1] : (\delta, R) \in U_q\} \qquad \text{for } 0 \le \delta \le 1 \tag{5.1}$$

and

$$\alpha_q^{\mathrm{lin}}(\delta) = \sup\{R \in [0,1] : (\delta, R) \in U_q^{\mathrm{lin}}\} \qquad \text{for } 0 \le \delta \le 1. \tag{5.2}$$

Here U_q (resp. U_q^{lin}) is the set of all ordered pairs $(\delta, R) \in [0,1]^2$ for which there exists a sequence $\{C_i\}_{i=1}^{\infty}$ of not necessarily linear (resp. linear) codes over $\mathbb{F}_q$ such that $n(C_i) \to \infty$ as $i \to \infty$ and

$$\delta = \lim_{i\to\infty} \frac{d(C_i)}{n(C_i)}, \quad R = \lim_{i\to\infty} \frac{\log_q |C_i|}{n(C_i)},$$

where $\log_q$ is the logarithm to the base q. The following basic properties of the functions α_q and α_q^{lin} can be found in [14, Section 1.3.1]. It is trivial that $\alpha_q(\delta) \ge \alpha_q^{\mathrm{lin}}(\delta)$ for $0 \le \delta \le 1$.

Proposition 5.1. *The functions α_q and α_q^{lin} have the following properties:*

(i) *α_q and α_q^{lin} are nonincreasing and continuous on $[0,1]$;*

(ii) *$\alpha_q(0) = \alpha_q^{\mathrm{lin}}(0) = 1$;*

(iii) *$\alpha_q(\delta) = \alpha_q^{\mathrm{lin}}(\delta) = 0$ for $(q-1)/q \le \delta \le 1$.*

Values of $\alpha_q(\delta)$ and $\alpha_q^{\mathrm{lin}}(\delta)$ are not known for $0 < \delta < (q-1)/q$. A central problem in the asymptotic theory of codes is to find lower bounds on $\alpha_q(\delta)$ and $\alpha_q^{\mathrm{lin}}(\delta)$ for $0 < \delta < (q-1)/q$. A classical lower bound dating from the 1950s is the following one (see [4] and [5]).

Proposition 5.2 (Asymptotic Gilbert-Varshamov Bound). *For any prime power q, we have*

$$\begin{aligned} &\alpha_q^{\mathrm{lin}}(\delta) \\ &\ge R_{\mathrm{GV}}(\delta) := 1 - \delta\log_q(q-1) + \delta\log_q\delta + (1-\delta)\log_q(1-\delta) \end{aligned} \tag{5.3}$$

for $0 < \delta < (q-1)/q$.

Since no improvement on (5.3) was obtained for a long time, there was speculation that maybe $\alpha_q^{\mathrm{lin}}(\delta) = R_{\mathrm{GV}}(\delta)$ for $0 < \delta < (q-1)/q$. However, this conjecture was disproved when the following bound was established in [15] by using algebraic-geometry codes.

Proposition 5.3 (Tsfasman-Vlăduţ-Zink Bound). *For any prime power q, we have*

$$\alpha_q^{\mathrm{lin}}(\delta) \geq 1 - \delta - \frac{1}{A(q)} \qquad \text{for } 0 \leq \delta \leq 1. \tag{5.4}$$

Here and in the following, we put

$$A(q) := \limsup_{g \to \infty} \frac{N_q(g)}{g},$$

where $N_q(g)$ denotes the maximum number of rational places that a global function field of genus g with full constant field $\mathbb{F}_q$ can have. We recall from [10, Chapter 5] that $A(q) > 0$ for all q and that $A(q) = \sqrt{q} - 1$ if q is a square. For nonsquares q the exact value of $A(q)$ is not known, but we have lower and upper bounds on $A(q)$ (see again [10, Chapter 5] and also the more recent paper [1]).

It was already proved in [15] that (5.4) yields a better lower bound on $\alpha_q^{\mathrm{lin}}(\delta)$ than (5.3) for all squares $q \geq 49$ and for certain ranges of the parameter δ. Later, analogous results were shown for other cases of sufficiently large composite q by using (5.4); see [10, Section 6.2] for a survey of such results.

Recently the Tsfasman-Vlăduţ-Zink bound (5.4) was improved for not necessarily linear codes by Elkies [2] and Xing [18]. Shortly thereafter, Niederreiter and Özbudak [7, Corollary 5.4] improved the bound in Xing [18] by showing that

$$\alpha_q(\delta) \geq 1 - \delta - \frac{1}{A(q)} + \log_q\left(1 + \frac{1}{q^3}\right) \qquad \text{for } 0 \leq \delta \leq 1. \tag{5.5}$$

Later, Stichtenoth and Xing [12] gave a simpler proof of (5.5). These results improve the Tsfasman-Vlăduţ-Zink bound on $\alpha_q(\delta)$ uniformly, i.e, for all values of q and δ.

The bound (5.4) for $\alpha_q^{\mathrm{lin}}(\delta)$ was already improved, although not uniformly in δ, by Vlăduţ [16] (see also [14, Chapter 3.4]) and Xing [17]. Recently, Niederreiter and Özbudak [8] improved the bound (5.5) for certain values of q and δ. Maharaj [6] refined the approach of Stichtenoth and Xing [12] and also obtained improvements on (5.5) for certain values of q and δ. Later, Niederreiter and Özbudak [9] refined and complemented the methods of [8] and improved all previous bounds on $\alpha_q^{\mathrm{lin}}(\delta)$ and $\alpha_q(\delta)$ for certain values of q and δ.

In this chapter we give an exposition of these results. In Section 5.2 we present some background and fix some notation. The uniform improve-

ments in [7] and [12] on the lower bound for $\alpha_q(\delta)$ yielding (5.5) are explained in Sections 5.3 and Section 5.4. The improvements on the lower bounds for $\alpha_q^{\mathrm{lin}}(\delta)$ and $\alpha_q(\delta)$ for certain values of q and δ given in [9] are explained in Section 5.5.

5.2. Preliminaries

In this section we recall some definitions, we explain essential terminology, and we give the basic background that we use throughout the chapter. We also recall some important existence results on sequences of global function fields over finite fields with certain properties that we will use in our constructions.

A *global function field* F over $\mathbb{F}_q$ is an extension field of $\mathbb{F}_q$ such that there exists an element $z \in F$ that is transcendental over $\mathbb{F}_q$ and for which F is a finite extension of the rational function field $\mathbb{F}_q(z)$. Moreover, we call $\mathbb{F}_q$ the *full constant field* of F if $\mathbb{F}_q$ is algebraically closed in F. A *place* of F is the maximal ideal of some valuation ring of F. Let $\mathbb{Z}$ denote the set of integers. A *normalized discrete valuation* of F is a surjective map $\nu : F \to \mathbb{Z} \cup \{\infty\}$ satisfying the following properties:

(i) $\nu(x) = \infty \iff x = 0$;
(ii) $\nu(xy) = \nu(x) + \nu(y)$ for all $x, y \in F$;
(iii) $\nu(x + y) \geq \min(\nu(x), \nu(y))$ for all $x, y \in F$;
(iv) $\nu(a) = 0$ for all $a \in \mathbb{F}_q \setminus \{0\}$.

There is a bijective correspondence between the places of F and the normalized discrete valuations of F. Let ν_P be the normalized discrete valuation of F corresponding to the place P of F. The valuation ring of P is

$$\mathcal{O}_P := \{x \in F : \nu_P(x) \geq 0\}$$

and the maximal ideal of $\mathcal{O}_P$ is

$$M_P := \{x \in \mathcal{O}_P : \nu_P(x) > 0\}.$$

If $\mathbb{F}_q$ is the full constant field of F, then the residue class field $\mathcal{O}_P/M_P$ can be identified with a finite extension of $\mathbb{F}_q$. The degree of this extension is called the *degree* of the place P, denoted by $\deg(P)$. A place of degree 1 is called *rational*. For detailed background on global function fields, we refer to the book of Stichtenoth [11].

Assume that F is a global function field with full constant field $\mathbb{F}_q$. Let $\mathbb{P}_F$ be the set of all places of F. For $f \in F \setminus \{0\}$,

$$(f) := \sum_{P \in \mathbb{P}_F} \nu_P(f) P$$

denotes the *principal divisor* of f and

$$(f)_0 := \sum_{\substack{P \in \mathbb{P}_F \\ \nu_P(f) > 0}} \nu_P(f) P$$

denotes the *zero divisor* of f. For an arbitrary divisor

$$G = \sum_{P \in \mathbb{P}_F} m_P P$$

of F, we write $\nu_P(G)$ for the coefficient m_P of P. Note that $\nu_P(G)$ is an integer. The *degree* of G is defined by

$$\deg(G) = \sum_{P \in \mathbb{P}_F} \nu_P(G) \deg(P).$$

The *support* of G is given by the finite set

$$\mathrm{supp}(G) = \{P \in \mathbb{P}_F : \nu_P(G) \neq 0\}.$$

We use the standard notation

$$\mathcal{L}(G) = \{f \in F : \nu_P(f) \geq -\nu_P(G) \text{ for all } P \in \mathbb{P}_F\}$$

for the *Riemann-Roch space* of G. Note that $\mathcal{L}(G)$ is a finite-dimensional vector space over $\mathbb{F}_q$. It is a well-known fact that

$$\mathcal{L}(G) = \{0\} \text{ if } \deg(G) < 0. \tag{5.6}$$

The following fundamental result provides information on the dimension $\dim(\mathcal{L}(G))$ of $\mathcal{L}(G)$ over $\mathbb{F}_q$.

Proposition 5.4 (Riemann-Roch Theorem). *Let F be a global function field with full constant field $\mathbb{F}_q$. Assume that g is the genus of F. Then for an arbitrary divisor G of F we have*

$$\dim(\mathcal{L}(G)) \geq \deg(G) + 1 - g,$$

and equality holds if $\deg(G) \geq 2g - 1$.

Remark 5.5. With the notation of Proposition 5.4, if $\dim(\mathcal{L}(G)) = \deg(G) + 1 - g$, then G is called a *nonspecial* divisor of F. It is well known that if G is a nonspecial divisor and $G' \geq G$, then the divisor G' is also nonspecial (cf. [11, Remark I.6.9]).

We also use the following simple lemma.

Lemma 5.6. *Let F be a global function field with full constant field $\mathbb{F}_q$. Let $P_1, \ldots, P_n$ be distinct rational places of F and r be an arbitrary integer. Then there exists a divisor G of F with* $\deg(G) = r$ *and* $supp(G) \cap \{P_1, \ldots, P_n\} = \varnothing$.

Proof. Let g be the genus of F. Using [11, Corollary V.2.10] we obtain that for an integer $\ell \geq 4g + 3$, there exist places $Q_{\ell+1}$ and Q_ℓ of F such that $\deg(Q_{\ell+1}) = \ell + 1$ and $\deg(Q_\ell) = \ell$. Hence putting, for example, $G = r\,(Q_{\ell+1} - Q_\ell)$ we complete the proof. □

The real-valued function $E_q(x)$ on the interval $[0, 1]$ defined below corresponds to the entropy function $H_2(x) = -x \log_2 x - (1 - x) \log_2(1 - x)$ when $q = 2$ (see [5, p. 308] and also (5.3)).

Definition 5.7. For a real number $0 < x < 1$, let

$$E_q(x) = -x \log_q x - (1 - x) \log_q(1 - x).$$

For $x \in \{0, 1\}$ we put $E_q(x) = 0$.

Using Stirling's formula, it is not difficult to show that for any real number $0 \leq x \leq 1$ we have the following asymptotic behavior of binomial coefficients:

$$\lim_{n\to\infty} \frac{\log_q \binom{n}{\lfloor xn \rfloor}}{n} = E_q(x).$$

Now we recall an important result on sequences of global function fields from the literature.

Remark 5.8. Assume that q is a square and a prime power. Let $\gamma = \sqrt{q}-1$. By [15] or [3], there exists a sequence $(F_i)_{i=1}^{\infty}$ of global function fields with full constant field $\mathbb{F}_q$ such that $g_i \to \infty$ as $i \to \infty$ and $\lim_{i\to\infty} \frac{n_i}{g_i} = \gamma$. Here n_i and g_i are the number of rational places and the genus of F_i, respectively, for $i \geq 1$. Note that for any such q and γ, we have that $1 - \frac{1}{\gamma} + \log_q \left(1 + \frac{1}{q^3}\right) > 0$.

5.3. Two Constructions of Asymptotically Good Codes

In this section we explain in detail two constructions by Niederreiter and Özbudak [7] of asymptotically good codes with excellent parameters. These constructions yield the currently best global improvement on the

Tsfasman-Vlăduţ-Zink bound (see (5.5) and Corollary 5.23). We present these constructions in a somewhat more general framework than that in [7].

We fix an arbitrary positive integer m. Let F be a global function field with full constant field $\mathbb{F}_q$ such that there exist at least $n \geq 1$ distinct rational places $P_1, \ldots, P_n$ of F. Let t_i be a local parameter of F at P_i for $1 \leq i \leq n$, that is, $\nu_{P_i}(t_i) = 1$ for $1 \leq i \leq n$.

For a nonnegative integer r, let G be a divisor of F with $\deg(G) = r$ and $\operatorname{supp}(G) \cap \{P_1, \ldots, P_n\} = \varnothing$ (cf. Lemma 5.6). For $f \in \mathcal{L}(G)$, the local expansion of f at P_i has the form

$$f = \sum_{\ell=0}^{\infty} f^{(\ell)}(P_i)\, t_i^{\ell}$$

with $f^{(\ell)}(P_i) \in \mathbb{F}_q$ for $1 \leq i \leq n$ and $\ell \geq 0$.

For each $i = 1, \ldots, n$, let

$$\begin{aligned} \phi_i : \mathcal{L}(G) &\to \mathbb{F}_q^m \\ f &\mapsto \left(f^{(m-1)}(P_i), \ldots, f^{(1)}(P_i), f^{(0)}(P_i)\right). \end{aligned}$$

Let Φ be the $\mathbb{F}_q$-linear map defined by

$$\begin{aligned} \Phi : \mathcal{L}(G) &\to \mathbb{F}_q^{mn} \\ f &\mapsto (\phi_1(f), \ldots, \phi_n(f)). \end{aligned} \tag{5.7}$$

Moreover, let ψ be the $\mathbb{F}_q$-linear map

$$\begin{aligned} \psi : \mathcal{L}(G) &\to \mathbb{F}_q^{n} \\ f &\mapsto \left(f^{(m)}(P_1), \ldots, f^{(m)}(P_n)\right). \end{aligned} \tag{5.8}$$

Remark 5.9. We note that the maps ψ and Φ (provided $m \geq 2$ for the map Φ) depend on the choice of the local parameters t_i.

Initially we will assume that G (or its degree r, cf. Lemma 5.6) is chosen such that Φ is a surjective map. We will relax this condition later in this section using an averaging argument (see also Remark 5.18).

We need to introduce certain weights on $\mathbb{F}_q^{(m+1)n}$ and $\mathbb{F}_q^{mn}$. Let $\boldsymbol{a} = \left(a^{(m)}, a^{(m-1)}, \ldots, a^{(0)}\right) \in \mathbb{F}_q^{m+1}$ and $\boldsymbol{b} = \left(b^{(m-1)}, \ldots, b^{(0)}\right) \in \mathbb{F}_q^{m}$. We define the weights $v_{m+1}(\boldsymbol{a})$ and $v_{(2,\ldots,m+1)}(\boldsymbol{b})$ as

$$v_{m+1}(\boldsymbol{a}) = \begin{cases} m+1 & \text{if } a^{(0)} \neq 0, \\ m & \text{if } a^{(1)} \neq 0 \text{ and } a^{(0)} = 0, \\ \vdots & \vdots \\ 1 & \text{if } a^{(m)} \neq 0, \text{ and } a^{(m-1)} = \cdots = a^{(0)} = 0, \\ 0 & \text{if } \boldsymbol{a} = \boldsymbol{0}, \end{cases} \tag{5.9}$$

and

$$v_{(2,\ldots,m+1)}(\boldsymbol{b}) = \begin{cases} m+1 \text{ if } b^{(0)} \neq 0, \\ \vdots \qquad \vdots \\ 2 \qquad \text{if } b^{(m-1)} \neq 0 \text{ and } b^{(m-2)} = \cdots = b^{(0)} = 0, \\ 0 \qquad \text{if } \boldsymbol{b} = \boldsymbol{0}. \end{cases} \tag{5.10}$$

If $\boldsymbol{A} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) \in \mathbb{F}_q^{(m+1)n}$ with $\boldsymbol{a}_i \in \mathbb{F}_q^{m+1}$ for $1 \leq i \leq n$, then the weight $V_{m+1}(\boldsymbol{A})$ is given by

$$V_{m+1}(\boldsymbol{A}) = \sum_{i=1}^{n} v_{m+1}(\boldsymbol{a}_i). \tag{5.11}$$

Similarly if $\boldsymbol{B} = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n) \in \mathbb{F}_q^{mn}$ with $\boldsymbol{b}_i \in \mathbb{F}_q^{m}$ for $1 \leq i \leq n$, then the weight $V_{(2,\ldots,m+1)}(\boldsymbol{B})$ is given by

$$V_{(2,\ldots,m+1)}(\boldsymbol{B}) = \sum_{i=1}^{n} v_{(2,\ldots,m+1)}(\boldsymbol{b}_i). \tag{5.12}$$

In the next lemma we show a useful relationship between the weights in (5.11) and (5.12).

Lemma 5.10. *Let* $\boldsymbol{C} = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n) \in \mathbb{F}_q^{(m+1)n}$ *with* $\boldsymbol{c}_i = \left(c_i^{(m)}, c_i^{(m-1)}, \ldots, c_i^{(0)}\right) \in \mathbb{F}_q^{m+1}$ *for* $1 \leq i \leq n$. *Consider the related codewords* $\boldsymbol{A} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) \in \mathbb{F}_q^{mn}$ *and* $\boldsymbol{b} = (b_1, \ldots, b_n) \in \mathbb{F}_q^{n}$ *given by*

$$\boldsymbol{a}_i = \left(c_i^{(m-1)}, \ldots, c_i^{(0)}\right) \in \mathbb{F}_q^{m} \text{ and } b_i = c_i^{(m)} \in \mathbb{F}_q$$

for $1 \leq i \leq n$. *Then we have*

$$V_{m+1}(\boldsymbol{C}) \leq V_{(2,\ldots,m+1)}(\boldsymbol{A}) + ||\boldsymbol{b}||,$$

where $||\boldsymbol{b}||$ *is the Hamming weight of* $\boldsymbol{b}$.

Proof. It suffices to prove that for each $1 \leq i \leq n$ we have

$$v_{m+1}(\boldsymbol{c}_i) \leq \begin{cases} v_{(2,\ldots,m+1)}(\boldsymbol{a}_i) & \text{if } b_i = 0, \\ v_{(2,\ldots,m+1)}(\boldsymbol{a}_i) + 1 & \text{if } b_i \neq 0. \end{cases}$$

Assume that $b_i = 0$. Then it follows from (5.9) and (5.10) that $v_{m+1}(\boldsymbol{c}_i) = v_{(2,\ldots,m+1)}(\boldsymbol{a}_i)$. Now assume that $b_i \neq 0$. If $\boldsymbol{a}_i \neq \boldsymbol{0}$, then again from the definitions in (5.9) and (5.10) we obtain $v_{m+1}(\boldsymbol{c}_i) = v_{(2,\ldots,m+1)}(\boldsymbol{a}_i)$. If $\boldsymbol{a}_i = \boldsymbol{0}$, then we get $v_{m+1}(\boldsymbol{c}_i) = v_{(2,\ldots,m+1)}(\boldsymbol{a}_i) + 1$. This completes the proof. □

The following notion is important for our constructions.

Definition 5.11. For a nonempty subset M of $\mathbb{F}_q^{mn}$, let $\hat{\rho}_{(2,\ldots,m+1)}(M)$ denote the integer

$$\hat{\rho}_{(2,\ldots,m+1)}(M) = \max\left\{V_{(2,\ldots,m+1)}(\boldsymbol{A}-\boldsymbol{B}) : \boldsymbol{A},\boldsymbol{B} \in M\right\}. \quad (5.13)$$

Now we are ready to give our basic construction. Let $M \subseteq \mathbb{F}_q^{mn}$ be nonempty and $N = \Phi^{-1}(M)$ be the inverse image of M under Φ. Let C be the q-ary code of length n defined by

$$C = \psi(N), \quad (5.14)$$

where ψ is given in (5.8).

Theorem 5.12. *Let m be an arbitrary positive integer. Assume that F is a global function field with full constant field $\mathbb{F}_q$ such that there exist at least $n \geq 1$ distinct rational places $P_1, \ldots, P_n$ of F. Let G be a divisor of F such that $\deg(G) = r$, $supp(G) \cap \{P_1, \ldots, P_n\} = \varnothing$, and the corresponding map Φ defined in (5.7) is surjective. Let $M \subseteq \mathbb{F}_q^{mn}$ be such that $|M| \geq 2$ and $(m+1)n - r - \hat{\rho}_{(2,\ldots,m+1)}(M) \geq 1$. Then C defined in (5.14) is an $(n, |C|, d(C))$ code over $\mathbb{F}_q$ with*

$$|C| = |\mathcal{L}(G - m(P_1 + \cdots + P_n))| \cdot |M| \quad (5.15)$$

and

$$d(C) \geq (m+1)n - r - \hat{\rho}_{(2,\ldots,m+1)}(M). \quad (5.16)$$

Also C is $\mathbb{F}_q$-linear if M is an $\mathbb{F}_q$-linear subset of $\mathbb{F}_q^{mn}$.

Proof. Let $f, g \in N$ with $f \neq g$. For $1 \leq i \leq n$ and $h \in \{f, g\}$, let

$$\begin{aligned}
\boldsymbol{c}_i(h) &= \left(h^{(m)}(P_i), h^{(m-1)}(P_i), \ldots, h^{(0)}(P_i)\right) \in \mathbb{F}_q^{m+1},\\
\boldsymbol{a}_i(h) &= \left(h^{(m-1)}(P_i), \ldots, h^{(0)}(P_i)\right) \in \mathbb{F}_q^{m},\\
b_i(h) &= h^{(m)}(P_i) \in \mathbb{F}_q.
\end{aligned}$$

Moreover, let

$$\begin{aligned}
\boldsymbol{C}(h) &= (\boldsymbol{c}_1(h), \ldots, \boldsymbol{c}_n(h)) \in \mathbb{F}_q^{(m+1)n},\\
\boldsymbol{A}(h) &= (\boldsymbol{a}_1(h), \ldots, \boldsymbol{a}_n(h)) \in \mathbb{F}_q^{mn},\\
\boldsymbol{b}(h) &= (b_1(h), \ldots, b_n(h)) \in \mathbb{F}_q^{n}.
\end{aligned}$$

Using Lemma 5.10, we obtain that

$$V_{m+1}(\boldsymbol{C}(f) - \boldsymbol{C}(g)) \leq V_{(2,\ldots,m+1)}(\boldsymbol{A}(f) - \boldsymbol{A}(g)) + ||\boldsymbol{b}(f) - \boldsymbol{b}(g)||. \quad (5.17)$$

Note that $\boldsymbol{A}(f), \boldsymbol{A}(g) \in M$ and hence we have (cf. Definition 5.11)

$$V_{(2,\ldots,m+1)}\left(\boldsymbol{A}(f)-\boldsymbol{A}(g)\right) \leq \hat{\rho}_{(2,\ldots,m+1)}(M). \tag{5.18}$$

By the definition of the weight in (5.9) we get

$$f-g \in \mathcal{L}\left(G-\sum_{i=1}^{n}\left(m+1-v_{m+1}\left(\boldsymbol{c}_i(f)-\boldsymbol{c}_i(g)\right)\right)P_i\right). \tag{5.19}$$

As $f \neq g$, the degree of the divisor in (5.19) is nonnegative (compare with (5.6)). This yields

$$r-(m+1)n+\sum_{i=1}^{n} v_{m+1}\left(\boldsymbol{c}_i(f)-\boldsymbol{c}_i(g)\right) \geq 0.$$

This means that

$$V_{m+1}\left(\boldsymbol{C}(f)-\boldsymbol{C}(g)\right) \geq (m+1)n-r. \tag{5.20}$$

Using (5.17), (5.18), and (5.20), we obtain that

$$\begin{aligned} ||\boldsymbol{b}(f)-\boldsymbol{b}(g)|| &\geq V_{m+1}\left(\boldsymbol{C}(f)-\boldsymbol{C}(g)\right)-V_{(2,\ldots,m+1)}(\boldsymbol{A}(f)-\boldsymbol{A}(g)) \\ &\geq (m+1)n-r-\hat{\rho}_{(2,\ldots,m+1)}(M). \end{aligned}$$

Therefore, as $\boldsymbol{b}(f)-\boldsymbol{b}(g)=\psi(f)-\psi(g)$, we obtain (5.16). Using (5.16) and the assumption $(m+1)n-r-\hat{\rho}_{(2,\ldots,m+1)}(M) \geq 1$, we conclude that the map ψ is one-to-one when restricted to N.

It remains to prove that $|C|$ is as given in (5.15). Note that $|C|=|N|$, and since Φ is a surjective $\mathbb{F}_q$-linear map, we have

$$|C|=|N|=|\Phi^{-1}(M)|=|\mathrm{Ker}(\Phi)|\cdot|M|.$$

Now $\mathrm{Ker}(\Phi)=\mathcal{L}\left(G-m(P_1+\cdots+P_n)\right)$, which completes the proof. □

Remark 5.13. Under the notation and assumptions of Theorem 5.12, let g be the genus of F. If

$$mn+2g-1 \leq r \leq (m+1)n-\hat{\rho}_{(2,\ldots,m+1)}(M)-1,$$

then both of the conditions of the theorem that Φ is surjective and that $(m+1)n-r-\hat{\rho}_{(2,\ldots,m+1)}(M) \geq 1$ are satisfied. Indeed, if $r-mn \geq 2g-1$, then using Proposition 5.4 we obtain that $\dim(\mathcal{L}(G)) = r+1-g$ and $\dim(\mathrm{Ker}(\Phi)) = r-mn+1-g$, and hence Φ is surjective.

Remark 5.14. Note that N is a nonlinear subset of $\mathcal{L}(G)$ in general. Nevertheless, there exists a nonempty subset $N_0 \subseteq N$ such that N is the disjoint union of $\mathcal{L}\left(G - m\left(P_1 + \cdots + P_n\right)\right)$-cosets of the elements of N_0, i.e.,

$$N = \bigcup_{g \in N_0} \left\{ f + g : f \in \mathcal{L}\left(G - m\left(P_1 + \cdots + P_n\right)\right) \right\}.$$

Let U_0 be the subset of C defined by

$$U_0 = \psi(N_0).$$

Then C is a nonlinear code in general, but C is the disjoint union of $\psi\big(\mathcal{L}(G - m(P_1 + \cdots + P_n))\big)$-cosets of the elements of U_0, namely

$$C = \bigcup_{\boldsymbol{u} \in U_0} \left\{ \boldsymbol{c} + \boldsymbol{u} : \boldsymbol{c} \in \psi\left(\mathcal{L}\left(G - m\left(P_1 + \cdots + P_n\right)\right)\right) \right\}.$$

Next we obtain an asymptotic version of Theorem 5.12. In the following theorem, for a fixed positive integer m, we assume the existence of suitable sequences of global function fields and suitable sequences of subsets of $\mathbb{F}_q^{mn}$ as $n \to \infty$. Recall that in Remark 5.8 we gave concrete examples of such sequences of global function fields. Moreover, in Proposition 5.20 below we will construct suitable sequences of subsets explicitly. These sequences of global function fields and sequences of subsets will imply excellent lower bounds on the function $\alpha_q(\delta)$ (see Corollaries 5.22 and 5.23).

Theorem 5.15. *Let m be an arbitrary positive integer. Assume that $(F_i)_{i=1}^{\infty}$ is a sequence of global function fields with full constant field $\mathbb{F}_q$ such that $g_i \to \infty$ as $i \to \infty$ and $\lim_{i\to\infty} \frac{n_i}{g_i} = \gamma > 2$. Here n_i and g_i are the number of rational places and the genus of F_i, respectively, for $i \geq 1$. For each $i \geq 1$, let M_i be a nonempty subset of $\mathbb{F}_q^{mn_i}$. We also assume that $(\log_q |M_i|)/n_i$ and $\hat{\rho}_{(2,\ldots,m+1)}(M_i)/n_i$ converge as $i \to \infty$, and moreover $\lim_{i\to\infty} \frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i} < 1 - \frac{2}{\gamma}$. Then for $0 < \delta \leq 1 - \frac{2}{\gamma} - \lim_{i\to\infty} \frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i}$ we have*

$$\alpha_q(\delta) \geq 1 - \delta - \frac{1}{\gamma} + \lim_{i\to\infty} \frac{\log_q |M_i|}{n_i} - \lim_{i\to\infty} \frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i}. \quad (5.21)$$

Proof. Throughout this proof, if not stated otherwise, i denotes a sufficiently large integer. Fix m and note that by the continuity of α_q (see Proposition 5.1(i)) we can assume that

$$0 < \delta < 1 - \frac{2}{\gamma} - \lim_{i\to\infty} \frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i}.$$

Let $\hat{r}_i$ be the real number satisfying

$$\delta = (m+1) - \frac{\hat{r}_i}{n_i} - \frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i}. \tag{5.22}$$

Note that as $\delta < 1 - 2\lim_{i\to\infty} \frac{g_i}{n_i} - \lim_{i\to\infty} \frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i}$ by assumption, we have

$$2g_i \le (1-\delta)n_i - \hat{\rho}_{(2,\ldots,m+1)}(M_i). \tag{5.23}$$

Putting $\hat{r}_i$ into (5.23), we get

$$2g_i \le -mn_i + \hat{r}_i. \tag{5.24}$$

Let $r_i = \lfloor \hat{r}_i \rfloor$. Then using (5.24), we obtain that

$$r_i \ge mn_i + 2g_i - 1 \tag{5.25}$$

and

$$\delta \ge (m+1) - \frac{r_i}{n_i} - \frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i} > \delta - \frac{1}{n_i}. \tag{5.26}$$

Recall that $\delta > 0$ and note that $\lim_{i\to\infty} n_i = \infty$. Therefore we have $\delta n_i > 1$. Using also (5.26), we obtain that the integer

$$(m+1)n_i - r_i - \hat{\rho}_{(2,\ldots,m+1)}(M_i) \ge 1. \tag{5.27}$$

Let $P_{i,1}, \ldots, P_{i,n_i}$ be distinct rational places of the global function field F_i. Moreover, using Lemma 5.6 we obtain a divisor G_i of F_i such that $\deg(G_i) = r_i$ and $\operatorname{supp}(G_i) \cap \{P_{i,1}, \ldots, P_{i,n_i}\} = \varnothing$.

Using G_i and $P_{i,1}, \ldots, P_{i,n_i}$ (corresponding to G and $P_1, \ldots, P_n$ of Theorem 5.12), we define the $\mathbb{F}_q$-linear maps Φ_i and ψ_i as in (5.7) and (5.8). It follows from (5.25) and Remark 5.13 that Φ_i is a surjective map.

Using the map Φ_i and the subset $M_i \subseteq \mathbb{F}_q^{mn_i}$, we define the subset N_i of $\mathcal{L}(G_i)$ as $N_i = \Phi_i^{-1}(M_i)$. Then we define the code $C_i \subseteq \mathbb{F}_q^{n_i}$ as $C_i = \psi_i(N_i)$.

We will prove that

$$\liminf_{i\to\infty} \frac{d(C_i)}{n_i} \ge \delta \tag{5.28}$$

and

$$\begin{aligned} &\lim_{i\to\infty} \frac{\log_q |C_i|}{n_i} \\ &= 1 - \delta - \frac{1}{\gamma} + \lim_{i\to\infty} \frac{\log_q |M_i|}{n_i} - \lim_{i\to\infty} \frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i}. \end{aligned} \tag{5.29}$$

In view of the fact that α_q is a nonincreasing function (see Proposition 5.1(i)), (5.28) and (5.29) will imply the bound (5.21).

First we prove (5.28). Recall that Φ_i is a surjective map. Note also that (5.27) holds. Therefore the conditions of Theorem 5.12 are satisfied. Using Theorem 5.12, we obtain that

$$\frac{d(C_i)}{n_i} \geq (m+1) - \frac{r_i}{n_i} - \frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i}. \tag{5.30}$$

By (5.22) and $r_i = \lfloor \hat{r}_i \rfloor$, we observe that

$$\delta = m + 1 - \lim_{i\to\infty} \frac{r_i}{n_i} - \lim_{i\to\infty} \frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i}, \tag{5.31}$$

and hence (5.30) implies (5.28).

It remains to prove (5.29). Using Theorem 5.12, (5.25), and Proposition 5.4, we obtain that

$$\frac{\log_q |C_i|}{n_i} = \frac{r_i - mn_i + 1 - g_i}{n_i} + \frac{\log_q |M_i|}{n_i}. \tag{5.32}$$

Hence (5.31) and (5.32) imply that

$$\begin{aligned}\lim_{i\to\infty} \frac{\log_q |C_i|}{n_i} &= \lim_{i\to\infty} \frac{r_i}{n_i} - m - \frac{1}{\gamma} + \lim_{i\to\infty} \frac{\log_q |M_i|}{n_i} \\ &= 1 - \delta - \frac{1}{\gamma} + \lim_{i\to\infty} \frac{\log_q |M_i|}{n_i} - \lim_{i\to\infty} \frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i}.\end{aligned}$$

This completes the proof. □

The condition $\delta \leq 1 - \frac{2}{\gamma} - \lim_{i\to\infty} \frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i}$ in Theorem 5.15 is quite restrictive. This follows from the condition that the map Φ in Theorem 5.12 is to be surjective. Now using an averaging argument, we extend Theorem 5.12 and Theorem 5.15 by relaxing these conditions. We note that although the averaging argument extends the range of δ to its full range, it uses a less constructive method (see Remark 5.18).

We assume that F is a global function field with full constant field $\mathbb{F}_q$. For $n \geq 1$, we assume that $P_1, \ldots, P_n$ are distinct rational places of F. Let G be a divisor of F such that $\deg(G) = r$ and $\operatorname{supp}(G) \cap \{P_1, \ldots, P_n\} = \varnothing$ (cf. Lemma 5.6). We define the $\mathbb{F}_q$-linear map Φ as in (5.7). Moreover, we assume that M is a nonempty subset of $\mathbb{F}_q^{mn}$ such that

$$(m+1)n - r - \hat{\rho}_{2,\ldots,m+1}(M) \geq 1. \tag{5.33}$$

For any $\boldsymbol{C} \in \mathbb{F}_q^{mn}$, let

$$M(\boldsymbol{C}) = \{\boldsymbol{A} + \boldsymbol{C} : \boldsymbol{A} \in M\}.$$

It is clear that

$$|M| = |M(\boldsymbol{C})| \tag{5.34}$$

and

$$\hat{\rho}_{(2,\ldots,m+1)}(M) = \hat{\rho}_{(2,\ldots,m+1)}(M(\boldsymbol{C})) \tag{5.35}$$

for any $\boldsymbol{C} \in \mathbb{F}_q^{mn}$. Using a standard averaging argument, we will obtain the existence of $M(\boldsymbol{C})$ with suitable properties. Let $\mathcal{S}$ be the subset of the cartesian product $\mathcal{L}(G) \times \mathbb{F}_q^{mn}$ defined by

$$\mathcal{S} = \left\{(f, \boldsymbol{C}) \in \mathcal{L}(G) \times \mathbb{F}_q^{mn} : f \in \mathcal{L}(G),\ \Phi(f) \in M(\boldsymbol{C})\right\}.$$

From (5.34) we obtain that

$$|\mathcal{S}| = |\mathcal{L}(G)| \cdot |M|.$$

For each $\boldsymbol{C} \in \mathbb{F}_q^{mn}$, let $N_{\boldsymbol{C}} \subseteq \mathcal{L}(G)$ and $\mathcal{S}_{\boldsymbol{C}} \subseteq \mathcal{S}$ be the subsets defined as

$$N_{\boldsymbol{C}} = \{f \in \mathcal{L}(G) : \Phi(f) \in M(\boldsymbol{C})\} \tag{5.36}$$

and

$$\mathcal{S}_{\boldsymbol{C}} = \{(f, \boldsymbol{C}) \in \mathcal{S} : f \in N_{\boldsymbol{C}}\}.$$

Note that

$$\mathcal{S} = \bigcup_{\boldsymbol{C} \in \mathbb{F}_q^{mn}} \mathcal{S}_{\boldsymbol{C}}$$

and for each $\boldsymbol{C} \in \mathbb{F}_q^{mn}$ we have

$$|\mathcal{S}_{\boldsymbol{C}}| = |N_{\boldsymbol{C}}|.$$

Hence there exists $\boldsymbol{C} \in \mathbb{F}_q^{mn}$ such that

$$|N_{\boldsymbol{C}}| = |\mathcal{S}_{\boldsymbol{C}}| \geq \frac{|\mathcal{S}|}{q^{mn}} = \frac{|\mathcal{L}(G)| \cdot |M|}{q^{mn}}. \tag{5.37}$$

For such $\boldsymbol{C} \in \mathbb{F}_q^{mn}$ we define the q-ary code C of length n by

$$C = \psi(N_{\boldsymbol{C}}), \tag{5.38}$$

where ψ is given by (5.8).

The following theorem extends Theorem 5.12.

Theorem 5.16. *Let m be an arbitrary positive integer. Assume that F is a global function field with full constant field $\mathbb{F}_q$ such that there exist at least $n \geq 1$ distinct rational places $P_1, \ldots, P_n$ of F. Let G be a divisor of F with $\deg(G) = r$ and $supp(G) \cap \{P_1, \ldots, P_n\} = \varnothing$. Let M be a subset of*

$\mathbb{F}_q^{mn}$ *such that* $|\mathcal{L}(G)| \cdot |M| > q^{mn}$ *and* $(m+1)n - r - \hat{\rho}_{(2,\ldots,m+1)}(M) \geq 1$. *Then* C *defined in* (5.38) *is an* $(n, |C|, d(C))$ *code over* $\mathbb{F}_q$ *with*

$$|C| \geq \left\lceil \frac{|\mathcal{L}(G)| \cdot |M|}{q^{mn}} \right\rceil \tag{5.39}$$

and

$$d(C) \geq (m+1)n - r - \hat{\rho}_{(2,\ldots,m+1)}(M). \tag{5.40}$$

Also C *is* $\mathbb{F}_q$*-linear if* M *is an* $\mathbb{F}_q$*-linear subset of* $\mathbb{F}_q^{mn}$.

Proof. Let $f, g \in N_{\boldsymbol{C}}$ with $f \neq g$. For $h \in \{f, g\}$, let $\boldsymbol{C}(h) \in \mathbb{F}_q^{(m+1)n}$, $\boldsymbol{A}(h) \in \mathbb{F}_q^{mn}$, and $\boldsymbol{b}(h) \in \mathbb{F}_q^n$ be as defined in the proof of Theorem 5.12. Note that

$$\boldsymbol{A}(f), \boldsymbol{A}(g) \in M(\boldsymbol{c})$$

by definition. Therefore we have

$$V_{(2,\ldots,m+1)}(\boldsymbol{A}(f) - \boldsymbol{A}(g)) \leq \hat{\rho}_{(2,\ldots,m+1)}(M(\boldsymbol{C})).$$

Using (5.35) and following the proof of Theorem 5.12, we obtain that

$$||\boldsymbol{b}(f) - \boldsymbol{b}(g)|| \geq (m+1)n - r - \hat{\rho}_{(2,\ldots,m+1)}(M),$$

which implies (5.40). As $(m+1)n - r - \hat{\rho}_{(2,\ldots,m+1)}(M) \geq 1$ by assumption, the restriction of ψ to $N_{\boldsymbol{C}}$ is one-to-one. Therefore $|C| = |N_{\boldsymbol{C}}|$ and (5.39) follows from (5.37). □

Now we are ready to extend Theorem 5.15 to the full range of δ.

Theorem 5.17. *Let* m *be an arbitrary positive integer. Assume that* $(F_i)_{i=1}^{\infty}$ *is a sequence of global function fields with full constant field* $\mathbb{F}_q$ *such that* $g_i \to \infty$ *as* $i \to \infty$ *and* $\lim_{i\to\infty} \frac{n_i}{g_i} = \gamma > 0$. *Here* n_i *and* g_i *are the number of rational places and the genus of* F_i*, respectively, for* $i \geq 1$. *For each* $i \geq 1$*, let* M_i *be a nonempty subset of* $\mathbb{F}_q^{mn_i}$. *We also assume that* $(\log_q |M_i|)/n_i$ *and* $\hat{\rho}_{(2,\ldots,m+1)}(M_i)/n_i$ *converge as* $i \to \infty$*, and*

$$1 - \frac{1}{\gamma} + \lim_{i\to\infty} \frac{\log_q |M_i|}{n_i} - \lim_{i\to\infty} \frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i} > 0. \tag{5.41}$$

For $0 < \delta \leq 1 - \frac{1}{\gamma} + \lim_{i\to\infty} \frac{\log_q |M_i|}{n_i} - \lim_{i\to\infty} \frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i}$ *we have*

$$\alpha_q(\delta) \geq 1 - \delta - \frac{1}{\gamma} + \lim_{i\to\infty} \frac{\log_q |M_i|}{n_i} - \lim_{i\to\infty} \frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i}. \tag{5.42}$$

Proof. We follow the proof of Theorem 5.15 and we assume that i is sufficiently large throughout this proof, if not stated otherwise. We can assume that

$$1-\delta-\frac{1}{\gamma}+\lim_{i\to\infty}\frac{\log_q|M_i|}{n_i}-\lim_{i\to\infty}\frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i}>0, \tag{5.43}$$

since otherwise the statement in (5.42) is obvious. Let $\hat{r}_i$ be the real number satisfying

$$\delta=(m+1)-\frac{\hat{r}_i}{n_i}-\frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i}, \tag{5.44}$$

and put $r_i=\lfloor \hat{r}_i \rfloor$. As $\delta>0$, following the arguments of the proof of Theorem 5.15 we obtain that

$$(m+1)n_i-r_i-\hat{\rho}_{(2,\ldots,m+1)}(M_i)\geq 1.$$

As in the proof of Theorem 5.15, let $P_{i,1},\ldots,P_{i,n_i}$ be distinct rational places of the global function field F_i. By Lemma 5.6 we obtain a divisor G_i of F_i such that $\deg(G_i)=r_i$ and $\operatorname{supp}(G_i)\cap\{P_{i,1},\ldots,P_{i,n_i}\}=\varnothing$. Similarly we define the $\mathbb{F}_q$-linear maps Φ_i and ψ_i from $\mathcal{L}(G_i)$ to $\mathbb{F}_q^{mn_i}$ and $\mathbb{F}_q^{n_i}$, respectively.

We obtain $\boldsymbol{C}_i\in\mathbb{F}_q^{mn_i}$ such that for $N_{\boldsymbol{C}_i}\subseteq\mathcal{L}(G_i)$ as given in (5.36) we have (cf. (5.37))

$$|N_{\boldsymbol{C}_i}|\geq\frac{|\mathcal{L}(G_i)|\cdot|M_i|}{q^{mn_i}}.$$

By Proposition 5.4 we have

$$\dim(\mathcal{L}(G_i))\geq r_i+1-g_i. \tag{5.45}$$

Then using (5.43) and (5.44) we get that

$$|\mathcal{L}(G_i)|\cdot|M_i|>q^{mn_i}.$$

Therefore the conditions of Theorem 5.16 are satisfied.

We define the code $C_i\subseteq\mathbb{F}_q^{n_i}$ as $\psi_i(N_{\boldsymbol{C}_i})$. Using Theorem 5.16 and (5.44), we obtain that

$$\liminf_{i\to\infty}\frac{d(C_i)}{n_i}\geq\delta. \tag{5.46}$$

From Theorem 5.16 we also obtain that

$$\liminf_{i\to\infty}\frac{\log_q|C_i|}{n_i}\geq\liminf_{i\to\infty}\frac{\dim(\mathcal{L}(G_i))}{n_i}+\lim_{i\to\infty}\frac{\log_q|M_i|}{n_i}-m. \tag{5.47}$$

Using (5.44) and (5.45), we get that

$$\begin{aligned}&\liminf_{i\to\infty}\frac{\dim(\mathcal{L}(G_i))}{n_i}\\&\geq m+1-\delta-\lim_{i\to\infty}\frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i}-\lim_{i\to\infty}\frac{g_i}{n_i}.\end{aligned}\tag{5.48}$$

Then from (5.47) and (5.48) we obtain that

$$\liminf_{i\to\infty}\frac{\log_q|C_i|}{n_i}\geq 1-\delta-\frac{1}{\gamma}+\lim_{i\to\infty}\frac{\log_q|M_i|}{n_i}-\lim_{i\to\infty}\frac{\hat{\rho}_{(2,\ldots,m+1)}(M_i)}{n_i}.$$

Using (5.46) and the fact that α_q is a nonincreasing function (see Proposition 5.1(i)), we complete the proof. □

Remark 5.18. The constructions of the sequences of codes in Theorem 5.15 and Theorem 5.17 are similar. However, while Theorem 5.15 uses the subset M_i in the sequence $(M_i)_{i=1}^{\infty}$ as it is given, Theorem 5.17 uses a nonconstructive existence argument for choosing $\boldsymbol{C}_i \in \mathbb{F}_q^{mn_i}$ in order to obtain the corresponding subset $M(\boldsymbol{C}_i)$. Therefore Theorem 5.17 is less constructive than Theorem 5.15.

We present a concrete example of a sequence $(M_i)_{i=1}^{\infty}$ of subsets as in Theorem 5.17 satisfying the required properties. First we give a definition.

Definition 5.19. Let $m \geq 1$ be an integer. For an integer $n \geq 1$ and

$$\boldsymbol{A}=\left(a_1^{(1)},\ldots,a_m^{(1)},a_1^{(2)},\ldots,a_m^{(2)},\ldots\ldots,a_1^{(n)},\ldots,a_m^{(n)}\right)\in\mathbb{F}_q^{mn},$$

let $I_m(\boldsymbol{A})$, $I_{m-1}(\boldsymbol{A}),\ldots,I_1(\boldsymbol{A})$ be the subsets of $\{1,\ldots,n\}$ defined by

$$\begin{aligned}I_m(\boldsymbol{A}) &= \left\{i\in\{1,\ldots,n\}: a_m^{(i)}\neq 0\right\},\\ I_{m-1}(\boldsymbol{A}) &= \left\{i\in\{1,\ldots,n\}: a_m^{(i)}=0,\ a_{m-1}^{(i)}\neq 0\right\},\\ &\vdots\\ I_1(\boldsymbol{A}) &= \left\{i\in\{1,\ldots,n\}: a_m^{(i)}=\cdots=a_2^{(i)}=0,\ a_1^{(i)}\neq 0\right\}.\end{aligned}$$

We obtain some properties of a subset of $\mathbb{F}_q^{mn}$ that we will use together with Theorem 5.17. Recall that the real-valued function E_q on the interval $[0,1]$ was defined in Definition 5.7.

Proposition 5.20. *Let $m \geq 1$ be an integer. Let Δ be the region in $\mathbb{R}^m$ consisting of the m-tuples $(x_1,\ldots,x_m)$ of real numbers such that*

$0 \le x_1, \dots, x_m$ *and* $x_1 + \cdots + x_m \le 1$. *For* $n \ge 1$ *and* $(x_1, \dots, x_m) \in \Delta$, *let* $S(n, x_1, \dots, x_m)$ *be the subset of* $\mathbb{F}_q^{mn}$ *defined by*

$$S(n, x_1, \dots, x_m) = \left\{ \boldsymbol{A} \in \mathbb{F}_q^{mn} : |I_1(\boldsymbol{A})| = \lfloor x_1 n \rfloor, \dots, |I_m(\boldsymbol{A})| = \lfloor x_m n \rfloor \right\}.$$

Then

$$\begin{aligned} &|S(n, x_1, \dots, x_m)| \\ &\quad = \binom{n}{\lfloor x_m n \rfloor} (q-1)^{\lfloor x_m n \rfloor} q^{(m-1)\lfloor x_m n \rfloor} \\ &\qquad \times \binom{n - \lfloor x_m n \rfloor}{\lfloor x_{m-1} n \rfloor} (q-1)^{\lfloor x_{m-1} n \rfloor} q^{(m-2)\lfloor x_{m-1} n \rfloor} \\ &\qquad \times \cdots \\ &\qquad \times \binom{n - (\lfloor x_m n \rfloor + \lfloor x_{m-1} n \rfloor + \cdots + \lfloor x_2 n \rfloor)}{\lfloor x_1 n \rfloor} (q-1)^{\lfloor x_1 n \rfloor} \end{aligned} \tag{5.49}$$

and

$$\begin{aligned} &\hat{\rho}_{(2,\dots,m+1)} \left(S(n, x_1, \dots, x_m) \right) \\ &\qquad \le 2 \left(2\lfloor x_1 n \rfloor + 3\lfloor x_2 n \rfloor + \cdots + (m+1)\lfloor x_m n \rfloor \right). \end{aligned} \tag{5.50}$$

Let $V(x_1, \dots, x_m)$ *be the real-valued function on* $\Delta \subseteq \mathbb{R}^m$ *defined by*

$$\begin{aligned} &V(x_1, \dots, x_m) \\ &= E_q(x_m) + \textstyle\sum_{j=1}^{m-1} (1 - x_m - x_{m-1} - \cdots - x_{j+1}) E_q \left(\frac{x_j}{1 - x_m - x_{m-1} - \cdots - x_{j+1}} \right) \\ &\quad + \left(\textstyle\sum_{j=1}^{m} x_j \right) \log_q (q-1) + \textstyle\sum_{j=2}^{m} (j-1) x_j - 2 \textstyle\sum_{j=1}^{m} (j+1) x_j. \end{aligned}$$

For $(x_1, \dots, x_m) \in \Delta$ *we have*

$$\begin{aligned} &\liminf_{n \to \infty} \left(\frac{\log_q |S(n, x_1, \dots, x_m)|}{n} - \frac{\hat{\rho}_{(2,\dots,m+1)} \left(S(n, x_1, \dots, x_m) \right)}{n} \right) \\ &\qquad \ge V(x_1, \dots, x_m). \end{aligned} \tag{5.51}$$

Moreover, $V(x_1, \dots, x_m)$ *attains its maximum at*

$$\begin{aligned} &(x_1, \dots, x_m) \\ &= \left(q^{m-1} \tfrac{q-1}{q^{m+3} + q^m - 1}, q^{m-2} \tfrac{q-1}{q^{m+3} + q^m - 1}, \dots, \tfrac{q-1}{q^{m+3} + q^m - 1} \right) \end{aligned} \tag{5.52}$$

and we have

$$V\left(q^{m-1}\frac{q-1}{q^{m+3}+q^m-1},\ldots,\frac{q-1}{q^{m+3}+q^m-1}\right) = \log_q\left(1+\tfrac{q^m-1}{q^{m+3}}\right). \tag{5.53}$$

Proof. The statements (5.49) and (5.50) follow from the definition of $S(n,x_1,\ldots,x_m)$. Using (5.49) and some manipulations, we obtain that

$$\lim_{n\to\infty}\frac{\log_q|S(n,x_1,\ldots,x_m)|}{n} = V(x_1,\ldots,x_m)+2\sum_{j=1}^{m}(j+1)x_j. \tag{5.54}$$

Then (5.51) follows from (5.50) and (5.54).

Let $1\le\ell\le m$ be an integer. For the partial derivative $\dfrac{\partial V(x_1,\ldots,x_m)}{\partial x_\ell}$ of $V(x_1,\ldots,x_m)$ we have

$$\frac{\partial V(x_1,\ldots,x_m)}{\partial x_\ell} = \log_q\left(\frac{1-(x_1+\cdots+x_m)}{x_\ell}\right)+\log_q(q-1)-(\ell+3). \tag{5.55}$$

The identity in (5.55) implies that the point defined in (5.52) is the unique critical point of $V(x_1,\ldots,x_m)$, and moreover it is easy to show that $V(x_1,\ldots,x_m)$ attains its maximum at the point given in (5.52). The result in (5.53) follows from an algebraic manipulation using the definition of $V(x_1,\ldots,x_m)$. □

Using Theorem 5.17 and Proposition 5.20, we obtain the following corollary.

Corollary 5.21. *Let $m\ge1$ be an integer. Assume that $(F_i)_{i=1}^{\infty}$ is a sequence of global function fields with full constant field $\mathbb{F}_q$ such that $g_i\to\infty$ as $i\to\infty$ and $\lim_{i\to\infty}\frac{n_i}{g_i}=\gamma>0$. Here n_i and g_i are the number of rational places and the genus of F_i, respectively, for $i\ge1$. Then we have*

$$\alpha_q(\delta)\ge1-\delta-\frac{1}{\gamma}+\log_q\left(1+\frac{q^m-1}{q^{m+3}}\right) \qquad \textit{for } 0\le\delta\le1.$$

Proof. We can assume that

$$1-\frac{1}{\gamma}+\log_q\left(1+\frac{q^m-1}{q^{m+3}}\right)>0$$

and

$$0<\delta\le1-\frac{1}{\gamma}+\log_q\left(1+\frac{q^m-1}{q^{m+3}}\right),$$

for otherwise the result is trivial. Let $(x_1, \ldots, x_m)$ be the m-tuple of rational numbers given in (5.52). Under the notations of Theorem 5.17 and Proposition 5.20, for each $i \geq 1$ let $M_i = S(n_i, x_1, \ldots, x_m) \subseteq \mathbb{F}_q^{mn_i}$. By passing, if necessary, to a subsequence, we see that the conditions of Theorem 5.17 are satisfied. We obtain the lower bound on $\alpha_q(\delta)$ using Theorem 5.17 and Proposition 5.20. □

Letting $m \to \infty$ in Corollary 5.21, we obtain the following result.

Corollary 5.22. *Assume that $(F_i)_{i=1}^{\infty}$ is a sequence of global function fields with full constant field $\mathbb{F}_q$ such that $g_i \to \infty$ as $i \to \infty$ and $\lim_{i\to\infty} \frac{n_i}{g_i} = \gamma > 0$. Here n_i and g_i are the number of rational places and the genus of F_i, respectively, for $i \geq 1$. Then we have*

$$\alpha_q(\delta) \geq 1 - \delta - \frac{1}{\gamma} + \log_q\left(1 + \frac{1}{q^3}\right) \qquad \text{for } 0 \leq \delta \leq 1.$$

Corollary 5.23. *For any prime power q, we have*

$$\alpha_q(\delta) \geq 1 - \delta - \frac{1}{A(q)} + \log_q\left(1 + \frac{1}{q^3}\right) \qquad \text{for } 0 \leq \delta \leq 1.$$

Proof. By the definition of $A(q)$ in Section 5.1, we can apply Corollary 5.22 with $\gamma = A(q)$. □

5.4. The Stichtenoth-Xing Construction

In [12] Stichtenoth and Xing gave a simpler construction of asymptotically good codes which yields the same lower bound on $\alpha_q(\delta)$ as Corollary 5.22, though on a shorter interval. In this section we present an exposition of their result.

Let F be a global function field with full constant field $\mathbb{F}_q$ such that there exist at least $n \geq 1$ distinct rational places $P_1, \ldots, P_n$ of F. Let G be a nonspecial divisor of F with $\deg(G) = r$, $\dim(\mathcal{L}(G)) \geq 1$, and $\operatorname{supp}(G) \cap \{P_1, \ldots, P_n\} = \varnothing$ (see Remark 5.5).

Let $1 \leq t \leq s$ be integers with $t \leq n$. Let $\mathcal{E}(s,t)$ be the set consisting of the divisors E of F such that

$$E = a_{i_1}P_{i_1} + a_{i_2}P_{i_2} + \cdots + a_{i_t}P_{i_t}, \tag{5.56}$$

where $1 \leq i_1 < i_2 < \cdots < i_t \leq n$ and $a_{i_1}, a_{i_2}, \ldots, a_{i_t}$ are positive integers with $a_{i_1} + a_{i_2} + \cdots + a_{i_t} = s$. It is easy to see that

$$|\mathcal{E}(s,t)| = \binom{n}{t}\binom{s-1}{t-1}. \tag{5.57}$$

For each $E \in \mathcal{E}(s,t)$ in the form (5.56), let $S(G,E)$ be the subset of F given by

$$S(G,E) = \{f \in \mathcal{L}(G+E) : \nu_{P_{i_1}}(f) = -a_{i_1}, \nu_{P_{i_2}}(f) = -a_{i_2}, \ldots, \nu_{P_{i_t}}(f) = -a_{i_t}\}.$$

Assume that $E_1, E_2 \in \mathcal{E}(s,t)$ are distinct divisors. Then

$$S(G,E_1) \cap S(G,E_2) = \varnothing. \tag{5.58}$$

Indeed, otherwise there exist integers $1 \le i \le n$ and $a \ne b$ such that $f \in S(G,E_1) \cap S(G,E_2)$ with $\nu_{P_i}(f) = -a$, $\nu_{P_i}(f) = -b$, which is a contradiction.

Lemma 5.24. *Under the notation and assumptions as above, for each $E \in \mathcal{E}(s,t)$ we have*

$$|S(G,E)| = q^{r+s-g+1}\left(1 - \frac{1}{q}\right)^t.$$

Proof. In this proof we follow the arguments of Maharaj in [6, Section 2]. Let $E = a_1P_1 + a_2P_2 + \cdots + a_tP_t$ for simplicity, where $a_1, a_2, \ldots, a_t$ are positive integers with $a_1 + a_2 + \cdots + a_t = s$. For integers $1 \le i \le t$ and $1 \le j \le a_i$, the fact that G is nonspecial (cf. Remark 5.5) implies that

$$\dim(\mathcal{L}(G + jP_i)) = \dim(\mathcal{L}(G + (j-1)P_i)) + 1,$$

and hence we can choose and fix $f_{i,j} \in \mathcal{L}(G + jP_i) \setminus \mathcal{L}(G + (j-1)P_i)$. Note that $\{f_{i,j} : 1 \le i \le t,\ 1 \le j \le a_i\}$ is a linearly independent set over $\mathbb{F}_q$. Moreover, $f \in S(G,E)$ if and only if there exist a uniquely determined $f_0 \in \mathcal{L}(G)$ and for $1 \le i \le t$ and $1 \le j \le a_i$ uniquely determined

$$\beta_{i,j} \in \begin{cases} \mathbb{F}_q & \text{if } 1 \le j < a_i, \\ \mathbb{F}_q \setminus \{0\} & \text{if } j = a_i, \end{cases}$$

such that $f = f_0 + \sum_{i=1}^{t}\sum_{j=1}^{a_i} \beta_{i,j} f_{i,j}$. Hence

$$|S(G,E)| = |\mathcal{L}(G)| q^s \left(\frac{q-1}{q}\right)^t = q^{r+1-g} q^s \left(1 - \frac{1}{q}\right)^t,$$

which is the desired result. □

Let $\mathcal{S}(G;s,t) \subseteq F$ be the set defined by

$$\mathcal{S}(G;s,t) = \bigcup_{E \in \mathcal{E}(s,t)} S(G,E).$$

Using Lemma 5.24, (5.57), and (5.58), we obtain that

$$\begin{aligned}&|\mathcal{S}(G;s,t)|\\&=\sum_{E\in\mathcal{E}(s,t)}|S(G,E)|=q^{r+s-g+1}\left(1-\frac{1}{q}\right)^t\binom{n}{t}\binom{s-1}{t-1}.\end{aligned}\tag{5.59}$$

Now we are ready to describe the basic construction of Stichtenoth and Xing [12]. For each $f\in\mathcal{S}(G;s,t)$, there exists a uniquely determined $E\in\mathcal{E}(s,t)$ such that $f\in S(G,E)$. If $P\in\{P_1,\dots,P_n\}\setminus\operatorname{supp}(E)$, then the evaluation $f^{(0)}(P)$ of f at P is a well-defined element of $\mathbb{F}_q$ (cf. Section 5.3). For $1\le i\le n$ and $f\in\mathcal{S}(G;s,t)$, put

$$\lambda_i(f)=\begin{cases}f^{(0)}(P_i) & \text{if } P_i\notin\operatorname{supp}(E),\\ 0 & \text{if } P_i\in\operatorname{supp}(E).\end{cases}$$

Then we define the map $\lambda:\mathcal{S}(G;s,t)\to\mathbb{F}_q^n$ as

$$\lambda(f)=(\lambda_1(f),\lambda_2(f),\dots,\lambda_n(f)).$$

Let C be the q-ary code of length n defined by

$$C=\lambda\left(\mathcal{S}(G;s,t)\right).\tag{5.60}$$

Theorem 5.25. *Assume that F is a global function field with full constant field $\mathbb{F}_q$ such that there exist at least $n\ge 1$ distinct rational places $P_1,\dots,P_n$ of F. Let G be a nonspecial divisor of F such that* $\deg(G)=r$, $\dim(\mathcal{L}(G))\ge 1$, *and $\operatorname{supp}(G)\cap\{P_1,\dots,P_n\}=\varnothing$. Let $1\le t\le s$ be integers with $t\le n$ and $n-r-2s-2t\ge 1$. Then C defined in* (5.60) *is an $(n,|C|,d(C))$ code over $\mathbb{F}_q$ with*

$$|C|=q^{r+s-g+1}\left(1-\frac{1}{q}\right)^t\binom{n}{t}\binom{s-1}{t-1}$$

and

$$d(C)\ge n-r-2s-2t.$$

Proof. Let $f_1,f_2\in\mathcal{S}(G;s,t)$ with $f_1\ne f_2$. Let $E_1,E_2\in\mathcal{E}(s,t)$ be the divisors corresponding to f_1 and f_2, respectively. Let $U=\{P_1,\dots,P_n\}\setminus(\operatorname{supp}(E_1)\cup\operatorname{supp}(E_2))$. Since $|\operatorname{supp}(E_1)|=|\operatorname{supp}(E_2)|=t$, we have

$$|U|\ge n-2t.\tag{5.61}$$

Let $f=f_1-f_2$ and $Z=\{P\in U: f^{(0)}(P)=0\}$. Then $f\in\mathcal{L}\left(G+E_1+E_2-\sum_{P\in Z}P\right)\setminus\{0\}$ and hence by (5.6),

$$\deg\left(G+E_1+E_2-\sum_{P\in Z}P\right)=r+2s-|Z|\ge 0.\tag{5.62}$$

Using (5.61) and (5.62), we obtain that

$$|U \setminus Z| = |U| - |Z| \geq n - r - 2s - 2t,$$

which implies for the Hamming weight $||\lambda(f_1) - \lambda(f_2)||$ of $\lambda(f_1) - \lambda(f_2) \in \mathbb{F}_q^n$ that

$$||\lambda(f_1) - \lambda(f_2)|| \geq n - r - 2s - 2t.$$

As $n - r - 2s - 2t \geq 1$, we obtain that the map λ is one-to-one and $|C| = |\mathcal{S}(G; s, t)|$. We complete the proof using (5.59). □

Theorem 5.26. *Assume that $(F_i)_{i=1}^{\infty}$ is a sequence of global function fields with full constant field $\mathbb{F}_q$ such that $g_i \to \infty$ as $i \to \infty$ and $\lim_{i\to\infty} \frac{n_i}{g_i} = \gamma > 0$. Here n_i and g_i are the number of rational places and the genus of F_i, respectively, for $i \geq 1$. Let $0 \leq x \leq y \leq 1$ be real numbers. We also assume that $1 - \frac{2}{\gamma} - 2y - 2x > 0$. For $0 < \delta \leq 1 - \frac{2}{\gamma} - 2y - 2x$ we have*

$$\alpha_q(\delta) \geq 1 - \delta - \frac{1}{\gamma} - y - 2x + x \log_q \left(1 - \frac{1}{q}\right) + E_q(x) + yE_q\left(\frac{x}{y}\right).$$

Proof. As in the proofs of Theorems 5.15 and 5.17, we assume that i is a sufficiently large integer throughout this proof. By the continuity of α_q (see Proposition 5.1(i)) we can assume that $\delta < 1 - \frac{2}{\gamma} - 2y - 2x$.

Let $\hat{r}_i$ be the real number satisfying

$$\delta = 1 - \frac{\hat{r}_i}{n_i} - 2y - 2x. \tag{5.63}$$

Note that as $\delta < 1 - \frac{2}{\gamma} - 2y - 2x$ and $\lim_{i\to\infty} \frac{n_i}{g_i} = \gamma$ by assumption, we have $\hat{r}_i \geq 2g_i$. Putting $r_i = \lfloor \hat{r}_i \rfloor$ we obtain that

$$r_i \geq 2g_i. \tag{5.64}$$

Let $P_{i,1}, \ldots, P_{i,n_i}$ be distinct rational places of F_i. Let G_i be a divisor of F_i such that $\deg(G_i) = r_i$ and $\operatorname{supp}(G_i) \cap \{P_{i,1}, \ldots, P_{i,n_i}\} = \varnothing$ (cf. Lemma 5.6). Then G_i is a nonspecial divisor of F_i (cf. Proposition 5.4 and Remark 5.5) and using (5.64) we obtain that $\dim(\mathcal{L}(G_i)) = r_i + 1 - g_i \geq g_i + 1 \geq 1$.

Let $s_i = \lfloor yn_i \rfloor$ and $t_i = \lfloor xn_i \rfloor$. As $\delta > 0$ we obtain that $n_i - r_i - 2s_i - 2t_i \geq 1$. Therefore using Theorem 5.25 for the global function field F_i, we obtain a q-ary code C_i of length n_i such that

$$\begin{aligned} &\log_q |C_i| \\ &= r_i + s_i - g_i + 1 + t_i \log_q \left(1 - \frac{1}{q}\right) + \log_q \binom{n_i}{t_i} + \log_q \binom{s_i - 1}{t_i - 1} \end{aligned} \tag{5.65}$$

and

$$d(C_i) \geq n_i - r_i - 2s_i - 2t_i. \tag{5.66}$$

Using (5.63) and (5.66), we obtain that

$$\liminf_{i\to\infty} \frac{d(C_i)}{n_i} \geq \delta.$$

Using Definition 5.7, (5.63), and (5.65), we get

$$\begin{aligned}&\lim_{i\to\infty} \frac{\log_q |C_i|}{n_i}\\ &= (1-\delta-2y-2x) + y - \frac{1}{\gamma} - x\log_q\left(1-\frac{1}{q}\right) + E_q(x) + yE_q\left(\frac{x}{y}\right).\end{aligned}$$

This completes the proof. □

Let Δ be the region in $\mathbb{R}^2$ consisting of the points (x, y) such that $0 \leq x \leq y \leq 1$. Let $V(x, y)$ be the real-valued function on Δ defined by

$$V(x,y) = -y - 2x + x\log_q\left(1-\frac{1}{q}\right) + E_q(x) + yE_q\left(\frac{x}{y}\right).$$

Then for the partial derivatives $\frac{\partial V(x,y)}{\partial x}$ and $\frac{\partial V(x,y)}{\partial y}$ we have

$$\begin{aligned}&\frac{\partial V(x,y)}{\partial x}\\ &= -2 + \log_q\left(1-\frac{1}{q}\right) + \log_q(1-x) + \log_q(y-x) - 2\log_q x\end{aligned} \tag{5.67}$$

and

$$\frac{\partial V(x,y)}{\partial y} = -1 + \log_q y - \log_q(y-x). \tag{5.68}$$

Using (5.67) and (5.68), we obtain that $V(x, y)$ attains its maximum on Δ at

$$(x,y) = \left(\frac{1}{q^3+1}, \frac{q}{(q-1)(q^3+1)}\right).$$

By straightforward manipulations we get

$$V\left(\frac{1}{q^3+1}, \frac{q}{(q-1)(q^3+1)}\right) = \log_q\left(1+\frac{1}{q^3}\right).$$

Therefore using $x = \frac{1}{q^3+1}$ and $y = \frac{q}{(q-1)(q^3+1)}$ in Theorem 5.26, we obtain the following corollary.

Corollary 5.27. *Assume that $(F_i)_{i=1}^{\infty}$ is a sequence of global function fields with full constant field $\mathbb{F}_q$ such that $g_i \to \infty$ as $i \to \infty$ and $\lim_{i\to\infty} \frac{n_i}{g_i} =$*

$\gamma > 0$. Here n_i and g_i are the number of rational places and the genus of F_i, respectively, for $i \geq 1$. Then for

$$0 < \delta \leq 1 - \frac{2}{\gamma} - \frac{4q-2}{(q-1)(q^3+1)}$$

we have

$$\alpha_q(\delta) \geq 1 - \delta - \frac{1}{\gamma} + \log_q\left(1 + \frac{1}{q^3}\right).$$

Proof. Note that for $x = \frac{1}{q^3+1}$ and $y = \frac{q}{(q-1)(q^3+1)}$, the upper bound $\delta \leq 1 - \frac{2}{\gamma} - 2y - 2x$ in Theorem 5.26 becomes $\delta \leq 1 - \frac{2}{\gamma} - \frac{4q-2}{(q-1)(q^3+1)}$. Then the corollary follows immediately from Theorem 5.26 and the fact that $V\left(\frac{1}{q^3+1}, \frac{q}{(q-1)(q^3+1)}\right) = \log_q\left(1 + \frac{1}{q^3}\right)$. □

5.5. Improved Bounds Using Distinguished Divisors

We recall that Vlăduţ [16] (see also [14, Chapter 3.4]) and Xing [17] improved the Tsfasman-Vlăduţ-Zink bound (5.4) for $\alpha_q^{\mathrm{lin}}(\delta)$, although not uniformly in δ. Similarly, Niederreiter and Özbudak [8] and Maharaj [6] improved the bound (5.5) for certain values of q and δ. Later, Niederreiter and Özbudak [9] refined and complemented the methods of [8] and they further improved the previous bounds on $\alpha_q^{\mathrm{lin}}(\delta)$ and $\alpha_q(\delta)$, including the bound in [6], for certain values of q and δ. In this section we give an exposition of the results of Niederreiter and Özbudak in [9].

We fix some notation as in Section 5.3. We choose an arbitrary positive integer m. Let F be a global function field with full constant field $\mathbb{F}_q$ such that there exist at least $n \geq 1$ distinct rational places $P_1, \ldots, P_n$ of F. Let t_i be a local parameter of F at P_i for $1 \leq i \leq n$. Moreover, we assume that the class number of F is h (see, for example, [11, Section V.1]).

We introduce and recall some definitions.

Definition 5.28. For a positive divisor D of F, let $\overline{D}$ be the divisor

$$\overline{D} = a_1 P_1 + \cdots + a_n P_n,$$

where $a_i = \min(m+1, \nu_{P_i}(D))$ for $1 \leq i \leq n$.

Definition 5.29. For a positive divisor D of F, let

$$\begin{aligned} \jmath_0(D) &= |\{i \in \{1,\dots,n\} : \nu_{P_i}(D) = m\}|, \\ \jmath_1(D) &= |\{i \in \{1,\dots,n\} : \nu_{P_i}(D) = m-1\}|, \\ &\vdots \\ \jmath_m(D) &= |\{i \in \{1,\dots,n\} : \nu_{P_i}(D) = 0\}|. \end{aligned}$$

Moreover, we define

$$J_m(D) = 2\jmath_1(D) + 3\jmath_2(D) + \cdots + (m+1)\jmath_m(D). \tag{5.69}$$

Let G be a divisor of F with $\mathrm{supp}(G) \cap \{P_1,\dots,P_n\} = \varnothing$ and $\dim(\mathcal{L}(G)) \geq 1$. Using the local expansions at $P_1,\dots,P_n$ and the divisor G, we define the $\mathbb{F}_q$-linear maps Φ and ψ from $\mathcal{L}(G)$ to $\mathbb{F}_q^{mn}$ and $\mathbb{F}_q^n$ as in (5.7) and (5.8), respectively.

Recall that for $\boldsymbol{A} \in \mathbb{F}_q^{mn}$, the subsets $I_m(\boldsymbol{A}), I_{m-1}(\boldsymbol{A}),\dots,I_1(\boldsymbol{A})$ of $\{1,\dots,n\}$ are defined in Definition 5.19.

Lemma 5.30. *For a divisor G of F with $\mathrm{supp}(G) \cap \{P_1,\dots,P_n\} = \varnothing$ and $\dim(\mathcal{L}(G)) \geq 1$, let $f \in \mathcal{L}(G) \setminus \{0\}$. Moreover, let $E = (f)_0$ be the zero divisor of f and $\boldsymbol{A} := \Phi(f) \in \mathbb{F}_q^{mn}$. Then*

$$\jmath_1(E) = |I_1(\boldsymbol{A})|,\ \jmath_2(E) = |I_2(\boldsymbol{A})|,\ \dots\ ,\ \jmath_m(E) = |I_m(\boldsymbol{A})|,$$

and

$$J_m(E) = 2\,|I_1(\boldsymbol{A})| + 3\,|I_2(\boldsymbol{A})| + \cdots + (m+1)\,|I_m(\boldsymbol{A})|.$$

Proof. For each $1 \leq i \leq n$ and $1 \leq \ell \leq m$, using Definition 5.19 we obtain that $i \in I_\ell(\boldsymbol{A}) \iff \nu_{P_i}(E) = m - \ell$. Hence by Definition 5.29 we have

$$\jmath_m(E) = |I_m(\boldsymbol{A})|,\ \jmath_{m-1}(E) = |I_{m-1}(\boldsymbol{A})|,\ \dots\ ,\ \jmath_1(E) = |I_1(\boldsymbol{A})|.$$

Using (5.69) we complete the proof. □

Next we define an important set of positive divisors.

Definition 5.31. For integers $r \geq s \geq 0$ and nonnegative integers $X_1, X_2, \dots, X_m$, let $\mathcal{V}_m(r, s; X_1, X_2, \dots, X_m)$ be the set consisting of the positive divisors D of the global function field F satisfying all of the following:

- Condition 1: $\deg(D) = r$ and $\deg\left(\overline{D}\right) \geq s$;

- Condition 2:

$$\begin{aligned} \jmath_m(D) &\le 2X_m, \\ \jmath_{m-1}(D) &\le 2X_{m-1} + X_m, \\ \jmath_{m-2}(D) &\le 2X_{m-2} + (X_{m-1} + X_m), \\ &\vdots \\ \jmath_1(D) &\le 2X_1 + (X_2 + X_3 + \cdots + X_m); \end{aligned}$$

- Condition 3: $J_m(D) \le 2\,(2X_1 + 3X_2 + \cdots + (m+1)X_m)$.

In the following proposition, which can be viewed as a refinement of Lemma 5.6, we use the Weak Approximation Theorem [11, Theorem I.3.1].

Proposition 5.32. *For integers $r \ge s \ge 0$ and nonnegative integers $X_1, \ldots, X_m$, if*

$$|\mathcal{V}_m(r, s; X_1, \ldots, X_m)| < h,$$

then there exists a divisor G of F of degree r such that $supp(G) \cap \{P_1, \ldots, P_n\} = \varnothing$ and for each $f \in \mathcal{L}(G) \setminus \{0\}$, if $E = (f)_0$ satisfies Conditions 2 and 3 of Definition 5.31 with the given $X_1, \ldots, X_m$, then $\deg\left(\overline{E}\right) \le s - 1$.

Proof. As $|\mathcal{V}_m(r, s; X_1, \ldots, X_m)| < h$, there exists a degree r divisor G of F such that G is nonequivalent to V for any $V \in \mathcal{V}_m(r, s; X_1, \ldots, X_m)$. Using the Weak Approximation Theorem, we can assume that $\mathrm{supp}(G) \cap \{P_1, \ldots, P_n\} = \varnothing$ without loss of generality (compare with [8, proof of Corollary 2.2]). Let $f \in \mathcal{L}(G) \setminus \{0\}$, $D = G + (f)$, and $E = (f)_0$. Since $\mathrm{supp}(G) \cap \{P_1, \ldots, P_n\} = \varnothing$ and D is positive, we have $\overline{D} = \overline{E}$. Assume that Conditions 2 and 3 of Definition 5.31 are satisfied by E. If $\deg\left(\overline{E}\right) \ge s$, then $D \in \mathcal{V}_m(r, s; X_1, \ldots, X_m)$ and hence D is nonequivalent to G, which is a contradiction. Thus, we must have $\deg\left(\overline{E}\right) \le s - 1$. □

The following two lemmas are useful for the basic construction of this section. They are technical lemmas and we refer to [9] for their proofs.

Lemma 5.33. *For $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{F}_q^{mn}$, we have*

$$\begin{aligned} &2\,|I_1(\boldsymbol{A} - \boldsymbol{B})| + 3\,|I_2(\boldsymbol{A} - \boldsymbol{B})| + \cdots + (m+1)\,|I_m(\boldsymbol{A} - \boldsymbol{B})| \\ &\le 2\,|I_1(\boldsymbol{A})| + 3\,|I_2(\boldsymbol{A})| + \cdots + (m+1)\,|I_m(\boldsymbol{A})| \\ &\quad + 2\,|I_1(\boldsymbol{B})| + 3\,|I_2(\boldsymbol{B})| + \cdots + (m+1)\,|I_m(\boldsymbol{B})|. \end{aligned}$$

Lemma 5.34. *For $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{F}_q^{mn}$, we have the following containment relations:*

$$\begin{aligned}
I_m(\boldsymbol{A}-\boldsymbol{B}) &\subseteq I_m(\boldsymbol{A}) \cup I_m(\boldsymbol{B}),\\
I_{m-1}(\boldsymbol{A}-\boldsymbol{B}) &\subseteq I_{m-1}(\boldsymbol{A}) \cup I_{m-1}(\boldsymbol{B}) \cup \{I_m(\boldsymbol{A}) \cap I_m(\boldsymbol{B})\},\\
I_{m-2}(\boldsymbol{A}-\boldsymbol{B}) &\subseteq I_{m-2}(\boldsymbol{A}) \cup I_{m-2}(\boldsymbol{B}) \cup \{I_{m-1}(\boldsymbol{A}) \cap I_{m-1}(\boldsymbol{B})\}\\
&\qquad\qquad \cup \{I_m(\boldsymbol{A}) \cap I_m(\boldsymbol{B})\},\\
&\vdots\\
I_1(\boldsymbol{A}-\boldsymbol{B}) &\subseteq I_1(\boldsymbol{A}) \cup I_1(\boldsymbol{B}) \cup \textstyle\bigcup_{2\le\nu\le m} \{I_\nu(\boldsymbol{A}) \cap I_\nu(\boldsymbol{B})\}.
\end{aligned}$$

For $\boldsymbol{C} \in \mathbb{F}_q^{mn}$ and nonnegative real numbers $x_1, \dots, x_m$ with $x_1 + \cdots + x_m \le 1$, let $M(x_1, \dots, x_m; \boldsymbol{C})$ be the subset of $\mathbb{F}_q^{mn}$ defined by

$$\begin{aligned}
&M(x_1, \dots, x_m; \boldsymbol{C})\\
&= \left\{\boldsymbol{A} \in \mathbb{F}_q^{mn} : |I_1(\boldsymbol{A}-\boldsymbol{C})| \le \lfloor x_1 n \rfloor, \dots, |I_m(\boldsymbol{A}-\boldsymbol{C})| \le \lfloor x_m n \rfloor\right\}.
\end{aligned}$$

Note that $|M(x_1, \dots, x_m; \boldsymbol{C})| = |M(x_1, \dots, x_m; \boldsymbol{0})|$ for each $\boldsymbol{C} \in \mathbb{F}_q^{mn}$ and

$$|M(x_1, \dots, x_m; \boldsymbol{0})| \ge |S(n, x_1, \dots, x_n)|, \tag{5.70}$$

where $S(n, x_1, \dots, x_n)$ is the subset of $\mathbb{F}_q^{mn}$ defined in Proposition 5.20.

Now we are ready to give the basic construction of this section. Assume that $r \ge s \ge 0$ are integers and $x_1, \dots, x_m \ge 0$ are real numbers with $x_1 + \cdots + x_m \le 1$ such that

$$|\mathcal{V}_m(r, s; \lfloor x_1 n \rfloor, \lfloor x_2 n \rfloor, \dots, \lfloor x_m n \rfloor)| < h. \tag{5.71}$$

Let G be a divisor of F of degree r obtained using (5.71) and Proposition 5.32. If

$$|\mathcal{L}(G)| \cdot |M(x_1, \dots, x_m; \boldsymbol{0})| > q^{mn}, \tag{5.72}$$

then there exists $\boldsymbol{C} \in \mathbb{F}_q^{mn}$ such that for the set

$$N_{\boldsymbol{C}} := \{f \in \mathcal{L}(G) : \Phi(f) \in M(x_1, \dots, x_m; \boldsymbol{C})\} \tag{5.73}$$

we have

$$|N_{\boldsymbol{C}}| \ge \frac{|\mathcal{L}(G)| \cdot |M(x_1, \dots, x_m; \boldsymbol{0})|}{q^{mn}} > 1. \tag{5.74}$$

Theorem 5.35. *Assume that $r \ge s \ge 0$ are integers and that $x_1, \dots, x_m$ are nonnegative real numbers with $x_1 + \cdots + x_m \le 1$ satisfying* (5.71)*. Let*

G be a divisor of F of degree r obtained using (5.71) *and Proposition* 5.32. *Assume also that* (5.72) *holds and that*

$$(m+1)n \geq s + 2\sum_{l=1}^{m}(l+1)\lfloor x_l n \rfloor. \tag{5.75}$$

Using the chosen divisor G and (5.72), *let $\boldsymbol{C} \in \mathbb{F}_q^{mn}$ be such that the set $N_{\boldsymbol{C}}$ satisfies* (5.74). *Let C be the q-ary code of length n given by $C = \psi(N_{\boldsymbol{C}})$. Then for the cardinality $|C|$ of C we have*

$$|C| \geq \left\lceil \frac{\mathcal{L}(G) \cdot |M(x_1, \ldots, x_m; \mathbf{0})|}{q^{mn}} \right\rceil$$

and for the minimum distance $d(C)$ of C we have

$$d(C) \geq (m+1)n + 1 - s - 2\sum_{l=1}^{m}(l+1)\lfloor x_l n \rfloor.$$

Proof. Let $f_1, f_2 \in N_{\boldsymbol{C}}$ be such that $f_1 \neq f_2$ and put $f = f_1 - f_2 \in \mathcal{L}(G)$. Let E be the zero divisor of f and

$$\overline{E} = a_1 P_1 + \cdots + a_n P_n$$

be the divisor defined in Definition 5.28. Let $\Phi(f_1) =: \boldsymbol{A}$ and $\Phi(f_2) =: \boldsymbol{B}$. We have

$$\Phi(f) = \boldsymbol{A} - \boldsymbol{B}. \tag{5.76}$$

As $\boldsymbol{A}, \boldsymbol{B} \in M(x_1, \ldots, x_m; \boldsymbol{C})$, we also have

$$|I_i(\boldsymbol{A} - \boldsymbol{C})| \leq \lfloor x_i n \rfloor \quad \text{and} \quad |I_i(\boldsymbol{B} - \boldsymbol{C})| \leq \lfloor x_i n \rfloor \quad \text{for } 1 \leq i \leq n. \tag{5.77}$$

Using (5.76), (5.77), Lemmas 5.30 and 5.33, we obtain that

$$\begin{aligned} J_m(E) &= 2\,|I_1(\boldsymbol{A} - \boldsymbol{B})| + 3\,|I_2(\boldsymbol{A} - \boldsymbol{B})| + \cdots + (m+1)\,|I_m(\boldsymbol{A} - \boldsymbol{B})| \\ &\leq 2\,|I_1(\boldsymbol{A} - \boldsymbol{C})| + 3\,|I_2(\boldsymbol{A} - \boldsymbol{C})| + \cdots + (m+1)\,|I_m(\boldsymbol{A} - \boldsymbol{C})| \\ &\quad + 2\,|I_1(\boldsymbol{B} - \boldsymbol{C})| + 3\,|I_2(\boldsymbol{B} - \boldsymbol{C})| + \cdots + (m+1)\,|I_m(\boldsymbol{B} - \boldsymbol{C})| \\ &\leq 2\,(2\lfloor x_1 n \rfloor + 3\lfloor x_2 n \rfloor + \cdots + (m+1)\lfloor x_m n \rfloor). \end{aligned}$$

Moreover, using (5.76), (5.77), Lemmas 5.30 and 5.34, we further obtain that

$$\begin{aligned}\jmath_m(E) &= |I_m((\boldsymbol{A}-\boldsymbol{C})-(\boldsymbol{B}-\boldsymbol{C}))|\\ &\le |I_m(\boldsymbol{A}-\boldsymbol{C})|+|I_m(\boldsymbol{B}-\boldsymbol{C})|\\ &\le 2\lfloor x_m n\rfloor,\end{aligned}$$

$$\begin{aligned}\jmath_{m-1}(E) &= |I_{m-1}((\boldsymbol{A}-\boldsymbol{C})-(\boldsymbol{B}-\boldsymbol{C}))|\\ &\le |I_{m-1}(\boldsymbol{A}-\boldsymbol{C})|+|I_{m-1}(\boldsymbol{B}-\boldsymbol{C})|+|I_m(\boldsymbol{A}-\boldsymbol{C})\cap I_m(\boldsymbol{B}-\boldsymbol{C})|\\ &\le 2\lfloor x_{m-1} n\rfloor+\lfloor x_m n\rfloor,\end{aligned}$$

$$\vdots$$

$$\begin{aligned}\jmath_1(E) &= |I_1((\boldsymbol{A}-\boldsymbol{C})-(\boldsymbol{B}-\boldsymbol{C}))|\\ &\le |I_1(\boldsymbol{A}-\boldsymbol{C})|+|I_1(\boldsymbol{B}-\boldsymbol{C})|+\textstyle\sum_{\nu=2}^{m}|I_\nu(\boldsymbol{A}-\boldsymbol{C})\cap I_\nu(\boldsymbol{B}-\boldsymbol{C})|\\ &\le 2\lfloor x_1 n\rfloor+\textstyle\sum_{\nu=2}^{m}\lfloor x_\nu n\rfloor.\end{aligned}$$

Hence by the choice of the divisor G (cf. Proposition 5.32), we have

$$\deg\left(\overline{E}\right)\le s-1. \tag{5.78}$$

Moreover, we obtain

$$\begin{aligned}&\sum_{i=1}^{n}(m+1-a_i)\\ &=(m+1)n-\sum_{i=1}^{n}a_i=(m+1)n-\deg\left(\overline{E}\right)\ge(m+1)n-s+1,\end{aligned}$$

where we used (5.78). Let $||\psi(f)||$ denote the Hamming weight of the vector $\psi(f)\in\mathbb{F}_q^n$. Then using Definition 5.29 and (5.69), we have

$$\begin{aligned}\sum_{i=1}^{n}(m+1-a_i)&=\sum_{\substack{i=1\\0\le a_i\le m}}^{n}(m+1-a_i)\le||\psi(f)||+\sum_{\substack{i=1\\0\le a_i\le m-1}}^{n}(m+1-a_i)\\ &=||\psi(f)||+J_m(E).\end{aligned}$$

Therefore we obtain

$$\begin{aligned}||\psi(f)||&\ge(m+1)n-s+1-J_m(E)\\ &\ge(m+1)n-s+1-2\left(2\lfloor x_1 n\rfloor+3\lfloor x_2 n\rfloor+\cdots+(m+1)\lfloor x_m n\rfloor\right).\end{aligned}$$

This yields the desired lower bound on $d(C)$. Using (5.75) we obtain that $d(C)\ge 1$, and so the map ψ is one-to-one on $N_{\boldsymbol{C}}$. Therefore $|C|=|N_{\boldsymbol{C}}|$, and hence the lower bound on $|C|$ follows from (5.74). This completes the proof. □

A special case of Theorem 5.35 gives linear codes.

Corollary 5.36. *Assume that $r \geq s \geq 0$ are integers and that $x_1 = x_2 = \cdots = x_m = 0$ satisfy* (5.71)*. Let G be a divisor of F of degree r obtained using* (5.71) *and Proposition* 5.32*. Assume also that*

$$|\mathcal{L}(G)| > q^{mn} \tag{5.79}$$

and that $(m+1)n \geq s$. Using the chosen divisor G and the kernel of the corresponding map Φ, put $C = \psi(\operatorname{Ker} \Phi)$. Then C is a linear code over $\mathbb{F}_q$ of length n. Moreover, for the dimension of C we have

$$\dim(C) \geq \dim(\mathcal{L}(G)) - mn$$

and for the minimum distance $d(C)$ of C we have

$$d(C) \geq (m+1)n + 1 - s.$$

Proof. The kernel of Φ is an $\mathbb{F}_q$-linear subspace of $\mathcal{L}(G)$ and is the Riemann-Roch space given by

$$\operatorname{Ker} \Phi = \mathcal{L}(G - m(P_1 + \cdots + P_n)).$$

As $\dim(\mathcal{L}(G - m(P_1 + \cdots + P_n))) \geq \dim(\mathcal{L}(G)) - mn$, using (5.79) we obtain that $\operatorname{Ker} \Phi \neq \{0\}$. The maps Φ and ψ are $\mathbb{F}_q$-linear, and hence C is a linear code over $\mathbb{F}_q$. We obtain the bounds on the dimension and the minimum distance of C using similar methods as in the proof of Theorem 5.35. □

For the asymptotic bounds we need the following assumption.

Assumption 1 Assume that $(F_i)_{i=1}^{\infty}$ is a sequence of global function fields with full constant field $\mathbb{F}_q$, with $g_i \to \infty$ as $i \to \infty$, and with $\limsup_{i\to\infty} \frac{n_i}{g_i} = \gamma > 0$, where n_i and g_i denote the number of rational places and the genus of F_i, respectively. For each $l \geq 1$, let $\gamma_l \geq 0$ be a real number with $\liminf_{i\to\infty} \frac{B_{i,l}}{g_i} \geq \gamma_l$, where $B_{i,l}$ is the number of degree l places of F_i. Using a suitable subsequence of $(F_i)_{i=1}^{\infty}$, we can take $\gamma_1 = \gamma$.

In the following definitions (Definitions 5.37 and 5.39), the following proposition (Proposition 5.38), and the following lemma (Lemma 5.40), we obtain an upper bound on the cardinality of some sets of the form $\mathcal{V}_m(r, s; X_1, \ldots, X_m)$. This bound will be useful for the asymptotic version of Theorem 5.35. Since the proofs of Proposition 5.38 and Lemma 5.40 are

rather technical, we prefer not to include them here and we refer to [9] for these proofs.

Definition 5.37. Under Assumption 1, let $y > 0$, $x_1, x_2, \ldots, x_m, \sigma \geq 0$ be real numbers satisfying

$$y + 2(2x_1 + 3x_2 + \cdots + (m+1)x_m) + \frac{\sigma}{\gamma} < 1. \tag{5.80}$$

For real numbers $0 \leq x \leq \frac{\sigma}{\gamma}$ and $t_1, t_2, \ldots, t_m$ satisfying

$$\begin{array}{l} 0 \leq t_m \leq 2x_m,\ 0 \leq t_{m-1} \leq 2x_{m-1} + x_m, \ldots, \\ 0 \leq t_1 \leq 2x_1 + (x_2 + x_3 + \cdots + x_m), \end{array} \tag{5.81}$$

and

$$2t_1 + 3t_2 + \cdots + (m+1)t_m \leq 2(2x_1 + 3x_2 + \cdots + (m+1)x_m), \tag{5.82}$$

let $S(\sigma, y, x, t_1, t_2, \ldots, t_m)$ be the real-valued function

$$\begin{array}{l} S(\sigma, y, x, t_1, t_2, \ldots, t_m) \\ \\ = E_q(t_m) + (1 - t_m) E_q\left(\frac{t_{m-1}}{1-t_m}\right) \\ + \cdots + (1 - (t_2 + \cdots + t_m))\, E_q\left(\frac{t_1}{1-(t_2+\cdots+t_m)}\right) \\ \\ + (1 - (t_1 + t_2 + \cdots + t_m))\, E_q\left(\dfrac{y + x + (t_1 + 2t_2 + \cdots + mt_m)}{1 - (t_1 + t_2 + \cdots + t_m)}\right) \\ \\ + \begin{cases} \left(y + \frac{\sigma}{\gamma} + (t_1 + 2t_2 + \cdots + mt_m)\right) E_q\left(\frac{y+x+(t_1+2t_2+\cdots+mt_m)}{y+\frac{\sigma}{\gamma}+(t_1+2t_2+\cdots+mt_m)}\right) \\ \qquad\qquad \text{if } \frac{y+x+(t_1+2t_2+\cdots+mt_m)}{y+\frac{\sigma}{\gamma}+(t_1+2t_2+\cdots+mt_m)} \geq 1 - \frac{1}{q}, \\ \\ \left(y + \frac{\sigma}{\gamma} + (t_1 + 2t_2 + \cdots + mt_m)\right) \\ -(y + x + (t_1 + 2t_2 + \cdots + mt_m)) \log_q(q-1) \\ \qquad\qquad \text{if } \frac{y+x+(t_1+2t_2+\cdots+mt_m)}{y+\frac{\sigma}{\gamma}+(t_1+2t_2+\cdots+mt_m)} \leq 1 - \frac{1}{q}. \end{cases} \end{array}$$

Note that by (5.80) we have $2\,(2x_1 + 3x_2 + \cdots (m+1)x_m) < 1$, and hence using (5.82) we obtain $t_1 + t_2 + \cdots + t_m \leq t_1 + \frac{3}{2}t_2 + \cdots + \frac{m+1}{2}t_m < \frac{1}{2}$.

Proposition 5.38. *Under Assumption* 1*, let* $y > 0$ *and* $x_1, x_2, \ldots, x_m, \sigma \geq 0$ *be real numbers satisfying* (5.80)*. For each integer* $i \geq 1$*, let* $r_i = \left\lfloor \left(m + y + \frac{\sigma}{\gamma} \right) n_i \right\rfloor$*,* $s_i = \lfloor (m+y) n_i \rfloor$*,* $X_1^{(i)} = \lfloor x_1 n_i \rfloor$*,* $X_2^{(i)} = \lfloor x_2 n_i \rfloor$*,* $\ldots$*,* $X_m^{(i)} = \lfloor x_m n_i \rfloor$*, and* $\mathcal{V}_m^{(i)}(r_i, s_i; X_1^{(i)}, X_2^{(i)}, \ldots, X_m^{(i)})$ *be the set of positive divisors of* F_i *defined in Definition* 5.31*. Then for the cardinalities of these sets we have*

$$\limsup_{i \to \infty} \frac{\log_q \left| \mathcal{V}_m^{(i)}(r_i, s_i; X_1^{(i)}, X_2^{(i)}, \ldots, X_m^{(i)}) \right|}{n_i} \leq \max S(\sigma, y, x, t_1, t_2, \ldots, t_m),$$

where the maximum is over all real numbers x *and* $t_1, t_2, \ldots, t_m$ *satisfying* $0 \leq x \leq \frac{\sigma}{\gamma}$ *and the conditions in* (5.81) *and* (5.82)*.*

Definition 5.39. Under Assumption 1, let $y > 0$ and $x_1, x_2, \ldots, x_m \geq 0$ be real numbers such that $y + 2\,(2x_1 + 3x_2 + \cdots + (m+1)x_m) < 1$. For $\sigma \geq 0$ and $y + 2\,(2x_1 + 3x_2 + \cdots + (m+1)x_m) + \frac{\sigma}{\gamma} < 1$, let $I_{y, x_1, x_2, \ldots, x_m}(\sigma)$ be the real-valued function of σ defined by

$$I_{y, x_1, x_2, \ldots, x_m}(\sigma) = \max S(\sigma, y, x, t_1, t_2, \ldots, t_m),$$

where the maximum is over all real numbers $x, t_1, t_2, \ldots, t_m$ with $0 \leq x \leq \frac{\sigma}{\gamma}$ and $t_1, t_2, \ldots, t_m$ satisfying conditions (5.81) and (5.82).

Lemma 5.40. *Under the assumptions of Definition* 5.39*, the real-valued function* $I_{y, x_1, x_2, \ldots, x_m}(\sigma)$ *is a strictly increasing function of* σ *on its domain of definition, which is the interval of* σ *such that* $\sigma \geq 0$ *and* $y + 2\,(2x_1 + 3x_2 + \cdots + (m+1)x_m) + \frac{\sigma}{\gamma} < 1$*.*

In the following proposition we compute $I_{y, x_1, x_2, \ldots, x_m}(\sigma)$ under some conditions. Again we refer to [9] for its proof.

Proposition 5.41. *We keep the assumptions of Definition* 5.39*. If* $m = 1$*, then let* $\bar{t}_1 = t_1^* = 2x_1$*. If* $m \geq 2$*, then let*

$$\bar{t}_m = 2x_m \quad \text{and} \quad \bar{t}_\ell = 2x_\ell + \sum_{\nu = \ell+1}^{m} x_\nu \qquad \text{for } 1 \leq \ell \leq m-1,$$

let t_1^* *be the real number defined by*

$$2t_1^* + \sum_{\ell=2}^{m} (\ell+1)\bar{t}_\ell = 2 \sum_{\ell=1}^{m} (\ell+1) x_\ell,$$

and for each $2 \leq \ell \leq m$*, let* t_ℓ^* *be the real number defined inductively using* $t_{\ell-1}^*$ *by*

$$(\ell+1)t_\ell^* - (\ell+1)\bar{t}_\ell = \ell t_{\ell-1}^* - \ell \bar{t}_{\ell-1}. \tag{5.83}$$

Moreover, let u be the real number depending on $x_1, \ldots, x_m$ defined by

$$u = t_1^* + \sum_{\ell=2}^{m} \ell \bar{t}_\ell,$$

where $u = t_1^*$ if $m = 1$. Assume also that all of the following conditions hold:

C1: $\frac{\sigma}{\gamma} \leq \frac{y}{q-1}$;
C2: for each $1 \leq \ell \leq m$,

$$\bar{t}_\ell \left(y + \frac{\sigma}{\gamma} + u \right)^{2\ell} < \left(1 - y - \frac{\sigma}{\gamma} - 2 \sum_{\nu=1}^{m} (\nu+1) x_\nu \right)^{\ell+1} \left(y + \frac{\sigma}{\gamma} \right)^{\ell};$$

C3: $\frac{\sigma}{\gamma}(1-y) < y^2$;
C4: if $m \geq 2$, *then for each* $1 \leq \ell \leq m-1$,

$$(\bar{t}_{\ell+1})^{\ell+1} \left(y + \frac{\sigma}{\gamma} + u \right) \leq (t_\ell^*)^{\ell+2}.$$

Then we have $I_{y,x_1,x_2,x_3,\ldots,x_m}(\sigma) = S(\sigma, y, 0, t_1^*, \bar{t}_2, \bar{t}_3, \ldots, \bar{t}_m)$.

Now we introduce an important function based on the function $I_{y,x_1,\ldots,x_m}(\sigma)$ defined in Definition 5.39. In the next definition we use the fact that $I_{y,x_1,\ldots,x_m}(\sigma)$ is an increasing function on its domain of definition, see Lemma 5.40.

Definition 5.42. Under Assumption 1 and for real numbers $y > 0$ and $x_1, \ldots, x_m \geq 0$ with $y + 2(2x_1 + 3x_2 + \cdots + (m+1)x_m) < 1$, let $\Psi(y, x_1, \ldots, x_m)$ be the real-valued function of $y, x_1, \ldots, x_m$ defined by

$$\Psi(y, x_1, \ldots, x_m) = \begin{cases} I^{-1}_{y,x_1,\ldots,x_m} \left(\frac{1}{\gamma} \left[1 + \sum_{l=1}^{\infty} \gamma_l \log_q \frac{q^l}{q^l - 1} \right] \right) \\ \qquad \text{if } \lim_{\sigma \to \theta^-} I_{y,x_1,\ldots,x_m}(\sigma) > \frac{1}{\gamma} \left[1 + \sum_{l=1}^{\infty} \gamma_l \log_q \frac{q^l}{q^l - 1} \right], \\ 0 \quad \text{otherwise,} \end{cases}$$

where $\theta = \gamma\,(1 - y - 2(2x_1 + 3x_2 + \cdots (m+1)x_m))$.

Now we are ready to establish our main results of this section.

Theorem 5.43. *Under Assumption* 1*, let* $x_1, \ldots, x_m \geq 0$ *be real numbers with* $2(2x_1 + 3x_2 + \cdots + (m+1)x_m) < 1$*. For each real number* $0 < \delta < 1 - 2(2x_1 + 3x_2 + \cdots + (m+1)x_m)$ *we have*

$$\begin{aligned}\alpha_q(\delta) \geq R_{\{\gamma_l\},x_1,\ldots,x_m}(\delta) &:= 1 - \delta - \frac{1}{\gamma} + (x_1 + \cdots + x_m)\log_q(q-1)\\ &-(x_1 \log_q x_1 + \cdots + x_m \log_q x_m)\\ &-(1 - (x_1 + \cdots + x_m))\log_q(1 - (x_1 + \cdots + x_m))\\ &-(4x_1 + 5x_2 + \cdots + (m+3)x_m)\\ &+\frac{1}{\gamma}\Psi\Big(1 - \delta - 2(2x_1 + 3x_2 + \cdots (m+1)x_m), x_1, x_2, \ldots, x_m\Big).\end{aligned}$$

Proof. Let $y = 1 - \delta - 2(2x_1 + 3x_2 + \cdots + (m+1)x_m)$ and $\sigma = \Psi(y, x_1, \ldots, x_m)$. If $R_{\{\gamma_l\},x_1,\ldots,x_m}(\delta) \leq 0$, then the statement of the theorem is trivial. If $\sigma = 0$ and $R_{\{\gamma_l\},x_1,\ldots,x_m}(\delta) > 0$, then the theorem follows from [7, Theorem 5.1]. Indeed, in this case let $r_i = \lfloor (m+y)n_i \rfloor$ for $i \geq 1$. As $R_{\{\gamma_l\},x_1,\ldots,x_m}(\delta) > 0$, the conditions of [7, Theorem 5.1] are satisfied. Then using [7, Theorem 5.1] for sufficiently large i with a divisor of degree r_i of F_i, we obtain a sequence of q-ary codes proving the theorem in this case. The computation of the parameters is similar to the case where $\sigma > 0$ and $R_{\{\gamma_l\},x_1,\ldots,x_m}(\delta) > 0$, which is explained in detail below. Note that Theorem 5.17 together with Proposition 5.20 can be used instead of [7, Theorem 5.1] above.

Now we consider the remaining case where $\sigma > 0$ and $R_{\{\gamma_l\},x_1,\ldots,x_m}(\delta) > 0$. Let $0 < \epsilon < \sigma$ be a real number satisfying

$$\begin{aligned}&y + (x_1 + \cdots + x_m)\log_q(q-1) - (x_1 \log_q x_1 + \cdots + x_m \log_q x_m)\\ &-(1 - (x_1 + \cdots + x_m))\log_q(1 - (x_1 + \cdots + x_m))\\ &+(x_2 + 2x_3 + \cdots + (m-1)x_m)\\ &> \frac{1 - (\sigma - \epsilon)}{\gamma}.\end{aligned} \tag{5.84}$$

For $i \geq 1$, let

$$\begin{aligned}&r_i = \left\lfloor \left(m + y + \tfrac{\sigma - \epsilon}{\gamma}\right) n_i \right\rfloor,\ s_i = \lfloor (m+y)n_i \rfloor,\\ &X_1^{(i)} = \lfloor x_1 n_i \rfloor,\ X_2^{(i)} = \lfloor x_2 n_i \rfloor, \ldots,\ X_m^{(i)} = \lfloor x_m n_i \rfloor.\end{aligned} \tag{5.85}$$

It follows from [13, Corollary 2] (see also [14, Exercise 2.3.27]) that we have

$$\liminf_{i\to\infty} \frac{\log_q h_i}{n_i} \geq \frac{1}{\gamma}\left[1+\sum_{l=1}^{\infty}\gamma_l \log_q \frac{q^l}{q^l-1}\right], \tag{5.86}$$

where h_i is the class number of F_i. For sufficiently large i, by Proposition 5.38 and (5.86), the hypotheses of Proposition 5.32 for the global function field F_i with r_i, s_i, and $X_1^{(i)},\ldots,X_m^{(i)}$ as in (5.85) are satisfied. Let G_i be the divisor of F_i given by Proposition 5.32 with these parameters for sufficiently large i.

Note that

$$\begin{aligned}
&\liminf_{i\to\infty} \frac{\log_q |M(x_1,\ldots,x_m;\mathbf{0})|}{n_i} \\
&\geq (x_1+\cdots+x_m)\log_q(q-1) - (x_1\log_q x_1+\cdots+x_m\log_q x_m) \\
&\quad -(1-(x_1+\cdots+x_m))\log_q(1-(x_1+\cdots+x_m)) \\
&\quad +(x_2+2x_3+\cdots+(m-1)x_m)
\end{aligned} \tag{5.87}$$

(see (5.70) and Proposition 5.20). Since we have (5.84), using the divisor G_i of the global function field F_i for sufficiently large i, Theorem 5.35, and (5.87), we obtain a sequence of q-ary codes $\{C_i\}_{i=1}^{\infty}$ of lengths $\{n_i\}_{i=1}^{\infty}$, respectively, such that $n_i\to\infty$ as $i\to\infty$ by Assumption 1 as well as

$$\begin{aligned}
&\liminf_{i\to\infty} \frac{\log_q |C_i|}{n_i} \\
&\geq y+\frac{\sigma-\epsilon}{\gamma}-\frac{1}{\gamma} \\
&\quad +(x_1+\cdots+x_m)\log_q(q-1)-(x_1\log_q x_1+\cdots+x_m\log_q x_m) \\
&\quad -(1-(x_1+\cdots+x_m))\log_q(1-(x_1+\cdots+x_m)) \\
&\quad +(x_2+2x_3+\cdots+(m-1)x_m) \\
&= 1-\delta-2(2x_1+3x_2+\cdots+(m+1)x_m)+\frac{\sigma-\epsilon}{\gamma}-\frac{1}{\gamma} \\
&\quad +(x_1+\cdots+x_m)\log_q(q-1)-(x_1\log_q x_1+\cdots+x_m\log_q x_m) \\
&\quad -(1-(x_1+\cdots+x_m))\log_q(1-(x_1+\cdots+x_m)) \\
&\quad +(x_2+2x_3+\cdots+(m-1)x_m) \\
&= R_{\{\gamma_l\},x_1,\ldots,x_m}(\delta)-\frac{\epsilon}{\gamma}
\end{aligned}$$

and

$$\liminf_{i\to\infty}\frac{d(C_i)}{n_i}\geq\delta.$$

Using the fact that α_q is a nonincreasing function (see Proposition 5.1(i)), we get

$$\alpha_q(\delta) \geq R_{\{\gamma_l\},x_1,\ldots,x_m}(\delta) - \frac{\epsilon}{\gamma}.$$

Letting $\epsilon \to 0^+$ completes the proof. □

Corollary 5.44. *Under Assumption* 1*, for each real number* $0 < \delta < 1$ *we have*

$$\alpha_q^{\mathrm{lin}}(\delta) \geq R^{\mathrm{lin}}_{\{\gamma_l\}}(\delta) := 1 - \delta - \frac{1}{\gamma} + \frac{1}{\gamma}\Psi\Big(1-\delta, 0\Big).$$

Proof. Taking $m = 1$ and using similar methods as in the proof of Theorem 5.43, but applying Corollary 5.36 instead of Theorem 5.35, we obtain the desired result. □

In the rest of this section we give numerical examples in order to demonstrate that Theorem 5.43 and Corollary 5.44 yield improvements on the lower bounds for $\alpha_q(\delta)$ and $\alpha_q^{\mathrm{lin}}(\delta)$ at least for certain values of q and certain values of δ.

Let $R_{NO2,\gamma,x}(\delta)$ and $R^{\mathrm{lin}}_{X,\gamma}(\delta)$ denote the lower bound in [8, Theorem 5.1] and Xing's lower bound for $\alpha_q^{\mathrm{lin}}(\delta)$ in [17], respectively (see also [8, Theorem 4.6]).

Let $q = 2^6$, $\gamma = \gamma_1 = \sqrt{q} - 1$ (see Remark 5.8), $\gamma_l = 0$ for $l \geq 2$, and

$$\delta = \frac{13763868443250238929521503984833381597731412559044}{46065097831342932365531985486767649347321318605709}$$

$$= 0.29879169026501515839\ldots.$$

In [8, Example 5.2], using $x = 10^{-13}$ it has been obtained that

$$\alpha_q(\delta) \geq R_{NO2,\gamma,x}(\delta) = 0.55835371587781529071\ldots,$$

and it has been demonstrated that $R_{NO2,\gamma,x}(\delta) - R^{\mathrm{lin}}_{X,\gamma}(\delta) \geq 7.3387 \cdot 10^{-15}$.

By Corollary 5.44 we obtain that

$$\alpha_q^{\mathrm{lin}}(\delta) \geq R^{\mathrm{lin}}_{\gamma}(\delta) = 0.55835395724081743804\ldots.$$

Note that $R^{\mathrm{lin}}_{\gamma}(\delta) - R_{NO2,\gamma,x}(\delta) \geq 2.4136300214732 \cdot 10^{-7}$, and $R^{\mathrm{lin}}_{\gamma}(\delta)$ is better than $R^{\mathrm{lin}}_{X,\gamma}(\delta)$. Hence we have an improvement on the lower bound for $\alpha_q^{\mathrm{lin}}(\delta)$ compared to Xing's bound in [17].

By Theorem 5.43 with $x_1 = 3.41 \cdot 10^{-16}$, $x_2 = 1.0634 \cdot 10^{-23}$, and $x_3 = 1.93 \cdot 10^{-31}$, we obtain $\alpha_q(\delta) \geq R_{\gamma,x_1,x_2,x_3}(\delta)$, where

$$R_{\gamma,x_1,x_2,x_3}(\delta) - R^{\mathrm{lin}}_{\gamma}(\delta) \geq 2.711029 \cdot 10^{-17}.$$

Hence $R_{\gamma,x_1,x_2,x_3}(\delta)$ gives a further improvement on the lower bound for $\alpha_q(\delta)$. Note that $R_{\gamma,x_1,x_2,x_3}(\delta)$ yields an improvement on $R_{NO2,\gamma,x}(\delta)$ of the order 10^{-7}, whereas Maharaj [6] obtained only an improvement of the order 10^{-15}.

Now let

$$\delta = \frac{3230122938809269343601048150193426774958990604666 5}{4606509783134293236553198548676764934732131860570 9}$$

$$= 0.70120830973498484160\ldots\ .$$

In [8, Example 5.2], using $x = 10^{-13}$ it has been obtained that

$$\alpha_q(\delta) \geq R_{NO2,\gamma,x}(\delta) = 0.15593709640785805503\ldots\ ,$$

and it has been demonstrated that $R_{NO2,\gamma,x}(\delta) - R^{\mathrm{lin}}_{X,\gamma}(\delta) \geq 1.97862 \cdot 10^{-14}$.

By Corollary 5.44 we obtain that

$$\alpha^{\mathrm{lin}}_q(\delta) \geq R^{\mathrm{lin}}_{\gamma}(\delta) = 0.15593754394482448829\ldots.$$

Note that $R^{\mathrm{lin}}_{\gamma}(\delta) - R_{NO2,\gamma,x}(\delta) \geq 4.4753696643325 \cdot 10^{-7}$, hence $R^{\mathrm{lin}}_{\gamma}(\delta)$ is better than $R^{\mathrm{lin}}_{X,\gamma}(\delta)$. Hence we have an improvement on the lower bound for $\alpha^{\mathrm{lin}}_q(\delta)$ compared to Xing's bound in [17].

By Theorem 5.43 with $x_1 = 3.89 \cdot 10^{-18}$, $x_2 = 1.98 \cdot 10^{-26}$, and $x_3 = 5.87 \cdot 10^{-35}$, we obtain $\alpha_q(\delta) \geq R_{\gamma,x_1,x_2,x_3}(\delta)$, where

$$R_{\gamma,x_1,x_2,x_3}(\delta) - R^{\mathrm{lin}}_{\gamma}(\delta) \geq 2.592642 \cdot 10^{-19}.$$

Hence $R_{\gamma,x_1,x_2,x_3}(\delta)$ gives a further improvement on the lower bound for $\alpha_q(\delta)$. Note that $R_{\gamma,x_1,x_2,x_3}(\delta)$ yields again an improvement on $R_{NO2,\gamma,x}(\delta)$ of the order 10^{-7}, whereas Maharaj [6] obtained only an improvement of the order 10^{-14}.

The numerical examples above were already given in [9, Section 7]. For further examples we refer to [9, Section 7].

Acknowledgments

The research of the first author was carried out while he was hosted by CNRS-FR2291 (FRUMAM) at the Université de la Méditerranée in Marseille-Luminy. The research of the second author was supported by TÜBİTAK under Grant No. TBAG-107T826.

References

[1] J. Bezerra, A. Garcia, and H. Stichtenoth, An explicit tower of function fields over cubic finite fields and Zink's lower bound, *J. Reine Angew. Math.*, vol. 589, pp. 159–199, 2005.

[2] N.D. Elkies, Excellent nonlinear codes from modular curves, in: *STOC'01, Proceedings of the 33rd Annual ACM Symposium on Theory of Computing* (Hersonissos, Greece, 2001), ACM Press, New York, 2001, pp. 200–208.

[3] A. Garcia and H. Stichtenoth, A tower of Artin-Schreier extensions of function fields attaining the Drinfeld-Vladut bound, *Invent. Math.*, vol. 121, pp. 211–222, 1995.

[4] J.H. van Lint, *Introduction to Coding Theory*, 3rd ed., Springer, Berlin, 1999.

[5] F.J. MacWilliams and N.J.A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland, Amsterdam, 1977.

[6] H. Maharaj, A note on further improvements of the TVZ-bound, *IEEE Trans. Inform. Theory*, vol. 53, pp. 1210–1214, 2007.

[7] H. Niederreiter and F. Özbudak, Constructive asymptotic codes with an improvement on the Tsfasman-Vlăduţ-Zink and Xing bounds, in: *Coding, Cryptography and Combinatorics* (K.Q. Feng, H. Niederreiter, and C.P. Xing, eds.), Progress in Computer Science and Applied Logic, vol. 23, Birkhäuser, Basel, 2004, pp. 259–275.

[8] H. Niederreiter and F. Özbudak, Further improvements on asymptotic bounds for codes using distinguished divisors, *Finite Fields Appl.*, vol. 13, pp. 423–443, 2007.

[9] H. Niederreiter and F. Özbudak, Improved asymptotic bounds for codes using distinguished divisors of global function fields, *SIAM J. Discrete Math.*, vol. 21, pp. 865–899, 2007.

[10] H. Niederreiter and C.P. Xing, *Rational Points on Curves over Finite Fields: Theory and Applications*, Cambridge University Press, Cambridge, 2001.

[11] H. Stichtenoth, *Algebraic Function Fields and Codes*, Springer, Berlin, 1993.

[12] H. Stichtenoth and C.P. Xing, Excellent nonlinear codes from algebraic function fields, *IEEE Trans. Inform. Theory*, vol. 51, pp. 4044–4046, 2005.

[13] M.A. Tsfasman, Some remarks on the asymptotic number of points, in: *Coding Theory and Algebraic Geometry* (H. Stichtenoth and M.A. Tsfasman, eds.), Lecture Notes in Mathematics, vol. 1518, Springer, Berlin, 1992, pp. 178–192.

[14] M.A. Tsfasman and S.G. Vlăduţ, *Algebraic-Geometric Codes*, Kluwer, Dordrecht, 1991.

[15] M.A. Tsfasman, S.G. Vlăduţ, and T. Zink, Modular curves, Shimura curves, and Goppa codes, better than Varshamov-Gilbert bound, *Math. Nachr.*, vol. 109, pp. 21–28, 1982.

[16] S.G. Vlăduţ, An exhaustion bound for algebro-geometric "modular" codes, *Problems Inform. Transmission*, vol. 23, pp. 22–34, 1987.

[17] C.P. Xing, Algebraic-geometry codes with asymptotic parameters better than the Gilbert-Varshamov and the Tsfasman-Vlăduţ-Zink bounds, *IEEE Trans. Inform. Theory*, vol. 47, pp. 347–352, 2001.

[18] C.P. Xing, Nonlinear codes from algebraic curves improving the Tsfasman-Vlăduţ-Zink bound, *IEEE Trans. Inform. Theory*, vol. 49, pp. 1653–1657, 2003.

Chapter 6

Algebraic Curves with Many Points over Finite Fields

Fernando Torres

Institute of Mathematics and Computer Sciences
P.O. Box 6065, University of Campinas, SP, Brazil
ftorres@ime.unicamp.br

> As long as Algebra and Geometry proceeded along separate paths, their advance was slow and their applications limited.
>
> But when these sciences joined company they drew from each other fresh vitality and thenceforward marched on at a rapid pace towards perfection.
>
> J.L. Lagrange
>
> (Cited in Goppa's book [38])

Contents

Introduction

The purpose of this chapter is to survey some results concerning the number of rational points of curves over finite fields. A remarkable motivation which is intimately related to mathematicians like Fermat, Euler, Lagrange, Legengre, Gauss, Jacobi, ... is the following question (cf. [17], [85], [81]). Let p be a prime and $m \geq 2$ an integer such that p does not divide m. Let $\mathbb{F}_p$ denote the finite field with p elements. How many solutions

in the projective plane $\mathbb{P}^2(\mathbb{F}_p)$ exist for the curve

$$X^m + Y^m + Z^m = 0?$$

In the early years of the 19th century, Gauss considered finite sums of powers of pth root of unity (now known as *Gauss sums*) to give a proof of one of the great theorems in mathematics: the Quadratic Reciprocity Law (cf. $m = 2$); the proof suggests an approach to Higher Reciprocity Law (cf. $m > 2$). Let N be the number of $\mathbb{F}_p$-solutions of the curve above. It turns out that N is a *Jacobi sum*; i.e., a finite sum of sums closely related to Gauss sums. Gauss calculated N for $m = 2$ and $m = 3$; see e.g. [81, Ch. 6]. If $m > 3$ however, things get progressively more complicated and in general there is only an estimate, namely

$$|N - (p+1)| \leq \lfloor 2g\sqrt{p} \rfloor ,$$

where $g = (m-1)(m-2)/2$ is the genus of the curve, see Weil [102]. This result is indeed a particular case of a deep result in Algebraic Curve Theory, namely the so-called *Hasse-Weil bound* (HW-bound) (or the *Riemann-Hypothesis*) for curves over finite fields. Throughout, let $\mathcal{X}$ be a curve (nonsingular, projective, geometrically irreducible) of genus g over the finite field $\mathbb{F}_q$ with q elements. The HW-bound assert that

$$|\#\mathcal{X}(\mathbb{F}_q) - (q+1)| \leq \lfloor 2g\sqrt{q} \rfloor ;$$

Hasse (around 1932) showed the case $g = 1$ via complex multiplication on elliptic curves and Weil (around 1940) showed the general case via the theory of the correspondences [101]. The key starting point was a conjecture of Artin (Ph.D. thesis, 1924) on the complex module of the zeroes of a *zeta-function* of a curve, see Theorem 6.2. Such a function was introduced by Artin himself in analogy with Dedekind's zeta-function of numerical fields and the aforementioned conjecture was inspired by the well-known classical Riemann hypothesis.

Bombieri [8] gave an elementary proof of the HW-bound by following ideas of Stepanov, Postnikov, Stark and Manin; his proof uses the Riemann-Roch theorem only. Now, once the HW-bound was available, some sharp upper bounds were obtained in the context of questions associated to curves; e.g. exponential sums [89], [70] and the number of elements of plane arcs [49], [50] and [48] (see also the references therein).

Let $N_q(g)$ be the maximal number of rational points that a curve of genus g over $\mathbb{F}_q$ can have. In the last years, due mainly to applications in Coding Theory and Cryptography, there has been considerable interest in

computing the actual value of $N_q(g)$. It is a classical result that $N_q(0) = q + 1$. Deuring [16] and Serre [88] computed $N_q(1)$ and $N_q(2)$; we quote these computations in Example 6.5. For $g = 3$ we have the Voloch's bound which says that $N_q(3) \leq 2q + 6$ whenever $q \neq 8, 9$, see Example 6.6. Serre computed $N_q(3)$ for $q < 25$ [88] and Top [94] extended these computations to $q < 100$; see Remarks 6.7, 6.8. The tables in [34] describe what is known about $N_q(g)$ for $g \leq 50$ and $q \in \{2, 3, 4, 8, 9, 16, 27\}$. By using narrow ray class extensions, Niederreiter and Xing found bounds on $N_q(g)$ for $q = 2, 3, 4, 5, 8, 16$ and $1 \leq g \leq 50$ [75]; see also [76] and the references therein.

In general, a closed formula for $N_q(g)$ seems still to be a long way off. An upper bound on $N_q(g)$ is clearly the HW-bound; Serre [87] observed that this bound may be sharpened in several cases via the HWS-bound in (6.4) or the "explicit formulas" method in Proposition 6.9. Osterlé used tools from linear programing to optimize this method [88] by selecting the "best" trigonometric polynomial in (6.6); this is called the Osterlé bound. Currently, powerful tools related to Abelian Varieties are used to investigate $N_q(g)$; cf. Howe, Lauter, Serre [59], [60], [61], [62], [63], [64], [65]; we will not survey these results here.

In order to find lower bounds on $N_q(g)$ we look for curves $\mathcal{X}$ "with many points" in the sense that $\#\mathcal{X}(\mathbb{F}_q)$ has to be as close as possible to the best upper bound known on $N_q(g)$. In most cases, the best known bound comes from Osterlé's (cf. [62]). If $\#\mathcal{X}(\mathbb{F}_q) = N_q(g)$, the curve is called *Optimal*. In Section 6.5 we investigate a particular family of optimal curves, the so-called *Maximal Curves*; i.e., those whose number of rational points attains the upper HW-bound. A distinguished example here is the *Hermitian curve* which is intrinsically determined by its genus and number of rational points [82]; see Theorem 6.25 here. There are also two important families of optimal curves, namely the *Suzuki curves* and the *Ree curves*; each curve in each family is intrinsically determined by the data: (1) the genus, (2) the number of rational points and (3) the automorphism group (see Hansen [39], Hansen-Pedesrsen [40], Hansen-Stichtenoth, [41], Heen [46]). An important result is Theorem 6.15, where we show that the Suzuki curve is characterized by properties (1) and (2) only; it seems that this property is an open problem for the Ree curve. It is worthwhile to point out that the Hermitian curve, the Suzuki curve and the Ree curve are respectively the Deligne-Lusztig varieties of positive genus associated to connected reduction algebraic group of type ${}_2A^2$, 2^2_B and ${}_2R^2$ [15].

Apart from Bombieri's work in simplifying the proof of the HW-bound

and the bounds on exponential sums and plane arcs mentioned above, qualitative aspects of the study of the HW-bound in 1940 was similar to that in 1977. The interest on this matter was renewed after Goppa (around 1977) constructed error-correcting codes from linear series on curves, the so-called *Geometric Goppa Codes* (GG-codes) (as they currently are known); see [37], [38]. These codes generalize the well-known Reed-Solomon codes, BCH-codes and the "classical" Goppa codes (see van Lint [67], van Lint-van der Geer [68]). Goppa's idea showed for the first time how two totally different areas of mathematics: Algebraic Curve Theory ("pure" subject) and Coding Theory ("applied" subject) can be related to each other.

Next we briefly describe (the dual construction) of a GG-code. Let g_e^r be a r-dimencional linear series on $\mathcal{X}$ of degree e defined over $\mathbb{F}_q$ and whose sections are contained in a Riemann-Roch space $\mathcal{L}(G)$. For simplicity we shall assume that $g_e^r = |G|$ is complete. Let $P_1, \ldots, P_n$ be pairwise distinct $\mathbb{F}_q$-rational points of the curve such that $P_i \notin \mathrm{Supp}(G)$ for any i. Consider the $\mathbb{F}_q$-linear map

$$\mathrm{e_v} : f \in \mathcal{L}(G) \mapsto (f(P_1), \ldots, f(P_n)) \in \mathbb{F}_q^n \,.$$

Then the following q-ary linear code, namely

$$C_{\mathcal{X}}(G, D) := \mathrm{e_v}(\mathcal{L}(G))$$

is the Goppa code defined by the triple $(\mathcal{X}, G, D)$, where $D := P_1 + \ldots + P_n$. Let k and d be respectively the dimension and minimum distance of the code. Then

(1) $k = \ell(G) - \ell(G - D)$, where $\ell(.)$ denotes the $\mathbb{F}_q$-dimension of the corresponding Riemann-Roch space;
(2) $d \geq n - \deg(G)$.

We observe that k and d can be handled by means of the Riemann-Roch theorem. In addition, (2) is only meaningful, if (fixed $\deg(G)$), $\mathcal{X}$ is a curve with many points. With respect to the dimension k, if $n > \deg(G)$, then

$$k = \ell(G) = \deg(G) + 1 - g + \ell(K - G) \geq \deg(G) + 1 - g \,,$$

where K is a canonical divisor on $\mathcal{X}$; in particular,

$$n + 1 \geq k + d \geq n + 1 - g \,. \qquad (*)$$

Thus we are ready to appreciate an amazing asymptotic property of families of GG-codes and to understand the first remarkable application of these codes in the context of asymptotic problems in Coding Theory. As a matter

of fact, Tsfasman, Vlăduţ and Zink [97] (see also [96], [70]) showed that, for $q \geq 49$ a square, the Gilbert-Varshamov bound can be improved via a sequence of GG-codes; roughly speaking, this is done as follows:

(A) They show that there is a family of GG-codes $(\mathcal{X}_i)$ such that the sequence of their relative parameters (r_i, δ_i) has a limit point (R, δ). Here the sequence of genus $g_i \to \infty$ and $\limsup_i \frac{n_i}{g_i} = \sqrt{q} - 1$;
(B) Then inequality $(*)$ implies $R + \delta = 1 - 1/(\sqrt{q} - 1)$; this improves the Gilbert-Varshamov.

For Items (A) and (B) above, one studies values $N_q(g)$ (q fixed and g large enough) and ask for the limit

$$A(q) := \limsup_g \frac{N_q(g)}{g}$$

to be as large as possible. We consider this question in Section 6.2, where our main references were the papers by Kresch et al. [56] and Elkies et al. [19].

Coming back to the study of the HW-bound for a single curve, Stöhr and Voloch (around 1982) development a geometric method to bound $\#\mathcal{X}(\mathbb{F}_q)$ based on $\mathbb{F}_q$-linear series on the curve [91]; such a bound will be denoted by SV-bound. We report some features on this theory in the Appendix. The SV-bound gives a new proof of the HW-bound and improvements in several cases. For example, via the SV-bound, Voloch obtained the best upper bound known so far on the order of complete arcs in projective planes over prime fields [99], [100].

There is a natural link between the arithmetic and geometry of a curve which comes from a linear series naturally defined from the zeta-function of the curve (see Section 6.3). This linear series is simple and its existence implies the uniqueness of the Suzuki curve. In the case of maximal curves, the linear series is very ample (Theorem 6.23) and thus we can study maximal curves embedded in projective spaces and apply classical results from Algebraic Curve Theory or Finite Geometry such as:

- The Castelnuovo genus bound for curves in projective spaces [10], [6], [78], [42];
- Halphen's bound on the genus of the curve taking into consideration the degree of a surface where the curve is contained [11];
- Properties of quadratic surfaces in $\mathbb{P}^3(\overline{\mathbb{F}}_q)$ [48].

We recall that Castelnuovo and Halphen bounds are valid in positive characteristic by Hartshorne [42] (space curves) and Rathmann [78].

From the interplay of these properties with the Stöhr-Voloch theory (Appendix) we deduce quantitave and qualitative properties of maximal curves (see Hirschfeld et al. [51]); we will mention a few of them in Section 6.5.

Tafasolian [92] (Ph. D. Thesis, 2008) investigated properties of maximal curves via Cartier Operators; among other things, he characterized certain HWS-maximal curves, HW-maximal Fermat curves and HW-maximal Artin-Schreier curves. His results improve on previous work in [2], [5], [3], [1], [22].

Standard references are the books by Fulton [25], Arbarello et al. [6], Hartshorne [42], Namba [73], Stichtenoth [90], Moreno, [70], Stepanov [89], Goldschmidt [36], Goppa [38], Tsfasman and Vladut [96], Hirschfeld et all. [51]. For the convenience of the reader we include an Appendix on the Theory of Stöhr-Voloch [91].

Throughout this chapter, by a curve over $\mathbb{F}_q$ (the finite field with q elements) we mean a nonsingular, projective, geometrically irreducible algebraic curve defined over $\mathbb{F}_q$.

6.1. The Function $N_q(g)$

In this section we discuss curves with many points. Our references on zeta-functions are e.g. the books [90], [89] or [70]. Let $\mathcal{X}$ be a curve of genus g over $\mathbb{F}_q$. Let $N_i = \#\mathcal{X}(\mathbb{F}_{q^i})$ be the number of $\mathbb{F}_{q^i}$-rational points of $\mathcal{X}$. Thanks to Riemann, Dedekind, Artin, Hasse, Weil, ... all the information about the N_i is contained in the zeta-function

$$Z(t) = Z(\mathcal{X}, q; t) := \exp(\sum_{i=1}^{\infty} N_i t^i / i) \tag{6.1}$$

of $\mathcal{X}$ over $\mathbb{F}_q$. By the Riemann-Roch theorem, there is a polynomial $L(t) = L(\mathcal{X}, q; t)$ of degree $2g$ satisfying:

Proposition 6.1.

(1) $L(t) = Z(t)(1-t)(1-qt)$;

(2) $L(t) = \pi_{j=1}^{2g}(1 - \alpha_j t)$ *where the* α_j *are algebraic integers which can be arranged in such a way that* $\alpha_j \bar{\alpha}_j = q$.

Thus from (6.1) we obtain

$$N_i = q^i + 1 - \sum_{j=1}^{2g} \alpha_j^i = q^i + 1 - \sum_{j=1}^{g} (\alpha_j^i + \bar{\alpha}_j^i)\,. \tag{6.2}$$

The main result to bound $\#\mathcal{X}(\mathbb{F}_q)$ is the following.

Theorem 6.2. (Riemann hypothesis) *The complex value of each α_j is $\sqrt{q}$.*

Therefore (6.2) implies the Hasse-Weil bound (HW-bound) mentioned in the introduction (for $i = 1$), namely

$$|\#\mathcal{X}(\mathbb{F}_{q^i}) - (q^i + 1)| \le \lfloor 2g\sqrt{q^i} \rfloor\,. \tag{6.3}$$

Example 6.3. (The Hermitian curve) If $q = \ell^2$, the HW-bound is sharp as the following curve, known as *the Hermitian curve*

$$\mathcal{H}: \quad X^{\ell+1} + Y^{\ell+1} + Z^{\ell+1} = 0$$

shows. The genus of $\mathcal{H}$ is $g = \ell(\ell-1)/2$ and $\#\mathcal{H}(\mathbb{F}_{\ell^2}) = \ell^3 + 1$. Rück and Stichtenoth [82] noticed that $\mathcal{H}$ is the unique curve having these properties; we will improve this result in Theorem 6.25.

Example 6.4. (The Klein quartic over $\mathbb{F}_8$) In general the HW-bound is not sharp: Consider the curve

$$\mathcal{K}: \quad X^3Y + Y^3Z + Z^3X = 0\,,$$

known as *the Klein quartic*. If $q = 8$, the curve is nonsingular of genus $g = 3$. The HW-bound is 25; however, $\#\mathcal{K}(\mathbb{F}_8) = 24$.

In Remark 6.26 we will see that the HW-bound is not necessarily sharp even if q is a square. Set

$$N_q(g) := \max\{\#\mathcal{Y}(\mathbb{F}_q) : \mathcal{Y} \text{ a curve of genus } g \text{ defined over } \mathbb{F}_q\}\,.$$

Example 6.5. (Deuring [16], Serre [88]) Write $q = p^\alpha$ and $m = \lfloor 2\sqrt{q} \rfloor$. Thus

- $N_q(1) = q + 1 + m$ except when $\alpha \ge 3$ is odd, and p divides m; in this case, $N_q(1) = q + m$.
- $N_q(2) = q + 1 + 2m$ except in the following cases: (1) $N_4(2) = 10$, $N_9(2) = 20$; (2) α is odd, p divides m; (3) α is odd and q of the form $x^2 + 1$, $x^2 + x + 1$ or $x^2 + x + 2$ ($x \in \mathbb{Z}$).

In cases (2) and (3) above we have $N_q(2) = q+2m$ if $2\sqrt{q}-m > (\sqrt{5}-1)/2$ or $N_q(2) = q+2m-1$ otherwise.

As a nice application of the Appendix we prove the Voloch's bound for curves of genus 3; cf. Serre [88], Top [94, Prop.1].

Example 6.6. For $q \neq 8, 9$, $N_q(3) \leq 2q+6$. Indeed, let $\mathcal{X}$ be a curve of genus 3 over $\mathbb{F}_q$ with $\#\mathcal{X}(\mathbb{F}_q) > 2q+6$. Then $\mathcal{X}$ is nonhyperelliptic. We apply the Appendix to the canonical linear series $\mathcal{D}$. Let $0 = \nu_0 < \nu_1$ be the $\mathbb{F}_q$-Frobenius orders and S the $\mathbb{F}_q$-divisor of $\mathcal{D}$ respectively. Thus

$$2q+6 < \deg(S)/2 = (4\nu_1 + (q+2)4)/2$$

so that $\nu_1 > 1$. Thus the order sequence of $\mathcal{D}$ is $0, 1, \nu_1$ and $j_2(P) \geq \nu_1 + 1$ for any $P \in \mathcal{X}(\mathbb{F}_q)$. The Hefez-Voloch theorem (Appendix) gives $\#\mathcal{X}(\mathbb{F}_q) = 4(q-2)$ and thus

$$\deg(R) = (1+\nu_1)4 + 12 \geq \#\mathcal{X}(\mathbb{F}_q) = 4(q-2)\,,$$

and hence $q < \nu_1 + 6$; i.e. $q \in \{2, 3, 4, 5, 7, 8, 9\}$ as $\nu_1 \leq 4$. On the other hand, $\#\mathcal{X}(\mathbb{F}_q) = 4q - 8 > 2q + 6$ so that $q = 8, 9$.

Remark 6.7. We have that $28 \leq N_9(3)$ and $24 \leq N_8(3)$ due to the Hermitian curve and the Klein quartic above.

Case: $q = 9$**.** Following the example above we find that $\nu_1 = 3$ whenever $\#\mathcal{X}(\mathbb{F}_9) \leq \deg(S)/2 = 28$ In particular, $N_9(3) = 28$. We observe that there is just one curve of genus $g = 3$ over $\mathbb{F}_9$ with 28 rational points, namely the Hermitian $X^4 + Y^4 + Z^4 = 0$, cf. [82].

Case. $q = 8$**.** As in Case 1 we find that $\nu_1 = 2$ and $N_8(3) = 24$. From [94, Prop1.1(a)] the Klein quartic over $\mathbb{F}_8$ is the unique curve of genus 3 with 24 rational points.

Remark 6.8. From the table in [94] we observe that Voloch's bound is sharp for $q = 4, 5, 7, 11, 13, 16, 17, 19, 25$. Let $q = p^{2a}$ with p an odd prime and $a \geq 1$ an integer. Ibukiyama [52] showed that there exist a curve of genus 3 over $\mathbb{F}_q$ whose number of rational points attains the HW-bound. Thus $N_q(3) = p^{2a} + 1 + 6p^a$.

Serre [87] noticed that the HW-bound (6.3) may be improved to the following HWS-bound:

$$|\#\mathcal{X}(\mathbb{F}_{q^i}) - (q^i + 1)| \leq g\lfloor 2\sqrt{q} \rfloor\,. \tag{6.4}$$

This bound is sharp as Example 6.4 above shows. Now we remark the Serre "explicit formulas" method; cf. [88], [39].

From Theorem 6.2 we can write $\alpha_j = \sqrt{q}\exp(\sqrt{-1}\theta_j)$. From (6.2)

$$N_i = q^i + 1 - 2\sqrt{q}^i \sum_{j=1}^{g} \cos(j\theta_j)\,. \tag{6.5}$$

Let $f(\theta)$ be a trigonometric polynomial of the form

$$f(\theta) = 1 + 2\sum_{n\geq 1} c_n \cos(n\theta)\,.$$

Set

$$\psi_d(t) := \sum_{n\geq 1} c_{nd} t^{nd} \quad d \geq 1\,.$$

After some computation, (6.5) implies

$$\sum_{j=1}^{g} f(\theta_j) + + \sum_{d\geq 1} d a_d \psi_d(q^{-1/2}) = g + \psi_1(q^{-1/2}) + \psi_1(q^{1/2})\,, \tag{6.6}$$

where a_d is the number of points of degree d. Notice that $N_i = \sum_{d|i} d a_d$. Whence we obtain the following.

Proposition 6.9. *Suppose that the c_i's are non-negative real number not all zero. Suppose that $f(\theta) \geq 0$ for all θ. Then*

$$N_1 = \#\mathcal{X}(\mathbb{F}_q) \leq \frac{g}{\psi_1(q^{-1/2})} + 1 + \frac{\psi_1(q^{1/2})}{\psi_1(q^{-1/2})}\,;$$

equality holds if and only if

$$\sum_{j=1}^{g} f(\theta_j) = 0\,, \quad \textit{and} \quad \sum_{d\geq 2} d a_d \psi_d(q^{-1/2}) = 0\,.$$

Set

$$h(t) = h(\mathcal{X}, q; t) := t^{2g} L(\mathcal{X}, q; t^{-1})\,. \tag{6.7}$$

The following result is the key starting point for the characterization of the Suzuki curve given in Section 6.4.

Proposition 6.10. (cf. [88]) *Let $q = 2q_0^2$ with q_0 a power of two. Let $\mathcal{X}$ be a curve of genus $g = q_0(q-1)$ with $N_1 = q^2 + 1$ rational points. Then*

$$h(t) = (t^2 + 2q_0 t + q)^g\,.$$

Proof. Let $h(t) = \prod_{j=1}^{g}(t-\alpha_j)(t-\bar{\alpha}_j)$ with $\alpha_j = \sqrt{q}\exp(\sqrt{-1}\theta_j)$. We let

$$f(\theta) := 1 + \sqrt{2}\cos(\theta) + \frac{1}{2}\cos(2\theta) = \frac{1}{2}(1+\sqrt{2}\cos(\theta))^2\,.$$

Thus $\psi_1(t) = \frac{\sqrt{2}}{2}t + \sqrt{1}4t^2$ and $\psi_2(t) = \frac{1}{4}t^2$. After some computations from Proposition 6.9 we have $\sum_{j=1}^{g} f(\theta_j) = 0$. It follows that $\cos(\theta_j) = -\frac{1}{\sqrt{2}}$ and hence $\alpha_j + \bar{\alpha}_j = -2q_0$; the result follows. □

6.2. Asymptotic Problems

In this section we survey a few results related to Tsfasman-Vlăditţ-Zink improvement on the Gilbert-Varshamov bound. The key matter is to find a family of curves $(\mathcal{X}_g)$ (indexed by its genus and defined over a fixed field $\mathbb{F}_q$) such that

$$A(q) := \limsup_g \frac{N_q(g)}{g}$$

be as large as possible. This number was introduced by Ihara [53] (and the inverse value was considered by Manin, loc. cit.). Ihara showed

$$N_q(g) \le q + 1 + \frac{1}{2}\sqrt{(8q+1)g^2 + 4(q^2-q)g} - g$$

and thus if $g > \sqrt{q}(\sqrt{q}-1)/2$, $N_q(g)$ is less than the HW-bound. From the upper bound on $N_q(g)$ above it follows that

$$A(q) \le \frac{1}{2}(\sqrt{8q+1}-1)\,.$$

Vlăduţ and Drinfeld [98] improve this bound and show that indeed

$$A(q) \le \sqrt{q} - 1\,.$$

To find lower bounds on $A(q)$ one needs to produce families of curves with many points. Serre used class field theory [87], [88] to show that $A(q) \ge \gamma_q$ with γ_q a positive constant depending of q (see also [74]). We have a stronger result, namely $N_q(g) > \gamma_q g$ for any g (see Elkies et al. [19]). Ihara used supersingular points on a family of modular curves $(\mathcal{X}_g)$ to show that, when q is an square, one can take $\gamma_q = \sqrt{q} - 1$ and hence

$$A(q) = \sqrt{q} - 1\,. \tag{6.8}$$

The GG-codes constructed on the respective curves $(\mathcal{X}_g)$ above have the best asymptotic parameters that can be constructed so far; for practical

applications one needs an explicit description of the aforementioned codes; this task seems to be very hard in the case of modular curves. Garcia and Stichtenoth proved (6.8) via curves defined by "explicit equations" (see [26], [27]). It is an intrigued fact that Garcia and Stichtenoth curves are also modular curves (see Elkies [18]).

For $q = p^{2m+1}$, it seems that the true value of $A(q)$ is unknown. Zink showed $A(p^3) \geq 2(p^2-1)/(p+2)$ (curves with no explicit equations). van der Geer and van der Vlugt [33] for $q = 8$ and Bezerra et al. [7] for any q as above generalized Zink's bound (curves with explicit equations).

Further asymptotic results on $N_q(g)$ which implies consequence both for $A(q)$ and $A^-(q) := \liminf_g N_q(g)/g$ can be found in the quite nice references [56] and [19] (see also the references therein).

6.3. Zeta-functions and Linear Series

Let $\mathcal{X}$ be a curve of genus g over $\mathbb{F}_q$ such that $\#\mathcal{X}(\mathbb{F}_q) > 0$. Let $L(t) = L(\mathcal{X}, q; t)$ be the enumerator of the zeta-function of $\mathcal{X}$ over $\mathbb{F}_q$. We consider the function $h(t)$ defined in (6.7), namely

$$h(t) = t^{2g} L(t^{-1}) = \prod_{j=1}^{g} (t - \alpha_j)(t - \bar{\alpha}_j)\,,$$

where the α_j are defined in Proposition 6.1. Then $h(t)$ is monic, of degree $2g$ whose independent term is non-zero; moreover, $h(t)$ is the characteristic polynomial of the Frobenius morphism $\mathbf{\Phi}_{\mathcal{J}}$ on the Jacobian $\mathcal{J}$ of the curve $\mathcal{X}$ (here we consider $\mathbf{\Phi}_{\mathcal{J}}$ as an endomorphism on a Tate module). Let

$$h(t) = \prod_j h_j^{r_j}(t)$$

be the factorization of $h(t)$ in $\mathbb{Z}[t]$. Since $\mathbf{\Phi}_{\mathcal{J}}$ is semisimple and the representation of endomorphisms of $\mathcal{J}$ on the Tate module is faithfully, see [93, Thm. 2], [58, VI§3], it follows that

$$\prod_j h_j(\mathbf{\Phi}_{\mathcal{J}}) = 0\,. \tag{6.9}$$

Let $\mathbf{\Phi}$ denote the Frobenius morphism on $\mathcal{X}$. Let $\pi : \mathcal{X} \to \mathcal{J}$ be the natural morphism $P \mapsto [P - P_0]$, where $P_0 \in \mathcal{X}(\mathbb{F}_q)$. We have

$$\pi \circ \mathbf{\Phi} = \mathbf{\Phi}_{\mathcal{J}} \circ \pi$$

and thus (6.9) implies the following linear equivalence of divisors on $\mathcal{X}$

$$\prod_j h_j(\mathbf{\Phi}) \sim mP_0\,, \quad \text{where } P \in \mathcal{X} \text{ and } m := \prod_j h_j(1)\,. \tag{6.10}$$

This suggests the study of the linear series

$$\mathcal{D} := |mP_0|\,.$$

Let us write

$$\prod_j h_j(t) = t^U + \alpha_1 t^{U-1} + \alpha_2 t^{U-2} + \ldots + \alpha_{U-1}t + \alpha_U\,.$$

We assume:

(A) $\alpha_1 \geq 1$, $\alpha_j \geq 0$ for $j = 2, \ldots, U$ (we already known that $\alpha_U \neq 0$);
(B) $\alpha_{j+1} \geq \alpha_j$ for $j = 1, \ldots, U-1$.

Remark 6.11. There are curves which do not satisfy conditions (A) and (B) above; cf. [9].

Next we compute some invariants of the linear series $\mathcal{D}$ above according to the results in the Appendix; we use the notation of that Appendix. Let r be the dimension of $\mathcal{D}$. For $P \in \mathcal{X}(\mathbb{F}_q)$ we have the following sequence of non-gaps at P:

$$0 = m_0(P) < m_1(P) < \ldots < m_{r-1}(P) < m_r(P) = m\,.$$

Lemma 6.12.

(1) If $P \in \mathcal{X}(\mathbb{F}_q)$, then the $(\mathcal{D}, P)$-orders are

$$0 = m - m_r(P) < m - m_{r-1}(P) < \ldots < m - m_1(P) < m - m_0(P)\,;$$

(2) If

$$P \notin \mathcal{X}(\mathbb{F}_q) \cup \mathcal{X}(\mathbb{F}_{q^2}) \cup \ldots \cup \mathcal{X}(\mathbb{F}_{q^U})$$

then $j_1(P) = 1$;

(3) The numbers $1, \alpha_1, \ldots, \alpha_U$ are orders of $\mathcal{D}$;

(4) If $\mathbf{\Phi}^i(P) \neq P$ for $i = 1, 2, \ldots, U+1$, then α_U is a non-gap at P. In particular, α_U is a generic non-gap of X;

(5) If $\mathbf{\Phi}^i(P) \neq P$ for $i = 1, 2, \ldots, U$ and $\mathbf{\Phi}^{U+1}(P) = P$, then $\alpha_U - 1$ is also a non-gap at P.

Proof. The proof of (1), (2) or (3) is similar to [22, Thm. 1.4, Prop. 1.5]. To show the other statements, let us apply $\mathbf{\Phi}_*$ in (6.10); thus

$$\alpha_U P \sim \mathbf{\Phi}^{U+1}(P)+(\alpha_1-1)\mathbf{\Phi}^{U}(P)+(\alpha_2-\alpha_1)\mathbf{\Phi}^{U-1}(P)+\ldots+(\alpha_U-\alpha_{U-1})\mathbf{\Phi}(P)\,.$$

Then (4) and (5) follow from hypothesis (A) and (B) above. □

We finish this section with some properties involving rational points.

Proposition 6.13. *Suppose that* $\mathrm{char}(\mathbb{F}_q)$ *does not divide* m.

(1) If $\#\mathcal{X}(\mathbb{F}_q) \geq 2g+3$, *then there exists* $P \in \mathcal{X}(\mathbb{F}_q)$ *such that* $(m-1)$ *and* m *are non-gaps at* P;
(2) The linear series $\mathcal{D}$ *is simple*; *i.e., the morphism* $\pi : \mathcal{X} \to \pi(\mathcal{X}) \subseteq \mathbb{P}^r(\bar{\mathbb{F}}_q)$ *defined by* $\mathcal{D}$ *is birational*.

Proof. (1) Following [103], let $P \neq P_0$ be a rational point. We have $mP \sim mP_0$ by (6.10). Let $x : \mathcal{X} \to \mathbb{P}^1(\bar{\mathbb{F}}_q)$ be a rational function with $\mathrm{div}(x) = mP - mP_0$. Let n be the number of rational points wchich are unramified for x. Then by Riemann-Hurwitz $2g-2 \geq m(-2)+2(m-1)+(\#\mathcal{X}(\mathbb{F}_q)-n-2)$ so that $n \geq \#\mathcal{X}(\mathbb{F}_q)-(2g+2) \geq 1$. Thus there exists $Q \in \mathcal{X}(\mathbb{F}_q)$, $Q \neq P, P_0$ such that $\mathrm{div}(x-a) = Q + D - mP_0$ with $D \in \mathrm{Div}(\mathcal{X})$, $P_0, Q \notin \mathrm{Supp}(D)$. Let y be a rational function such that $\mathrm{div}(y) = mP_0 - mQ$. Then $\mathrm{div}((x-a)y) = D-(m-1)Q$ and the proof is complete.

(2) Let $Q \in \mathcal{X}(\mathbb{F}_q)$ be the point in (1) and $x, y \in \mathbb{F}_q(X)$ be such that $\mathrm{div}_\infty(x) = (m-1)Q$ and $\mathrm{div}_\infty(y) = mQ$. Then $\mathbb{F}_q(\mathcal{X}) = \mathbb{F}_q(x,y)$ and we are done. □

Proposition 6.14.

(1) $\epsilon_r = \nu_{r-1}$;
(2) Let $P \in \mathcal{X}(\mathbb{F}_q)$ *and suppose that* $\#\mathcal{X}(\mathbb{F}_q) \geq q(m-\alpha_U)+2$. *Then* $j_{r-1}(P) < \alpha_U$; *in particular,* $\epsilon_r = \alpha_U$ *and* P *is a* $\mathcal{D}$*-Weierstrass point*;
(3) If $\#\mathcal{X}(\mathbb{F}_q) \geq q\alpha_U + 1$, *then* $\#\mathcal{X}(\mathbb{F}_q) = q\alpha_U + 1$ *and* $m_1(P) = \alpha_U$ *for any* $P \in \mathcal{X}(\mathbb{F}_q)$.

Proof. (1) We have $\#\mathcal{X}(\mathbb{F}_q) \leq qm_1(P)+1$ by Lewittes [66, Thm. 1(b)]. Then the result follows from Lemma 6.12.

(2) Let $P \in \mathcal{X}(\mathbb{F}_q)$. We have $m_1(P) \leq m_1(Q)$, where Q is a generic point of $\mathcal{X}$ (apply the Appendix to the canonical linear series on $\mathcal{X}$). Therefore, $m_1(Q) \leq \alpha_U$ by Lemma 6.12 and hence $q\alpha_U + 1 \leq \mathcal{X}(\mathbb{F}_q) \leq qm_1(P)+1 \leq q\alpha_U + 1$. □

6.4. A Characterization of the Suzuki Curve

This section is based on [24]; it is a nice application of the interplay of Section 6.3 and the Appendix. Throughout, we let $q_0 = 2^s > 2$ be a power of two and set $q := 2q_0^2$. As we mentioned in the Introduction, the Suzuki curve $\mathcal{S}$ is the unique curve over $\mathbb{F}_q$ defined by the following data:

(I) genus: $g = q_0(q-1)$;
(II) number of $\mathbb{F}_q$-rational points: $N_1 = q^2 + 1$;
(III) $\mathbb{F}_q$-automorphism group equals the Suzuki group.

Our aim is to show the following.

Theorem 6.15. *Let $\mathcal{X}$ be a curve of genus $g = q_0(q-1)$ over $\mathbb{F}_q$ such that $N_1 = \#\mathcal{X}(\mathbb{F}_q) = q^2 + 1$. Then $\mathcal{X}$ is isomorphic to the Suzuki curve $\mathcal{S}$.*

We first show some lemmas. The reference "Lemma A" below is placed in the Appendix.

Let $\mathcal{X}$ be as in the theorem. Let $h(t) = t^{2g}L(t^{-1})$ be the polynomial defined in (6.7). The starting point of the proof is Proposition 6.10; thus

$$h(t) = (t^2 + 2q_0t + q)^g\,.$$

Let $\mathbf{\Phi} : \mathcal{X} \to \mathcal{X}$ be the Frobenius morphism on $\mathcal{X}$. From Section 6.3 we conclude that $\mathcal{X}$ is equipped with the linear series

$$\mathcal{D} := |(1 + 2q_0 + q)P_0|\,, \quad P_0 \text{ a rational point}$$

such that for any $P \in \mathcal{X}$

$$\mathbf{\Phi}^2(P) + 2q_0\mathbf{\Phi}(P) + q_0P \sim (1 + 2q_0 + q)P_0\,. \tag{6.11}$$

Let r denote the dimension of $\mathcal{D}$. We already know that $m = m_r(P) = 1 + 2q_0 + q$ for any $P \in \mathcal{X}(\mathbb{F}_q)$. Lemma 6.12 and Proposition 6.14 imply the following properties:

(1) $m_1(P) = q$ and $j_{r-1}(P) = 1 + 2q_0$ for any $P \in \mathcal{X}(\mathbb{F}_q)$;
(2) $\epsilon_1 = 1$ and $\epsilon_r = \nu_{r-1} = q$.

Lemma 6.16. *$r \geq 3$ and $\epsilon_{r-1} = 2q_0$.*

Proof. By Lemma 6.12 the numbers 1, $2q_0$ and q are orders of $\mathcal{D}$ and thus $r \geq 3$. Since $\epsilon_{r-1} \leq j_{r-1}(P) = 1 + 2q_0$ (Lemma A) and $\epsilon_r = q$ we have

$$2q_0 \leq \epsilon_{r-1} \leq 1 + 2q_0\,.$$

Suppose that $\epsilon_{r-1} = 1 + 2q_0$ (observe that $2q_0$ is also an order of $\mathcal{D}$). Let $P \in \mathcal{X}(\mathbb{F}_q)$. By Lemma A

$$\nu_{r-2} \le j_{r-1}(P) - j_1(P) \le \epsilon_{r-2} = 2q_0 \,.$$

Thus the sequence of Frobenius orders of $\mathcal{D}$ would be $\epsilon_0, \epsilon_1, \ldots, \epsilon_{r-2}, \epsilon_r$. Now for any $P \in \mathcal{X}(\mathbb{F}_q)$ (Lemma A)

$$\begin{aligned} v_P(S) &\ge \sum_{i=0}^{r-1} (j_{i+1}(P) - \nu_i) \\ &= \sum_{i=0}^{r-2} (j_{i+1}(P) - \nu_i) + (j_r(P) - \nu_{r-1}) \ge (r-1) j_1(P) + 1 + 2q_0 \end{aligned}$$

so that

$$\deg(S) \ge (r + 2q_0) N_1 \,. \tag{6.12}$$

From the following identities

- $2g - 2 = (2q_0 - 2)(1 + 2q_0 + q) = (2q_0 - 2) m_r(P)$,
- $N_1 = (1 - 2q_0 + q)(1 + 2q_0 + q) = (1 - 2q_0 + q) m_r(P)$,

inequality (6.12) becomes

$$(2q_0 - 2) \sum_{i=0}^{r-1} \nu_i + (r + q) \ge (r + 2q_0)(1 - 2q_0 + q) \,.$$

Since $\nu_{r-1} = q$ it follows that

$$\sum_{i=0}^{r-2} \epsilon_i = \sum_{i=0}^{r-2} \nu_i \ge (r-1) q_0 \,.$$

Now we use a property involving the orders of $\mathcal{D}$ (see [20]): $\epsilon_i + \epsilon_j \le \epsilon_{i+j}$ for $i + j \le r$. We apply this in the form $\epsilon_i + \epsilon_j \le \epsilon_{r-2}$ with $i + j = r - 2$. Thus

$$2 \sum_{i=0}^{r-2} \epsilon_i \le (r-1) \epsilon_{r-2} = (r-1) 2q_0 \,.$$

We conclude that $\epsilon_i + \epsilon_{r-2-i} = \epsilon_{r-2}$ for $i = 0, 1, \ldots, r-2$. In particular, $\epsilon_{r-3} = 2q_0 - 1$ and the p-adic criterion (cf. [91, Cor. 1.9]) would imply $\epsilon_i = i$ for $i = 0, 1, \ldots, r-3$. These facts imply $r = 2q_0 + 2$. Finally, we are going to see that this is a contradiction according to Castelnuovo's genus bound applied to $\mathcal{D}$; we must have

$$2g = 2q_0(q-1) \le \frac{(q + 2q_0 - (r-1)/2)^2}{r-1} \,.$$

For $r = 2q_0 + 2$ this gives $2q_0(q-1) < (q+q_0)^2/2q_0 = q_0 q + q/2 + q_0/2$, a contradiction. □

Remark 6.17. We write an alternative proof of the previous lemma. We have $2q_0 \le \epsilon_{r-1} \le j_{r-1}(P) = 2q_0 + 1$. Suppose $\epsilon_{r-1} = 2q_0 + 1$ and thus $\epsilon_{r-2} = 2q_0$. For any $P \in \mathcal{X}(\mathbb{F}_q)$, $\epsilon_{r-2} \le j_{r-2}(P) < j_{r-1}(P) = 1 + 2q_0$; thus $j_{r-2}(P) = 2q_0$ and $1 + q \in H(P)$. If we take $\tilde{P} \in \mathcal{X}(\mathbb{F}_q)$ such that $1+2q_0+q, 2q_0+q \in H(\tilde{P})$ (Proposition 6.13), $H(\tilde{P})$ contains the semigroup

$$H := \langle q, q+1, 2q_0+q, 1+2q_0+q \rangle$$

and hence $g \le g(H) := (\mathbb{N}_0 \setminus H)$. However, one shows that $g > g(H)$ as in Remark 6.20 below.

Lemma 6.18. *There exists $P \in \mathcal{X}(\mathbb{F}_q)$ such that the following properties hold true*:

(1) $j_1(P) = 1$;
(2) $j_i(P) = \nu_{i-1} + 1$ *for* $i = 2, \ldots, r-1$.

Proof. Let $P \in \mathcal{X}(\mathbb{F}_q)$. In the proof of Lemma 6.16 we obtained the following inequality

$$v_P(S) \ge \sum_{i=0}^{r-2} (j_{i+1}(P) - \nu_i) + 1 + 2q_0 \ge (r-1) j_1(P) + 1 + 2q_0 \ge r + 2q_0 \,.$$

Thus it is enough to show that $v_P(S) = r + q_0$ for some point $P \in \mathcal{X}(\mathbb{F}_q)$. Suppose on the contrary that $v_P(S) \ge r + 2q_0 + 1$ for any $P \in \mathcal{X}(\mathbb{F}_q)$. Then arguing as in the proof of Lemma 6.16 we would have

$$\sum_{i=0}^{r-2} \nu_i \ge r q_0 + 1 \,.$$

As $\nu_i \le \epsilon_{i+1}$, then

$$1 + \sum_{i=0}^{r-2} \nu_i \le \sum_{i=0}^{r-1} \epsilon_i \le r \epsilon_{r-1}/2$$

and thus

$$r q_0 + 2 \le r \epsilon_{r-1}/2$$

so that $\epsilon_{r-1} > 2q_0$ which is a contradiction according to Lemma 6.16. □

Lemma 6.19.

(1) ϵ_2 *is a power of two*;
(2) $\nu_1 > \epsilon_1 = 1$.

Proof. (1) It is a consequence of the p-adic criterion [91, Cor. 1.9].

(2) Suppose that $\nu_1 = 1$. Let P be a $\mathbb{F}_q$-rational point satisfying Lemma 6.18. Then $j_2(P) = 2$ and thus by Lemma 6.12 the Weierstrass semigroup $H(P)$ at P contains the semigroup

$$H := \langle q, -1 + 2q_0 + q, 2q_0 + q, 1 + 2q_0 + q \rangle .$$

Therefore $g \le g(H) := \#(\mathbb{N}_0 \setminus H)$. This is a contradiction as we will see in the remark below. □

Remark 6.20. Let H be the semigroup defined above. We are going to show that $g(H) = g - q_0^2/4$. To begin with we notice that $L := \cup_{i=1}^{2q_0-1} L_i$ is a complete system of residues module q, where

$$\begin{aligned}
L_i &= \{iq + i(2q_0 - 1) + j : j = 0, \dots, 2i\} \quad \text{if } \ 1 \le i \le q_0 - 1,\\
L_{q_0} &= \{q_0 q + q - q_0 + j : j = 0, \dots, q_0 - 1\},\\
L_{q_0+1} &= \{(q_0 + 1)q + 1 + j : j = 0, \dots, q_0 - 1\},\\
L_{q_0+i} &= \{(q_0 + i)q_0 + (2i - 3)q_0 + i - 1 + j : j = 0, \dots, q_0 - 2i + 1\} \cup\\
&\quad \{(q_0 + i)q + (2i - 2)q_0 + i + j : j = 0, \dots q_0 - 1\} \quad \text{if } \ 2 \le i \le q_0/2,\\
L_{3q_0/2+i} &= \{(3q_0/2 + i)q + (q_0/2 + i - 1)(2q_0 - 1) + q_0 + 2i - 1 + j :\\
&\quad j = 0, \dots, q_0 - 2i - 1\} \quad \text{if } \ 1 \le i \le q_0/2 - 1.
\end{aligned}$$

Moreover, for each $\ell \in L$, $\ell \in H$ and $\ell - q \notin H$. Hence $g(H)$ can be computed by summing up the coefficients of q from the above list (see e.g. [86, Thm. p.3]); i.e.

$$\begin{aligned}
g(H) = &\textstyle\sum_{i=1}^{q_0-1} i(2i + 1) + q_0^2 + (q_0 + 1)q_0 + \sum_{i=2}^{q_0/2}(q_0 + i)(2q_0 - 2i + 2) +\\
&\textstyle\sum_{i=1}^{q_0/2-1}(3q_0/2 + i)(q_0 - 2i) = q_0(q - 1) - q_0^2/4 .
\end{aligned}$$

In the remaining part of this chapter we let P_0 be a point satisfying Lemma 6.18. We set $m_i := m_i(P_0)$ and denote by $v = v_{P_0}$ the valuation at P_0.

By Lemma 6.19 the Frobenius orders of $\mathcal{D}$ are $\nu_0 = 0, \nu_1 = \epsilon_2, \dots, \nu_{r-1} = \epsilon_r$ and thus

$$\begin{cases}
m_i = 2q_0 + q - \epsilon_{r-i} \text{ if } i = 1, \dots, r - 2,\\
m_{r-1} = 2q_0 + q,\\
m_r = 1 + 2q_0 + q.
\end{cases} \tag{6.13}$$

Let $x, y_2, \dots, y_r \in \mathbb{F}_q(\mathcal{X})$ be rational functions such that $\mathrm{div}_\infty(x) = m_1 P_0$, and $\mathrm{div}_\infty(y_i) = m_i P_0$ for $i = 2, \dots, r$. The fact $\nu_1 > 1$ means that the

following matrix

$$\begin{pmatrix} 1 & x^q & y_2^q & \dots & y_r^q \\ 1 & x & y_2 & \dots & y_r \\ 0 & 1 & D_x^1 y_2 & \dots & D_x^1 y_r \end{pmatrix}$$

has rank two (cf. [91, Sect. 2]). Here $D_x^j y_i$ denotes the jth Hasse derivative (see e.g. [83], [84], [44]). In particular,

$$y_i^q - y_i = D_x^1 y_i (x^q - x) \quad \text{for } i = 2, \dots, r\,. \tag{6.14}$$

Lemma 6.21.

(1) For $P \in \mathcal{X}(\mathbb{F}_q)$, the divisor $(2g-2)P$ is canonical; in particular, the Weierstrass semigroup at P is symmetric;
(2) Let $n \in H(P_0)$. If $n < 2q_0 + q$, then $n \leq q_0 + q$;
(3) For $i = 2, \dots, r$ there exists $g_i \in \mathbb{F}_q(\mathcal{X})$ such that $D_x^1 y_i = g_i^{\epsilon_2}$. Furthermore, $\mathrm{div}_\infty(g_i) = \frac{qm_i - q^2}{\epsilon_2} P_0$.

Proof. (1) Let $P \in \mathcal{X}(\mathbb{F}_q)$. We have $m_r P \sim m_r P_0$ by (6.11) and $2g-2 = (2q_0-2)m_r$. Thus we can assume $P = P_0$. Let t be a local parameter at P_0. We shall show that $v(\frac{dx}{dt}) = 2g-2$. The equation $i = r$ in (6.14) by $\frac{dx}{dt}$ and the product rule give

$$\frac{dx}{dt}(y_r^q - y_r) = \frac{dy_r}{dt}(x^q - x)\,;$$

from properties of valuations: $v(\frac{dx}{dt}) - qm_r = -m_r - (q^2+1)$; i.e.,

$$v(\frac{dx}{dt}) = (q-1)m_r - (1 - 2q_0 + q)m_r = (2q_0 - 2)m_r = 2g - 2\,.$$

(2) We know that the elements q, $2q_0 + q$ and $1 + 2q_0 + q$ belong to the Weierstrass semigroup $H(P_0)$ at P_0. Then the numbers

$$kq + j(2q_0 + q) + i(1 + 2q_0 + q) = (k + j + i)q + (j + i)q_0 + i$$

are also non-gaps at P_0 where $k, j, i \in \mathbb{N}_0$. Let $k = 2q_0 - 2$, $j + i = q_0 - 2$. Hence,

$$(2q_0 - 2)q + q - 4q_0 + j \quad \text{for } j = 0, \dots, q_0 - 2$$

are also non-gaps at P_0. Therefore, by the symmetry of $H(P_0)$, the elements below

$$1 + q_0 + q + j \quad \text{with } j = 0, \dots, q_0 - 2$$

are gaps at P_0 and the proof follows.

(3) Set $f_i := D^1_x y_i$. By Hasse-Schmidt [43, Satz 10] it is enough to show that

$$D^j_x f_i = 0 \ , \quad \text{for } 1 \le j < \epsilon_2 \, .$$

From Eqs 6.14 it is clear that $D^1_x f_i = 0$. Now as $\epsilon_2 > 2$ each matrix below has rank two (cf. [91, Sect. 1])

$$\begin{pmatrix} 1 & x & y_2 & \dots & y_r \\ 0 & 1 & D^1_x y_2 & \dots & D^1_x y_r \\ 0 & 0 & D^j_x y_2 & \dots & D^j_x y_r \end{pmatrix} \ , \quad 2 \le j < \epsilon_2 \, ;$$

consequently $D^j_x f_i = 0$ for $2 \le j < \epsilon_2$. Finally from the computations $v(g_i) = v(f_i)/\epsilon_2$ and $-qm_i = v(f_i) - q^2$ by (6.14) we find $v(f_i) = -qm_i + q_0$. If $P \neq P_0$, $\frac{\mathrm{d}f_i}{\mathrm{d}t} = \frac{\mathrm{d}y_i}{\mathrm{d}t}$ where $t = x - x(P)$ is a local parameter at P by Item (1). □

Lemma 6.22. *$\epsilon_2 = q_0$ and $r = 4$.*

Proof. By Lemma 6.16, $r \ge 3$. We claim that $r \ge 4$; otherwise, let g_2 be the rational function in Lemma 6.21(3). We have $v(g_2) = -q$ since $m_2 = 2q_0 + q$ and $\epsilon_2 = 2q_0$. Therefore there exist elements $a \neq 0$ and b in $\mathbb{F}_q$ such that $x = ag_2 + b$ (notice that $v(x) = -q$). The case $i = 2$ in (6.14) reads

$$\frac{y_2}{a}^q - \frac{y_2}{a} = g_2^{2q_0}(g_2^q - g_2) \, ;$$

let $v := y_2/a$, $u := g_2$ and set $w := v^{q_0} - u^{q_0+1}$. Thus

$$w^q - w = u^{q_0}(u^q - u)$$

and we find that $q_0 + q$ is a non-gap at P_0 (cf. [41, Lemma 1.8]). This contradiction eliminates the case $r = 3$.

Let $r \ge 4$ and $2 \le i < r$. We show that $\epsilon_2 = q_0$. The element $(qm_{r-2} - q^2)/\epsilon_2$ is a positive non-gap at P_0 and hence at least $m_1 = q$. Thus $m_{r-2} - q \ge \epsilon_2$ $(*)$ and $2q_0 - \epsilon_2 \ge \epsilon_2$ by (6.13); it follows that $q_0 \ge \epsilon_2$. Now by Lemma 6.21(2) $m_{r-2} \le q_0 + q$; from $m_{r-2} = 2q_0 + q - \epsilon_2$, $q_0 \le \epsilon_2$.

Finally we show that $r = 4$. As in $(*)$ we deduce that $m_2 - q \ge \epsilon_2$ and from (6.13) $2q_0 - \epsilon_{r-2} \ge \epsilon_2 = \ell$; i.e, $q_0 \ge \epsilon_{r-2} \ge \epsilon_2 = q_0$ so that $\epsilon_{r-2} = \epsilon_2$ and the proof follows. □

Proof of Theorem 6.15. Let $P_0 \in \mathcal{X}(\mathbb{F}_q)$ be as above. The case $i = 2$ in (6.14) and Lemma 6.21 give

$$y_2^q - y_2 = g_2^{q_0}(x^q - x) \ ;$$

moreover, $m_2 = q_0 + q$ and so $v(g_2) = -q$. Thus $x = ag_2 + b$ with a and b in $\mathbb{F}_q$, $a \neq 0$; in particular,

$$\frac{y_2}{a}^q - \frac{y_2}{a} = g_2^{q_0}(g_2^q - g_2)\,.$$

It follows that $\mathcal{X}$ is defined by the plane equation

$$v^q - v = u^{q_0}(u^q - u)\,,$$

where $v := y_2/a$ and $u := g_2$, and thus its automorphism group (over $\bar{\mathbb{F}}_q$) is the Suzuki group (Henn [46]). As the Suzuki group is simple it follows that it is also defined over $\mathbb{F}_q$. We conclude that $\mathcal{X}$ is isomorphic to the Suzuki curve by the statements (I), (II) and (III) stated at the beginning of this section.

6.5. Maximal Curves

Let $\mathcal{X}$ be a curve of genus $g > 0$ over $\mathbb{F}_q$ with $q = \ell^2$. The curve is called *maximal* if its number of rational points attains the HW-bound. By (6.2), the α_j's in Proposition 6.1 satisfies $\alpha_j = -\ell$ for any j. Thus the polynomial $h(t)$ in (6.7) is of the form

$$h(t) = (t + \ell)^{2g}\,.$$

Let $\mathbf{\Phi} : \mathcal{X} \to \mathcal{X}$ be the Frobenius morphism over $\mathbb{F}_q$. By Section 6.3 the curve $\mathcal{X}$ is equipped with the linear series

$$\mathcal{D} = |(1 + \ell)P_0|\,, \quad P_0 \text{ a rational point}$$

such that

$$\mathbf{\Phi}(P) + \ell P \sim (1 + \ell)P_0 \tag{6.15}$$

for any $P \in \mathcal{X}$; see the picture below

We already know that the Hermitian curve $\mathcal{H}$ is maximal. We can obtain many examples of maximal curves by taking into consideration the following Serre's remark (cf. [57]). Let $\mathcal{Y}$ be a curve over $\mathbb{F}_{\ell^2}$ and $\mathcal{X} \to \mathcal{Y}$ a non-constant morphism over $\mathbb{F}_{\ell^2}$; then $P(\mathcal{Y}, \ell^2; t)$ divides $P(\mathcal{X}, \ell^2; t)$. In particular, if $\mathcal{X}$ is maximal, $\mathcal{Y}$ is so. Therefore if G is a subgroup of the automorphisms group of $\mathcal{H}$, the quotiont curve $\mathcal{H}/G$ is also maximal; we remark that there exists maximal curves which are not covered by the Hermitian curve (see Example 6.31). van der Geer and van der Vlught constructed maximal curves via methods coming from linear codes. See Hirschfeld et al. [51] for a complete bibliography on maximal curves.

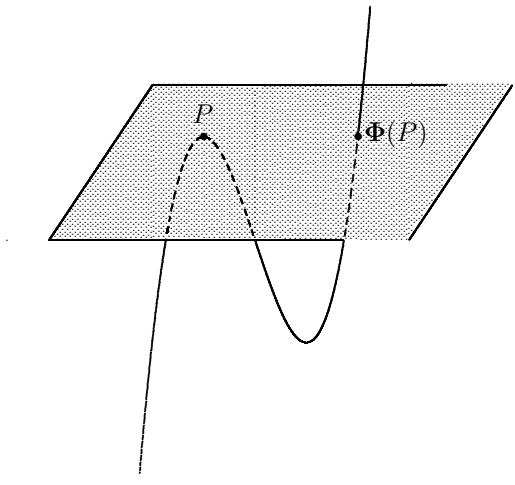


Fig. 6.1. Maximal curve

5.1 The linear series $\mathcal{D}$. Let r and

$$\pi = (f_0 : f_1 : \ldots : f_r)$$

be respectively the dimension and the morphism defined by $\mathcal{D}$. We use the notation of the Appendix. Set $\mathbb{P}^r := \mathbb{P}^r(\bar{\mathbb{F}}_{\ell^2})$, $\mathbb{P}^M := \mathbb{P}^M(\bar{\mathbb{F}}_{\ell^2})$.

By Proposition 6.14, $\epsilon_r = \ell$ which is equivalent (see e.g. [31]) to the existence of rational functions $w_0, w_1, \ldots, w_r$ (not all zero) such that

$$w_0^\ell f_0 + w_1^\ell f_1 + \ldots + w_r^\ell f_r = 0\,. \tag{6.16}$$

For $P \in \mathcal{X}$ let $v = v_P$ and $t = t_P$ denote respectively the valuation and a local parameter at P. We let $e = e_P := \min\{v(w_0), v(w_1), \ldots, v(w_r)\}$ and $z_i := t^{-e} w_i$.

Then for any $P \in \mathcal{X}$, the $\mathcal{D}$-osculating hyperplane at P is defined by

$$(z_0^\ell(P), z_1^\ell(P), \ldots, z_r^\ell(P))\,.$$

Hence from (6.15) and (6.16) we obtain the following dual relation

$$z_0 f_0^\ell + z_1 f_1^\ell + \ldots + z_r f_r^\ell = 0\,. \tag{6.17}$$

A natural question is the following: Is π an embedding?. Since $j_1(P) = 1$ for any P we have just to investigate whether or not π is injective. Let us consider the morphism

$$\phi := (w_0 : w_1 : \ldots : w_r) = (z_0 : z_1 : \ldots : z_r)\,.$$

Let M be the dimension of the linear series $\mathcal{D}'$ associated to ϕ. By (6.17) $\mathcal{D}'$ satisfies (6.15) in the sense that all the divisor of type $\mathbf{\Phi}(P)+qP \in \mathcal{D}'$; we notice that we may have $M < r$ since the w_i's may be linearly dependent. We obtain the following qualitative and quantitative properties of maximal curves [54].

Theorem 6.23.

(1) The morphism $\pi : \mathcal{X} \to \mathbb{P}^r$ is an embedding;
(2) The morphism $\phi : \mathcal{X} \to \mathbb{P}^r$ is an embedding; thus $\mathcal{X}$ is isomorphic to $\phi(\mathcal{X}) \subseteq \mathbb{P}^M$;
(3) Let us identify the curve $\mathcal{X}$ with its image $\pi(\mathcal{X}) \subseteq \mathbb{P}^r$. The curve is contained in an Hermitian variety;
(4) Let $\mathcal{Y} \subseteq \mathbb{P}^r$ be a curve of degree $\ell+1$ over $\mathbb{F}_q$. If $\mathcal{Y}$ is contained in an Hermitian variety, then $\mathcal{Y}$ is a maximal curve.

Proof. (sketch) (1) If $\pi(P)=\pi(Q)$, by (6.15) $\{P, \mathbf{\Phi}(P)\} = \{Q, \mathbf{\Phi}(Q)\}$. Let $P = \mathbf{\Phi}(Q)$ (and one shows that Q is rational). Let $\tilde{\mathbf{\Phi}} : \mathbb{P} \to \mathbb{P}$ denote the Frobenius morphism on $\mathbb{P}^r$. We have $\pi \circ \mathbf{\Phi} = \tilde{\mathbf{\Phi}} \circ \pi$ and hence $\pi(P)$ is rational; that is $\tilde{\mathbf{\Phi}}(\pi(P)) = \pi(P)$. After a change of coordinates we can assume $\pi(P) = (1 : 0 : \ldots : 0)$ with $f_0 = 1$ and $v(f_i) \geq 1$. Let $z_i(t) = z_i(P) + a_i^{(1)} t + \ldots$ for $i = 0, 1, \ldots, r$. From (6.16):

$$D = (z_0(P)f_0 + z_1(P)f_1 + \ldots + z_r(P)f_r) = -\sum_{i=0}^{r} f_i((a_i^1)^\ell t^\ell + \ldots)$$

We have to show that $v_P(D) = \ell + 1$. From the equation above,

$$v(D) = \ell + v\left(\sum_{i=0}^{r} f_i \left((a_i^{(1)})\ell + ..\right)\right)$$

As $v_P(f_i) \geq 1$ for $i \geq 1$ we just have to check that $a_0^{(1)} = 0$. This comes from (6.17).

(2) The proof is similar to (1).

(3) The linear series $\mathcal{D}'$ is a sub linear series of $\mathcal{D}$; in particular each z_j is a $\mathbb{F}_{\ell^2}$-linear combination of type $z_j = \sum_{i=1}^{r} a_{ij} f_i$. After some linear computations, the result follows from (6.17).

(4) See [54, Thm. 4.1]. □

Remark 6.24. The minimum dimension of the Hermitian variety which contains a maximal curve is $M = \dim(\mathcal{D}')$.

5.2 The Hermitian Curve. Notation as above. We notice that $r \geq 2$ by (6.15). We shall prove the following. We recall that the Hermitian curve can be also defined by the equation $y^\ell + y = x^{\ell+1}$.

Theorem 6.25. ([24], [72]) *Let $\mathcal{X}$ be a maximal curve over $\mathbb{F}_{\ell^2}$ of genus $g > 0$. The following statements are equivalent*:

(1) $\mathcal{X}$ is the Hermitian curve;
(2) $g > (\ell-1)^2/4$;
(3) $r = 2$.

Proof. The genus in (1) is $\ell(\ell-1)/2$ and (2) follows. Assume (2). Since $\mathcal{D}$ is simple we apply Castelnuovo's genus bound; i.e.,

$$2g \leq (2\ell - r + 1)^2/4(r-1)\,.$$

If $r \geq 3$, then $2g \leq (\ell-1)^2/4$, a contradiction. Now assume (3). To proof (1) we proceed as in Theorem 6.15. Let $x, y \in \mathbb{F}_{\ell^2}(\mathcal{X})$ whose pole divisor are respectively $\mathrm{div}_\infty(x) = \ell P_0$ and $\mathrm{div}_\infty y = (\ell+1)P_0$ (Lemma 6.12). Since $\nu_1 = \ell$ we have a relation of type

$$(y^{q^2} - y)D_x^1 x = (x^{\ell^2} - x)D_x^1 y\,, \tag{6.18}$$

Let $f := D_x^1 y$. Then $D_x^1 f = 0$. Now since $\epsilon_2 = \nu_1 = \ell$ (Proposition 6.14), for $i = 2, \ldots, < \epsilon_2 = \ell$ the rank of the following matrices is two:

$$\begin{pmatrix} 1 & x & y \\ 0 & 1 & f \\ 0 & 0 & D_x^i y \end{pmatrix}\,.$$

Thus $D_x^i y = 0$ for $i = 2, \ldots, \ell$ and from (6.18), $D_x^i f = 0$ for $i = 1, \ldots, \ell-1$. So by [43, Satz 10], f is a ℓ^2-th power, says $f = f_1^{\ell^2}$. From (6.18), $v_{P_0}(f) = -\ell^2$ and so $v_{P_0}(f_1) = -\ell$; thus $f_1 = ax + b$ with $a, b \in \mathbb{F}_{\ell^2}$, $a \neq 0$. If $x_1 := ax + b$ and $y_1 := ay$, the equation (6.18) becomes

$$y_1^{\ell^2} - y_1 = x_1^\ell (x_1^{\ell^2} - x_1);$$

therefore

$$(y_1^\ell + y_1 - x_1^{\ell+1})^\ell = y_1^\ell + y_1 - x_1^{\ell+1}$$

and the proof is complete. □

5.3. The genus. Here we discuss some properties concerning the genus g of a maximal curve over $\mathbb{F}_{\ell^2}$. First of all we notice that Theorem 6.25 implies the following restriction on g which was conjectured by Xing and

Stichtenoth [103]; see [21], [23]. (This gives a partial answer of a question of Serre [88].) We have

$$g \le g_2 := \lfloor (\ell-1)^2/4 \rfloor\,, \quad \text{or} \quad g = g_1 = \ell(\ell-1)/2\,. \tag{6.19}$$

Remark 6.26. Thus $N_{\ell^2}(g) < \ell^2+1+2\ell g$ for $g_2 < g < g_1$ (cf. Lauter [59]).

We already know that $g = g_1$ occurs only for the Hermitian curve. A similar property holds for $g = g_2$: the unique maximal curves of genus g_2 is the quotient of the Hermitian curve by certain involutions; these curves are defined by the following plane curves [22], [2], [55]

- $y^q + y = x^{(q+1)/2}$ if q is odd;
- $y^{q/2} + \ldots + y^2 + y = x^{q+1}$ if q is even.

We can improve (6.19) as follows. Let $g_3 := h(\ell+1,3) = \lfloor (\ell^2-\ell+4)/6 \rfloor$ denote the Halphen's number which asserts that any non-degenerate curve in $\mathbb{P}^3(\overline{\mathbb{F}}_{\ell^2})$ of degree $\ell+1$ of genus $g > g_3$ is contained in a quadratic surface. Thus, as the curve has many rational points, $g \ge g_2$.

Theorem 6.27. ([55]) *The genus g of a maximal curve over $\mathbb{F}_{\ell^2}$ satisfies*

$$g \le g_3 = \lfloor (\ell^2-\ell+4)/6 \rfloor\,, \quad \textit{or} \quad g = g_2 = \lfloor (\ell-1)^2/4 \rfloor \quad \textit{or} \quad g = g_1 = \ell(\ell-1)/2\,.$$

There exist examples of maximal curves of $g = g_3$: for example the quotient curves of the Hermitian curve by certain subgroups of order three; they are defined by the following plane equations [28], [13], [14]

- $x^{(\ell+1)/3} + x^{2(\ell+1)/3} + y^{\ell+1} = 0$ if $\ell \equiv 2 \pmod 3$;
- $\omega x^{(\ell-1)/3} - y x^{2(\ell-1)/3} + y^{\ell} = 0$ if $\ell \equiv 1 \pmod 3$, where $\omega \in \mathbb{F}_{\ell^2}$ such that $\omega^{\ell-1} = -1$;
- $y^{\ell} + y = (\sum_i^t x^{\ell/3})^2$ if $\ell = 3^t$.

Question 6.28. *There is a unique maximal curve of genus g_3 which is Galois covered by the Hermitian curve, namely the examples above [14, Prop. 2.1]. Is there exist a maximal curve of genus g_3 which is not covered by the Hermitian curve?*

For $\ell \not\equiv 0 \pmod 3$, we can improve Theorem 6.27 as follows.

Theorem 6.29. ([95]) *Let $\mathcal{X}$ be a maximal curve over $\mathbb{F}_{\ell^2}$ of genus g. Assume $\ell \not\equiv 0 \pmod 3$ and $r = 3$. If $(4\ell-1)(2g-2) > (\ell+1)(\ell^2-5\ell-2)$, then*

$$g \ge (q^2 - 2q + 3)/6\,.$$

Proof. First we show that $\epsilon_2 = 2$; on the contrary, $\epsilon_3 \geq 4$, by the p-adic criterion (here we use the hypothesis on ℓ). Let R and S be the ramification and $\mathbb{F}_{\ell^2}$-Frobenius divisor of $\mathcal{D}$ respectively. We have (Lemma A)

$$v_P(S) \geq j_2(P) + (j_3(P) - \epsilon_2) \geq 5 \quad \text{for any } P \in \mathcal{X}(\mathbb{F}_{\ell^2})$$

and so the maximality of $\mathcal{X}$ implies

$$\deg(S) = (\ell+1)(2g-2) + (\ell+3)(\ell+1) \geq 5(\ell+1)^2 + 5\ell(2g-2)\,.$$

It follows that

$$(\ell+1)(\ell^2 - 5\ell - 2) \geq (4\ell - 1)(2g-2)\,,$$

a contradiction. Now we use the ramification divisor R:

$$\deg(R) = (\ell+2+1)(2g-2) + 4(\ell+1) \geq (\ell+1)^2 + \ell(2g-2)$$

and thus $g \geq (\ell^2 - 2\ell + 3)/6$. □

Corollary 6.30. *Let $\mathcal{X}$, g and ℓ be as in the theorem above. If $g > (\ell - 1)(\ell-2)/6$, then*

$$g \geq (\ell^2 - 2\ell + 3)/6\,.$$

Proof. The hypothesis on g implies $r \leq 3$. If $r = 2$, then $g = \ell(\ell-1)/2$ by Theorem 6.25. Let $r = 3$; the hypothesis on g is equivalent to $(2g-2) > (\ell+1)(\ell-4)/3$ and hence

$$(4\ell-1)(2g-2) > (4\ell-1)(\ell+1)(\ell-4)/3 > (\ell+1)^2(\ell^2 - 5\ell - 2)$$

and the result follows. □

5.4 Examples. Throughout, by a maximal curve we mean a maximal curve over $\mathbb{F}_{\ell^2}$.

Example 6.31. (Curves covered by the Hermitian curve, I) We have already noticed that any curve covered by the Hermitian curve is also maximal. However, there exist maximal curves that cannot arise in this way. The first example of such a situation was given by Giulietti and Korchmáros [35]; their example is the case $m = 3$ of the nonsingular model of the curve defined in $\mathbb{P}^3(\overline{\mathbb{F}}_{\ell^{2m}})$ (m odd) by the equations

$$\begin{cases} z^{(\ell^m+1)/(\ell+1)} = yh(x) \\ (x^\ell + x)^{N/\ell} = y^{\ell+1} \end{cases}$$

where $h(x) = \sum_{i=0}^{N} (-1)^{i+1} x^{(\ell-1)i}$ and $N(\ell-1)+1 = (\ell^m+1)/(\ell+1)$. After some computations one shows that the curve is contained in an Hermitian

variety and that any irreducible component is defined over $\mathbb{F}_{\ell^{2m}}$; it follows that each irreducible component is maximal according to Theorem 6.23. In addition the genus of such components is $(\ell^m+1)(L\ell-2)/2+1$. By using the Riemann-Hurwitz genus formula and by counting rational points one concludes that such components cannot be covered by the Hermitian curve. We should say that we have no a theoretically explanation on the existence of these examples. We shall start with the question below.

Example 6.32. (Curves covered by the Hermitian curve, II) Let $\mathcal{X}$ be a maximal of genus g. By Theorem 6.27, $\mathcal{X}$ is covered by the Hermitian curve provided that

$$g > c(\ell) = (\ell^2 - \ell + 4)/6\,.$$

Question 6.33. *Shall we improve the bound $c(\ell)$?*

Notice that $c(\alpha)$ is the Halphen' s bound related to quadratic surfaces in $\mathbb{P}^3$; we may obtain further improvements on $c(\ell)$ by taking into considerations constraints that curves with many rational points may impose on surface of arbitrary degree.

Example 6.34. (On the uniqueness of maximal curves, I) Let $\mathcal{X}$ be a maximal curve of genus g. Let d be a divisor of $\ell+1$. The curve is defined by the plane curve

$$y^\ell + y = x^{(\ell+1)/d}$$

whenever there exists a rational point P of $\mathcal{X}$ such that $(\ell+1)/d$ belongs to the Weierstrass semigroup at P [22] (see also [1], [3] for analogous results).

Example 6.35. (On the uniqueness of maximal curves, II) Let d be a divisor of $\ell+1$. The previous example suggests to consider the uniqueness of maximal curves $\mathcal{X}$ of genus

$$g = \frac{1}{2}(\ell-1)(\frac{\ell+1}{d} - 1)\,.$$

If $d=2$, g coincides with Castelnuovo genus bound. In this case, the geometry of the curve equipped with the linear series $|2\mathcal{D}|$ implies the hypothesis on non-gaps above; thus there is a unique maximal curve of genus $(\ell-1)^2/4$ as we have pointed out above.

If $d=3$, g also coincides with Castelnuovo genus bound and as in the case above, the hypothesis on non-gaps hold true and there exists a unique maximal curve of genus $(\ell-1)(\ell-2)/6$.

Now observe that $g = (\ell-1)(\ell-2)/6$ is an integer for $\ell \equiv 2 \pmod 3$. However, for $\ell \geq 13$, there is no maximal curves having such a genus [55]. Here one uses a beautiful theorem due to Accola [4] concerning further constraints on curves whose genus equals Castelnuovo's genus bound.

Question 6.36. *Shall we exclude the hypothesis on non-gaps in Example 6.34?*

Example 6.37. (On the uniqueness of maximal curves, III) A maximal curve is not necessarily characterized via its genus.

(1) Let $\ell \equiv 3 \bmod 4$. Consider the maximal curve $\mathcal{X}$ and $\mathcal{Y}$ defined respectively by the plane curves:

$$x^{(\ell+1)/2} + y^{(\ell+1)/2} + 1 = 0\,, \quad \text{and} \quad y^\ell + y = x^{(\ell+1)/4}\,.$$

They have the same genus $g = (\ell-1)(\ell-3)/8$ but they are not isomorphic because the semigroup $\langle (\ell-1)/2, , (\ell+1)/2\rangle$ is a Weierstrass semigroup at some point of $\mathcal{X}$ but there is no point on $\mathcal{Y}$ satisfying this property [30], [12]. Moreover, in the last reference it is shown that the unique plane maximal curve of degree $(\ell+1)/2$ ($\ell \geq 11$ odd) is the curve $\mathcal{X}$ above.

(2) Let us consider maximal curves over $\mathbb{F}_{64}$. Let ϵ be a primitive 3th-root of unity.

Curve $\mathcal{X}$: The Hermitian curve is given by $x^9 + y^9 + 1 = 0$. Consider $T_1 : (x, y) \mapsto (x, \epsilon y)$. Thus the quotient curve $\mathcal{X}_1 := \mathcal{H}/ < T_1 >$ is defined by $u^9 + v^3 + 1 = 0\,(*)$. Now consider $T_2 : (u, v) \mapsto (\epsilon u, \epsilon^{-1} v)$. Then $\mathcal{X}_1/\langle T_2\rangle$ is defined by $z^4 + z = w^3$ (to see this we just multiply $(*)$ by u^3); clearly its genus is $g = 3$ and it is maximal since it is covered by the Hermitian curve (cf. Rodriguez [79], Luengo et al. [80]).

Curve $\mathcal{Y}$: Consider the maximal curve $\mathcal{Y}_1 : x^4 + x^2 + x = y^9$. We can use the automorphism $T_1 : (x, y) \to (x, \epsilon y)$ to obtain the maximal curve $\mathcal{Y} := \mathcal{Y}_1/\langle T_1\rangle$ of genus 3 defined by $u^4 + u^2 + u = w^3$.

Claim. The curves $\mathcal{X}$ and $\mathcal{Y}$ above are non-isomorphic over $\bar{\mathbb{F}}_{64}$ (cf. [29], [95]). There is just one point P_0 over $x = \infty$ or $u = \infty$. The number 5 does not belong to the Weierstrass semigroup at P_0 and so for both curves $\mathcal{D} = 4P_0$ is the canonical linear series. We apply the Appendix to $\mathcal{D}$ and one shows that the curve $\mathcal{X}$ and $\mathcal{Y}$ have 5 and 17 Weierstrass points respectively.

Let $\ell \not\equiv 0 \pmod 3$. Then by Corollary 6.30 the genus g of a maximal curve does not belong to the interval

$$[\lfloor \frac{1}{6}(\ell-1)(\ell-2)\rfloor + 1, \lceil \frac{1}{6}(\ell^2 - 2\ell + 3)\rceil - 1]\,. \tag{6.20}$$

Let $S(\ell)$ be the set of numbers that arise as the genus of maximal curves over $\mathbb{F}_{\ell^2}$. For $\ell \leq 5$, the set $S(\ell)$ is complete determined [28, Remark 6.1]; by taking into consideration such a remark we work out the following.

Example 6.38. Case $\ell = 7$. $g \leq g_3 = 7$ or $\{0, 1, 2, 3, 5, 7, 9, 21\} \subseteq S(7)$; $6 \notin S(7)$ by (6.20).

Case $\ell = 8$. $g \leq g_3 = 10$ or $\{0, 1, 2, 3, 4, 6, 7, 9, 10, 12, 28\} \subseteq S(8)$; $8 \notin S(8)$ by (6.20),

Case $\ell = 11$. $g \leq g_3 = 19$ or $\{0, 1, 2, 3, 4, 5, 7, 9, 10, 11, 13, 15, 18, 19, 25, 55\} \subseteq S(11)$; $16 \notin S(11)$ by (6.20),

Case $\ell = 13$. $g \leq g_3 = 26$ or $\{0, 1, 2, 3, 6, 9, 12, 15, 18, 26, 36, 78\} \subseteq S(13)$; $23, 24 \notin S(13)$ by (6.20). Moreover, $22 \notin S(13)$ (cf. Example 6.35).

Case $\ell = 16$. $g \leq g_3 = 40$ or $\{0, 1, 2, 4, 6, 8, 12, 24, 28, 56, 120\} \subseteq S(16)$; $36, 37 \notin S(16)$ by (6.20. Moreover, $35 \notin S(16)$ (cf. Example 6.35).

Question 6.39.

(1) Does 4 *(resp. 5) belong to* $S(7)$ *(resp.* $S(8)$*)?*
(2) Does $g = g_4 := \lceil \frac{1}{6}(\ell^2 - 2\ell + 3) \rceil$ *belong to* $S(\ell)$ *for infinitely many* ℓ*? (In each case above such a* g *exists).*
(3) What about the genus of a maximal curves in the interval $[g_4, g_3 - 1]$*?*

Example 6.40. (Plane maximal curves) Here we consider (nonsingular) plane maximal curves (over $\mathbb{F}_{\ell^2}$)

(1) Fermat curves: $X^m + Y^m + Z^m = 0$. Clearly the curve is maximal if $m \mid (\ell + 1)$. Tafazolian [92] proved that in fact the curve is maximal only if this condition holds.

(2) Hurwitz curves (cf. [5], [29]). Let $\mathcal{H}_n : X^n Y + Y^n Z + Z^n X = 0$. This curve is covered by the Fermat curve

$$U^{n^2-n+1} + V^{n^2-n+1} + W^{n^2-n+1} = 0$$

(via an unramified morphism). In particular, if $(n^2 - n + 1) \mid (\ell + 1)$, $\mathcal{H}_n$ is maximal. Conversely, if $\mathcal{H}_n$ is maximal, $\ell + 1$ belongs to the Weierstrass semigroup at any rational point. After some computations via the Weierstrass semigroup at $P = (0 : 1 : 0)$, which is generated by the set

$$S = \{s(n-1) + 1 : s = 1, \ldots, n\},$$

one shows that $(q+1)$ is a multiple of $(n^2 - n + 1)$. As a numerical example we choose $n = 3$ and conclude that the Klein curve is maximal over $\mathbb{F}_{\ell^2}$ if and only if $\ell \equiv 6 \pmod 7$.

Appendix: On the Stöhr-Voloch theory.

In this appendix, we recall some results of Stöhr-Voloch paper [91] concerning Weierstrass points and Frobenius orders. Let $\mathcal{X}$ be a curve of genus g defined over $\bar{\mathbb{F}}_q$.

Let $\mathcal{D} \subseteq |E|$ be a base-point-free linear series of dimension r and degree D on $\mathcal{X}$. For $P \in \mathcal{X}$ and $i \geq 0$ an integer, we define sub-sets of $\mathcal{D}$ which will provide with geometric information on $\mathcal{X}$. Let $\mathcal{D}_i(P) := \{D \in \mathcal{D} : v_P(D) \geq i\}$ (here $D = \sum_P v_P(D)P$). We have $\mathcal{D}_i(P) = \emptyset$ for $i > D$,

$$\mathcal{D} \supseteq \mathcal{D}_0(P) \supseteq \mathcal{D}_1(P) \supseteq \ldots \supseteq \mathcal{D}_{d-1}(P) \supseteq \mathcal{D}_D(P)\,,$$

and each $\mathcal{D}_i(P)$ is a sub-linear series of $\mathcal{D}$ such that the codimension of $\mathcal{D}_{i+1}(P)$ in $\mathcal{D}_i(P)$ is at most one. If $\mathcal{D}_i(P) \supsetneqq \mathcal{D}_{i+1}(P)$, then the integer i is called a $(\mathcal{D}, P)$-*order*; thus by Linear Algebra we have a sequence of $(N+1)$ orders at P:

$$0 = j_0(P) < j_1(P) < \ldots < j_r(P) \leq d\,.$$

Notice that $\mathcal{D} = \mathcal{D}_0(P)$ since $\mathcal{D}$ is base-point-free by hypothesis. It is a fundamental result the fact that the sequence above is the same for all but finitely many points P of $\mathcal{X}$, see [91, Thm. 1.5]. This constant sequence is called the *order sequence* of $\mathcal{D}$ and will be denoted by

$$0 = \epsilon_0 < \epsilon_1 < \ldots < \epsilon_r\,.$$

The finitely many points P, where exceptional $(\mathcal{D}, P)$-orders occur, are called the $\mathcal{D}$-*Weierstrass points* of $\mathcal{X}$. There exists a divisor R on $\mathcal{X}$, the *ramification divisor* of $\mathcal{D}$, whose support is exactly the set of $\mathcal{D}$-Weierstrass points:

$$R = \mathrm{div}\left(\det\left(D_t^{\epsilon_i} f_j\right)\right) + \left(\sum_{i=0}^{r} \epsilon_i\right)\mathrm{div}(\mathrm{dt}) + (r+1)E\,,$$

where $\pi = (f_0 : f_1 : \ldots : f_r)$ is the morphism defined by $\mathcal{D}$, t a separating element of $\bar{\mathbb{F}}_\ell(\mathcal{X})|\bar{\mathbb{F}}_\ell$ and the operator D_t^i is the ith Hasse derivative (properties of these operators can be seen in Hefez's paper [44]). Moreover, the number of $\mathcal{D}$-Weierstrass points of $\mathcal{X}$ (counted with multiplicity) is the degree of R.

Now to deal with rational points over $\mathbb{F}_q$ we require that both $\mathcal{X}$ and $\mathcal{D}$ be defined over this field. Choose the coordinates f_i's above in such a way that $v_P(f_i) + v_P(E) = j_i(P)$, where v_P denotes the valuation at P. Set $L_i(P) = \langle f_i, \ldots, f_r \rangle$. Thus

$$\mathcal{D}_i(P) = \{\mathrm{div}(f) + E : f \in L_i(P)\}\,.$$

For $i = 0, \ldots, r-1$ set

$$S_i(P) := \mathcal{D}_{j_{i+1}}(P) \cap \ldots \cap \mathcal{D}_{j_r}(P) \quad \text{and}$$

$$T_i(P) := \cap_{D \in S_i} \operatorname{Supp}(D)\,.$$

This is a subspaces of the dual of $\mathbb{P}^r(\bar{\mathbb{F}}_q)$ whose projective dimension is i. Notice that

$$\{P\} = T_0(P) \subsetneq T_1(P) \subsetneq \ldots \subsetneq T_{r-1}(P)\,.$$

The spaces $T_{r-1}(P)$ and $T_1(P)$ are usually called the *$\mathcal{D}$-osculating hyperplane* and the *$\mathcal{D}$-tangent line* at P respectively.

Let $\mathbf{\Phi} : \mathcal{X} \to \mathcal{X}$ be the Frobenius morphism on $\mathcal{X}$. Suppose that for a generic P, $\mathbf{\Phi}(P) \in T_{N-1}(P)$. Then there exists an integer $1 \leq I \leq r-1$ such that $\phi(P) \in T_I(P) \backslash T_{I-1}(P)$. Define $\nu_j := \epsilon_j$ for $0 \leq j \leq I-1$ and $\nu_j = \epsilon_{j+1}$ for $j = I, \ldots, r-1$. The sequence $0 = \nu_0 < \nu_1 < \ldots < \nu_{N-1}$ is called the *Frobenius order sequence* of $\mathcal{D}$ (with respect to $\mathbb{F}_q$; cf. [91, Sect. 2]). The key property related with rational points in [91] is the existence of a divisor S, *the Frobenius divisor* of $\mathcal{X}$ (over $\mathbb{F}_q$) satisfying Lemma A(3)(4)(5)(6) below. This divisor is defined as follows. Let $\tilde{L}$ denote the determinant of the matrix whose rows are:

$$(f_0^\ell, f_1^\ell, \ldots, f_r^\ell)\,, \quad (D_t^{\nu_i} f_0, D_t^{\nu_i} f_1, \ldots, D_t^{\nu_i} f_r)\,, \quad i = 0, 1, \ldots, r-1\,.$$

Then

$$S := \operatorname{div}(\tilde{L}) + (\sum_{i=0}^{r-1} \nu_i)\operatorname{div}(\mathrm{dt}) + (q+r)E\,.$$

We notice that $\mathcal{X}(\mathbb{F}_q) \subseteq \operatorname{Supp}(S)$ and $v_P(S) \geq r$ for $P \in \mathcal{X}(\mathbb{F}_q)$ (Lemma below). Thus

$$\#\mathcal{X}(\mathbb{F}_q) \leq \deg(S)/r\,.$$

We subsume some properties of the ramification divisor and Frobenius divisor of $\mathcal{D}$.

Lemma A. Let $P \in \mathcal{X}$ and q be a power of a prime p.

(1) For each i, $j_i(P) \geq \epsilon_i$;
(2) $v_P(R) \geq \sum_{i=0}^{r} (j_i(P) - \epsilon_i)$; equality holds if and only if $\det\left(\binom{j_i(P)}{\epsilon_j}\right) \not\equiv 0 \pmod{p}$;
(3) If $P \in \mathcal{X}(\mathbb{F}_q)$, then for each i, $\nu_i \leq j_{i+1}(P) - j_1(P)$;
(4) If $P \in \mathcal{X}(\mathbb{F}_q)$, then $v_P(S) \geq \sum_{i=0}^{r-1} (j_{i+1}(P) - \nu_i)$; equality holds if and only if $\det\left(\binom{j_{i+1}(P)}{\nu_j}\right) \not\equiv 0 \pmod{p}$;

(5) If $P \in \mathcal{X}(\mathbb{F}_q)$, then $v_P(S) \geq rj_1(P)$;
(6) If $P \notin \mathcal{X}(\mathbb{F}_q)$, then $v_P(S) \geq \sum_{i=0}^{r-1}(j_i(P) - \nu_i)$.

Frobenius non classical plane curves. (Hefez-Voloch [45]) Let $\mathcal{X}$ be a plane curve of degree d defined over $\mathbb{F}_q$. We consider the linear series $\mathcal{D} := g_d^2$ (whose elements cuts out the curve by lines). Let $0 < \nu$ be the $\mathbb{F}_q$-Frobenius order sequence of $\mathcal{D}$. Assume that $\nu > 1$ (one usually says that $\mathcal{X}$ *is non-classical*). Thus the order sequence of $\mathcal{D}$ is $0 < 1 < \nu$ and hence

$$\deg(R) = (1+\nu)(2g-2) + 3d\,, \quad \text{and} \quad \deg(S) = \nu(2g-2) + (q+2)d\,.$$

The Hefez-Voloch used in this paper affirm

$$\#\mathcal{X}(\mathbb{F}_q) = \deg(S) - \deg(R) = d(q-d+2)\,.$$

References

[1] M. Abdón and A. Garcia, "On a characterization of certain maximal curves", Finite Fields Appl. **10**(2) (2004), 133–158).
[2] M. Abdón and F. Torres, "On maximal curves in characteristic two", Manuscripta Math. **99** (1999), 39–53.
[3] M. Abdón and F. Torres, "On $\mathbb{F}_{q^2}$-maximal curves of genus $(q-3)9/6$", Beitr. Algebra Geom., **46**(1) (2005), 241–260.
[4] R.D. Accola, "On Castelnuovo's inequality for algebraic curves I", Trans. Amer. Math. Soc. **251** (1979), 357–373.
[5] A. Aguglia, G. Korchmáros and F. Torres, "Plane maximal curves", Acta Arithm. **98**(2) (2001), 165–179.
[6] E. Arbarello, M. Cornalba, P.A. Griffiths and J. Harris, "Geometry of Algebraic Curves", Vol I. Springer-Verlag, New York, 1985.
[7] J. Bezerra, A. Garcia and H. Stichtenoth, "An explicit tower on function fields over cubic finite fields and Zink's lower bound", J. Reine Angew. Math. **589** (2005), 159–199.
[8] E. Bombieri, "Hilbert's 8th problem: An analogue", Proc. Symp. Pure Math. **28** (1976), 269–274.
[9] P. Carbonne and T. Henocq, "Décomposition de la Jacobienne sur les corps finis", Bull. Polish Acad. Sci. Math. **42**(3) (1994), 207–215.
[10] G. Castelnuovo, "Ricerche di geometria sulle curve algebriche", Atti. R. Acad. Sci. Torino **24** (1889), 196–223.
[11] L. Chiantini, C. Ciliberto, "Towards a Halphen theory of linear series on curves", Trans. Amer. Math. Soc. **356**(6) (1999), 2197–2212.
[12] A. Cossidente, J.W.P. Hirschfeld, G. Korchmáros and F. Torres, "On plane maximal curves", Compos. Math. **121** (2000), 163–181.
[13] A. Cossidente, G. Korchmáros and F. Torres, "On curves covered by the Hermitian curve", J. Algebra **216** (1999), 56–76.

[14] A. Cossidente, G. Korchmáros and F. Torres, "Curves of large genus covered by the Hermitian curve", Comm. Algebra **28**(10) (2000), 4707–4728.
[15] P. Deligne and G. Lusztig, "Representations of reductive groups over finite fields", Ann. of Math. **103** (1976), 103–161.
[16] M. Deuring, "Die Typen der Multiplicatorenringe elliptishher Funktionen Körper", Abh. Math. Sem. Univ. Hamburg **14** (1941), 197–272.
[17] L.E. Dickson, "History of the Theory of Numbers", Vol. II, Chelsea Publ. Comp., New York 1971.
[18] N.D. Elkies, "Explicit modular towers", Proceeding of the Thirty-Fifth Annual Allerton Conference on Communication, Control and Computing , Univ. of Illinois at Urbana -Champaign, E.T. Basar and A. Vardy Eds. (1998), 23–32.
[19] N.D. Elkies, E.W. Howe, A. Kresch, B. Poonen, J.L. Wetherell and M.E. Zieve, "Curves of every genus with many points, II: Asymptotically good families", Duke Math. J. **122**(2) (2004), 399–421.
[20] E. Esteves, "A geometric proof of an inequality of order sequences", Comm. Algebra **21**(1) (1993), 231–238.
[21] R. Fuhrmann, "Algebraische Funktionenkörper über endlichen Körpern mit maximaler Anzahl rationaler Stellen, Disseration, Universität GH Essen, 1995.
[22] R. Fuhrmann, A. Garcia and F. Torres, "On maximal curves", J. Number Theory **67**(1) (1997), 29–51.
[23] R. Fuhrmann and F. Torres, "The genus of curves over finite fields with many rational points", Manuscripta Math. **89** (1996), 103–106.
[24] R. Fuhrmann and F. Torres, "On Weierstrass points and optimal curves", Rend. Circ. Mat. Palermo, Suppl. **51** (Recent Progress in Geometry, E. Ballico and G. Korchmáros Eds.) (1998), 25–46,
[25] W. Fulton, "Algebraic Curves", Benjamin, New York, 1969.
[26] A. Garcia and H. Stichtenoth, "A tower of Artin-Schreier extensions of functions fields attaining the Drinfeld-Vlăduţ bound", Invent. Math. **121**(1) (1995), 211–233.
[27] A. Garcia and H. Stichtenoth, "On the asymptotic behavior of some towers of function fields over finite fields", J. Number Theory **6** (1996), 248–273.
[28] A. Garcia, H. Stichtenoth and C.P. Xing, "On subfields of the Hermitian function field", Compos. Math. **120** (2000), 137–170.
[29] A. Garcia and F. Torres"On unramified coverings of maximal curves", Proceedings AGCT-10, to appear.
[30] A. Garcia and P. Viana, "Weierstrass points on certain non-classical curves", Arch. Math. **46** (1986), 315–322.
[31] A. Garcia and J.F. Voloch, "Wronskians and linear independence in fields of prime characteristic", Manuscripta Math. **59** (1987), 457–469.
[32] G. van der Geer, "Error-Correcting Codes and Curves over Finite Fields", Mathematics Unlimited - 2001 and Beyond, Springer, 1115–1138, 2000.
[33] G. van der Geer and M. van der Vlugt, "An asymptotically good tower of curves over the finite field with eight elements", Bull. London Math. **24** (2002), 291–300.

[34] G. van der Geer and M. van der Vlugt, "Tables of curves with many points", Math. Comp. **69** (2000), 797–810. Updates of these tables are found in, Tables for the function $N_q(g)$, available from http://www.wins.uva.nl/ geer.
[35] M. Giulietti and G. Korchmáros, "A new family of $\mathbb{F}_{q^2}$-maximal curves", preprint, 2007.
[36] D.M. Goldschmidt, "Algebraic Functions and Projective Curves", Springer-Verlag, New York, 2003.
[37] V.D. Goppa, "Codes associated with divisors", Problems of Information Transmition **I**, 1977.
[38] V.D. Goppa, "Geometry and Codes", Mathematics and its applications, Vol. 24, Kluwer Academic Publisher, Dordrecht-Boston-London, 1988.
curves", Regional Conference Series in Math.
[39] J.P. Hansen, "Deligne-Lusztig varieties and group codes", Lect. Notes Math. **1518** (1992), 63–81.
[40] J.P. Hansen and J.P. Pedersen, "Automorphism group of Ree type, Deligne-Lusztig curves and function fields", J. Reine Angew. Math. **440** (1993), 99–109.
[41] J.P. Hansen and H. Stichtenoth, "Group codes on certain algebraic curves with many rational points", AAECC **1** (1990), 67–77.
[42] R. Hartshorne, "Algebraic Geometry", Grad. Texts in Math. Vol. 52, Springer-Verlag, New York/Berlin, 1977.
[43] H. Hasse and F.K. Schmidt, "Noch eine Begründung der höheren Differentialquotienten in einem algebraischen Funktionenkörper einer Unbestimmten; Zusatz bei der Korrektur", J. Reine Angew. Math. **177** (1937), 215–237.
[44] A. Hefez, "Non reflexive curves", Compos. Math. **69** (1989), 3–35.
[45] A. Hefez and J.F. Voloch, "Frobenius non classical curves", Arch. Math. **54** (1990), 263–273.
[46] H.W. Henn, "Funktionenkörper mit groβer Automorphismengruppe", J. Reine Angew. Math. **302** (1978), 96–115.
[47] A. Hefez and J.F. Voloch, ""Frobenius non classical curves", Arch. Math. **71** (1990), 263–273.
[48] J.P.W. Hirschfeld, "Projective Geometries Over Finite Spaces", 2nd ed. Oxford University Press, Oxford,1998.
[49] J.W.P. Hirschfeld and G. Korchmáros, "On the embedding of an arc into a conic in a finite plane", Finite Fields Appl. **2** (1996), 274–292.
[50] J.W.P. Hirschfeld and G. Korchmáros, "Arcs and curves over a finite field", Finite Fields Appl. **5** (1999), 393–408.
[51] J.W.P. Hirschfeld, G. Korchmáros and F. Torres, "Algebraic Curves Over a Finite Field", Princenton University Press, Princenton and Oxford, 2008.
[52] T. Ibukiyama, "On rational points of curves of genus 3 over finite fields", Tôhuku Math. J. **45** (1993), 311–329.
[53] Y. Ihara, "Some remarks on the number of rational points of algebraic curves over finite fields", J. Fac. Sci. Tokio **28** (1981), 721–724.
[54] G. Korchmáros and F. Torres, "Embedding of a maximal curve in a Hermitian variety", Compos. Math. **128** (2001), 95–113.

[55] G. Korchmáros and F. Torres, "On the genus of a maximal curve", Math. Ann. **323**(3) (2002), 589–608.
[56] A. Kresch, J.L. Wetherell and M. Zieve, "Curves of every genus with many points, I: Abelian and toric families", J. Algebra **250** (2002), 353–370.
[57] G. Lachaud, "Sommes d'Eisenstein et nombre de points de certaines courbes algébriques sur les corps finis", C.R. Acad. Sci. Paris **305** Série I (1987), 729–732.
[58] S. Lang, "Abelian Varieties", Interscience Pub., New York, 1950.
[59] K. Lauter, "Improved upper bounds for the number of rational points on algebraic curves over finite fields", C.R. Acad. Sci. Paris **328**(12) Série I (1999), 1181–1185.
[60] K. Lauter, "A formula for constructing curves over finite fields with many points", J. Number Theory **74** (1999), 56–72.
[61] K. Lauter, "Non-existence of a curve over $\mathbb{F}_3$ of genus 5 with 14 rational points", Proc. Amer. Math. Soc. **128** (2000), 369–374.
[62] K. Lauter, "Zeta-functions of curves over finite fields with many rational points", Buchmann, Johannes (ed.) et al., Coding Theory, Cryptography and Related Areas. Proceedings of an international conference, Guanajuato, Mexico, April 1988. Berlin: Springer 167–174 (2000).
[63] K. Lauter (with an Appendix by J.P. Serre), "Geometric methods for improving the upper bounds on the number of rational points on algebraic curves over finite fields", J. Alg. Geometry **10**(1) (2001), 19–36.
[64] K. Lauter, "The maximum or minimum number of rational points on curves of genus three over finite fields (with an appendix by J.P. Serre)", Compos. Math. **134** (2002), 87–111.
[65] E.W. Howe and K. Lauter, "Improved upper bounds for the number of points on curves over finite fields", Ann. Inst. Fourier **53**(6) (2003), 1677–1737.
[66] J. Lewittes, "Places of degree one in function fields over finite fields", J. Pure Appl. Alg. **69** (1990), 177–183.
[67] J.H. van Lint, "Introduction to Coding Theory", Springer-Verlag, third edition, 1999.
[68] J.H. van Lint and G. van der Geer, "Introduction to Coding Theory and Algebraic Geometry, Birkhäuser, 1988.
[69] J.S. Milne, Abelian Varieties, "Arithmetic Geometry" (G. Cornell and J.H. Silverman Eds.), 103–150, Springer-Verlag, New York, 1986.
[70] C. Moreno, "Algebraic curves over finite fields", Cambridge Tracts in Math. Vol. 97, Cambridge Univ. Press, Cambridge, 1991.
[71] D. Mumford, "Abelian Varieties", Tata Inst. Fund. Res., Oxford University Press, Bombay, 1994.
[72] C. Munuera and F. Torres, "Sobre curvas algebraicas y códigos correctores", La Gaceta de la RSME, **9**(1) (2006), 203–222.
[73] N. Namba, "Geometry of algebraic projective curves", Marcel Dekkers INC, New York and Basel, 1984.
[74] H. Niederreited and C. Xing,"Towers of global function fields with asymptotically many rational places and an improvement of the Gilbert-

Varshamov bound", Math. Nachr. **195** (1998), 171–186.
[75] H. Niederreited and C. Xing,"Global function fields with many rational places and their applications", Contemporary Math. **225** (1999), 87–111.
[76] H. Niederriter and C. Xing,"Rational Points on Curves over Finite Fields: Theory and Applications", London Mathematical Society Lecture Notes, London Mathematical Society Lecture Note Series **285**, Cambridge Univ. Press, Cambridge, 2001.
[77] J.P. Pedersen, "A function field related to the Ree group", Lect. Notes Math. **1518** (1992), 122–131.
[78] J. Rathmann, "The uniform position principle for curves in characteristic p", Math. Ann. **276**, (1987), 565–579.
[79] M.C. Rodriguez-Palánquex, "Arithmética de curvas quasihermíticas. Aplicaciones a los códigos geométricos de Goppa", Ph.D. Thesis, UCM, 2000.
[80] M.C. Rodriguez-Palánquex, I.J. García-Villalba and L. Luengo-Velasco, "Computing the genus of a class of curves", Applied Algebra, Algebraic Algorithms and Error Correcting Codes, Lectures Notes in Comput. Sci. **2227**, Springer, Berlin, 2001.
[81] H.E. Rosen, "A course in Number Theory", Clarendon Press - Oxford, 1988.
[82] H.G. Rück and H. Stichtenoth, "A characterization of Hermitian function fields over finite fields", J. Reine Angew. Math. **457** (1994), 185–188.
[83] F.K. Schmidt, "Die Wronskisch Determinante in belebigen differenzierbaren Funktionenkörpern", Math. Z. **45** (1939), 62–74.
[84] F.K. Schmidt, "Zur arithmetischen Theorie der algebraischen Funktionen II, Allgemeine Theorie der Weierstrasspunkte",
[85] W. Scharlau and H. Opolka, "From Fermat to Minkowski: Lectures on the Theory of Numbers and Its Historical Development", Springer-Verlag, New York, 1985.
[86] E.S. Selmer, "On the linear diophantine problem of Frobenius", J. Reine Angew. Math. **293/294** (1977), 1–17.
[87] J.P. Serre, "Sur le nombre des points rationnels d'une courbe algébrique sur un corps fini", C.R. Acad. Sci. Paris **296** Série I, (1983), 397–402. (Oeuvres III, 128, 658–663).
[88] J.P. Serre, "Rational points on curves over finite fields", Notes by F. Gouvea of lectures at Harvard University, 1985.
[89] S.A. Stepanov, "Arithmetic of Algebraic Curves", Monographs in Contemporary Mathematics, Consultans Bureau, New York - London, 1994.
[90] H. Stichtenoth, "Algebraic function fields and codes", Springer-Verlag, Berlin, 1993.
[91] K.O. Stöhr and J.F. Voloch, "Weierstrass points and curves over finite fields", Proc. London Math. Soc. (3) **52** (1986), pp. 1–19.
[92] S. Tafazolian, "On Supersingular Curves Over Finite Fields", Ph. D., IMPA. 2008.
[93] J. Tate, "Endomorphisms of abelian varieties over finite fields", Invent. Math. **2** (1966), 134–144.
[94] J. Top, "Curves of genus 3 over small finite fields", Indag. Mathem. N.S.

14(2) (2003), 275–283.
[95] F. Torres, "Maximal curves over $\mathbb{F}_{64}$", (2007), preprint.
[96] M.A. Tsfasman and S.G. Vlăduţ, "Algebraic-Geometric Codes", Mathematics and its applications, Vol. 58, Kluwer Academic Publisher, Dordrecht-Boston-London, 1991.
[97] M.A. Tsfasman, S.G. Vlăduţ and T. Zink, "On Goppa codes which are better than the Varshamov-Gilbert bound", Math. Nachr. **109** (1982), 21–28.
[98] S.G. Vlăduţ and V.G. Drinfeld, "Number of points of an algebraic curve", Funct. Anal. **17**(1) (1983), 68–69.
[99] J.F. Voloch, "Arcs in projective planes over prime fields", J. Geom. **38** (1990), 198–200.
[100] J.F. Voloch, "Complete arcs in Galois planes of non square order", Advances in Finite Geometries and Designs, (J.W.P. Hirschfeld et al., Eds.) Oxford Univ. Press, Oxford, 401–405, 1991.
[101] A. Weil, "Courbes Algébriques et Variétés Abeliennes", Hermann, Paris, 1971.
[102] A. Weil, *Number of solutions of equations in finite fields*, Bull. Amer. Math. Soc. **55** (1949), 497–508.
[103] C. Xing and H. Stichtenoth, "The genus of maximal functions fields", Manuscripta Math. **86** (1995), 217–224.

Chapter 7

Algebraic Geometry Codes from Higher Dimensional Varieties

John B. Little

Department of Mathematics and Computer Science,
College of the Holy Cross,
Worcester, MA 01610 USA
little@mathcs.holycross.edu

This chapter gives a general survey of work on Goppa-type codes from higher dimensional algebraic varieties. The construction and several techniques for estimating the minimum distance are described first. Codes from various classes of varieties, including Hermitian hypersurfaces, Grassmannians, flag varieties, ruled surfaces over curves, and Deligne-Lusztig varieties are considered. Connections with the theories of toric codes and order domains considered elsewhere in this volume are also briefly indicated.

Contents

7.1. Introduction

The codes considered in this chapter can all be understood as examples of *evaluation codes* produced from a finite set $\mathcal{S} = \{P_1, \ldots, P_n\}$ of $\mathbb{F}_q$-rational points on an algebraic variety X and an $\mathbb{F}_q$-vector space of functions $\mathcal{F}$ defined on $\mathcal{S}$. The set of codewords is the image of an evaluation mapping

$$\begin{aligned} ev_{\mathcal{S}} : \mathcal{F} &\longrightarrow \mathbb{F}_q^n \\ f &\mapsto (f(P_1), \ldots, f(P_n)). \end{aligned} \tag{7.1}$$

X will usually be assumed smooth, but in fact many of the constructions also make sense for normal varieties.

The Goppa $C_L(D, G)$ codes from curves X where $\mathcal{F} = L(G)$ for some divisor G on X were the first examples of codes of this type to be considered. Relatively early in the history of applications of algebraic geometry to coding theory, however, Tsfasman and Vladut proposed in Chapter 3.1 of [55] that higher dimensional varieties might also be used to construct codes. By the results of [46], *every* linear code can be obtained by the construction of Definition 7.1 below, starting from some $\mathcal{S} \subseteq X(\mathbb{F}_q)$ for some variety X and some line bundle $\mathcal{L}$ on X; indeed *curves* suffice for this (see Section 7.8). Hence the question is whether one can identify specific higher dimensional varieties X, spaces of functions $\mathcal{F}$, and sets of rational points $\mathcal{S}$ that yield particularly interesting codes using algebraic geometric constructions. There has been a fairly steady stream of articles since the 1990's studying such codes and our first main goal here is to survey the methods that have been developed and the results that have been obtained.

In a sense, the first major difference between higher dimensional varieties and curves is that points on X of dimension ≥ 2 are subvarieties of codimension ≥ 2, not divisors. This means that many of the familiar tools used for Goppa codes (e.g. Riemann-Roch theorems, the theory of differentials and residues, etc.) do not apply in exactly the same way.

A second difference is the possibility of performing *birational modifications* such as blowing up points or other subvarieties on a variety of higher dimension. For instance, if p is a point in a smooth algebraic variety X of dimension $\delta \geq 2$, there is another smooth variety $Y = \mathrm{Bl}_p(X)$, a proper morphism $\pi : Y \to X$, and an *exceptional divisor* $E \simeq \mathbb{P}^{\delta-1}$ in Y such that $\pi(E) = \{p\}$, and $\pi|_{Y-E} : Y - E \simeq X - \{p\}$ as varieties. Because Y and X

have isomorphic nonempty Zariski-open subsets, they have isomorphic function fields. Such varieties Y and X are said to be *birationally isomorphic*. This says that function fields in two or more variables always have many different nonisomorphic smooth models, and the connection with function fields is not as tight as in the curve case.

It must be said that the theory of Goppa-type codes from higher dimensional varieties is much less advanced at this point than the theory for Goppa codes from curves, perhaps because of these differences. There is still no clear understanding of how best to harness the properties of higher dimensional varieties in coding theory. Indeed, as we will see, most of the work that has appeared to date has been devoted to case studies of the *structural properties* of codes constructed from certain particular families of varieties X – their parameters, their weight distributions, their hierarchies of higher Hamming weights, and so forth. A few general ideas for estimating the minimum distance d have been developed. However, in many of the cases where the exact weight distributions are known, other algebraic constructions yield better codes. In addition, the development of efficient encoding and decoding algorithms for these codes has not really begun (see Section 7.8 on this point, though). The theory of order domains should yield tools here as well as for codes from curves. Nevertheless, the universality of this construction offers hope that good examples can be constructed this way, and our second main goal is to encourage others to explore this area.

This survey is organized as follows. In Section 7.2, we give two variants of Tsfasman and Vladut's code construction, one starting from an abstract variety X and line bundle $\mathcal{L}$ on X, the other starting from an embedded variety $X \subseteq \mathbb{P}^m$. We also present some first examples. Four general methods for estimating the minimum distance are presented in Section 7.3. Two appeared first in S.H. Hansen's article [26]. For the first of these, it is assumed that all of the $\mathbb{F}_q$-rational points of interest are contained in a family of curves on X and intersection products of divisors with those curves are used to bound d. The second method is based on the Seshadri constant of the line bundle $\mathcal{L}$ with respect to the set of $\mathbb{F}_q$-rational points on X. A third method from [17] can be used when the set of $\mathbb{F}_q$-rational points is itself a complete intersection in $\mathbb{P}^m$. Finally, we present another, more arithmetic, method based on the Weil conjectures developed by Lachaud in [37].

The next sections 7.4 and 7.5 present a selection of the examples of these codes that have appeared in the literature, codes constructed from quadric hypersurfaces, Hermitian hypersurfaces, Grassmannians and flag varieties,

Del Pezzo surfaces, ruled surfaces, and Deligne-Lusztig varieties. Finally, we present some comparisons between codes in section 7.7.

Where practicable, we have provided brief proofs of the results we state, in order to show the methods involved in the study of these codes.

As we proceed through these examples, the prerequisites from algebraic geometry steadily increase. Our intended audience includes both coding theorists familiar with the theory of Goppa codes on curves but not higher dimensional geometry and algebraic geometers curious about how higher dimensional varieties might be used in the coding theory context. The text [28] by Hartshorne is a good general reference for most of the algebraic geometry we need. The construction of Grassmannians via exterior algebra, Schubert varieties, and the intersection theory on Grassmannians are covered in Griffiths and Harris, [19]. A full understanding of the Deligne-Lusztig varieties also depends on the theory of reductive algebraic groups G over fields of characteristic p and the classification of their finite subgroups G^F by root systems and Dynkin diagrams with an action of the Frobenius endomorphism, F. The book [5] of Carter contains all the information needed for this.

Because of space limitations, it has not been possible to discuss all the results of every paper in this area in detail. Pointers to all of the literature of which the author is aware are provided in the bibliographic notes in Section 7.8, the references, and their bibliographies.

Any omissions or errors are entirely due to the author. Any comments or suggestions are welcome.

7.1.1. *Notation*

We will use the following general notational and terminological conventions.

- The number of elements in a finite set $\mathcal{T}$ will be denoted by $\#\mathcal{T}$.
- The *parameters* of a linear code are denoted $[n, k, d]$ as usual, where n is the block length, k is the dimension, and d is the minimum distance.
- The *generalized Hamming weights* are denoted d_r, $1 \leq r \leq k$. As in [58], d_r is the size of the minimal support of an r-dimensional subcode of C, extending the usual minimum distance $d = d_1$.
- We denote an algebraically closed field of characteristic p by $\mathbb{F}$ and all finite fields $\mathbb{F}_q$ for $q = p^m$ are considered as subfields of $\mathbb{F}$.
- The *projective spaces* $\mathbb{P}^m$, *Grassmannians* $\mathbb{G}(\ell, m)$, and so forth are

considered as varieties over the algebraically closed field $\mathbb{F}$ in order to "do geometry." The $\mathbb{F}_q$-rational points used in the construction of the codes are finite subsets of these varieties.

- If f is a homogeneous polynomial in $\mathbb{F}_q[x_0, \ldots, x_m]$, $\mathbf{V}(f)$ is the zero locus of f in $\mathbb{P}^m$.
- A *line bundle* is a locally free sheaf of rank one. At several points, it will be convenient to use the *sheaf cohomology* groups $H^i(X, \mathcal{L})$ for a line bundle $\mathcal{L}$. The space of global sections will also be written $\Gamma(X, \mathcal{L})$.

7.2. The General Construction

Several apparently different, but essentially equivalent, versions of the construction are commonly encountered in the literature. For instance, one description starts from a smooth projective variety X defined over $\mathbb{F}_q$, a set $\mathcal{S} \subseteq X(\mathbb{F}_q)$ of $\mathbb{F}_q$-rational points of X, and a line bundle $\mathcal{L}$ on X, also defined over $\mathbb{F}_q$. Let P be an $\mathbb{F}_q$-rational point of X. The stalk $\mathcal{L}_P$, modulo sections vanishing at P, denoted $\overline{\mathcal{L}}_P$, is isomorphic to $\mathbb{F}_q$ by a choice of local trivialization.

Definition 7.1. The choice of such local trivializations at each point in $\mathcal{S}$ defines a linear mapping (called the *germ map* in [55])

$$\alpha : \Gamma(X, \mathcal{L}) \longrightarrow \bigoplus_{i=1}^{n} \overline{\mathcal{L}}_{P_i} \simeq \mathbb{F}_q^n, \tag{7.2}$$

and the image is the code denoted $C(X, \mathcal{L}; \mathcal{S})$, or $C(X, \mathcal{L})$ if the set of points $\mathcal{S}$ is understood from the context.

If $\mathcal{L} = \mathcal{O}_X(G)$ for an $\mathbb{F}_q$-rational divisor G on X whose support is disjoint from $\{P_1, \ldots, P_n\}$, then up to monomial equivalence, this is the same as the evaluation code as in Equation (7.1) from the subspace $\mathcal{F}$ of the field of rational functions of X given by

$$\mathcal{F} = \{f \in \mathbb{F}_q(X)^* : \operatorname{div}(f) + G \geq 0\} \cup \{0\}.$$

For instance, when X is a smooth algebraic curve and $\mathcal{L} = \mathcal{O}_X(G)$ for some divisor G defined over $\mathbb{F}_q$ whose support is disjoint from the support of $D = P_1 + \cdots + P_n$, then this is the same as the algebraic geometric Goppa code $C_L(D, G)$ from X.

For explicit constructions of codes from embedded varieties $X \subseteq \mathbb{P}^m$, another more elementary description is also available using *homogeneous coordinates* $(a_0 : a_1 : \cdots : a_m)$ for points in $\mathbb{P}^m$.

Definition 7.2. Choosing any one such homogeneous coordinate vector defined over $\mathbb{F}_q$ for each of the points P_i in the set $\mathcal{S}$, define an evaluation map $ev_{\mathcal{S}}$ and a code as in Equation (7.1) using the vector space $\mathcal{F}_1$ of linear forms (homogeneous polynomials of degree 1) in $\mathbb{F}_q[x_0, \ldots, x_m]$. The code obtained as the image of this mapping is denoted $C(X)$, or $C(X; \mathcal{S})$ if it is important to specify the set of points. Similarly, the space of linear forms can be replaced by the vector space $\mathcal{F}_h$ of homogeneous polynomials of any degree $h \geq 1$, and corresponding codes denoted $C_h(X; \mathcal{S})$ or $C_h(X)$ are obtained.

Example 7.3. Let $X = \mathbb{P}^m$ itself, and let $\mathcal{S}$ be the set of *affine* $\mathbb{F}_q$-rational points of X, that is, points in the complement of the hyperplane $\mathbf{V}(x_0)$, having homogeneous coordinate vectors of the form $(1 : a_1 : \ldots : a_m)$. With these particular coordinate vectors, the code $C_h(X; \mathcal{S})$ is the well-known q-ary hth order (generalized) *Reed-Muller* code, denoted $\mathcal{R}_q(h, m)$. (When $m = 1$, this is the same as an *extended Reed-Solomon* code.) The block length is $n = q^m$. If $h < q$, then the monomials $x^\beta = x_0^{\beta_0} \cdots x_m^{\beta_m}$ where $|\beta| = \beta_0 + \cdots + \beta_m = h$ are linearly independent on $\mathcal{S}$, so the dimension of $\mathcal{R}_q(s, m)$ is $k = \binom{m+h}{h}$. If $\mathcal{S} = \mathbb{P}^m(\mathbb{F}_q)$, the resulting *projective Reed-Muller codes* have block length $n = q^m + \cdots + q + 1$. $\diamondsuit$

There is, of course, a tight connection between Definition 7.1 and Definition 7.2. If X is embedded in $\mathbb{P}^m$ and $\mathcal{L} = \mathcal{O}_X(1)$ is the hyperplane bundle, then $C(X, \mathcal{O}_X(1))$ and $C(X)$ are monomially equivalent codes (they differ at most by constant multiples in each component depending on how the isomorphisms of the fibers with $\mathbb{F}_q$ are chosen). Similarly, $C_h(X)$ is equivalent to $C(X, \mathcal{O}_X(h))$. Also, in theory it suffices to consider the $C(X) = C_1(X)$ codes, since the $C_h(X)$ code on X is the same as the C_1 code on the variety $\nu_h(X)$, where ν_h is the degree-h *Veronese mapping*

$$\begin{aligned} \nu_h : \mathbb{P}^m &\longrightarrow \mathbb{P}^{\binom{m+h}{h}-1} \\ (x_0 : x_1 : \cdots : x_m) &\mapsto (\cdots : x^\beta : \cdots), \end{aligned}$$

and $x^\beta = x_0^{\beta_0} \cdots x_m^{\beta_m}$ ranges over all monomials of total degree h. The image $\nu_h(\mathbb{P}^m)$ has dimension m, degree h^m, and is isomorphic to $\mathbb{P}^m$.

7.3. Estimating the Parameters

7.3.1. *Elementary bounds*

Suppose Definition 7.2 is used to construct a code $C_h(X;\mathcal{S})$ from a variety X. The block length of the code is $n = \#\mathcal{S}$. Using a standard linear algebra result, the dimension is

$$k = \dim \mathcal{F}_h - \dim \ker ev_{\mathcal{S}}.$$

Forms of degree h vanishing on X always give elements of $\ker ev_{\mathcal{S}}$. The dimension of the space of such forms can be computed using the long exact cohomology sequence of

$$0 \longrightarrow \mathcal{I}_X(h) \longrightarrow \mathcal{O}_{\mathbb{P}^m}(h) \longrightarrow \mathcal{O}_X(h) \longrightarrow 0. \tag{7.3}$$

But if the points in $\mathcal{S}$ are not in general position, there can be other elements of the kernel as well and it may be necessary to take the properties of $\mathcal{S}$ as a 0-dimensional algebraic set into account to understand the parameters of the $C_h(X;\mathcal{S})$ codes. See Section 7.3.4 below.

Since each codeword is $ev_{\mathcal{S}}(f) = (f(P_1), \ldots, f(P_n))$ for some form f, the codeword weight is $n - \#(\mathbf{V}(f) \cap \mathcal{S})$, the number of P_i in $\mathcal{S}$ where f is not zero. Therefore,

$$d = \min_{f \neq 0 \in \mathcal{F}_h} \left(n - \#(\mathbf{V}(f) \cap \mathcal{S})\right). \tag{7.4}$$

Along similarly general lines, let $\dim Y = \delta$ and let the degree of Y be $s < q+1$ in $\mathbb{P}^m$. Let E be an $\mathbb{F}_q$-rational linear subspace of dimension $m - \delta - 1$ with $E \cap Y = \emptyset$. By projection from E onto a linear subspace $L \simeq \mathbb{P}^\delta$, each $\mathbb{F}_q$-rational point of L corresponds to at most s such points of Y, so

$$\#Y(\mathbb{F}_q) \leq s \cdot \#\mathbb{P}^\delta(\mathbb{F}_q) = s(q^\delta + \cdots + q + 1). \tag{7.5}$$

Applying Equation (7.5) to $Y = X \cap H$ for a hyperplane, Lachaud obtains the following elementary bound in [37].

Theorem 7.4. *Let X be a projective variety of dimension δ and degree $s < q+1$. Then the $C(X)$ code has*

$$d \geq n - s(q^{\delta-1} + \cdots + q + 1).$$

A more refined estimate of the number of $\mathbb{F}_q$ rational points on a projective hypersurface establishes the following result for the projective Reed-Muller codes introduced in Example 7.3.

Theorem 7.5. *Let $h \leq q$. The projective Reed-Muller code of order h has parameters*

$$\left[q^m + \cdots + q + 1, \binom{m+h}{h}, (q+1-h)q^{m-1}\right].$$

Proof. Write $\mathcal{S} = \mathbb{P}^m(\mathbb{F}_q)$. The evaluation mapping is injective and $k = \dim \mathcal{F}_s = \binom{m+h}{h}$ provided that $d > 0$. By [52], if f is a homogeneous polynomial of degree $h \leq q$, then (improving the bound of Equation (7.5))

$$\#(\mathbf{V}(f) \cap \mathcal{S}) \leq hq^{m-1} + q^{m-2} + \cdots + q + 1.$$

Moreover, if $\mathbf{V}(f)$ is the union of h $\mathbb{F}_q$-rational hyperplanes meeting along a common $(m-2)$-dimensional linear subspace, this bound is attained. Hence

$$d = (q^m + q^{m-1} + \cdots + q + 1) - (hq^{m-1} + q^{m-2} + \cdots + q + 1) = (q+1-h)q^{m-1}$$

as claimed. □

The reducible f featuring in the proof of Theorem 7.5 give a first indication of a general theme related to these codes.

Observation 7.6. The minimum weight codewords of a $C_h(X; \mathcal{S})$ code tend to come from "maximally reducible" $\mathbf{V}(f) \cap X$ for $f \in \mathcal{F}_h$.

The underlying reason for this is the fact that reducible varieties (especially those that are unions of linear subspaces or other rational varieties) can have many more $\mathbb{F}_q$-rational points than other varieties of the same degree. The Weil-type bounds discussed below can be used to quantify this remark.

7.3.2. *Bounds from covering families of curves*

For the following discussion, it will be most convenient to use the code construction given in Definition 7.1. In many concrete cases, it can be seen that the points in the set $\mathcal{S}$ are distributed on a collection of *curves* C_i (subvarieties of dimension 1) on the variety X. Since each section $f \in \Gamma(X, \mathcal{L})$ on X defines a divisor of zeroes $Z(f)$, a subvariety of codimension 1 on X, determining the minimum distance of the $C(X, \mathcal{L})$ code reduces to understanding how many times the divisors $Z(f)$ can intersect the curves C_i at points of $\mathcal{S}$. To prepare, let C be any irreducible curve in X. Observe that the divisors $Z(f)$ for $f \in \Gamma(X, \mathcal{L})$ all cut out divisors on C of the same degree. This degree will be denoted by $\mathcal{L} \cdot C$. In this situation, Hansen derives a lower bound for d in [26].

Theorem 7.7. *Let X be a normal projective variety defined over $\mathbb{F}_q$, of dimension $\dim X \geq 2$. Let $\mathcal{S} \subseteq X(\mathbb{F}_q)$ and assume $\mathcal{S} \subset \bigcup_{i=1}^{a} C_i$ where C_i are irreducible curves on X, also defined over $\mathbb{F}_q$. Assume that $\#(C_i \cap \mathcal{S}) \leq N$ for all i. Let $\mathcal{L}$ be a line bundle on X defined over $\mathbb{F}_q$ such that*

$$0 \leq \mathcal{L} \cdot C_i \leq \eta \leq N$$

for all i. Let

$$\ell = \max_{f \neq 0 \in \Gamma(X,\mathcal{L})} \#\{i : Z(f) \text{ contains } C_i\}.$$

Then the code $C(X, \mathcal{L}; \mathcal{S})$ has

$$d \geq \#\mathcal{S} - \ell N - (a - \ell)\eta.$$

Proof. Let $f \in \Gamma(X, \mathcal{L})$, let $D = Z(f)$, and let $E = Z(f) \cap \bigcup_{i=1}^{a} C_i$. Suppose E contains $\ell' \leq \ell$ of the C_i. The number points of $\mathcal{S}$ that are contained in E is estimated as follows:

$$\begin{aligned} \#(E \cap \mathcal{S}) &\leq \ell' N + (a - \ell')\eta \\ &\leq \ell N + (a - \ell)\eta \end{aligned}$$

(since by hypothesis $\eta \leq N$). Hence $ev_{\mathcal{S}}(f)$ has at least $\#\mathcal{S} - \ell N - (a - \ell)\eta$ nonzero entries. □

Example 7.8. Let $X = \mathbb{P}^1 \times \mathbb{P}^1$. Let $\mathcal{S} = X(\mathbb{F}_q)$, which consists of $(q+1)^2$ points, equally distributed over the lines $C_1, \ldots, C_{q+1}$ of one of the rulings. The Picard group of line bundles modulo isomorphism is $\mathrm{Pic}(X) \simeq \mathbb{Z} \oplus \mathbb{Z}$, so the lines C_i may be taken as the divisors of zeros of sections of a line bundle of type $(1, 0)$. Let $\mathcal{L}$ have type (α, β) where $0 \leq \alpha, \beta \leq q + 1$, and apply Theorem 7.7 to estimate d for the $C(X, \mathcal{L})$ code. Because of the description of $\mathcal{S}$ above, $N = q + 1$. The divisor $Z(f)$ for $f \in \Gamma(X, \mathcal{L})$ contains at most α of the C_i, so $\ell = \alpha$. Moreover, $\mathcal{L} \cdot C_i = \beta$ for each i, so $\eta = \beta$. The bound is

$$d \geq (q+1)^2 - \alpha(q+1) - (q+1-\alpha)\beta = (q+1-\alpha)(q+1-\beta).$$

It is easy to construct codewords of this weight via bihomogeneous polynomials on $\mathbb{P}^1 \times \mathbb{P}^1$. So this is the exact minimum distance. ♢

7.3.3. *Bounds using Seshadri constants*

A second general method for estimating the minimum distance of the $C(X, \mathcal{L}; \mathcal{S})$ codes is based on the *Seshadri constant* of $\mathcal{L}$ relative to the set $\mathcal{S}$. This is potentially useful but requires some significantly more sophisticated birational geometry to state and apply. Let $\pi : Y \to X$ be the blow up of the X at the points in $\mathcal{S}$ and call the exceptional divisor E. Then the Seshadri constant is defined as

$$\varepsilon(\mathcal{L}, \mathcal{S}) = \sup\{\varepsilon \in \mathbb{Q} : \pi^*\mathcal{L} - \varepsilon E \text{ is nef on } Y\}.$$

(Here, "nef" means *numerically effective*, that is, $(\pi^*\mathcal{L} - \varepsilon E) \cdot C \geq 0$ for all irreducible curves C on Y.) Hansen proves the following estimate for the minimum distance of the $C(X, \mathcal{L}; \mathcal{S})$ codes in [26].

Theorem 7.9. *Let X be a nonsingular projective variety of dimension ≥ 2 over $\mathbb{F}_q$. If $\mathcal{L}$ is ample with Seshadri constant $\varepsilon(\mathcal{L}, \mathcal{S}) \geq e \in \mathbb{N}$, and $n > e^{1-\dim(X)}\mathcal{L}^{\dim(X)}$, then $C(X, \mathcal{L}; \mathcal{S})$ has minimum distance $d \geq n - e^{1-\dim(X)}\mathcal{L}^{\dim(X)}$.*

This is particularly well-suited for analyzing certain codes from Deligne-Lusztig varieties to be defined in Section 7.5 below.

7.3.4. *Bounds from $\mathcal{S}$ itself*

All of the $C_h(X; \mathcal{S})$ codes introduced in Section 7.2 can be viewed as punctures of the projective Reed-Muller code of order h on $\mathbb{P}^m$ (delete the components corresponding to points in the complement of $\mathcal{S}$). For this reason, in addition to making use of the properties of the variety X, it is also possible to use properties of the 0-dimensional algebraic set (or scheme) $\mathcal{S}$ itself to study these codes.

Let $\mathcal{I}_\mathcal{P}$ be the sheaf of ideals defining any 0-dimensional $\mathcal{P}$. From the long exact cohomology sequence of the exact sequence of sheaves

$$0 \longrightarrow \mathcal{I}_\mathcal{P} \longrightarrow \mathcal{O}_{\mathbb{P}^m} \longrightarrow \mathcal{O}_\mathcal{P} \longrightarrow 0,$$

it follows that for all $h \geq 0$, the following sequence is exact:

$$0 \to H^0(\mathcal{I}_\mathcal{P}(h)) \to H^0(\mathcal{O}_{\mathbb{P}^m}(h)) \to H^0(\mathcal{O}_\mathcal{P}(h)) \to H^1(\mathcal{I}_\mathcal{P}(h)) \to 0. \quad (7.6)$$

The term $H^0(\mathcal{I}_\mathcal{P}(h))$ gives the space of homogeneous forms of degree h vanishing on $\mathcal{P}$. The term $H^1(\mathcal{I}_\mathcal{P}(h))$ measures the failure of the points in

$\mathcal{P}$ to impose independent conditions on forms of degree h. The dimension of the $C_h(\mathcal{S})$ code is given by the *Hilbert function* of $\mathcal{S}$:

$$H_{\mathcal{S}}(h) = \dim H^0(\mathcal{O}_{\mathbb{P}^m}(h)) - \dim H^0(\mathcal{I}_{\mathcal{S}}(h)) = \#\mathcal{S} - \dim H^1(\mathcal{I}_{\mathcal{S}}(h)).$$

In the case that $\mathcal{S}$ is a *complete intersection* of hypersurfaces of degrees $d_1, \ldots, d_m$ defined by homogeneous polynomials $f_1, \ldots, f_m$, the Hilbert function can be computed explicitly from the Koszul complex of the regular sequence $f_1, \ldots, f_m$ (see [23]). Moreover, there are particularly nice techniques from commutative algebra and algebraic geometry related to the classical *Cayley-Bacharach Theorem* that apply. A modern version of this result due to Davis, Geramita, and Orecchia can be stated as follows in the situation at hand.

Theorem 7.10. *Let $\mathcal{S} \subset \mathbb{P}^m$ be a reduced complete intersection of hypersurfaces of degrees $d_1, \ldots, d_m$. Let Γ', Γ'' be disjoint subsets of $\mathcal{S}$ with $\mathcal{S} = \Gamma' \cup \Gamma''$. Let $s = \sum_{i=1}^m d_i - m - 1$. Then for all $h \geq 0$,*

$$\dim H^0(\mathcal{I}_{\Gamma'}(h)) - \dim H^0(\mathcal{I}_{\mathcal{S}}(h)) = \dim H^1(\mathcal{I}_{\Gamma''}(s-h)).$$

Applied to the corresponding codes from a complete intersection $\mathcal{S}$ consisting of $d_1 d_2 \ldots d_m$ distinct $\mathbb{F}_q$-rational points, this result implies the following.

Theorem 7.11. *Let $\mathcal{S}$ be a reduced complete intersection of hypersurfaces of degrees $d_1, \ldots, d_m$ in $\mathbb{P}^m$. Let $s = \sum_{i=1}^m d_i - m - 1$ as in Theorem 7.10. If $1 \leq h \leq s$, the code $C_h(\mathcal{S})$ has minimum distance*

$$d \geq \sum_{i=1}^{m} d_i - h - (m-1) = s - h + 2.$$

The proof is accomplished by showing that under these hypotheses, any form of degree h that is zero on a subset Γ' that is too large must be zero at all points in $\mathcal{S}$ because the $H^1(\mathcal{I}_{\Gamma''}(s-h))$ group vanishes.

The bound on d given here was improved rather strikingly by Ballico and Fontanari to $d \geq m(s-h)+2$ under the assumption that all subsets of $m+1$ of the points in $\mathcal{S}$ span $\mathbb{P}^m$ – see [2] for this.

Bounds derived by these methods are usually interesting only for h close to s. Moreover some, but not all, interesting examples of $\mathcal{S}$ satisfy the complete intersection hypothesis. For instance the affine $\mathbb{F}_q$-rational points in $\mathbb{P}^m$ form a complete intersection for all m. The $\mathbb{F}_8$-rational points on the Klein quartic and the $\mathbb{F}_{r^2}$ points on the Hermitian curve are other examples.

7.3.5. *General Weil-type bounds*

From Equation (7.4) above, and the proof of Theorem 7.5, the minimum distance of a $C(X)$ code as in Definition 7.2 is determined by the numbers of $\mathbb{F}_q$-rational points on the subvarieties $Y = X \cap \mathbf{V}(f)$. Hence, another possible approach to estimate d is to apply general bounds for $\#Y(\mathbb{F}_q)$, for instance bounds derived from the statements of the Weil conjectures, or refined versions of these.

We very briefly recall the deep mathematics behind this approach. Thinking of X as a variety over the algebraic closure of the finite field, the number of $\mathbb{F}_q$-rational points on X can be computed by an analog of the Lefschetz trace formula for the action of the Frobenius endomorphism F on the ℓ-adic étale cohomology groups of X, $H^i(X)$ (where ℓ is any prime not dividing q):

$$\#X(\mathbb{F}_q) = \sum_{i=0}^{2m} (-1)^i \operatorname{Tr}(F|H^i(X)). \tag{7.7}$$

Moreover, the eigenvalues of F on $H^i(X)$ are algebraic numbers of absolute value $q^{i/2}$. When X is obtained from a variety Y defined over the ring of integers R of some number field by reduction modulo some prime ideal in R, then the dimensions of the $H^i(X)$ are the same as the topological Betti numbers of the variety over $\mathbb{C}$ corresponding to Y.

Thus, for instance, if X is a smooth curve of genus g which is the reduction of a smooth curve Y, then

$$\#X(\mathbb{F}_q) = q + 1 - \sum_{j=0}^{2g} \alpha_j,$$

where $|\alpha_j| = q^{1/2}$ for all j. The Hasse-Weil bound often used in the theory of Goppa codes from curves is a direct consequence:

$$|\#X(\mathbb{F}_q) - (q+1)| \leq 2g\sqrt{q}.$$

There is a correspondingly concrete Weil-type bound for hypersurfaces in $\mathbb{P}^m$, and this can be used to derive bounds on the numbers of $\mathbb{F}_q$-rational points in hyperplane sections as well. A hypersurface is said to be *nondegenerate* if it not contained in any linear subspace of $\mathbb{P}^m$.

Theorem 7.12. *Let X be a smooth nondegenerate hypersurface of degree s in $\mathbb{P}^m$, $m \geq 2$. Then*

$$|\#X(\mathbb{F}_q) - (q^{m-1} + \cdots + q + 1)| \leq b(s) q^{(m-1)/2}, \tag{7.8}$$

where $b(s) = \frac{s-1}{s}((s-1)^m - (-1)^m)$ *is the middle Betti number of a smooth hypersurface of degree* s *when* m *is even, and one less than that number when* m *is odd.*

Equation (7.8) follows from the shape of the cohomology groups $H^i(X)$ of a smooth hypersurface in $\mathbb{P}^m$, which (by the Lefschetz hyperplane theorem and Poincaré duality) look like the corresponding groups for $\mathbb{P}^{m-1}$, except possibly in the middle dimension $i = m-1$.

Example 7.13. If $m = 2$ and X is a smooth curve of degree s in $\mathbb{P}^2$, then

$$b(s) = \frac{s-1}{s}((s-1)^2 - 1) = (s-1)(s-2) = 2g(X)$$

as expected. In order to obtain long codes over $\mathbb{F}_q$, the *maximal curves*, that is, curves attaining the maximum $\#X(\mathbb{F}_q)$ from Equation (7.8), have been especially intensively studied. For instance, when $q = r^2$, the Hermitian curve of degree $s = r+1$ over $\mathbb{F}_{r^2}$, $X = \mathbf{V}(x_0^{r+1} + x_1^{r+1} + x_2^{r+1})$, has $\#X(\mathbb{F}_{r^2}) = r^3 + 1 = 1 + r^2 + r(r-1)r$. $\diamondsuit$

Example 7.14. When $m = 3$ and $q = r^2$, the analogous *Hermitian surfaces* $X = \mathbf{V}(x_0^{r+1} + x_1^{r+1} + x_2^{r+1} + x_3^{r+1})$ also attain the upper bound from Equation (7.8), which reads

$$\#X(\mathbb{F}_{r^2}) \leq 1 + r^2 + r^4 + \frac{r}{r+1}(r^3+1)r^2 = (r^2+1)(r^3+1).$$

The Hermitian surface contains this many distinct $\mathbb{F}_{r^2}$-rational points because, for instance, it is possible to take the defining equation to the affine form

$$y_1^r + y_1 = y_2^{r+1} + y_3^{r+1}$$

by a linear change of coordinates that puts a plane tangent to the surface as the plane at infinity. Then there are r^5 affine $\mathbb{F}_{r^2}$-rational points (r for each pair $(y_2, y_3) \in (\mathbb{F}_{r^2})^2$). There are also $(r+1)r^2 + 1$ rational points at infinity since the intersection of the surface with each of its tangent planes at an $\mathbb{F}_{r^2}$-rational point is the union of $r+1$ concurrent lines in that plane. This yields $r^5 + (r+1)r^2 + 1 = (r^3+1)(r^2+1)$ points as claimed. $\diamondsuit$

The following result of Lachaud appears in [37].

Theorem 7.15. *Let* X *be a smooth nondegenerate hypersurface of degree* s *in* $\mathbb{P}^m$ *for* $m \geq 3$. *Let* $H = \mathbf{V}(f)$ *for a linear form in* $\mathbb{F}_q[x_0, \ldots, x_m]$, *and*

let X_H denote the intersection $X \cap H$ (with the reduced scheme structure). Then

$$|\#X_H(\mathbb{F}_q) - (q^{m-2} + \cdots + q + 1)| \leq (s-1)^{m-1} q^{(m-1)/2}, \tag{7.9}$$

and

$$|q\#X_H(\mathbb{F}_q) - \#X(\mathbb{F}_q)| \leq (s-1)^{m-1}(q+s-1)q^{(m-1)/2}. \tag{7.10}$$

These bounds are proved by comparing the cohomology of X and X_H, taking into account possible singularities of X_H. For a proof, see Corollary 4.6 and preceding results of [37].

When $\mathcal{S}$ is the full set of $\mathbb{F}_q$-rational points on X, so $n = \#\mathcal{S}$ for the $C(X;\mathcal{S})$ code and H is a general hyperplane, these imply the following bounds on $\#(H \cap \mathcal{S})$. Equation (7.10) implies

$$\left|(n - \#(H \cap \mathcal{S})) - \frac{(q-1)}{q} n\right| \leq (s-1)^{m-1}(q+s-1)q^{(m-3)/2} \tag{7.11}$$

and

$$\left|(n - \#(H \cap \mathcal{S})) - q^{m-2}\right| \leq s(s-1)^{m-1} q^{(m-1)/2}. \tag{7.12}$$

These, together with Equation (7.9), give universally applicable lower bounds on d by applying Equation (7.4).

As is perhaps to be expected, it is often possible to derive tighter bounds in specific cases by taking the properties of X into account.

7.4. Examples

This section will consider codes produced according to the constructions from Section 7.2 from various special classes of varieties. By analogy with the case of Goppa codes from curves, much work has focused in identifying varieties with *many rational points* over finite fields $\mathbb{F}_q$ and studying the codes constructed from those X.

7.4.1. *Quadrics*

First consider the $C(X)$ codes from quadric hypersurfaces $X = \mathbf{V}(f)$ for homogeneous f of degree 2 in $\mathbb{F}_q[x_0, \ldots, x_m]$. The following statements are proved, for instance, in Chapter 22 of [30]. Up to projective equivalence over $\mathbb{F}_q$, such X are completely described by a positive integer called the *rank* and a second integer called the *character*, which takes values in the finite set $\{0, 1, 2\}$. The rank, denoted ρ, can be described as the minimum

number of variables needed to express f after a linear change of coordinates in $\mathbb{P}^m$. X is said to be *nondegenerate* if $\rho = m+1$. Nondegenerate quadrics are always smooth varieties. Degenerate quadrics are singular, but they are cones over nondegenerate quadrics in a linear subspace of $\mathbb{P}^m$. Hence in principle it suffices to study nondegenerate quadrics and we will consider only that case here. The character, denoted w, is most easily described by considering a finite set of possible normal forms for f.

If m is even, then every nondegenerate quadric can be taken to the form

$$x_0^2 + x_1x_2 + x_3x_4 + \cdots + x_{m-1}x_m.$$

$\mathbf{V}(f)$ is called a *parabolic* quadric in this case, and the character w is defined to be 1.

On the other hand, if m is odd, there are two distinct possible forms:

$$x_0x_1 + x_2x_3 + \cdots + x_{m-1}x_m \quad \text{or}$$
$$q(x_0, x_1) + x_2x_3 + \cdots + x_{m-1}x_m.$$

In the first case, $\mathbf{V}(f)$ is called a *hyperbolic* quadric and $w = 2$. In the second, $q(x_0, x_1)$ is a quadratic form in two variables which can be further reduced to slightly different normal forms depending on whether q is even or odd. For both even and odd q, in the second case, $\mathbf{V}(f)$ is called a *elliptic* quadric and $w = 0$.

Theorem 7.16. *A nondegenerate quadric X in $\mathbb{P}^m$ with character w has*

$$\#X(\mathbb{F}_q) = q^{m-1} + \cdots + q + 1 + (w-1)q^{(m-1)/2}.$$

In particular, this result says that hyperbolic and parabolic quadrics attain the upper bound from Equation (7.8) with $s = 2$, and elliptic quadrics attain the lower bound.

Because each linear section of X is also a quadric in a lower-dimensional space, Theorem 7.16 can be used to determine the full weight distributions of the $C(X)$ codes. In particular,

Theorem 7.17. *The $C(X)$ code from a smooth quadric X in $\mathbb{P}^m$ has n given in Theorem 7.16, $k = m+1$ and*

$$d = \begin{cases} q^{m-1} & \text{if } w = 2 \\ q^{m-1} - q^{(m-2)/2} & \text{if } w = 1 \\ q^{m-1} - q^{(m-1)/2} & \text{if } w = 0. \end{cases} \tag{7.13}$$

For instance, if m is even, so $w = 1$ (the parabolic case), the hyperplane section of X containing the most $\mathbb{F}_q$-rational points will be a hyperbolic section and d is as above. When $w = 2$ (for example, for codes from hyperbolic quadrics in $\mathbb{P}^3$), the minimum weight codewords come from hyperplane sections that are degenerate quadrics.

The same sort of reasoning has also been used by Nogin and Wan to determine the complete hierarchy of generalized Hamming weights $d_1(C(X)), \ldots, d_k(C(X))$. The results are somewhat intricate to state, though, so we refer the interested reader to the articles [44, 57] and the notes in Section 7.8.

For the $C_h(X)$ codes with $h \geq 2$, the dimension can be estimated using Equation (7.3), where $\mathcal{I}_X(h) \simeq \mathcal{O}_{\mathbb{P}^m}(h-2)$. This yields

$$k \leq \binom{m+h}{h} - \binom{m+h-2}{h-2}.$$

7.4.2. *Hermitian hypersurfaces*

For the $C(X)$ codes constructed from the Hermitian surfaces of Example 7.14 with $q = r^2$, Equation (7.9) gives

$$d \geq (r^2+1)(r^3+1) - (r^2+1+r^4) = r^5 - r^4 + r^3.$$

However, closer examination of the hyperplane sections of the Hermitian surface yields the following statement.

Theorem 7.18. *Let $X = \mathbf{V}(x_0^{r+1} + x_1^{r+1} + x_2^{r+1} + x_3^{r+1})$ be the Hermitian surface over $\mathbb{F}_{r^2}$. The $C(X)$ code on $\mathcal{S} = X(\mathbb{F}_{r^2})$ has parameters*

$$\left[(r^2+1)(r^3+1), 4, r^5\right].$$

Proof. Every $\mathbb{F}_{r^2}$-rational plane in $\mathbb{P}^3$ intersects X either in a Hermitian curve containing r^3+1 points over $\mathbb{F}_{r^2}$, or else in $r+1$ concurrent lines containing $(r+1)r^2+1$ points. Hence by Equation (7.4),

$$d = n - ((r+1)r^2 + 1) = r^5.$$

□

The $C_h(X)$ codes with $h > 1$ are more subtle here.

Theorem 7.19. *Let X and $\mathcal{S}$ be as in Theorem 7.18. If $h < r+1$, the $C_h(X)$ code has parameters*

$$\left[(r^2+1)(r^3+1), \binom{4+h}{h}, d \geq n - h(r+1)(r^2+1)\right].$$

Proof. This bound follows from Theorem 7.4 by the fact that if f is a form of degree h, then $\mathbf{V}(f) \cap X$ is a curve of degree $\delta = h(r+1)$ in $\mathbb{P}^3$. The hypothesis on h implies that the evaluation mapping is injective. For larger h, Equation (7.3) would be used to determine the dimension of the space of forms of degree h vanishing on the Hermitian variety. □

An even tighter bound

$$d \geq n - (h(r^3 + r^2 - r) + r + 1) \tag{7.14}$$

has been conjectured by Sørensen for these codes in [53].

The Hermitian curve and surface codes can be generalized as follows (see Chapter 23 of [30]). Over a field of order $q = r^2$, consider the Hermitian hypersurface in $\mathbb{P}^m$ defined by

$$X = \mathbf{V}(x_0^{r+1} + x_1^{r+1} + \cdots + x_m^{r+1}). \tag{7.15}$$

The mapping $F(x) = x^r$ is a involutory field automorphism of $\mathbb{F}_{r^2}$. The defining polynomial of X may be understood as $H(x, x)$ for the mapping $H : \mathbb{F}_{r^2}^{m+1} \times \mathbb{F}_{r^2}^{m+1} \to \mathbb{F}_{r^2}$ given by

$$H(x, y) = x_0 y_0^r + \cdots + x_m y_m^r.$$

It is clear that H is additive in each variable and satisfies $H(\lambda x, y) = \lambda H(x, y)$ and $H(x, \lambda y) = \lambda^r H(x, y) = F(\lambda) H(x, y)$ for the automorphism F above. Hence H is an example of what is known as a *sesquilinear form* on $\mathbb{F}_{r^2}^{m+1} \times \mathbb{F}_{r^2}^{m+1}$. After a linear change of coordinates defined over $\mathbb{F}_{r^2}$, any such sesquilinear H on $V \times V$, where V is a finite-dimensional $\mathbb{F}_{r^2}$-vector space, can be expressed as

$$H(x, y) = x_0 y_0^r + \cdots + x_\ell y_\ell^r \tag{7.16}$$

for some $\ell \leq \dim V$. H is said to be nondegenerate if $\ell = \dim V$ and degenerate otherwise.

It follows that every linear section $L \cap X$ of a Hermitian hypersurface is also a Hermitian variety in the linear subspace $L = \mathbb{P}W$ for some vector subspace W. Moreover, if the section is degenerate (i.e. $\ell < \dim W$ in Equation (7.16)), then the section is a cone over a nondegenerate Hermitian variety in a linear subspace of L. Thus, the properties of the codes $C(X)$ from the Hermitian hypersurfaces are formally quite similar to (and even somewhat simpler than) the properties of codes from quadrics discussed above. The main ingredient is the following statement for the nondegenerate Hermitian hypersurfaces.

Theorem 7.20. *Let X be the nondegenerate Hermitian hypersurface from Equation (7.15). Then*

$$\#X(\mathbb{F}_{r^2}) = r^{2m-2} + \cdots + r^2 + 1 + b(r+1)r^{m-1},$$

where $b(r+1) = \frac{r}{r+1}(r^m - (-1)^m)$.

In other words, for all m, the nondegenerate Hermitian hypersurfaces meet the upper bound from Equation (7.8) for a hypersurface of degree $s = r + 1$.

Theorem 7.21. *Let $\mathcal{S} = X(\mathbb{F}_{r^2})$ for the nondegenerate Hermitian hypersurface X in $\mathbb{P}^m$. The $C(X;\mathcal{S})$ code has n given in Theorem 7.20, $k = m + 1$, and*

$$d = \begin{cases} r^{2m-1} - r^{m-1} & \text{if } m \equiv 0 \bmod 2 \\ r^{2m-1} & \text{if } m \equiv 1 \bmod 2. \end{cases}$$

When m is even, the minimum weight codewords of the $C(X)$ come from nondegenerate Hermitian variety hyperplane sections. On the other hand, if m is odd, then the minimum weight codewords of $C(X)$ come from hyperplane sections that are degenerate Hermitian varieties. In both cases, the nonzero codewords of $C(X)$ have only two distinct weights:

$$r^{2m-1} + (-1)^{m-1}r^{m-1} \text{ and } r^{2m-1}.$$

The hierarchies of generalized Hamming weights d_r are also known for the $C(X)$ codes by work of Hirschfeld, Tsfasman, and Vladut, [31]. The same sort of techniques used in Theorem 7.18 above can be applied to the $C_h(X)$ codes for $h \geq 2$ here. However, much less is known about the exact Hamming weights of these codes.

7.4.3. *Grassmannians and flag varieties*

The *Grassmannian* $\mathbb{G}(\ell, m)$ is a projective variety whose points are in one-to-one correspondence with the ℓ-dimensional vector subspaces of an m-dimensional vector space (or equivalently the $(\ell-1)$-dimensional linear subspaces of $\mathbb{P}^{m-1}$). We very briefly recall the construction.

Let $\mathbb{F}$ denote an algebraic closure of $\mathbb{F}_q$. Given any basis $B = \{v_1, \ldots, v_\ell\}$ for an ℓ-dimensional vector subspace W of $\mathbb{F}^m$, form the $\ell \times m$ matrix $M(B)$ with rows v_i. Consider the determinants of the maximal square $(\ell \times \ell)$ submatrices of $M(B)$. There is one such maximal minor for each subset $I \subset \{1, \ldots, m\}$ with $\#I = \ell$, so writing $p_I(W)$ for the maximal

minor in the columns corresponding to I, the *Plücker coordinate vector* of W is the homogeneous coordinate vector

$$p(W) = (\cdots : p_I(W) : \cdots) \in \mathbb{P}^{\binom{m}{\ell}-1}, \tag{7.17}$$

where I runs through all subsets of size ℓ in $\{1, \ldots, m\}$. The point $p(W)$ is a well-defined invariant of W because a change of basis in W multiplies the matrix $M(B)$ on the left by the change of basis matrix, an element of $\mathrm{GL}(\ell, \mathbb{F})$. All components of the Plücker coordinate vector are multiplied by the determinant of the change of basis matrix, an element of $\mathbb{F}^*$. Hence any choice of basis in W yields the same point $p(W)$ in $\mathbb{P}^{\binom{m}{\ell}-1}$.

The locus of all such points (for all W) forms the Grassmannian $\mathbb{G}(\ell, k)$, an algebraic variety whose defining ideal is generated by a collection of *Plücker quadrics*. Consider the set of W such that $p_{I_0}(W) \neq 0$, so the maximal minor with $I_0 = \{1, \ldots, \ell\}$ is invertible. The set of such W is one of the open subsets in the standard affine cover of $\mathbb{G}(\ell, m)$. In the row-reduced echelon form of $M(B)$, the entries in the columns complementary to I_0 (an $\ell \times (m-\ell)$ block) are arbitrary and uniquely determine W. Hence

$$\dim \mathbb{G}(\ell, m) = \ell(m - \ell).$$

To construct Grassmannian codes, one uses the $\mathbb{F}_q$-rational points of $\mathbb{G}(\ell, m)$, which come from subspaces W defined over $\mathbb{F}_q$. Nogin has established the following result.

Theorem 7.22. *Let $\mathcal{S}$ be the set of all the $\mathbb{F}_q$-rational points on $X = \mathbb{G}(\ell, m)$. Then the $C(X; \mathcal{S})$ code (from linear forms in the Plücker coordinates) has parameters*

$$\left[\begin{bmatrix} m \\ \ell \end{bmatrix}_q, \ \binom{m}{\ell}, \ q^{\ell(m-\ell)} \right],$$

where

$$\begin{bmatrix} m \\ \ell \end{bmatrix}_q = \frac{(q^m - 1)(q^m - q) \cdots (q^m - q^{\ell-1})}{(q^\ell - 1)(q^\ell - q) \cdots (q^\ell - q^{\ell-1})}.$$

Proof. The numerator in the formula for $\begin{bmatrix} m \\ \ell \end{bmatrix}_q$ is precisely the number of ways of picking a list of ℓ linearly independent vectors in $\mathbb{F}_q^m$ (a basis for a W defined over $\mathbb{F}_q$). Similarly, the denominator is the number of ways of picking ℓ linearly independent vectors in $\mathbb{F}_q^\ell$, hence the order of the group $\mathrm{GL}(\ell, \mathbb{F}_q)$. The quotient is the number of distinct ℓ-dimensional

subspaces of $\mathbb{F}_q^m$. This shows $n = \#\mathcal{S} = \begin{bmatrix} m \\ \ell \end{bmatrix}_q$. Assuming $d = q^{\ell(m-\ell)}$ for the moment, the fact that $d > 0$ says the evaluation mapping on the vector space of linear forms in $\mathbb{P}^{\binom{m}{\ell}-1}$ is injective, and the formula for k follows. Finally, we must prove that $d = q^{\ell(m-\ell)}$.

The complement of the hyperplane section $\mathbb{G}(\ell, m) \cap \mathbf{V}(p_{I_0})$ contains exactly $q^{\ell(m-\ell)}$ $\mathbb{F}_q$-rational points of $\mathbb{G}(\ell, m)$. Hence $d \leq q^{\ell(m-\ell)}$. The cleanest way to prove that this is an equality is to use the language of exterior algebra on $\mathbb{F}_q$-vector spaces, following Nogin in [45].

Let $V = \mathbb{F}_q^m$ and write e_i for the standard basis vectors in V. The $\mathbb{F}_q$-rational points of the Grassmannian $\mathbb{G}(\ell, m)$ can be identified with the subset of $\mathbb{P}\left(\bigwedge^\ell V\right) \simeq \mathbb{P}^{\binom{m}{\ell}-1}$ corresponding to the *completely decomposable* elements of the exterior product $\bigwedge^\ell V$ (that is, nonzero elements of the form $\omega = w_1 \wedge w_2 \wedge \cdots \wedge w_\ell$ for some $w_i \in V$ that form a basis for the subspace they span).

The hyperplanes in $\mathbb{P}\left(\bigwedge^\ell V\right)$ correspond to elements of $\mathbb{P}\left(\bigwedge^\ell V\right)^*$, hence to elements of $\bigwedge^{m-\ell} V$ (up to scalars) via the nondegenerate pairing

$$\wedge : \bigwedge^{m-\ell} V \times \bigwedge^\ell V \to \bigwedge^m V \simeq \mathbb{F}_q.$$

It follows that the hyperplanes in $\mathbb{P}\left(\bigwedge^\ell V\right)$ all have the form

$$H(\alpha) = \mathbb{P}\left\{\omega \in \bigwedge^\ell V : \alpha \wedge \omega = 0\right\}$$

for some nonzero $\alpha \in \bigwedge^{m-\ell} V$.

Under these identifications, each hyperplane $\mathbf{V}(f)$ for f a linear form in the Plücker coordinates corresponds to $H(\alpha)$ for some α. For instance, $\mathbf{V}(p_{I_0})$ corresponds to $H(\alpha_0)$ for the completely decomposable element $\alpha_0 = e_{\ell+1} \wedge \cdots \wedge e_m$. All completely decomposable $\alpha \in \bigwedge^{m-\ell} V$ define hyperplane sections of the Grassmannian with the same number of $\mathbb{F}_q$-rational points. Call this number N_ℓ.

What must be proved is that if $\beta \in \bigwedge^{m-\ell} V$ is arbitrary, then the linear forms f in the Plücker coordinates defining the hyperplane $H(\beta)$ satisfy

$$\mathrm{wt}(ev_{\mathcal{S}}(f)) \geq N_\ell.$$

This follows by induction on ℓ using the easily checked fact that if $e \in V$ and $\alpha \in \bigwedge^{m-\ell} V$, then

$$\alpha \wedge e = 0 \Longleftrightarrow \alpha = \alpha' \wedge e \tag{7.18}$$

for some $\alpha' \in \bigwedge^{m-\ell-1} V$.

If $\ell = 1$, there is nothing to prove because every element of $\bigwedge^{m-1} V$ is completely decomposable. If $\ell > 1$, writing $[\ell]_q = \#\mathrm{GL}(\ell, \mathbb{F}_q)$,

$$\begin{aligned}\mathrm{wt}(ev_{\mathcal{S}}(f)) &= \#\{W = \mathrm{Span}(w_1, \ldots, w_\ell) : \beta \wedge w_1 \wedge \cdots \wedge w_\ell \neq 0\} \\ &= \#\{(w_1, \ldots, w_\ell) : \beta \wedge w_1 \wedge \cdots \wedge w_\ell \neq 0\} \,/\, [\ell]_q.\end{aligned}$$

Hence by the induction hypothesis, if α is completely decomposable

$$\begin{aligned}[\ell]_q \cdot \mathrm{wt}(ev_{\mathcal{S}}(f)) &= \sum_{w_1 : \beta \wedge w_1 \neq 0} \#\{(w_2, \ldots, w_\ell) : (\beta \wedge w_1) \wedge w_2 \wedge \cdots \wedge w_\ell \neq 0\} \\ &\geq \sum_{w_1 : \beta \wedge w_1 \neq 0} N_{\ell-1} \cdot [\ell - 1]_q \\ &= N_{\ell-1} \cdot [\ell - 1]_q \cdot \#\{w_1 : \beta \wedge w_1 \neq 0\} \\ &\geq N_{\ell-1} \cdot [\ell - 1]_q \cdot \#\{w_1 : \alpha \wedge w_1 \neq 0\} \quad \text{by Equation (7.18)} \\ &= [\ell]_q \cdot N_\ell.\end{aligned}$$

□

The exterior algebra language can also be used to say more about the weight distribution of $C(\mathbb{G}(\ell, m); \mathcal{S})$. For instance, the number of minimum weight words of this code is equal to the number of linear forms corresponding to completely decomposable α. This number is exactly $q - 1$ times the number of $\mathbb{F}_q$-rational points of the dual Grassmannian $\mathbb{G}(m - \ell, m)$, or

$$(q-1) \begin{bmatrix} m \\ m - \ell \end{bmatrix}_q = (q-1) \begin{bmatrix} m \\ \ell \end{bmatrix}_q .$$

For further information on these codes see the bibliographic notes in Section 7.8.

Codes on certain subvarieties of Grassmannians, the so-called *Schubert varieties*, have also been studied in detail by Chen, Guerra and Vincenti, and Ghorpade and Tsfasman. Let $\alpha = (\alpha_1, \ldots, \alpha_\ell) \in \mathbb{Z}^\ell$, where $1 \leq \alpha_1 \leq \cdots \leq \alpha_\ell \leq m$. If $B = \{v_1, \ldots, v_m\}$ is a fixed basis of $\mathbb{F}_q^m$, let A_i be the span of the first i vectors in B. Then the Schubert variety Ω_α is defined as

$$\Omega_\alpha = \{p(W) \in \mathbb{G}(\ell, m) : \dim W \cap A_{\alpha_i} \geq i\}. \tag{7.19}$$

See Section 7.8 for some pointers to the literature here.

Just as Grassmannians parametrize linear subspaces in $\mathbb{F}^m$, the *flag varieties* parametrize flags of linear subspaces, that is nested sequences of subspaces

$$V_1 \subset V_2 \subset \cdots \subset V_s,$$

where $\dim V_i = \ell_i$ and $0 < \ell_1 < \ell_2 < \ldots < \ell_s < m$. The flag is said to have *type* $(\ell_1, \ell_2, \ldots, \ell_s)$. Also set $\ell_{s+1} = m$ and $\ell_0 = 0$ by convention. The group $G = \mathrm{GL}(m, \mathbb{F})$ acts on the set of flags of each fixed type and the isotropy subgroup of a particular flag is a parabolic subgroup P conjugate to the group of block upper-triangular matrices with diagonal blocks M_r of sizes $\ell_r - \ell_{r-1}$ for $1 \leq r \leq s+1$. Hence the quotient G/P, which is denoted $\mathcal{F}(\ell_1, \ell_2, \ldots, \ell_s; m)$, classifies flags of type $(\ell_1, \ell_2, \ldots, \ell_s)$. The set G/P has the structure of a projective variety, which can be described as follows. Each V_i corresponds to a point of $\mathbb{G}(\ell_i, m)$. So the flag corresponds to a point of the product variety $\mathbb{G}(\ell_1, m) \times \cdots \times \mathbb{G}(\ell_s, m)$ and $\mathcal{F}(\ell_1, \ell_2, \ldots, \ell_s; m)$ is the subset of this product defined by the conditions $V_i \subset V_{i+1}$ for all i. This can be embedded in $\mathbb{P}^{N_1} \times \cdots \times \mathbb{P}^{N_s}$, for $N_i = \binom{m}{\ell_i}$, by the Plücker coordinates as in Equation (7.17). Finally, the product

$$\mathbb{P}^{N_1} \times \cdots \times \mathbb{P}^{N_s} \hookrightarrow \mathbb{P}^N$$

for $N = (N_1 + 1) \cdot \cdots \cdot (N_s + 1) - 1$ by another standard construction called the *Segre map*.

As in the Grassmannian case, $\mathbb{F}_q$-rational points on the flag variety $\mathcal{F}(\ell_1, \ell_2, \ldots, \ell_s; m)$ correspond to flags that are defined over $\mathbb{F}_q$. As an example of codes from flag varieties, consider the code $C(X; \mathcal{S})$ from $X = \mathcal{F}(1, m-1; m)$ (that is, the variety parametrizing flags $V_1 \subset V_2$ consisting of a line V_1 and a hyperplane V_2 containing that line). In this case

$$\mathcal{F}(1, m-1; m) \subset \mathbb{G}(1, m) \times \mathbb{G}(m-1, m) \simeq \mathbb{P}^{m-1} \times \mathbb{P}^{m-1} \hookrightarrow \mathbb{P}^{m^2-1}.$$

Theorem 7.23. *Let $\mathcal{S}$ be the set of all the $\mathbb{F}_q$-rational points on $X = \mathcal{F}(1, m-1; m)$. Then the $C(X; \mathcal{S})$ code has parameters*

$$\left[\frac{(q^m - 1)(q^{m-1} - 1)}{(q-1)^2}, m^2 - 1, q^{2m-3} - q^{m-2} \right].$$

The proof is due to Rodier and appears in [48]. The evaluation mapping using linear forms on $\mathbb{P}^{m^2-1}$ is *not* injective in this case because the condition that $V_1 \subset V_2$ is expressed by a linear equation in the coordinates of the Segre embedding of $\mathbb{P}^{m-1} \times \mathbb{P}^{m-1}$.

7.4.4. *Blow-ups and Del Pezzo surfaces*

Consider the surface $X = \mathbb{P}^2$. Let

$$Y_k \to Y_{k-1} \to \cdots \to Y_1 \to Y_0 = X, \tag{7.20}$$

be a sequence of morphisms where for all j, $\pi_j : Y_j \to Y_{j-1}$ is the blow up of an $\mathbb{F}_q$-rational point of the surface Y_{j-1}. The result will be a surface $Y = Y_k$ containing divisors $E_1, \ldots, E_k$ that are all contracted to a point on X. Each E_j is isomorphic to $\mathbb{P}^1$, and each contributes q additional $\mathbb{F}_q$-rational points. Therefore

$$\#Y(\mathbb{F}_q) = q^2 + q + 1 + kq,$$

which also attains the upper Weil bound for a surface with the Betti numbers of these examples. Whether this construction gives interesting codes depends very much on the the embedding of the surface Y into $\mathbb{P}^m$ (that is, on the linear series of divisors forming the hyperplane sections).

One famous family of examples of such surfaces are the so-called *Del Pezzo* surfaces. Hartshorne's text [28] and Manin [40] are good general references for these. By definition, a Del Pezzo surface is a surface of degree m in $\mathbb{P}^m$ on which the anticanonical line bundle $\mathcal{K}^{-1}$ is ample. A classical result in the theory of algebraic surfaces is that every Del Pezzo surface over an algebraically closed field $\mathbb{F}$ is obtained either as the degree 2 Veronese image of a quadric in $\mathbb{P}^3$, or as follows. Let ℓ be one of the integers $0, 1, \ldots, 6$, and take points $p_1, \ldots, p_\ell$ in $\mathbb{P}^2$ in general position (no three collinear, and no six contained in a conic curve). The linear system of *cubic* curves in $\mathbb{P}^2$ containing the base points $\{p_1, \ldots, p_\ell\}$ gives a rational map $\rho : \mathbb{P}^2 - - \to \mathbb{P}^{9-\ell}$. The image is a surface X_ℓ of degree $9 - \ell$ on which the points p_i blow up to exceptional divisors $E_i \simeq \mathbb{P}^1$ as in the composition of all the maps in Equation (7.20). Since the canonical sheaf on $\mathbb{P}^2$ is $\mathcal{K} \simeq \mathcal{O}_{\mathbb{P}^2}(-3)$, the anticanonical divisors are precisely the divisors in the linear system of cubics containing $\{p_1, \ldots, p_\ell\}$. For instance, with $\ell = 6$, X_ℓ is a cubic surface in $\mathbb{P}^3$, and every smooth cubic surface is obtained by blowing up some choice of points $p_1, \ldots, p_6$. With $\ell = 0$, the surface X_0 is the degree 3 Veronese image of $\mathbb{P}^2$, a surface of degree 9 in $\mathbb{P}^9$.

To get a Del Pezzo surface defined over $\mathbb{F}_q$, the points p_i should be $\mathbb{F}_q$-rational points in $\mathbb{P}^2$. This means that the construction above can fail for certain small fields (there may not be enough points p_i in general position). It suffices to take $q > 4$, however in order to construct the Del Pezzo surfaces with $0 \leq \ell \leq 6$.

By considering the possible hyperplane sections of the Del Pezzo surface Boguslavsky derives the following result in [3].

Theorem 7.24. *Let X_ℓ be the Del Pezzo surface constructed as above and*

let $q > 4$. The parameters of the $C(X_\ell)$ code are

$$n = q^2 + q + 1 + \ell q, \quad k = 10 - \ell,$$

and d given in the following table

ℓ	0	1	2	3	4	5	6
$d(C(X_\ell))$	$q^2 - 2q$	$q^2 - 2q$	$q^2 - 2q$	$q^2 - 2q + 1$	q^2	$q^2 + 2q$	$q^2 + 4q + 1^*$

The case $\ell = 6$ corresponds to the code from a cubic surface in $\mathbb{P}^3$. Note the asterisk in the table above. In the generic case, there are plane sections of a cubic surface consisting of three lines forming a triangle, but no sections consisting of three concurrent lines. The triangle plane sections contain the maximum number of $\mathbb{F}_q$-rational points, namely $3q$. Hence $d(C(X_6)) = q^2 + 7q + 1 - 3q = q^2 + 4q + 1$, as claimed in this case. For some special configurations of points p_i, however, the corresponding cubic surface will have *Eckardt points* where there is a plane section consisting of three concurrent lines. For those surfaces, the minimum distance is $q^2 + 4q$ rather than $q^2 + 4q + 1$.

7.4.5. *Ruled surfaces and generalizations*

A *ruled surface* is a surface X with a mapping $\pi : X \to C$ to a smooth curve C, whose fibers over all points of C are $\mathbb{P}^1$'s. Moreover, it is usually required that π has a section, that is, a mapping $\sigma : C \to X$ such that $\pi \circ \sigma$ is the identity on C. For instance, over an algebraically closed field, quadric surfaces in $\mathbb{P}^3$ are isomorphic to the product ruled surface $\mathbb{P}^1 \times \mathbb{P}^1$. For background on these varieties, Chapter V of [28] is a good reference.

Starting from a curve C and a vector bundle of rank 2 (that is, a locally free sheaf of rank 2) $\mathcal{E}$ on C, the projective space bundle $X = \mathbb{P}(\mathcal{E})$ is a ruled surface. Conversely, every ruled surface $\pi : X \to C$ is isomorphic to $\mathbb{P}(\mathcal{E})$ for some locally free sheaf of rank 2 on C. Given a curve C and two vector bundles on C, the ruled surfaces $\mathbb{P}(\mathcal{E})$ and $\mathbb{P}(\mathcal{E}')$ are isomorphic if and only if $\mathcal{E} \simeq \mathcal{E}' \otimes \mathcal{L}$ for some line bundle $\mathcal{L}$ on C. By choosing $\mathcal{L}$ appropriately, it is possible to make $H^0(\mathcal{E}) \neq 0$ but $H^0(\mathcal{E} \otimes \mathcal{M}) = 0$ whenever $\mathcal{M}$ is a line bundle on C of negative degree and in this case we say $\mathcal{E}$ is *normalized*. Then there is a section C_0 of X with $C_0^2 = -e$ where $e = \deg(E)$ is the degree of the divisor E on C corresponding to the line bundle $\bigwedge^2 \mathcal{E}$. If $\mathcal{E}$ is decomposable (a direct sum of two line bundles) and normalized, then $e \geq 0$. If $\mathcal{E}$ is indecomposable, then it is known that $-g(C) \leq e \leq 2g(C) - 2$, where $g(C)$ is the genus.

Up to numerical equivalence, each divisor D on X is $D \sim b_1 C_0 + b_2 f$, where f is a fiber of the mapping π and $b_1, b_2 \in \mathbb{Z}$. The intersection product on divisors is determined by the relations $C_0^2 = -e$, $C_0 \cdot f = 1$, $f^2 = 0$. S.H. Hansen has shown the following result.

Theorem 7.25. *Let $\pi : X \to C$ be a normalized ruled surface with invariant $e \geq 0$. Let $\#C(\mathbb{F}_q) = a$, and let $\mathcal{S}$ be the full set of $\mathbb{F}_q$-rational points on X. Let $\mathcal{L} = \mathcal{O}_X(b_1 C_0 + b_2 f)$. Then the $C(X, \mathcal{L}; \mathcal{S})$ code has parameters*

$$[a(q+1), \dim\Gamma(X, \mathcal{L}), d \geq n - b_2(q+1) - (a - b_2)b_1],$$

(provided that $b_2 < a$ and the bound on d is positive).

Proof. Let $f_1, \ldots, f_a$ be the fibers of π over the $\mathbb{F}_q$-rational points of C. These are disjoint curves on X isomorphic to $\mathbb{P}^1$, hence contain $q+1$ $\mathbb{F}_q$-rational points each. Every $\mathbb{F}_q$-rational point of X lies on one of these lines, so $n = a(q+1)$. As usual, the statement for k follows if $d > 0$. The estimate for d comes from the method of Theorem 7.7 applied to the covering family of curves $f_1, \ldots, f_a$. In the notation of that theorem, we have $N = q + 1$ and $\eta = (b_1 C_0 + b_2 f) \cdot f = b_1$. At most $\ell = b_2$ of the fibers are contained in any divisor D corresponding to a global section of $\mathcal{O}_X(b_1 C_0 + b_2 f)$ since $D \cdot C_0 = (b_1 C_0 + b_2 f) \cdot C_0 = -eb_1 + b_2 \leq b_2$. The bound on d follows immediately. □

The dimension of the space of global sections of $\mathcal{L}$ can be computed via divisors on C because of general facts about sheaves on the projective space bundle $\mathbb{P}(\mathcal{E})$ (see [28], Lemma V.2.4). See the bibliographic notes in Section 7.8 for more information about these codes and for work on codes from projective bundles of higher fiber dimension.

7.4.6. *Other work on codes from surfaces*

One interesting recent contribution to the search for good codes from higher-dimensional varieties is described in the unpublished preprint [56] of Voloch and Zarzar and the article [61] of Zarzar. Following Observation 7.6, Voloch and Zarzar seek good surfaces for constructing codes by limiting the presence of reducible $\mathbf{V}(f) \cap X$ via control of the rank of the Néron-Severi group of X.

7.5. Codes from Deligne-Lusztig Varieties

Some of the most interesting varieties that have been used to produce codes by the constructions of Section 7.2 are the so-called *Deligne-Lusztig varieties* from representation theory. As we will see, their description involves several of the general processes on varieties involved in the examples above.

Let G be a connected reductive affine algebraic group over the algebraic closure $\mathbb{F}$ of $\mathbb{F}_q$, a closed subgroup of $\mathrm{GL}(n, \mathbb{F})$ for some n. We have the q-Frobenius endomorphism $F : G \to G$ whose fixed points are the $\mathbb{F}_q$-rational points of G.

A *Borel subgroup* of G is a maximal connected solvable subgroup of G. A *torus* is a subgroup of G isomorphic to $(\mathbb{F}^*)^s$ for some s. All Borel subgroups are conjugate, and each maximal torus T is contained in some Borel subgroup. Let $N(T)$ be the normalizer of T in G. The quotient $N(T)/T$ is a finite group called the *Weyl group* of G.

The set $\mathcal{B}$ of all Borel subgroups of G can be identified with the quotient G/B for any particular B via the mapping $G/B \to \mathcal{B}$ given by $g \mapsto g^{-1}Bg$. If $w \in W$, then the Deligne-Lusztig variety associated to w can be described as follows. Let B be an F-stable Borel subgroup, then

$$X(w) = \{x \in G : x^{-1}F(x) \in BwB\}/B \subset \mathcal{B}.$$

Theorem 7.26. *Let $w = s_1 \cdots s_n$ be a minimal factorization of w into simple reflections in W, the Weyl group of G as above. Then*

(1) $X(w)$ is a locally closed smooth variety of pure dimension n.
(2) The variety $X(w)$ is fixed by the action of the group G^F and is defined over $\mathbb{F}_{q^\delta}$, where δ is the smallest integer such that F^δ fixes w.
(3) The closure of $X(w)$ in $\mathcal{B}$ is the union of the $X(s_{i_1} \cdots s_{i_r})$ such that $1 \le i_1 < i_2 < \cdots < i_r \le n$ and $X(e)$.

We refer to [5] for the classification of reductive G in terms of Dynkin diagrams with action of F. In [21], J. Hansen studied the Hermitian curves over $\mathbb{F}_{q^2}$, the Suzuki curves over $\mathbb{F}_{2^{2n+1}}$ and the Ree curves over $\mathbb{F}_{3^{2n+1}}$, all well-known maximal curves, and all used to construct interesting Goppa codes with very large automorphism groups. Hansen showed that the underlying reason these particular curves are so rich in good properties is that they are the Deligne-Lusztig varieties for groups G for which there is just one orbit of simple reflections in the Weyl group under the action of F. The Hermitian curves come from groups of type ${}^2A_2$, the Suzuki curves come

from the groups of type ${}^2B_2$, and the Ree curves from the groups of type ${}^2G_2$.

It is known that there are seven cases in which there are two F-orbits in the set of reflections in W, so taking s_1, s_2 from the distinct orbits, the Deligne-Lusztig construction with $w = s_1 s_2$ leads to algebraic surfaces:

$$A_2, C_2, G_2, {}^2A_3, {}^2A_4, {}^3D_4, {}^2F_4.$$

One of these cases is relatively uninteresting. In [47], Rodier shows that the complete, smooth Deligne-Lusztig variety $\overline{X}(s_1, s_2)$ from the group of type A_2 is isomorphic to the blow-up of $\mathbb{P}^2$ at all of its $\mathbb{F}_q$-rational points.

For the group of type ${}^2A_3$, however, Rodier shows that $\overline{X}(s_1, s_2)$ is isomorphic to the blow-up of the Hermitian surface in $\mathbb{P}^3$ at its $\mathbb{F}_{q^2}$-rational points. Hence as in the discussion of the blow-ups of $\mathbb{P}^2$ above, and using Example 7.14, we get a surface with $(q^3+1)(q^2+1)^2$ points.

Similarly the $\overline{X}(s_1, s_2)$ from a group of type ${}^2A_4$ is isomorphic to the blow-up of the complete intersection Y of the two hypersurfaces

$$\begin{aligned} 0 &= x_0^{q+1} + x_1^{q+1} + \cdots + x_4^{q+1} \\ 0 &= x_0^{q^3+1} + x_1^{q^3+1} + \cdots + x_4^{q^3+1} \end{aligned} \tag{7.21}$$

in $\mathbb{P}^4$ at the $(q^5+1)(q^2+1)$ $\mathbb{F}_{q^2}$-rational points on that surface. (These are the same as the $\mathbb{F}_{q^2}$-rational points on the Hermitian 3-fold in $\mathbb{P}^4$ defined by the first equation.) It is easy to check that these points are all singular, and in fact they blow up to Hermitian curves (not $\mathbb{P}^1$'s) on the Deligne-Lusztig surface. Hence the Deligne-Lusztig surface X has a very large number of $\mathbb{F}_{q^2}$-rational points in this case,

$$\#X(\mathbb{F}_{q^2}) = (q^5+1)(q^2+1)(q^3+1).$$

Rodier determines the structure and number of $\mathbb{F}_{q^\delta}$-rational points in the $G_2, {}^3D_4$, and ${}^2F_4$ cases as well. Interestingly enough, his method is to realize the Deligne-Lusztig varieties as certain subsets of flag varieties as above, where the subspaces in the flags are related to each other using the Frobenius endomorphism.

Rodier and S.H. Hansen also discuss the properties of the $C_h(X)$ codes on these varieties. For instance in [26], Hansen shows the following result by relating codes on Y from Equation (7.21) and codes on the Deligne-Lusztig surface itself.

Theorem 7.27. *Let X be the Deligne-Lusztig surface of type ${}^2A_4$ over the*

field $\mathbb{F}_{q^2}$. *For* $1 \leq h \leq q^2$, *there exist codes over* $\mathbb{F}_{q^2}$ *with*

$$\begin{aligned} n &= (q^5+1)(q^3+1)(q^2+1), \\ k &= \binom{4+h}{h} - \binom{4+h-(q+1)}{t-(q+1)}, \text{ and} \\ d &\geq n - hP(q), \end{aligned}$$

where $P(q) = (q^3+1)(q^5+1) + (q+1)(q^3+1)(q^2-h+1)$.

Since $P(q)$ has degree 8 in q, this shows that $d + k \geq n - O(n^{4/5})$ with $n = O(q^{10})$, some very long codes indeed! Hansen also considers the codes obtained from the singular points on the complete intersection from Equation (7.21) (that is from the Hermitian 3-fold).

7.6. Connections with Other Code Constructions

In this section we point out some connections between the construction presented here and some other examples of algebraic geometric codes related to higher dimensional varieties in the literature. There is a close connection between the codes $C(X, \mathcal{L}; \mathcal{S})$ and the *toric codes* constructed from polytopes or fans in $\mathbb{R}^s$ as in [22] or Chapter 8 of this volume. A toric variety of dimension s over an algebraically closed field $\mathbb{F}$ is a variety X containing a Zariski-open subset isomorphic to the s-dimensional algebraic torus $T \simeq (\mathbb{F}^*)^s$ and on which T acts in a manner compatible with the multiplicative group structure on T. The combinatorial data in a fan Σ in $\mathbb{R}^s$ encodes the gluing information needed to produce a normal toric variety X_Σ from affine open subsets of the form $\mathrm{Spec}(\mathbb{F}[S_\sigma])$ where $\mathbb{F}[S_\sigma]$ is a semigroup algebra associated to the cone σ in the fan Σ. A polytope P in $\mathbb{R}^s$ determines a normal fan Σ_P and line bundle $\mathcal{L}_\mathcal{P}$ on X_{Σ_P}. The toric codes are codes $C(X, \mathcal{L}; \mathcal{S})$ for $X = X_{\Sigma_P}$, $\mathcal{L} = \mathcal{L}_P$ and $\mathcal{S} = T \cap \mathbb{F}_q^s = \left(\mathbb{F}_q^*\right)^s$. It is not difficult to see that toric codes are s-dimensional cyclic codes with certain other properties generalizing those of Reed-Solomon codes.

The study of decoding algorithms for one-point algebraic geometric Goppa codes has been unified and simplified by the theory of *order domains* discussed in [14, 32]. The article [38] shows how order domains can be constructed from many of the higher dimensional varieties discussed here.

7.7. Code Comparisons

It is instructive to compare codes constructed by the methods described in this chapter and the best currently known codes for the same n, k. We will focus on the minimum distance, although there are many other considerations too in deciding on codes for given applications. According to Observation 7.6, when highly reducible $\mathbf{V}(f) \cap X$ exist for some f in $\mathcal{F}_h$, the resulting codes may not be very good.

All comparisons will be made by means of the online tables of Markus Grassl, [18]. One initial observation is that many of the varieties X that we have discussed have *so many* $\mathbb{F}_q$-rational points that the $C(X)$ codes have extremely low information rates k/n and the n achieved are far beyond the ranges explored to date. When no explicit codes are known, it is still possible to make comparisons with general bounds. Since the k for most of the $C_h(X)$ codes we have seen are much smaller than n, the *Griesmer bound* yields some information. The usual form of the Griesmer bound (see [33]) says that for an $[n, k, d]$ code over $\mathbb{F}_q$,

$$n \geq \sum_{i=0}^{k-1} \left\lceil \frac{d}{q^i} \right\rceil .$$

Given n, k, this inequality can also be used to derive an upper bound on realizable d for $[n, k]$ codes that, in a sense, improves the Singleton bound $d \leq n - k + 1$. It should be noted, however, that there are many pairs n, k for which there are no codes attaining the Griesmer upper bound on d.

We begin by noting the following well-known fact.

Theorem 7.28. *The projective Reed-Muller codes with $h = 1$ from Theorem 7.5 attain the Griesmer upper bound for all m.*

This follows since $n = \#\mathbb{P}^m(\mathbb{F}_q) = q^m + \cdots + q + 1$, $d = q^m$, and $k = m + 1$.

For $h > 1$, however, the presence of *reducible* forms of degree h, which can have many more $\mathbb{F}_q$-rational zeroes than irreducible forms (see the proof of Theorem 7.5), tends to reduce the minimum distance relative to other code constructions. This is true for all q, although the difference shows up for smaller h the larger q is.

For instance, in the binary case, the $h = 2$ projective Reed-Muller code with $m = 5$ has parameters $[63, 21, 16]$, but there are binary $[63, 21, 18]$ codes known by [18]. Similarly, with $q = 4$, the $h = 2$ projective Reed-Muller code with $m = 3$ over $\mathbb{F}_4$ has parameters $[85, 10, 48]$, but there

are $[85, 10, 52]$ codes known over $\mathbb{F}_4$ by [18]. In the cases that have been explored in detail, the gap between the projective Reed-Muller codes and the best known codes seems to increase with m for fixed h, and also for h with fixed m (for the cases $h < q + 1$ considered here at least).

The minimum distance for the $C(X)$ codes from quadrics from Equation (7.13) also tend to be relatively close to the Griesmer bound for their n, k, although the bounds grow slightly faster than the actual d as $m \to \infty$ and slightly better codes are known in a number of cases. The codes from *elliptic* quadrics ($w = 0$) are superior in general to those from hyperbolic quadrics ($w = 2$) when m is odd. When $m = 3$, this is a reflection of the phenomenon noted in Observation 7.6. For larger odd m, this is an interesting example showing that the "greedy" approach of maximizing $n = \#X(\mathbb{F}_q)$ does not always yield the best codes.

For example, over $\mathbb{F}_8$, the $C(X)$ code from a hyperbolic quadric in $\mathbb{P}^3$ has parameters $[81, 4, 64]$, but there are $[81, 4, 68]$ codes known by [18]. (The Griesmer bound in this case gives $d \leq 69$.) By way of contrast, the $C(X)$ code from an elliptic quadric has parameters $[65, 4, 56]$, and this is the best possible by the Griesmer bound. Similar patterns hold over all of the small fields where systematic exploration has been done. For larger m, however, it is *not* always the case that $C(X)$ codes from elliptic quadrics meet the Griesmer bound, and there are slightly better known codes in some cases. The $C_2(X)$ codes from quadrics seem to be similar, at least in the case $m = 3$, where the results of Edoukou from [12] can be applied. Over $\mathbb{F}_8$ for instance, the $C_2(X)$ code from a hyperbolic quadric surface has parameters $[81, 9, 49]$, but there are $[81, 9, 58]$ codes known by [18]. On the other hand, the $C_2(X)$ code from an elliptic quadric has parameters $[65, 9, 47]$, and this matches the best known d for this n, k over $\mathbb{F}_8$. (The tightest known upper bound is $d \leq 50$.)

The Hermitian hypersurface codes seem to be similar to those from quadrics. The $C(X)$ codes are quite good, coming quite near the Griesmer bound. For instance, the Hermitian surface code from Theorem 7.18 over $\mathbb{F}_{16}$ has parameters $[1105, 4, 1024]$. This is far outside the range of n and fields for which tables are available, but by way of comparison, $d \leq 1034$ by the Griesmer bound. However, the $C_2(X)$ codes are nowhere near as good, and the gap grows with h.

The codes from Del Pezzo surfaces from Theorem 7.24 seem to be interesting only for $\ell = 0$ (the case $X \simeq \mathbb{P}^2$) and $\ell = 6$ (the case of the cubic surface in $\mathbb{P}^3$). The intermediate cases are quite inferior to the best known codes because hyperplane sections can contain many of the exceptional di-

visors (an instance of Observation 7.6).

For the other families of varieties we have considered (Grassmannians, flag varieties, Deligne-Lusztig varieties), once q or m get even moderately large, n is so huge that very little is known. On the basis of rather limited evidence, the Grassmannian and flag variety codes might be especially good only over very small fields, though. For example, the $C(X)$ code from $X = \mathbb{G}(2,4)$ over $\mathbb{F}_2$ has parameters $[35, 6, 16]$, which attains the Griesmer bound. Over $\mathbb{F}_3$, the corresponding Grassmannian code has $[130, 6, 81]$, but there are $[130, 6, 84]$ codes over $\mathbb{F}_3$ known by [18] and the Griesmer bound gives $d \leq 84$ in this case.

It is unrealistic to expect every code constructed from a variety of dimension ≥ 2 to be a world-beater. The study of error control codes constructed from higher dimensional varieties is an area where it is certainly true that we have just barely begun feeling out the lay of the land and just barely scratched the surface of what should be possible. If this survey of past work inspires further exploration, then one of its goals will have been achieved!

7.8. Bibliographic Notes

Section 7.1. The universality of the Goppa construction for producing linear codes is proved in [46]. This refers specifically to Pellikaan, Shen, and van Wee's result that every linear code is *weakly algebraic-geometric*: Given C, there exists a smooth projective curve X, a set $\mathcal{S}$ of $\mathbb{F}_q$-rational points on X, and a line bundle $\mathcal{L} = \mathcal{O}(G)$ for some divisor G with support disjoint from $\mathcal{S}$, such that C is isomorphic to $C(X, \mathcal{L}; \mathcal{S})$ (with no restriction on the degree of G).

Although very little work to date has been done on decoding methods, the large groups of automorphisms of some of the varieties considered here make the *permutation decoding* paradigm a possibility for certain of these codes. Some work along these lines has been done by Kroll and Vincenti, [34, 35].

Section 7.2. Both forms of the construction of codes from varieties (Definitions 7.1 and 7.2) come from [55], which was the first place where this idea was described in published form. The form in Definition 7.2 can be made even more concrete and less algebraic-geometric by the language of *projective systems* of points and their associated codes.

Section 7.3. Theorem 7.5 is taken from [37]. It does not include the codes

for $h > q$ because the evaluation mapping is no longer injective in those cases, The parameters of the C_h codes for $h > q$ have been studied by Lachaud in [36] and Sørensen in [54]. The generalized Hamming weights d_r for the Reed-Muller codes have been studied by Heijnen and Pellikaan in [29]. Some ideas about finding good subcodes of the C_2 codes have been presented by Brouwer in [4].

Theorem 7.7, the following example, and the bound using Seshadri constants in Theorem 7.9 are all due to S.H. Hansen and are taken from [26].

The results on bounds for the minimum distance when $\mathcal{S}$ is a complete intersection come from [17] and that article's bibliography gives several sources for the Cayley-Bacharach theorem and modern generalizations. The genesis for this was the observation that if $\mathcal{S}$ is a reduced complete intersection of two cubic curves in $\mathbb{P}^2$, and Γ' is any subset of eight of the nine points in $\mathcal{S}$, then every cubic that contains the eight points in Γ' also passes through the ninth point in $\mathcal{S}$. Related applications to coding theory were discussed by Duursma, Renteria and Tapia-Recillas in [10] and J. Hansen in [23]. The theorem stated here can also be extended to yield a criterion for MDS codes.

The Weil conjectures were originally stated in [59] and proved in complete generality by Deligne in [9] following three decades of work by Dwork, Serre, Artin, Grothendieck, Verdier, and many others. Weil's paper gives a different form for middle Betti number in Equation (7.8), but it can be seen that his form is equivalent to ours. The discussion of Weil-type bounds follows Lachaud's presentation in [37]. Because of space limitations and the significantly higher prerequisites needed to work with the ℓ-adic étale cohomology theory in any detail in higher codimension, we have focused only on the application of Lachaud's results to codes from hypersurfaces. The discussion in [37] is considerably more general. Edoukou has verified Sørensen's conjecture (see Equation (7.14)) on the Hermitian surface codes in the case $h = 2$ in [11].

Section 7.4. The codes from quadrics have been intensively studied since at least the 1975 article [60] of Wolfmann. They are especially accessible because so much is known about the sets of $\mathbb{F}_q$-rational points on quadrics as finite geometries; see Hirschfeld and Thas, [30]. The complete hierarchies of generalized Hamming weights d_r for the $C(X)$ codes were determined independently by Nogin in [44] and Wan in [57]. To aid in comparing these different sources, we note that Wan's invariant δ is related to Hirschfeld and Thas's (and our) character w by $\delta = 2 - w$. The character can also

be defined by $w = 2g - m + 3$ where g is the dimension of the largest linear subspace of $\mathbb{P}^m$ contained in the quadric X. Comparatively little has appeared in the literature concerning the $C_h(X)$ codes with $h > 1$ on quadrics following the work of Aubry in [1]. One recent article studying the $C_2(X)$ codes from quadrics in $\mathbb{P}^3$ is Edoukou, [12].

Hirschfeld and Thas also contains a wealth of information related to the codes on Hermitian hypersurfaces. The parameters of the $C(X)$ codes were established by Chakravarti in [6], and the generalized Hamming weights were determined in by Hirschfeld, Tsfasman, and Vladut in [31].

Grassmannian codes were studied first in the binary case by C. Ryan and K. Ryan in [49–51]. The material on Grassmannian codes presented here is taken from [45]. In that article, Nogin also determines the complete weight distribution for the codes from $X = \mathbb{G}(2, m)$ and shows that the generalized weights d_r of the Grassmann codes meet the generalized Griesmer bound when $r \leq \max\{\ell, m - \ell\} + 1$. More information on the generalized weights was established by Ghorpade and Lachaud in [15] and these codes are also discussed as a special case of the code construction from flag varieties by Rodier in [48]. That article also gives the proof of Theorem 7.23. Codes from the Schubert varieties defined in Equation (7.19) have been studied in [7, 16, 20].

The material on Del Pezzo surface codes is taken from Boguslavsky, [3]. That article also determines the complete hierarchy of generalized Hamming weights d_r for these codes.

Codes from ruled surfaces were studied by S.H. Hansen in [26] as an example of how the bound from Theorem 7.7 could be applied. That article also addresses the cases where the invariant $e < 0$, and presents some examples involving ruled surfaces over the Hermitian elliptic curve over $\mathbb{F}_4$. Codes from ruled surfaces were also considered in Lomont's thesis, [39]. The results for codes over ruled surfaces have been generalized to give corresponding results for codes on projective bundles $\mathbb{P}(\mathcal{E})$ for $\mathcal{E}$ of all ranks $r \geq 2$ by Nakashima in [43]. Nakashima also considers codes on Grassmann, quadric, and Hermitian bundles in [42].

Other work on codes from algebraic surfaces is contained in the Ph.D. theses of Lomont, [39], and Davis, [8].

Section 7.5. Rodier's article [47] is a gold mine of information and techniques for the Deligne-Lusztig surfaces and Deligne-Lusztig varieties more generally. The original article of Deligne and Lusztig and a number of other works devoted to this construction are referenced in the bibliography. The

Picard group and other aspects of the finer structure of Deligne-Lusztig varieties have been studied by S.H. Hansen in [24–26]. Hansen's thesis, [24] contains chapters corresponding to the other articles here.

Section 7.6. A standard reference for the theory of toric varieties over $\mathbb{C}$ is Fulton's text, [13]; the construction generalizes to fields of characteristic p with no difficulty. See Chapter 8 of this volume [41] and the references there for other studies of toric codes.

References

[1] Y. Aubry, Reed-Muller codes associated to projective algebraic varieties, in: *Coding Theory and Algebraic Geometry (Proceedings, Luminy 1991)*, H. Stichtenoth and M.A. Tsfasman, eds. Springer Lecture Notes in Mathematics 1518 (Springer, Berlin, 1992), 4–17.

[2] E. Ballico and C. Fontanari, The Horace method for error-correcting codes, *Appl. Algebra Engrg. Comm. Comput.* **17**, 135–139 (2006).

[3] M.I. Boguslavsky, Sections of Del Pezzo surfaces and generalized weights, *Probl. Inf. Transm.* **34**, 14–24 (1998).

[4] A. Brouwer, Linear spaces of quadrics and new good codes, *Bull. Belg. Math. Soc.* **5**, 177-180 (1998).

[5] R. Carter, *Finite Groups of Lie Type* (Wiley, New York, 1985).

[6] I.M. Chakravarti, Families of codes with few distinct weights from singular and nonsingular Hermitian varieties and quadrics in projective geometries and Hadamard difference sets and designs associated with two-weight codes, in: *Coding Theory and Design Theory, I*, IMA Vol. Math Appl. 20 (Springer, New York, 1990), 35–50.

[7] H. Chen, On the minimum distance of Schubert codes, *IEEE Trans. Inform. Theory*, **46**, 1535–1538 (2000).

[8] J. Davis, Algebraic geometric codes on anticanonical surfaces, Ph.D. thesis, University of Nebraska, 2007.

[9] P. Deligne, La conjecture de Weil, I, *Publ. Math. IHES* **43**, 273–307 (1974).

[10] I. Duursma, C. Renteria and H. Tapia-Recillas, Reed-Muller codes on complete intersections, *Algebra Engrg. Comm. Comput.* **11**, 455–462 (2001).

[11] F. Edoukou, Codes defined by forms of degree 2 on hermitian surfaces and Sørensen's conjecture, *Finite Fields Appl.* **13**, 616–627 (2007).

[12] F. Edoukou, Codes defined by forms of degree 2 on quadric surfaces, *IEEE Trans. Inform. Theory* **54**, 860–864 (2008).

[13] W. Fulton, *Introduction to toric varieties*, Annals of Mathematics Studies 131, (Princeton University Press, Princeton,1993).

[14] O. Geil and R. Pellikaan, On the structure of order domains, *Finite Fields Appl.* **8**, 369–396 (2002).

[15] S. Ghorpade and G. Lachaud, Higher weights of Grassmann codes, in: *Coding Theory, Crypotography, and Related Areas, Proceedings Guanaju-*

ato 1998, J. Buchmann, T. Høholdt, H. Stichtenoth, H. Tapia-Recillas eds. (Springer, Berlin, 2000), 122–131.

[16] S. Ghorpade and M. Tsfasman, Schubert varieties, linear codes and enumerative combinatorics, *Finite Fields Appl.* **11**, 684–699 (2005).

[17] L. Gold, J. Little and H. Schenck, Cayley-Bacharach and evaluation codes on complete intersections, *J. Pure Appl. Algebra* **196**, 91–99 (2005).

[18] M. Grassl, Bounds on minimum distance of linear codes, available online at `http://www.codetables.de`, accessed on 2008-02-02.

[19] P. Griffiths and J. Harris, *Principles of Algebraic Geometry* (Wiley, New York, 1978).

[20] L. Guerra and R. Vincenti, On the linear codes arising from Schubert varieties, *Des. Codes Cryptogr.* **33**, 173–180 (2004).

[21] J. Hansen, Deligne-Lusztig varieties and group codes, in: *Coding Theory and Algebraic Geometry (Proceedings, Luminy 1991)*, H. Stichtenoth and M.A. Tsfasman, eds. Springer Lecture Notes in Mathematics 1518 (Springer, Berlin, 1992), 63–81.

[22] J. Hansen, Toric surfaces and error correcting codes, *Coding Theory, Cryptography, and Related Areas, Proceedings Guanajuato 1998*, J. Buchmann, T. Høholdt, H. Stichtenoth, H. Tapia-Recillas eds. (Springer, Berlin, 2000), 132–142.

[23] J. Hansen, Linkage and codes on complete intersections, *Appl. Algebra Engrg. Comm. Comput.* **14**, 175–185 (2003).

[24] S.H. Hansen, The geometry of Deligne-Lusztig varieties: Higher dimensional AG codes, Ph.D. thesis University of Aarhus, 1999.

[25] S.H. Hansen, Canonical bundles of Deligne-Lusztig varieties. *Manuscripta Math.* **98** 363–375 (1999).

[26] S.H. Hansen, Error-correcting codes from higher-dimensional varieties, *Finite Fields Appl.* **7**, 530–552 (2001).

[27] S.H. Hansen, Picard groups of Deligne-Lusztig varieties—with a view toward higher codimensions, *Beiträge Algebra Geom.* **43**, 9–26 (2002).

[28] R. Hartshorne, *Algebraic Geometry* (Springer, New York, 1977).

[29] P. Heijnen and R. Pellikaan, Generalized Hamming weights of q-ary Reed Muller codes, *IEEE Trans. Inform. Theory* **44**, 181–196 (1998).

[30] J.W.P. Hirschfeld and J.A.Thas, *General Galois Geometries* (Oxford University Press, Oxford, 1991).

[31] J.W.P. Hirschfeld, M. Tsfasman and S.G. Vladut, The weight hierarchy of higher dimensional Hermitian codes, *IEEE Trans. Inform. Theory* **40**, 275–278 (1994).

[32] T. Høholdt, J. van Lint and R. Pellikaan, Algebraic geometry codes, in: *Handbook of Coding Theory*, W. Huffman and V. Pless, eds. (Elsevier, Amsterdam, 1998), 871–962.

[33] W. Huffman and V. Pless, *Fundamentals of Error-Correcting Codes*, (Cambridge University Press, Cambridge, 2003).

[34] H.-J. Kroll and R. Vincenti, PD-sets for the codes related to some classical varieties, *Discrete Math.* **301**, 89–105 (2005).

[35] H.-J. Kroll and R. Vincenti, PD-sets for binary RM-codes and the codes

related to the Klein quadric and to the Schubert variety of $PG(5,2)$, *Discrete Math.* **308**, 408–414 (2008).

[36] G. Lachaud, The parameters of projective Reed-Muller codes, *Discrete Math.* **81**, 217–221 (1990).

[37] G. Lachaud, Number of points of plane sections and linear codes defined on algebraic varieties, in: *Arithmetic, Geometry and Coding Theory, Proceedings Luminy 1993*, R. Pellikann, M. Perret, S.G. Vladut eds. (Walter de Gruyter, Berlin, 1996), 77–104.

[38] J. Little, The ubiquity of order domains for the construction of error control codes, *Adv. Math. Communications* **1**, 1–27 (2007).

[39] C. Lomont, *Error correcting codes on algebraic surfaces*, Ph.D. thesis, Purdue University, 2003, `arXiv:math/0309123`.

[40] Yu.I. Manin, *Cubic Forms: Algebra, Geometry, Arithmetic*, (North Holland, Amsterdam, 1986).

[41] E. Martínez-Moro and D. Ruano, Toric Codes. In eds. E. Martínez-Moro, C. Munuera, and D. Ruano, *Advances in Algebraic Geometry Codes*, chapter 8. pp. 295–322. World Scientific, (2008).

[42] T. Nakashima, Codes on Grassmann bundles, *J. Pure Appl. Algebra* **199**, 235-244 (2005).

[43] T. Nakashima, Error-correcting codes on projective bundles, *Finite Fields Appl.* **12**, 222–231 (2006).

[44] D.Yu. Nogin, Generalized Hamming weights of codes on multidimensional quadrics, *Probl. Inf. Transm.* **29**, 21–30 (1993).

[45] D.Yu. Nogin, Codes associated to Grassmannians, in: *Arithmetic, Geometry and Coding Theory, Proceedings Luminy 1993*, R. Pellikann, M. Perret, S.G. Vladut eds. (Walter de Gruyter, Berlin, 1996), 145–154.

[46] R. Pellikaan, B.-Z. Shen and G. van Wee, Which linear codes are algebraic-geometric? *IEEE Trans. Inform. Theory* **IT-37**, 583–602 (1991).

[47] F. Rodier, Nombre de points des surfaces de Deligne et Lusztig, *J. Algebra* **227**, 706–766 (2000).

[48] F. Rodier, Codes from flag varieties over a finite field, *J. Pure Appl. Algebra* **178**, 203–214 (2003).

[49] C.T. Ryan, An application of Grassmannian varieties to coding theory, *Congr. Numer.* **57**, 257–271 (1987).

[50] C.T. Ryan, Projective codes based on Grassmann varieties, *Congr. Numer.* **57**, 273–279 (1987).

[51] C.T. Ryan and K.M. Ryan, The minimum weight of Grassmannian codes $C(k,n)$, *Discrete Appl. Math.* **28**, 149–156 (1990).

[52] J.P. Serre, *Lettre à Tsfasman*, in: Journées Arithmetiques, 1989 (Luminy, 1989), *Asterisque* **198-200**, 351–353 (1991).

[53] A. Sørensen, Rational points on hypersurfaces, Reed-Muller codes, and algebraic-geometric codes, Ph.D. thesis, Aarhus, 1991.

[54] A. Sørensen, Projective Reed-Muller codes, *IEEE Trans. Inform. Theory* **37**, 1567–1576 (1991).

[55] M.A. Tsfasman and S.G.Vladut, *Algebraic-geometric codes* (Kluwer, Dordrecht, 1991).

[56] J.Voloch and M. Zarzar, *Algebraic geometric codes on surfaces*, preprint.
[57] Z. Wan, The weight hierarchies of the projective codes from nondegenerate quadrics, *Des. Codes Cryptogr.* **4**, 283–300 (1994).
[58] V.K. Wei, Generalized Hamming weights for linear codes, *IEEE Trans. Inform. Theory* **37**, 1412–1418 (1991).
[59] A. Weil, Numbers of solutions of equations in finite fields, *Bull. Amer. Math. Soc.* **55**, 497–508. (1949).
[60] J. Wolfmann, Codes projectifs a deux ou trois poids associés aux hyperquadriques d'une géométrie finie, *Discrete Math.* **13**, 185–211 (1975).
[61] M. Zarzar, Error-correcting codes on low-rank surfaces, *Finite Fields Appl.* **13**, 727–737 (2007).

This page intentionally left blank

Chapter 8

Toric Codes

Edgar Martínez-Moro and Diego Ruano*

Departamento de Matemática Aplicada, Universidad de Valladolid, Campus "Duques de Soria", E-42004, Soria, Castilla, Spain.
edgar@maf.uva.es

Department of Mathematics, Technical University of Denmark, Matematiktorvet, Building 303, DK-2800, Lyngby, Denmark.
D.Ruano@mat.dtu.dk

This chapter is an introduction to toric codes, algebraic geometry codes from toric varieties. We present a brief introduction to toric geometry. We describe the structure and properties of toric codes, compute their parameters and comment the bibliography.

Contents

*The authors were partially supported by the Spanish Ministry of Education through project MEC MTM2007-64704 and and by Junta de CyL VA065A07.

Introduction

Algebraic geometry codes can be defined using higher dimensional varieties, as shown in the previous chapter [27]. In this chapter we also study a family of codes from higher dimensional varieties, those arising from toric varieties. They were introduced by J.P. Hansen in 1998 [16].

In general, it is not possible to compute a basis of a code coming from a higher dimensional variety or to estimate its parameters. However, this is possible for toric codes since every definition and property for toric varieties has a combinatorial description.

We present an introduction to toric geometry, we assume some knowledge in algebraic geometry but not in toric geometry. Then we define the codes and study their parameters, structure and properties. We do not mention any decoding algorithm since it does not exist so far.

8.1. Toric geometry

In this section we present some results and definitions of toric geometry necessary to define toric codes.

Formally, a *toric variety* X is a normal variety containing an algebraic torus $T = \mathbb{K}^* \times \cdots \times \mathbb{K}^*$, where $\mathbb{K}$ is a field, as a dense subset for the Zariski topology and, furthermore, the torus acts on the variety X. The importance of these varieties is based on their correspondence with combinatorial objects, like cones and polytopes, which allows us to perform computations. The results of this section can be found in [8, 14, 31]. We mainly use the notation in [14].

8.1.1. *Cones, fans, polytopes and toric varieties*

Let N be a lattice isomorphic to $\mathbb{Z}^r$ for some $r \geq 1$. Let $M = \mathrm{Hom}(N, \mathbb{Z})$ be the dual lattice of N. One has the $\mathbb{Z}$-bilinear map $\langle\ ,\ \rangle : M \times N \to \mathbb{Z}$, $(u, v) \mapsto u(v)$. Let $N_{\mathbb{R}} = N \otimes \mathbb{R}$ and $M_{\mathbb{R}} = M \otimes \mathbb{R}$, where $M_{\mathbb{R}}$ is the dual vector space of $N_{\mathbb{R}}$ ($N_{\mathbb{R}}$ and $M_{\mathbb{R}}$ are isomorphic to $\mathbb{R}^r$). Then, $N_{\mathbb{R}}$ and $M_{\mathbb{R}}$ inherit the $\mathbb{R}$-bilinear map $\langle\ ,\ \rangle : M_{\mathbb{R}} \times N_{\mathbb{R}} \to \mathbb{R}$, $(u, v) \mapsto u(v)$.

A *convex polyhedral cone* σ is a set

$$\sigma = \{s_1 v_1 + \cdots + s_k v_k \in N_{\mathbb{R}} \mid s_i \geq 0\}$$

generated by a finite number of elements $v_1, \ldots, v_k \in N_{\mathbb{R}}$. The *dimension* of σ, $\dim(\sigma)$, is the dimension of the vector space $\sigma + (-\sigma) = \mathbb{R}\sigma$, where

$-\sigma$ is $\{-s \mid s \in \sigma\}$. The *dual cone* $\sigma^\vee \subset M_\mathbb{R}$ of a convex polyhedral cone is defined as

$$\sigma^\vee = \{u \in M_\mathbb{R} \mid \langle u, v\rangle \geq 0 \ \ \forall\ v \in \sigma\}$$

A *face* τ of a cone σ is the intersection of σ with a hyperplane defined by a non-negative linear form in σ, that is, $\tau = \sigma \cap u^\perp = \{v \in \sigma \mid \langle u, v\rangle = 0\}$ for a $u \in \sigma^\vee$. The convex cone σ is in fact a face of σ since it is the intersection with the linear form defined by 0. Moreover, each face is a convex polyhedral cone generated by the vectors of σ such that $\langle u, v_i\rangle = 0$. The one-dimensional faces are called *edges* and they will be denoted by ρ.

The *primitive element* $v(\rho) \in N$ of an edge ρ is the unique generator of $\rho \cap N$ as additive semigroup. One has a partial order in the set of faces of σ: let τ and τ' be two faces of σ, if $\tau \subset \tau'$ we denote it by $\tau < \tau'$.

A convex polyhedral cone σ is said to be *rational* if it has a generator system in the lattice N. A convex polyhedral cone σ is *strongly convex* if $\sigma \cap (-\sigma) = \{0\}$ or, equivalently, if $\sigma^\vee$ generates $M_\mathbb{R}$. Every rational cone is generated by a minimal number of elements in N; if the cone is strongly convex this minimal set of generators consists of the primitive elements of the edges. Moreover, if σ is a strongly convex rational cone then $\sigma^\vee$ is a rational polyhedral cone in $M_\mathbb{R}$ [31, Proposition 1.3]. For the sake of simplicity, in this chapter a strongly convex rational cone will be called *cone*.

Let σ be a cone, then $S_\sigma = \sigma^\vee \cap M$ is a finitely generated group by Gordan's lemma [31, Proposition 1.1]. We consider the $\mathbb{K}$-algebra associated to S_σ, $\mathbb{K}[S_\sigma] = \bigoplus_{u \in S_\sigma} \mathbb{K}\chi^u$ (where $\chi^u\chi^{u'} = \chi^{u+u'}$, and the zero element is χ^0). Therefore, we can define the affine variety U_σ as $U_\sigma = \mathrm{Spec}(\mathbb{K}[S_\sigma])$, called *affine toric variety associated to* σ.

A finitely generated commutative algebra A determines an affine variety $\mathrm{Spec}(A)$. However, it is possible to consider toric varieties without using the language of schemes: choosing generators of A, one can see that it is isomorphic to $\mathbb{K}[X_1, \ldots, X_r]/I$, where I is an ideal in $\mathbb{K}[X_1, \ldots, X_r]$. In this way one can identify the points of $Z(I) = \{p \in \mathbb{K}^n \mid f(p) = 0, \forall f \in I\}$ with maximal ideals of $\mathrm{Spec}(A)$, which are called closed points of the variety and are denoted by $\mathrm{Specm}(A)$. A morphism of algebras $A \to B$ determines a morphism of varieties $\mathrm{Spec}(B) \to \mathrm{Spec}(A)$. In particular closed points correspond to morphisms of algebras $A \to \mathbb{K}$.

For toric varieties, one has that closed points correspond to homomorphisms of semigroups $S_\sigma \to \mathbb{K}$, where $\mathbb{K} = \mathbb{K}^* \cup \{0\}$ is a commutative semigroup, and that $\mathrm{Specm}(\mathbb{K}[\mathrm{S}_\sigma]) \simeq \mathrm{Hom}(S_\sigma, \mathbb{K})$. One may also consider χ^u as

a Laurent monomial, $\chi^u(t) = t_1^{u_1} \cdots t_r^{u_r} \in \mathbb{K}[t_1, \ldots, t_r]_{(t_1 \cdots t_r)}$ (localization of $\mathbb{K}[t_1, \ldots, t_r]$ at $(t_1 \cdots t_r)$), moreover, χ^u defines a map $(\mathbb{K}^*)^r \to \mathbb{K}^*$. In the language of algebraic groups χ^u is called *character* [21].

We call $T = (\mathbb{K}^*)^r$, the *algebraic torus* of dimension r. We claim that T is contained as a dense subset in any toric variety and that it acts on the toric variety, extending the action on itself. Furthermore, as we claimed, these two properties characterize toric varieties [24].

Let us see that $T = \mathbb{K}^* \times \cdots \times \mathbb{K}^*$ is contained in U_σ. One has that S_σ is a subsemigroup of $S_{\{0\}} = M$. Let $v_1, \ldots, v_r$ be a basis of N and $u_1 = v_1^*, \ldots, u_r = v_r^*$ its dual basis in M. A generator system of M as a semigroup, is $u_1^*, \ldots, u_r^*, -u_1^*, \ldots, -u_r^*$, therefore, if we write $x_i = \chi^{u_i^*} \in \mathbb{K}[M]$, one has that

$$\mathbb{K}[M] = \mathbb{K}[x_1, x_2, \ldots, x_r, x_1^{-1}, x_2^{-1}, \ldots, x_r^{-1}] = \mathbb{K}[x_1, x_2, \ldots, x_r]_{(x_1 x_2 \cdots x_r)}$$

which is the ring of Laurent polynomials with r indeterminates and $U_0 = \operatorname{Spec}(\mathbb{K}[M]) = \mathbb{K}^* \times \cdots \times \mathbb{K}^* = (\mathbb{K}^*)^r = T$. Consequently, as any semigroup S_σ is a subsemigroup of M, $\mathbb{K}[S_\sigma]$ is a subalgebra of $\mathbb{K}[M]$. To summarize, $\mathbb{K}[S_\sigma]$ is a domain and $T \subset U_\sigma$. Moreover, we can write T without choosing coordinates

$$T = \operatorname{Spec}(\mathbb{K}[M]) = \operatorname{Hom}(M, \mathbb{K})$$

Let σ be a cone in N, the torus T acts over U_σ in the following way: a point $t \in T$ is identified with a homomorphism of groups $M \to \mathbb{K}$ and a point $x \in U_\sigma$ is identified with a homomorphism of semigroups $S_\sigma \to \mathbb{K}$. Then

$$\begin{aligned} T \times U_\sigma &\to U_\sigma \\ (t, x) &\mapsto t \cdot x \end{aligned}$$

where $t \cdot x$ is the homomorphism of semigroups

$$\begin{aligned} t \cdot x : S_\sigma &\to \mathbb{K} \\ u &\mapsto t(u)x(u) \end{aligned}$$

Example 8.1. Let σ be the cone generated by $v_1, \ldots, v_l$ with $1 \leq l \leq r$. One has that

$$S_\sigma = \mathbb{Z}_{\geq 0} u_1 + \cdots + \mathbb{Z}_{\geq 0} u_l + \mathbb{Z} u_{l+1} + \cdots + \mathbb{Z} u_r$$

Therefore, $\mathbb{K}[\sigma] = \mathbb{K}[x_1, \ldots, x_l, x_{l+1}, x_{l+1}^{-1}, \ldots, x_r, x_r^{-1}]$ and $U_\sigma = \mathbb{K}^l \times (\mathbb{K}^*)^{r-l}$.

From this example, one may infer that if σ is generated by l elements that can be completed to form a basis of N, then U_σ is the product of an

affine space of dimension l and an $(r-l)$-dimensional torus, and that U_σ is a non-singular variety (as we will see in theorem 8.5).

A *fan* $\triangle$ in N is a finite set of cones in $N_\mathbb{R}$ such that: each face of a cone in $\triangle$ is also a cone in $\triangle$ and the intersection of two cones in $\triangle$ is a face of each of them. For a fan $\triangle$ the *toric variety* $X_\triangle$ is constructed taking the disjoint union of the affine toric varieties U_σ for $\sigma \in \triangle$, and gluing the affine varieties with common faces: for the cones σ, $\sigma' \in \triangle$ one has that $\sigma \cap \sigma'$ is a face of each of them, and we can therefore identify $U_{\sigma\cap\sigma'}$ as an open subvariety of U_σ and of $U_{\sigma'}$. These identifications are compatible since the correspondence between cones and affine varieties preserves the order on the faces. Moreover, $X_\triangle$ is an scheme since $U_{\sigma\cap\sigma'} \to U_\sigma \times U_{\sigma'}$ is a closed immersion, where σ and σ' are two cones whose intersection is a common face. In particular, let σ be a cone in N and let $\triangle$ be the fan composed of the faces of σ, then $X_\triangle = U_\sigma$.

In next example we construct a toric variety from a fan.

Example 8.2. Let $\triangle$ be the fan of figure 8.1. The zero-dimensional cone is $(0,0)$. The one-dimensional cones are the 4 half-lines with origin the point $(0,0)$ and generated by $v(\rho_1) = (1,0)$, $v(\rho_2) = (0,1)$, $v(\rho_3) = (-1,0)$ and $v(\rho_4) = (0,-1)$. The two-dimensional cones are the 4 quadrants σ_i, $i = 1,\ldots,4$.

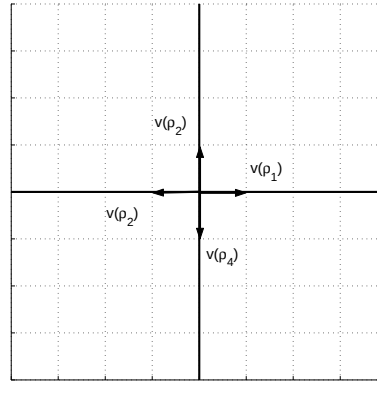


Fig. 8.1. Fan $\triangle$, example 8.2

The toric varieties $U_{\sigma_i} \simeq \mathbb{K}^2$ (corresponding to the algebras $\mathbb{K}[x_1, x_2]$, $\mathbb{K}[x_1^{-1}, x_2]$, $\mathbb{K}[x_1^{-1}, x_2^{-1}]$ and $\mathbb{K}[x_1, x_2^{-1}]$, respectively) glue in the usual way

to give $\mathbb{P}^1 \times \mathbb{P}^1$ (fixing coordinates, $x_1 = t_1/t_0$, and $x_2 = t'_1/t'_0$ where $(t_1 : t_0) \times (t'_1 : t'_0)$ are the coordinates of $\mathbb{P}^1 \times \mathbb{P}^1$).

A convex rational polytope in $M_{\mathbb{R}}$ is the convex hull of a finite set of points in M, for the sake of simplicity we will call it *polytope*. We can represent a polytope as the intersection of half-spaces. In the same way as in $N_{\mathbb{R}}$, one can consider faces of polytopes in $M_{\mathbb{R}}$. A *facet* F of a polytope P is a face of P of codimension 1 in M, therefore there exists a normal subspace to this face, generated by two elements of the lattice N, one inner and one outward. Let $v_F \in N$ be the primitive and inner element generating the normal face to F and a_F an integer such that

$$P = \bigcap_{F \text{ facet of } P} \{u \in M_{\mathbb{R}} \mid \langle u, v_F \rangle \geq -a_F\}$$

For a face τ of P, let σ_τ be the cone generated by v_F for all the facets F containing τ. Then

$$\triangle_P = \{\sigma_\tau \mid \tau \text{ is a face of } P\}$$

is a fan called *fan associated to* P. The toric variety defined by P is denoted by X_P.

Example 8.3. Let P be the plane polytope of $M_{\mathbb{R}}$ with vertices $(0,0)$, $(1,0)$, $(1,1)$, $(0,1)$, i.e. the convex hull of those points.

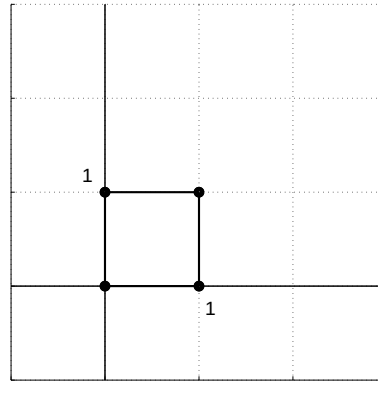


Fig. 8.2. Polytope P, example 8.3

P is the intersection of the following half-spaces,

$$P = \{u_1 \geq 0\} \cap \{-u_1 \geq -1\} \cap \{u_2 \geq 0\} \cap \{-u_2 \geq -1\}$$

The inner primitive elements of N normal to the facets are $v_1, -v_1, v_2, -v_2$. Each vertix of the polytope defines a cone of dimension 2 in the normal fan (for instance, $(0,0)$ defines a cone generated by v_1 and v_2). Therefore, one has that the normal fan $\triangle_P$ is the one in example 8.2, and hence $X_P = \mathbb{P}^1 \times \mathbb{P}^1$.

8.1.2. *Properties of toric varieties*

A toric variety X_P defined using a polytope P is a projective variety [7, Theorem 12.2], that is, we may consider X_P not only as an abstract variety, but also as being embedded in $\mathbb{P}^l$ for some l.

Theorem 8.4. *Let $\triangle$ be a fan. The toric variety $X_\triangle$ associated to $\triangle$ is projective if and only if $\triangle$ is the normal fan associated to a polytope.*

For a variety defined using a polytope P, the proof of the previous result considers the map

$$\begin{aligned} \varphi_P : T &\to \mathbb{P}^{l-1} \\ t &\mapsto (\chi^{u_1}(t), \ldots, \chi^{u_l}(t)) \end{aligned}$$

where $P \cap M = \{u_1, \ldots, u_l\}$. Then the map φ_Q is an embedding for s sufficiently large, where $Q = sP = \{sp \mid p \in P\}$. This gives an elementary way to define toric varieties coming from a polytope: X_P is the Zariski closure of the image of φ_Q for s sufficiently large. If the variety is non-singular, then $s = 1$ [31, Corollary 2.15].

A fan $\triangle$ is said to be *non-singular* if for every face $\sigma \in \triangle$ there exists a $\mathbb{Z}$-basis $\{v_1, \ldots, v_r\}$ of N such that σ is generated by $\{v_1, \ldots, v_s\}$, where $s \leq r$ is the dimension of σ (it is said to be singular otherwise). Thanks to this result, one has a combinatorial criterion to determine whether a toric variety is singular [14, section 2.1], [31, theorem 1.10].

Theorem 8.5. *The toric variety $X_\triangle$ associated to a fan $\triangle$ is singular if and only if $\triangle$ is singular.*

Let $\triangle$ be a fan, a fan $\triangle'$ is a *refinement* of $\triangle$ if every cone of $\triangle$ is union of cones of $\triangle'$. The morphism $X'_\triangle \to X_\triangle$ induced by the identity map from N to N is birational and proper. Furthermore, it is an isomorphism over the torus contained in both varieties [14, section 2.4].

This construction can be used to understand and compute the resolution of singularities of toric varieties. The fan $\triangle'$, which is non-singular and a

refinement of $\triangle$, is called *refined fan of* $\triangle$. For toric surfaces, this construction can be computed in terms of continued fractions with the complexity of the Euclidean algorithm [14, section 2.6].

Example 8.6. Let P be the polytope with vertices $(0,0)$, $(2,2)$ and $(0,4)$.

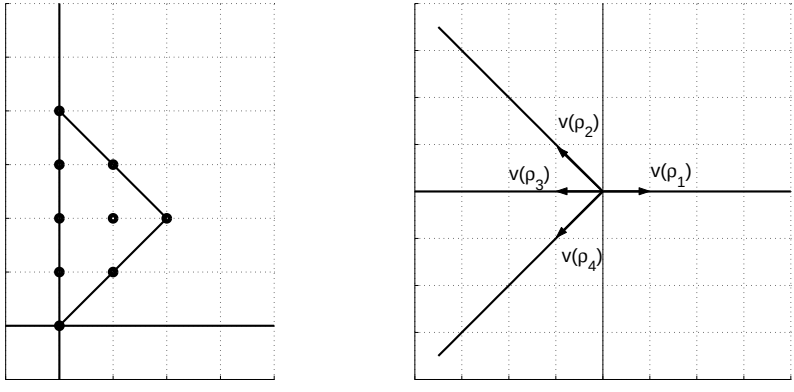


Fig. 8.3. P and its normal refined fan.

$$P = \{u_1 \geq 0\} \cap \{-u_1 + u_2 \geq 0\} \cap \{-u_1 - u_2 \geq -4\}$$

The edges of $\triangle_P$ are generated by $(1,0)$, $(-1,1)$ and $(-1,-1)$, hence $\triangle_P$ is a singular fan, since $(-1,1)$ and $(-1,-1)$ do no generate the lattice N.

We may refine the fan $\triangle_P$ just considering the edge generated by $(-1,0)$, that is, the edges of the refined fan $\triangle'$ are generated by $(1,0)$, $(-1,1)$, $(-1,0)$ and $(-1,-1)$. One has that $\triangle'$ is a refinement of $\triangle$, and that $X_{\triangle'}$ is a non-singular variety by theorem 8.5.

The toric varieties defined above are *normal* varieties, also called normal toric varieties. That is, all of their local rings are integrally closed domains in their fraction rings [14, section 2.1]. These varieties are normal because we have used all the points of the lattice to define them [7, theorem 7.2].

Let σ be a cone generated by $u_1, \ldots, u_t$, then $\mathbb{K}[S_\sigma] = \mathbb{K}[\chi^{u_1}, \ldots, \chi^{u_t}]$. One has that $\mathbb{K}[S_\sigma] = \mathbb{K}[Y_1, \ldots, Y_t]/I$, where I is generated by binomials of the form $Y_1^{a_1} \cdots Y_t^{a_t} - Y_1^{b_1} \cdots Y_t^{b_t}$, with $a_1, \ldots, a_t, b_1, \ldots, b_t$ non-negative integers such that

$$a_1 u_1 + \cdots + a_t u_t = b_1 u_1 + \cdots + b_t u_t$$

Nevertheless, not every quotient algebra defined by a binomial ideal defines a normal toric variety [35].

A fan $\triangle$ is said to be *complete* if the union of all its faces is $N_{\mathbb{R}}$. One has the following result that characterizes complete toric varieties [14, Section 2.4], [31, Section 1.4]

Theorem 8.7. *A toric variety $X_{\triangle}$ is complete (compact) if and only if $\triangle$ is complete.*

Let P be a polytope containing the origin, then the normal fan $\triangle_P$ is complete and hence the variety X_P is complete. Therefore, one usually considers polytopes containing the origin.

8.1.3. *Orbits and divisors*

Let P be a polytope and $\triangle_P$ its associated normal fan. Since X_P is a normal projective variety, we may consider the commutative group $\mathrm{Div}(X_P)$ of *Weil divisors* of X. A Weil divisor is a finite linear combination (over $\mathbb{Z}$) of irreducible varieties of codimension 1. We denote by T–$\mathrm{Div}(X_P)$ the subgroup of Weil divisors that are invariant under the action of T.

The following theorem characterizes the T-orbits, orbits by the action of T, in terms of $\triangle_P$ [31, proposition 1.6], using the bijective correspondence of the faces of a polytope and the cones of its normal fan.

Theorem 8.8. *Let P be a polytope and $\triangle_P$ its normal fan. For every $\sigma \in \triangle_P$ we consider*

$$\mathrm{orb}(\sigma) = \mathrm{Hom}(M \cap \sigma^{\perp}, \mathbb{K}^*)$$

Every T-orbit of X_P has this form above and, moreover, there exists a bijective correspondence between the cones of $\triangle_P$ and the T-orbits of X_P. Furthermore, one has the following properties:

- $\mathrm{orb}(\{0\}) = T$.
- *Let $\sigma \in \triangle$. Then* $\mathrm{orb}(\sigma)$ *is an open set in its own closure, which we denote by $V(\sigma)$. The variety $V(\sigma)$ is a closed toric subvariety of X_P of codimension* $\dim(\sigma)$*, that is,* $\dim\sigma + \dim V(\sigma) = r$.
- *If X_P is non-singular then, $V(\sigma)$ is non-singular.*

By the previous result, $\{V(\rho) \mid \rho \in \triangle_P(1)\}$ is a basis of T–$\mathrm{Div}(X_P)$ over $\mathbb{Z}$. We denote by $\mathrm{PDiv}(X_P)$ the subgroup of *principal divisors of* $\mathrm{Div}(X_P)$,

i.e., the divisors of the form

$$\operatorname{div}(f) = \sum_V v_V(f)V$$

with f a rational function on X_P, different from zero, and $v_V(f)$ the order of f in the closed subvarieties V of X_P of codimension 1. Each $u \in M$ corresponds to a character χ^u, which is a regular function in T and gives rise to a rational function on X_P.

The subgroup $\operatorname{CDiv}(X_P)$ of *Cartier divisors* of $\operatorname{Div}(X_P)$ are the locally principal Weil divisors, i.e., there exists an open covering $X_P = \bigcup U_j$ and non-zero rational functions, f_j, such that the Cartier divisor can be written as the divisor $\operatorname{div}(f_j^{-1})$ in U_j. If X_P is non-singular, then every Weil divisor is a Cartier divisor, i.e. $\operatorname{CDiv}(X_P) = \operatorname{Div}(X_P)$ [20, proposition 6.11].

A polytope P defines the following T-invariant Cartier Divisor

$$D_P = \sum_{F \text{ facet of } P} a_F V(\rho_F)$$

and given $u \in P$

$$\operatorname{div}(\chi^u) = \sum_{F \text{ facet of } P} \langle u, v_F \rangle V(\rho_F)$$

Example 8.9. Two polytopes with the same inner normal vectors define the same toric variety. For instance, the polytope $P_{a,b}$ with vertices $(0,0)$, $(a,0)$, (a,b) and $(0,b)$ defines the toric variety $\mathbb{P}^1 \times \mathbb{P}^1$ for all $a, b \in \mathbb{N}$ (see examples 8.2 and 8.3). In particular, for $a = b = 1$ one has the polytope in example 8.3. However, they define different Cartier divisors

$$D_{P_{a,b}} = aV(\rho) + bV(\rho')$$

where ρ and ρ' are the cones generated by $(-1,0)$ and $(0,-1)$ respectively.

A complete fan $\triangle$ and T-invariant Cartier divisor $D = \sum a_\rho V(\rho)$ define a polytope,

$$P_D = \{u \in M_{\mathbb{R}} \mid \langle u, v(\rho) \rangle \geq -a_\rho \ \forall\ \rho \text{ edge of } \triangle\}$$

Furthermore, $P_{D_P} = P$. Therefore, a polytope is the same datum as a complete normal toric variety and a Cartier divisor.

8.2. Definition of toric codes

Let $\mathbb{F}_q$ be the finite field with q elements. Let P be a rational polytope of dimension $r \geq 2$, X_P the toric variety defined by its refined normal fan and D_P its associated Cartier divisor over X_P, as in the previous section. In particular, one has that X_P is a non-singular toric variety.

We define toric codes evaluating rational functions at the $(q-1)^r$ points of $T = (\mathbb{F}_q^*)^r$. The *toric code* $\mathcal{C}_P$ *associated to* P, which is an algebraic geometry code in the sense of chapter 7 of this volume [27], is the image of the $\mathbb{F}_q$-linear evaluation map given by

$$\begin{array}{rcl} \mathrm{ev} : \mathrm{H}^0(X_P, \mathcal{O}(D_P)) & \to & (\mathbb{F}_q)^{\#T} \\ f & \mapsto & (f(t))_{t \in T} \end{array}$$

Since we evaluate in $\#T$ points, $\mathcal{C}_P$ has *length* $n = \#T = (q-1)^r$. The following result shows us a basis of the functions that we evaluate. That is, it allows us to compute a basis of $\mathcal{L}(D_P) = \mathrm{H}^0(X_P, \mathcal{O}(D_P))$, i.e., the rational functions -global sections- f over X_P such that $\mathrm{div}(f) + D_P \succeq 0$ (f has zeros and poles bounded by D_P).

Lemma 8.10. *Let X_P be the toric variety associated to a polytope P. A basis of the $\mathbb{F}_q$-vector space $\mathcal{L}(D)$ is $\{\chi^u \mid u \in P \cap M\}$.*

Then, for every t in $T = (\mathbb{F}_q^*)^r$, the rational functions of $\mathcal{L}(D_P)$ can be evaluated at t

$$\begin{array}{rcl} \mathcal{L}(D_P) & \to & \mathbb{F}_q \\ f & \mapsto & f(t) \end{array}$$

since f is a linear combination of characters χ^u that can be considered as Laurent monomials. This map is just the evaluation of a Laurent polynomial whose monomials have exponents in $P \cap M$ at points with non-zero coordinates.

Remark 8.11. For a polytope P, we can also define a toric code $\mathcal{C}_P$ using the embedding of the refined toric variety X_P in the projective space (see [27, Definition 7.2]). We consider the map

$$\begin{array}{rcl} \varphi_P : T & \to & \mathbb{P}^{l-1} \\ t & \mapsto & (\chi^{u_1}(t), \ldots, \chi^{u_l}(t)) \end{array}$$

where $P \cap M = \{u_1, \ldots, u_l\}$. The map φ_P fixes coordinates for the points of the torus in $\mathbb{P}^{l-1}$, let $T = \{t_1, \ldots, t_n\}$ and $P_i = \varphi(t_i)$. Let $\mathcal{F}$ be the vector space of linear forms (homogeneous polynomials of degree 1) in

$\mathbb{F}_q[x_1, \ldots, x_l]$. The toric code E_P is the image of the $\mathbb{F}_q$-linear evaluation map

$$\begin{aligned} \alpha : \mathcal{F} &\to \mathbb{F}_q^n \\ f &\mapsto (f(P_1), \ldots, f(P_n)) \end{aligned}$$

One has that $\mathcal{C}_P = \mathrm{E}_P$, in particular $\mathrm{ev}(\chi^{u_i}) = \alpha(x_i)$.

From lemma 8.10 it follows that $\mathrm{H}^0(X_P, \mathcal{O}(D_P))$ is a finite dimensional $\mathbb{F}_q$-vector space with basis $\{\chi^u \mid u \in M \cap P\}$, therefore a generator system of the code $\mathcal{C}_P$ is $\{(\chi^u(t))_{t\in T} \mid u \in M \cap P\}$. This generator system is a basis of the code if and only if the evaluation map ev is injective.

Example 8.12. Let P be polytope of example 8.3, $P \cap M$ is $\{(0,0), (1,0), (0,1), (1,1)\}$. The toric code C_P has length $(q-1)^2$ and is generated by

$$\mathcal{C}_P = \langle (1)_{t\in T}, (t_1)_{t\in T}, (t_2)_{t\in T}, (t_1 t_2)_{t\in T} \rangle \subset \mathbb{F}_q^{(q-1)^2},$$

where $t = (t_1, t_2) \in T$. We will see in theorem 8.13 that this generator system is injective for any finite field $\mathbb{F}_q$, therefore the dimension of $\mathcal{C}_P = \#(P \cap M) = 4$.

We consider the class of elements of $P \cap M$ modulo H, that is, $\overline{u} = (u_1 \bmod (q-1), \ldots, u_r \bmod (q-1)) \in H$, where $u = (u_1, \ldots, u_r) \in H$. This allows us to see whether the evaluation map is injective [32, theorem 3.3].

Theorem 8.13. *Let P be a polytope, the evaluation map* $\mathrm{ev} : \mathrm{H}^0(X_P, \mathcal{O}(D_P)) \to (\mathbb{F}_q)^{\#T}$ *is injective if and only if $\overline{u_1} \neq \overline{u_2}$ with $u_1 \neq u_2$, for all $u_1, u_2 \in P \cap M$.*

Proof. Let $f \in \mathrm{H}^0(X_P, \mathcal{O}(D_P))$ be non-zero, $f = \sum_{u\in P\cap M} \lambda_u X^u \in \mathbb{F}_q[X_1, \ldots, X_r]$. Let $\overline{f} = \sum_{u\in P\cap M} \lambda_u X^{\overline{u}}$. One has that $f(t) = \overline{f}(t)$ in $\mathbb{F}_q$ for all $t \in T$, although $f \neq \overline{f}$ in $\mathbb{F}_q[X_1, \ldots, X_r]$ (in general).

Let $\overline{u_1} \neq \overline{u_2}$ with $u_1 \neq u_2$, for all $u_1, u_2 \in P \cap M$. One has that $\overline{f}$ cannot vanish completely at T since it is a non-zero polynomial whose monomials only have exponents in $H = (\{0, \ldots, q-2\})^r \subset \mathbb{Z}^r$ (see [32, lemma 3.2]).

Let $u_1, u_2 \in P \cap M$ with $u_1 \neq u_2$, then $X^{u_1} - X^{u_2} \in \mathrm{H}^0(X_P, \mathcal{O}(D_P))$ is non-zero and belongs to the kernel of ev. □

The dimension of $\mathcal{C}_P$ is equal to $k = \#\overline{P} = \#\{\overline{u} \mid u \in P \cap M\}$. When the evaluation map is injective the dimension of $\mathcal{C}_P$ is equal to the number of lattice points of the polytope. There are algorithms that can compute this number [9]. Furthermore, there are formulas, that depend on the geometry

of X_P, to compute the number of lattice points of the polytope: let P be a polytope and $P_\sigma = P \cap (\sigma^\perp + u(\sigma))$, where $\sigma \in \triangle_P$ and $u(\sigma)$ is any element in $M/(\sigma^\perp \cap M)$, in other words P_σ is the intersection of P with the corresponding translation of the perpendicular subspace to σ [14, chaper 5]. Then, combining the Riemann-Roch theorem with a result about Todd classes, one has that (see [14, chaper 5])

$$\#(P \cap M) = \sum_{\sigma \in \triangle_P} r_\sigma \mathrm{vol}_r(P_\sigma)$$

where vol_r is the Lebesgue volume and r_σ is a rational number that depends only on the geometry of σ (see [5, 30] and their references). A particular case of the previous statement is the well-known Pick's formula for plane polytopes

$$\#(P \cap M) = \mathrm{vol}_2(P) + \frac{\mathrm{Perimeter}(P)}{2} + 1$$

Some computations of families of toric codes $(\mathcal{C}_{P_i})_{i=1}^{\infty}$ can be found in the bibliography, where P_{i+1} is a dilatation of P_i, that is, there is $\lambda \in \mathbb{N}$ such that $P_{i+1} = \lambda P_i = \{\lambda p \mid p \in P\}$. The Ehrhart function of P counts the lattice points of the polytope λP, with $\lambda \in \mathbb{N}$. This function is computed in [4].

Although there are toric codes with good parameters, see [23] for some families, the asymptotic behaviour of the families considered so far is not good. One important remark is that the length of a toric code n is fixed, $n = (q-1)^r$, by the base field and the dimension of the variety. If one wants to consider a infinite family of toric codes with strictly increasing length, the dimension of the polytopes should be increased as well, if the base field is fixed. Then, we cannot study these problems using the Ehrhart function.

In section 8.5 we will see some examples of toric codes. We refer the reader to [23] for a larger collection of examples, some of them with good parameters.

Example 8.14. We present some families of toric codes with known parameters (see [17]). One can compute their minimum distance using several approaches in the bibliography.

- Let P be the plane polytope with vertices $(0,0)$, (a,a), $(0,2a)$ with $2a < q-1$. Then $k = (a+1)^2$ and $d = (q-1)^2 - 2a(q-1)$.
- Let P be the plane polytope with vertices $(0,0)$, $(0,a)$, $(0,a)$ with $a < q-1$, $X_P = \mathbb{P}^2$. Then $k = (a+1)(a+2)/2$ and $d = (q-1)^2 - a(q-1)$.

- Let P be the plane polytope with vertices $(0,0)$, $(a,0)$, (a,b), $(0,b)$ with $a,b < q-1$, $X_P = \mathbb{P}^1 \times \mathbb{P}^1$. Then $k = (a+1)(b+1)$ and $d = (q-1-a)(q-1-b)$.
- Let P be the plane polytope with vertices $(0,0)$, $(a,0)$, $(a,b+ra)$, $(0,b)$ with $a,b < q-1$ and $b+ra < q-1$, X_P is the Hirzebruch surface F_r (see [14, section 1.1]). Then $k = (a+1)(b+1) + ra(a+1)/2$ and $d = \min\{(q-a-1)(q-b-1), (q-1)(q-b-ar-1)\}$.

8.3. Classification of toric codes

Let us consider a classification of monomially equivalent codes from [29]: two codes, $\mathcal{C}_1, \mathcal{C}_2$ with generator matrix G_1, G_2, are *monomially equivalent* if there is an invertible $n \times n$ diagonal matrix Δ and an $n \times n$ permutation matrix Π such that $G_2 = G_1 \Delta \Pi$ is a generator matrix for C_2. Two monomially equivalent codes have the same parameters.

One has that two polytopes P, Q are *lattice equivalent* if there exists an invertible map $T : M_{\mathbb{R}} \to M_{\mathbb{R}}$, $T(x) = M(x) + \lambda$, such that $T(P) = Q$, where M is an $r \times r$ matrix and λ is an r-dimensional vector. With the following result one has a characterization of monomially equivalent toric codes.

Theorem 8.15. *Let P, Q be two polytopes lattice equivalent, then the toric codes $\mathcal{C}_P$ and $\mathcal{C}_Q$ are monomially equivalent.*

One can also find the beginning of a classification of plane toric codes in [29], for small k, using this result.

Example 8.16. Let P be a polytope and $Q = \{p + u_0 \mid p \in P\}$ for $u_0 \in M$; Q is the polytope obtained shifting P by u_0. One can easily see that P and Q are lattice equivalent and therefore $\mathcal{C}_P$ and $\mathcal{C}_Q$ have the same parameters.

Moreover, one can also obtain this result, for this example, using a more elementary approach: let $q = p + u_0$, with $q \in Q$ and $p \in P$. Then $\mathrm{ev}(\chi^q) = \mathrm{ev}(\chi^p \chi^{u_0}) = \mathrm{ev}(\chi^p) * \mathrm{ev}(\chi^{u_0})$, where $*$ is the component-wise product. Then $\mathrm{ev}(\chi^q)$ and $\mathrm{ev}(\chi^p)$ have the same weight since every coordinate of $\mathrm{ev}(\chi^{u_0})$ is non-zero.

8.4. Structure of toric codes

In this section we will see that toric codes are multicyclic and we will also compute the dual of a toric code.

8.4.1. *Multicyclic structure*

Multicyclic codes are a natural extension of cyclic codes, in particular a cyclic code is a 1-D cyclic code. A code $\mathcal{C} \subset A = \mathbb{F}_q[X_1, \ldots, X_r]/(X_1^{N_1} - 1, \ldots, X_r^{N_r} - 1)$ is *multicyclic or r-D cyclic* if it is an ideal in A, with $N_1, \ldots, N_r \in \mathbb{N}$. Let $\mathbb{F}_q[X_1, \ldots, X_r]_{\leq (N_1-1,\ldots,N_r-1)}$ be the $\mathbb{F}_q$-vector space of polynomials in the variables $X_1, \ldots, X_r$ of degree lower than N_i in each variable X_i for all i. We can consider the following isomorphisms of vector spaces $\mathbb{F}_q^n \simeq \mathbb{F}_q[X_1, \ldots, X_r]_{\leq (N_1-1,\ldots,N_r-1)} \simeq A$ where $n = N_1 \cdots N_r$, and we can identify its elements.

Let $\mathcal{C}_P$ be the toric code, for a polytope P. Set α a primitive element of $\mathbb{F}_q$, i.e. $\mathbb{F}_q^* = \{\alpha^0, \alpha^1, \ldots, \alpha^{q-2}\}$ and therefore $T = \{\alpha^i = (\alpha^{i_1}, \ldots, \alpha^{i_r}) \mid i \in H\}$. Then $\mathcal{C}_P$ is the vector subspace of $\mathbb{F}_q^n$ generated by $\{(\chi^u(\alpha^i))_{i\in H} \mid u \in P\}$, where $\chi^u(\alpha^i) = \alpha^{\langle u,i\rangle} = \alpha^{u_1 i_1 + \cdots + u_n i_n}$. In order to study the multicyclic structure we shall use the previous isomorphism. We represent (with multi-index notation for X^i)

$$(\alpha^{\langle u,i\rangle})_{i\in H} \quad \text{by} \quad \sum_{i\in H} \alpha^{\langle u,i\rangle} X^i$$

Proposition 8.17. *Let P be a polytope, $\mathcal{C}_P$ is an r-D cyclic code with $N_1 = q-1$, …, $N_r = q-1$.*

Proof. Let $u \in P$, we use the polynomial notation for $\mathcal{C}_P$: $\sum_{i\in H} \alpha^{\langle u,i\rangle} X^i \in \mathcal{C}_P$. Hence $\sum_{i\in H} \alpha^{\langle u,i-a\rangle} X^i = \alpha^{-\langle u,a\rangle} \sum_{i\in H} \alpha^{\langle u,i\rangle} X^i$. And the result holds due to the linearity of $\mathcal{C}_P$. □

The previous result can be found in [11, 33]. Moreover, in [33] it is considered as an extension of toric codes, called *generalized toric codes*, that are evaluation codes of the algebra

$$\mathbb{F}_q[H] = \langle Y^u = Y_1^{u_1} \cdots Y_r^{u_r} \mid u = (u_1, \ldots, u_r) \in H\rangle \subset \mathbb{F}_q[Y_1, \ldots, Y_r]$$

at the points of T. With this setting one can prove that any r-D cyclic code, with $N_1 = q-1$, … , $N_r = q-1$, is a generalized toric code. Therefore, the generalized toric codes and the r-D cyclic codes with $N_1 = q-1$, …, $N_r = q-1$ are the same family of codes.

8.4.2. *Dual of a toric code*

The following result gives the dual of a toric code (see [6, 33]). The dual of a toric code is not a toric code in general. However, it is a generalized toric code.

Theorem 8.18.

Let P be a polytope and $\mathcal{C}_P$ its associated toric code. Let $u_1, u_2 \in P$, one has that

$$\langle \mathrm{ev}(\chi^{u_1}), \mathrm{ev}(\chi^{u_2}) \rangle = \begin{cases} 0 & \text{if } \overline{u_1 + u_2} \neq 0 \\ (-1)^r & \text{if } \overline{u_1 + u_2} = 0 \end{cases}$$

Let $u \in P$, $u' = \overline{-u}$ with $\overline{u}$ as in the previous section (see theorem 8.13) and $P' = \{u' \mid u \in P \cap M\} \subset \mathbb{Z}^r$, $\#P' = \#(P \cap M)$. Let $P^\perp = H \setminus P'$, then the dual code of $\mathcal{C}_P$ is $\mathcal{C}_P^\perp = \mathcal{C}_{P^\perp}$. One has that $\mathcal{C}_{P^\perp}$ is not a toric code in general, since $P^\perp$ is not a convex polytope, but just an evaluation code.

Proof. Let $u_1, u_2 \in P$, then one has that $\langle (\alpha^{\langle u_1, i\rangle})_{i\in H}, (\alpha^{\langle u_2, i\rangle})_{i\in H} \rangle = \sum_{i\in H} \alpha^{\langle u_1+u_2, i\rangle} =$

$$\sum_{i\in H} \alpha^{\langle \overline{u_1+u_2}, i\rangle} = \begin{cases} \frac{q(q-1)}{2}(\mathrm{supp}(\overline{u_1 + u_2})) = 0 & \text{if } \overline{u_1 + u_2} \neq 0 \\ \sum_{i\in H} 1 = (-1)^r & \text{if } \overline{u_1 + u_2} = 0 \end{cases}$$

where $\mathrm{supp}(\overline{u_1 + u_2})$ is the number of nonzero coordinates of $\overline{u_1 + u_2}$.

Then $\langle \mathrm{ev}(\chi^{u_1}), \mathrm{ev}(\chi^{u_2}) \rangle = 0$ for $u_1 \in P$, $u_2 \in P^\perp$ since $\overline{u_1 + u_2} \neq 0$. On account of the dimension of $\mathrm{H}^0(X_P, \mathcal{O}(D_P))$ and the linearity of the codes the proof is completed. □

The previous result shows that the dual of a toric code $\mathcal{C}_{P_1}$ is a toric code only when there exists a convex polytope P_2 such that $\overline{P_1}^\perp = \overline{P_2}$. However, the dual of a generalized toric code is a generalized toric code.

Let P be the convex hull of the points of H, $P \cap M = H$, the matrix M of the evaluation map $\mathrm{ev} : \mathrm{H}^0(X_P, \mathcal{O}(D_P)) \to \mathbb{F}_q^n$ is

$$M = \begin{pmatrix} \alpha^{\langle u_1, i_1\rangle} & \alpha^{\langle u_1, i_2\rangle} & \cdots\cdots & \alpha^{\langle u_1, i_n\rangle} \\ \alpha^{\langle u_2, i_1\rangle} & \alpha^{\langle u_2, i_2\rangle} & \cdots\cdots & \alpha^{\langle u_2, i_n\rangle} \\ \vdots & \vdots & \vdots\ \vdots & \vdots \\ \vdots & \vdots & \vdots\ \vdots & \vdots \\ \alpha^{\langle u_n, i_1\rangle} & \alpha^{\langle u_n, i_2\rangle} & \cdots\cdots & \alpha^{\langle u_n, i_n\rangle} \end{pmatrix}$$

where $\{u_1, \ldots, u_n\} = \{i_1, \ldots, i_n\} = H$ and if moreover $u_j = i_j$ then M is a symmetric matrix, therefore we assume $u_j = i_j$ $\forall j = 1, \ldots n$.

We have thus proved that a generator matrix of the code $\mathcal{C}_P$ with $k = \#\overline{P}$, is the $(k \times n)$-matrix consisting of the k rows $\alpha^{\langle u, i_1\rangle}, \ldots, \alpha^{\langle u, i_n\rangle}$ of M

with $u \in \overline{P}$ and a control matrix of $\mathcal{C}_P$ is the $(n-k \times n)$-matrix consisting of the $n-k$ rows $\alpha^{\langle u,i_1\rangle}, \ldots, \alpha^{\langle u,i_n\rangle}$ of M with $u \in P^{\perp}$.

Moreover, one has that there exist no *self-dual* toric codes: for q even, $n = (q-1)^r$ is odd, and then there are no self-dual codes. For q odd, n is even but there are 2^r elements in H, $u_1, \ldots, u_{2^r}$, such that $\langle \mathrm{ev}(\chi^{u_i}), \mathrm{ev}(\chi^{u_i})\rangle \neq 0$ ($2u_i = 0 \bmod (q-1)$ if and only if u_i is equal to 0 or $(q-1)/2$). Therefore, the maximum dimension of a self-orthogonal code ($\mathcal{C}^{\perp} \subset \mathcal{C}$) is $n/2 - 2^{r-1} < n/2$, and there exist no self dual generalized toric codes.

8.5. Minimum distance

In this section we describe some estimations for the minimum distance of a toric code. Namely, we show three lower bounds, the first two are valid for toric codes coming from a plane polytope and they are based on invariants of the polytope. The third one holds for toric codes coming from a polytope of arbitrary dimension.

Apart from these three bounds, a bound for the minimum distance of a toric code coming from a polytope of arbitrary dimension can be found in [29]. This bound is established considering a generalization of Vandermonde determinants for the matrix M presented in section 8.4.2, however, the computations only hold, so far, for rectangular polytopes and simplices. For these two families of polytopes the bounds are sharp, we refer the reader to [29] for details.

Furthermore, for a polytope P, one has an upper bound for the minimum distance of $\mathcal{C}_P$ just considering a polytope $Q \subset P$, such that the minimum distance of $\mathcal{C}_Q$ is known. Then $d(\mathcal{C}_P) \leq d(\mathcal{C}_Q)$. This has been used in the bibliography to check whether the lower bounds of the minimum distance are sharp.

8.5.1. *Bound with Minkowski sum*

One can find this bound for the minimum distance in [28], we refer the reader to [28] and its references for all results, missing definitions and further details. This bound is based on the Minkowski sum of two plane polytopes and the Hasse-Weil bound for the number of $\mathbb{F}_q$-rational points of a curve, this bound is valid for toric codes coming from plane polytopes.

The Minkowski sum of two polytopes P, Q is the pointwise sum of their points $P + Q = \{p + q \mid p \in P, q \in Q\}$.

The Hasse-Weil bounds: for a non-singular curve X over $\mathbb{F}_q$ and absolutely irreducible over $\mathbb{F}_q$ (irreducible over the algebraic closure of $\mathbb{F}_q$) one has the following bounds for the number of $\mathbb{F}_q$-rational points of X

$$1 + q - 2g\sqrt{q} \leq \#X \leq 1 + q + 2g\sqrt{q}$$

where g is the genus of X. This formula also holds for singular curves, where g is now the arithmetic genus of X.

Then, one can prove (see [28, proposition 5.2]) that for q sufficiently large the rational functions with more irreducible components have more zeros in T than those fewer irreducible components. This allows us to obtain the following lower bound, sharp for large base fields.

Theorem 8.19. *Let P be a polytope with $P \cap M \subset H = \{0, \ldots, q-2\}^r$ and q sufficiently large (see [28, proposition 5.2]). Let l be the largest positive integer such that there exists a polytope $Q \subset P$ such that it is the Minkowski sum of l non-trivial polytopes (positive dimension), $Q = P_1 + \cdots + P_l$. Then there exists some $Q \subset P$ of this form such that*

$$d(\mathcal{C}_P) \geq \sum_{i=1}^{l} d(\mathcal{C}_{P_i}) - (l-1)(q-1)^2$$

Example 8.20. Let P be the plane polytope with vertices $(0,0)$, $(4,1)$, $(1,4)$. The toric code $\mathcal{C}_P^t \subset \mathbb{F}_8^{49}$ has parameters [49,11,28]. This example was given by D. Joyner [23] and it had better parameters than any other known code when this article was written. We will also consider this example for the other bounds.

By theorem 8.19 one has that

$$d(\mathcal{C}_P) \geq (q-1)^2 - 3(q-1)$$

for q sufficiently large (larger than 729). One can easily see that this lower bound is sharp (for instance considering the upper bound [32, proposition 4.2]). However, the minimum distance of $\mathcal{C}_P$ with $q = 8$ cannot be obtained using this bound.

8.5.2. *Bound with Minkowski length*

We refer the reader to [34] and its references for all the results, missing definitions and further details in this section. This bound is valid for toric codes coming from plane polytopes. The main tools are the full Minkowski length of a plane polytope and the Hasse-Weil bounds.

Let P be a polytope and let $P = P_1 + \cdots + P_l$ for some non-trivial polytopes (positive dimension). The *Minkowski length* of P, $\ell(P)$, is the largest number of summands in such decompositions of P. Moreover, we can define the *full Minkowski length* of a polytope P as $L(P) = \max\{\ell(Q) \mid Q \subset P\}$. The full Minkowski length is invariant by monomial transformations. A subpolytope $Q \subset P$ is called maximal for P if $\ell(Q) = L(P)$.

A polytope is *strongly indecomposable* if its full Minkowski length is 1. One can prove that a strongly indecomposable polytope of dimension 1 is monomially equivalent to the polytope with vertices $(0,0)$, $(1,0)$ and that a a strongly indecomposable polytope P of dimension 2 is monomially equivalent to one of the following polytopes: the polytope with vertices $(0,0)$, $(1,0)$, $(0,1)$ called P unit triangle or the polytope with vertices $(1,0)$, $(2,2)$, $(0,1)$ called P exceptional triangle.

Classifying the plane polytopes with full Minkowski length 2 and using the Hasse-Weil bound the following lower bound for the minimum distance is obtained

Theorem 8.21. *Let $P \subset H$ be a polytope with area $A = A(P)$ and full Minkowski length $L = L(P)$. Then, for $q \geq \max\left(23, \left(c + \sqrt{c^2 + 5/2}\right)^2\right)$, where $c = A/2 - L + 9/4$, the minimum distance of the toric code $\mathcal{C}_P$ is bounded by*

$$d(\mathcal{C}_P) \geq (q-1)^2 - L(q-1) - 2\sqrt{q} + 1$$

If no maximal decomposition in P contains an exceptional triangle, then for $q \geq \max\left(37, \left(c + \sqrt{c^2 + 2}\right)^2\right)$, where $c = A/2 - L + 11/4$, the minimum distance of the toric code $\mathcal{C}_P$ is bounded by

$$d(\mathcal{C}_P) \geq (q-1)^2 - L(q-1)$$

In [34, Section 3] one can find two algorithms to compute the Minkowski length of a plane polytope and to determine whether there is a maximal Minkowski decomposition with an exceptional triangle as a summand. Those algorithms answer these two questions in polynomial time in the number of lattice points of the polytope.

This bound is sharp for large fields and can be considered as an improvement of the bound of the previous section, since it is sharp for smaller base fields.

Example 8.22. Let P be the plane polytope with vertices $(0,0)$, $(4,1)$, $(1,4)$. The toric code $\mathcal{C}_P^t \subset \mathbb{F}_8^{49}$ has parameters [49,11,28] (see example 8.20).

One has that $L(P) = 3$ and $A = 15/2$. By theorem 8.21 one has that

$$d(\mathcal{C}_P) \geq (q-1)^2 - 3(q-1)$$

for $q \geq 53$. This can be improved considering the zeros of an irreducible curve at $X_P \setminus T$ (see [34, Section 2.2]) and one has that the previous bound holds for $q \geq 37$. One can easily see that this lower bound is sharp (for instance considering the upper bound [32, proposition 4.2]). However, one cannot obtain the minimum distance of $\mathcal{C}_P$ with $q = 8$ using this bound.

8.5.3. *Bound with intersection theory*

One can estimate the minimum distance using intersection theory. This lower bound is the one in [19] for toric varieties. This bound is also presented in [27, theorem 7.7] in this book. One can find the computations for this lower bound in [17, 32]. We remark that the following results are only valid for a non-singular toric variety.

One has that T is contained in the following $(q-1)^{r-1}$ lines (and therefore irreducible curves):

$$C_{\eta_1,\ldots,\eta_{r-1}} = \mathrm{Z}(\{\chi^{u_i} - \eta_i : i = 1,\ldots,r-1\}), \qquad \eta_i \in \mathbb{F}_q^* \ \forall i$$

where $\{u_1, \ldots u_r\}$ is a basis of M. Following [27, theorem 7.7], for a polytope P, the toric code $\mathcal{C}_P$ has

$$d(\mathcal{C}_P) \geq (q-1)^r - l(q-1) - ((q-1)^{r-1} - l)(D_P \cdot C)$$

where l is the maximum number of lines where a function can vanish completely and $D_P \cdot C$ is the *intersection number* of the Cartier Divisor D_P and the 1-cycle $C = \mathrm{V}(\{\chi^{u_i} : i = 1,\ldots,r-1\})$ [12], since $D_P \cdot C_{\eta_1,\ldots,\eta_{r-1}}$ has the same value for any $\eta_1,\ldots,\eta_{r-1} \in \mathbb{F}_q$.

The bound for the minimum distance can be understood in the following way: the weight of a word in $\mathcal{C}_P$ is greater than or equal to the length minus the maximum number of points of T belonging to the zero locus of a rational function. The number of points of this zero locus is considered as the union of the lines where a rational function f vanishes completely and the points in the zero locus of the lines where f does not vanish; this second number is by definition $D_P \cdot C$ [10].
Following [14, Chapter 5] one has that

$$D_P \cdot C = D_P \cdot (\mathrm{div}(\chi^{u_1}))_0 \cdot \ldots \cdot (\mathrm{div}(\chi^{u_{r-1}}))_0$$

and this intersection number can be computed using the mixed volume of the associated polytopes

$$r!V_r(P, P_{(\mathrm{div}(\chi^{u_1}))_0}, \ldots, P_{(\mathrm{div}(\chi^{u_{r-1}}))_0})$$

The *mixed volume* V_r of r polytopes $P_1, \ldots, P_r$ is

$$V_r(P_1, \ldots, P_r) = \frac{1}{r!} \sum_{j=1}^{r} (-1)^{r-j} \sum_{1 \leq i_1 < \cdots < i_j \leq r} \mathrm{vol}_r(P_{i_1} + \cdots + P_{i_j})$$

where Vol_r is the Lebesgue volume. An algorithm to compute the Lebesgue volume of a polytope may be found in [3]. Moreover under certain hypothesis the mixed volume can be computed directly [25].

Let $f \in \mathrm{H}^0(X_P, \mathcal{O}(D_P))$, since $\mathcal{C}_P = \mathcal{C}_{P'}$ if and only if $\overline{P} = \overline{P'}$ we assume without loss of generality that $\deg_{t_i} f \leq q-2$.

$$f(t_1, \ldots, t_r) = f_0(t_1, \ldots, t_{r-1}) + f_1(t_1, \ldots, t_{r-1}) t_r + \cdots + f_{q-2}(t_1, \ldots, t_{r-1}) t_r^{q-2}$$

Let $C_{\eta_1, \ldots, \eta_{r-1}}$ be a line where f vanishes, $f(\eta_1, \ldots, \eta_{r-1}, t_r) \in \mathbb{F}_q[t_r]$ and $\deg f(\eta_1, \ldots, \eta_{r-1}, t_r) < t_r^{q-1}$. Therefore, since $f(\eta_1, \ldots, \eta_{r-1}, t_r) = 0 \, \forall \, t_r \in \mathbb{F}_q^*$, it follows $f_i(\eta_1, \ldots, \eta_{r-1}) = 0 \; \forall \; i$.

The number l is lower than or equal to the maximum number of zeros of a non-zero function $f \in \mathrm{H}^0(X_{P'}, \mathcal{O}(D_{P'}))$ where P' is the r-projection of the polytope P. This can be repeated until we reach dimension 2.

For a *plane polytope* we compute the minimum distance as in [17]. Let us consider P a plane polytope and let us bound the minimum distance. In dimension 2 we can improve the previous computation: as in this dimension a 1-cycle is a Weil divisor and f vanishes in l of the previous lines, one has that

$$\mathrm{div}(f) + D_P - l(\mathrm{div}(\chi^{u_1}))_0 \succeq 0$$

Or, equivalently, $f \in \mathrm{H}^0(X_P, \mathcal{O}(D_P - l(\mathrm{div}(\chi^{u_1}))_0))$, and the maximum number of zeros of f in the other $(q-1-l)$ lines is $D_P - l(\mathrm{div}(\chi^{u_1}))_0 \cdot (\mathrm{div}(\chi^{u_1}))_0$, which is lower than or equal to the previous one. This will probably allow us to to give a sharper bound.

From lemma 8.10 one has that

$$l \leq \max\{u_2 - u_2' \mid u_1 = u_1', (u_1, u_2) \in P, (u_1', u_2') \in P\}$$

since $\{\chi^u \mid u \in P \cap M\}$ is a basis of $\mathcal{L}(D)$.

One can find in [17, 32] several examples where this bound is sharp. Finally, we compute the intersection number of the two Cartier divisors

just in the same way as for $r > 2$, using the mixed volume of the associated polytopes.

Example 8.23. Let P be the plane polytope with vertices $(0,0)$, $(b_1,0)$, (b_1,b_2), $(0,b_2)$, with $b_1, b_2 \in \mathbb{N}$. The polytope P is a dilatation of the polytope in example 8.3, the variety X_P is $\mathbb{P}^1 \times \mathbb{P}^1$ (see 8.2). This example has been also considered in chapter 7 of this volume [27, Example 7.8]. From the polytope P it is also clear that the code $\mathcal{C}_P$ is the product of two doubly extended Reed-Solomon codes.

The fan $\triangle_P$ associated to P is generated by cones with edges generated by $v(\rho_1) = (1,0)$, $v(\rho_2) = (0,1)$, $v(\rho_3) = (-1,0)$ and $v(\rho_4) = (0,-1)$. The toric variety X_P is non-singular.

$$P = \bigcap_{i=1}^{4} \{\langle u, \rho_i \rangle \geq -a_i\}$$

where $a_1 = 0$, $a_2 = 0$, $a_3 = b_1$, $a_4 = b_2$. Therefore $D_P = \sum a_i V(\rho_i) = b_1 V(\rho_3) + b_2 V(\rho_4)$.

Then $\mathcal{C}_P$ has length $n = (q-1)^2$. By theorem 8.13 the evaluation map ev is injective if $b_1, b_2 < q-1$ and the dimension of $\mathcal{C}_P$ is

$$k = \max\{b_1 + 1, q-1\} \times \max\{b_2 + 1, q-1\}$$

By using the bound obtained by intersection theory, we get that the minimum distance is

$$d \geq n - l(q-1) + (q-1-l)(D_P - l(\mathrm{div}(\chi^{u_1}))_0 \cdot (\mathrm{div}(\chi^{u_1}))_0)$$

where $l \leq b_1$.
One has that $\mathrm{div}(\chi^{u_1}) = \sum \langle u_1, v(\rho_i) \rangle V(\rho_i) = V(\rho_1) - V(\rho_3)$. Therefore, $(\mathrm{div}(\chi^{u_1}))_0 = V(\rho_1)$.
The intersection number is (see [32] for details):

$$\begin{aligned}(D_P - l(\mathrm{div}(\chi^{u_1}))_0) \cdot (\mathrm{div}(\chi^{u_1}))_0 &= (b_1 V(\rho_3) + b_2 V(\rho_4)) \cdot V(\rho_1) \\ &\quad - l V(\rho_1) \cdot V(\rho_1) \\ &= b_2 - 0 = b_2\end{aligned}$$

Therefore the minimum distance is bounded by

$$d \geq (q-1)^2 - (b_1(q-1-b_2) + (q-1)b_2) = (q-1-b_1)(q-1-b_2)$$

The previous example shows the following fact: whenever the refined fan of a plane toric variety includes the cones generated by $(1,0)$, $(0,1)$, $(-1,0)$ and $(0,-1)$, the lower bound can be easily computed using the following result. This situation is quite common when P defines a singular variety and the normal fan needs to be refined using the algorithm in [14, section 2.2].

Proposition 8.24. *Let P be a polytope, and let $\triangle_P$ be the refined normal fan such that the cones generated by $(1,0)$, $(0,1)$, $(-1,0)$ and $(0,-1)$ are in $\triangle_P$. In other words, $\triangle_P$ is a refinement of the fan of $\mathbb{P}^1 \times \mathbb{P}^1$.*
Then, the bound with intersection theory is

$$d \geq (q-1)^2 - \text{width(P)}(q-1) - (q-1-\text{width(P)})A$$

where $A = \text{vol}_2(P + (-1,0)) - \text{vol}_2(P)$, and

$$\text{width}(P) = \max\{u_2 - u_2' \mid u_1 = u_1', (u_1,u_2) \in P, (u_1',u_2') \in P\}$$

Proof. Let $\triangle_P$ be the normal fan of P. Then $\triangle_P$ is generated by cones with edges generated by $v(\rho_1) = (1,0)$, $v(\rho_2) = (0,1)$, $v(\rho_3) = (-1,0)$, $v(\rho_4) = (0,-1)$, $v(\rho_5), \ldots, v(\rho_m)$, let $P = \bigcap_{i=1}^m \{\langle u, \rho_i\rangle \geq -a_i\}$.
One has that

$$(\text{div}(\chi^{u_1}))_0 = V(\rho_1) + V(\rho_3) + \sum_{i=5}^{m} a_i V(\rho_i)$$

We claim that $P_{(\text{div}(\chi^{u_1}))_0}$ is the one-dimensional polytope Q with vertices $(-1,0)$ and $(0,0)$. It is clear that $P_{(\text{div}(\chi^{u_1}))_0} \subset Q$, the other content holds because one can easily prove that there is no line in $M_{\mathbb{R}}$ defined by $\langle u, v(\rho_i)\rangle = -a_i$ through Q. Therefore, the Lebesgue volume of $P_{(\text{div}(\chi^{u_1}))_0}$ is zero.

Therefore

$$d \geq (q-1)^2 - l(q-1) - (q-1-l)((D_P - l(\text{div}(\chi^{u_1}))_0) \cdot (\text{div}(\chi^{u_1}))_0) =$$

$$(q-1)^2 - l(q-1) - (q-1-l)(D_P \cdot (\text{div}(\chi^{u_1}))_0)$$

where $l = \text{width}(P)$.

Moreover, one has that $(D_P \cdot (\text{div}(\chi^{u_1}))_0) = 2V_2(P, P_{(\text{div}(\chi^{u_1}))_0}) = \text{vol}_2(P + P_{(\text{div}(\chi^{u_1}))_0}) - \text{vol}_2(P) - \text{vol}_2(P_{(\text{div}(\chi^{u_1}))_0}) = \text{vol}_2(P+Q) - \text{vol}_2(P)$.
And the result holds □

In particular, the previous result shows that the parameters of the code, considering the designed distance obtained using intersection theory, are not

better than those of the codes in example 8.23. This result also explains why this bound is not sharp at all for Joyner's example.

Example 8.25. Let P be the plane polytope with vertices $(0,0)$, $(4,1)$, $(1,4)$ and let $q = 8$. The toric code $\mathcal{C}_P^t \subset \mathbb{F}_8^{49}$ has parameters [49,11,28] (see examples 8.20 and 8.22).

The fan $\triangle_P$ associated to P is generated by cones with edges generated by $(-1,4)$, $(-1,-1)$, $(4,-1)$. The variety X_P is singular since $(4,-1),(-1,4)$ is not a basis of $\mathbb{Z}^2$. We consider the algorithm in [14, section 2.6] to refine the normal fan and we obtain the minimal resolution of X_P. The refined fan is generated by 13 edges. Namely,

$$P = \bigcap_{i=1}^{13} \{u \mid \langle u, v(\rho_i) \geq -a_i \rangle\}$$

where $v(\rho_1) = (1,0)$, $v(\rho_2) = (0,1)$, $v(\rho_3) = (-1,4)$, $v(\rho_4) = (-1,3)$, $v(\rho_5) = (-1,2)$, $v(\rho_6) = (-1,1)$, $v(\rho_7) = (-1,0)$, $v(\rho_8) = (-1,-1)$, $v(\rho_9) = (0,-1)$, $v(\rho_{10}) = (1,-1)$, $v(\rho_{11}) = (2,-1)$, $v(\rho_{12}) = (3,-1)$, $v(\rho_{13}) = (4,-1)$ and $a_1 = 0$, $a_2 = 0$, $a_3 = 0$, $a_4 = 1$, $a_5 = 2$, $a_6 = 3$, $a_7 = 4$, $a_8 = 5$, $a_9 = 4$, $a_{10} = 3$, $a_{11} = 2$, $a_{12} = 1$, $a_{13} = 0$.

Therefore, we can use proposition 8.24 to compute the bound for the minimum distance. We have that the width of P is 3. Let Q be the polytope with vertices $(-1,0)$ and $(0,0)$, then $\mathrm{vol}_2(P+Q) - \mathrm{vol}_2(P) = 4$. Hence, the minimum distance is

$$d \geq (8-1)^2 - 3(8-1) - (8-1-3)4 = 12$$

where 12 is much smaller than the minimum distance, 28.

Let P be a polytope, the bound using intersection theory can be improved using the following approaches:

- One may obtain a sharper bound considering different lines $C_{\eta_1,\ldots,\eta_{r-1}}$ to bound the zeros of a rational function at T.
- One can consider an invertible map $T : M_{\mathbb{R}} \to M_{\mathbb{R}}$, $T(x) = M(x) + \lambda$, such that M is an $r \times r$ matrix and λ is an r-dimensional vector (see section 8.3). Then P and $Q = T(P)$ are lattice equivalent polytopes whose associated codes, $\mathcal{C}_P$ and $\mathcal{C}_Q$, have the same parameters. Then the intersection theory bound may give a sharper bound for the minimum distance of $\mathcal{C}_P$.

However, it is not clear so far how to consider these improvements in a canonical way for an arbitrary polytope.

8.6. Conclusion

This chapter has been an introduction to toric geometry and toric codes, we hope this will help the reader to start working with toric codes. There are two main research problems for toric codes: although there are examples of toric codes with good parameters, the problem of identifying polytopes that will give rise to good toric codes is still open. The second problem is to obtain an efficient decoding algorithm. Another interesting remark, is that new results for toric geometry and polytopes can give rise to obtaining new results for toric codes.

8.7. Bibliographical notes

In this section we give the reader some references for deeper study and the source of the results of this chapter. Moreover, we would like to mention that there is a generalization of toric codes: papers [1] and [2] generalize toric geometry to allow certain non-rational varieties. Algebraic geometry codes coming from these so-called T-Varieties are investigated in [22]. Estimates for the dimension and minimum distance of the codes can be calculated using a similar approach to the toric case (with intersection theory).

Section 8.1: In order to work with toric codes, one should be familiar with some concepts of algebraic geometry and toric geometry. For an introduction, we recommend the introduction to toric geometry by D. Cox in [7]. Even though, most of the bibliography for toric geometry is over the complex field, one can find in [8] a survey of toric geometry where the results hold for any field. The source of toric geometry [26] may be interesting for some readers. One can find the standard references for toric geometry in [14, 31], and a deep study of the mixed volume and the combinatorics of toric geometry in [15]. For intersection theory, we refer the reader to the book by Fulton [12], and we recommend for an introduction [13].

Section 8.2: Toric codes were introduced by J.P Hansen [16–18]. However, toric codes were introduced in [16] in a slightly different way to the one presented in this chapter. J.P. Hansen considers the algebraic closure of $\mathbb{F}_q$ as the base field and he evaluates the functions of $\mathrm{H}^0(X_P, \mathcal{O}(D_P))$ invariant under the action of Frobenious, that is, the functions that are $\mathbb{F}_q$-linear combination of $\{\chi^u \mid u \in P \cap M\}$. Therefore, this construction gives the same toric code as the one presented in this chapter.

D. Joyner defines in [23] a code for a toric variety coming from a com-

plete fan, a Cartier divisor and a 1-cycle, he uses the 1-cycle to evaluate rational functions at their support. Then, he considers the special case where the 1-cycle has support T and he calls these codes *standard toric codes*. However, since a complete fan and a Cartier divisor is the same data as a polytope P, the toric codes defined here are the standard toric codes [23, definition 4.5]. D. Joyner has implemented the construction of toric codes and some related procedures, including the desingularization, [36], and one can also find there an introduction to toric geometry and toric codes.

Section 8.3: The classification used for toric codes comes from [29]. One can find a classification for toric codes coming from plane polytopes of small dimension k.

Section 8.4: The multicylic structure was proved for plane toric codes in [11] by representing the words of the code by matrices, it is also claimed in there that toric codes coming from a polytope of higher dimension are multicylic as well. One can find the proof, representing the words of the code by polynomials, in [33]. The dual of a toric code was obtained indepently in [6] and [33].

Section 8.5: The different lower bounds for the minimum distance are from [17, 28, 29, 32, 34]. The bounds with the Minkowski sum and length are valid for toric codes coming from plane polytopes. The bound in [29] is valid for r-dimensional polytopes but the computation only holds for two families so far (rectangular polytopes and simplices). The bound in [17] is for plane polytopes and it was extended for arbitrary dimension in [32].

References

[1] K. Altmann and J. Hausen, Polyhedral divisiors and algebraic torus actions, *Math Ann.* **344**, 557–607, (2006).
[2] K. Altmann, J. Hausen, and H. Süß, Gluing affine torus actions via divisorial fans, *Transformation Groups.* **13**, (2008).
[3] A. I. Barvinok, Computing the volume, counting integral points, and exponential sums, *Discrete Comput. Geom.* **10**, 123–141, (1993).
[4] A. Barvinok, Computing the Ehrhart quasi-polynomial of a rational simplex, *Math. Comp.* **75**(255), 1449–1466 (electronic), (2006).
[5] N. Berline and M. Vergne, The equivariant Todd genus of a complete toric variety, with Danilov condition, *J. Algebra.* **313**(1), 28–39, (2007).
[6] M. Bras-Amorós and M. E. O'Sullivan, Duality for some families of correction capability optimized evaluation codes, *Adv. Math. Commun.* **2**(1), 15–33, (2008).
[7] D. Cox. What is a toric variety? In ed. R. K. R. Goldman, *Topics in Al-*

gebraic Geometry and Geometric Modeling, vol. 334. AMS Contemporary Mathematics, (2003).
[8] V. I. Danilov, The geometry of toric varieties, *Russian Math. Surverys.* **33** (2), 97–154, (1978).
[9] J. A. De Loera, The many aspects of counting lattice points in polytopes, *Math. Semesterber.* **52**(2), 175–195, (2005).
[10] O. Debarre, *Higher-Dimensional Algebraic Geometry.* Universitext, (Springer-Verlag, 2001).
[11] V. Díaz, C. Guevara, and M. Vath, Codes from n-dimensional polyhedra and n-dimensional cyclic codes, *Proceedings of SIMU summer institute.* (2001).
[12] W. Fulton, *Intersection Theory.* Ergebnisse der Mathematik und ihrer Grenzgebiete 3. Folge, Band 2, (Springer-Verlag, 1984).
[13] W. Fulton, *Introduction to Intersection Theory in Algebraic Geometry.* vol. 54, *Conference Board of the Mathematical Sciences*, (AMS, 1984).
[14] W. Fulton, *Introduction to Toric Varieties.* Annals of Mathematics Studies, (Princeton University Press, 1993).
[15] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky, *Discriminants, Resultants and Multidimensional Determinants.* Mathematics: Theory & Applications, (Birkhäuser, 1994).
[16] J. P. Hansen, Toric surfaces and error-correcting codes, *Coding theory, cryptography and related areas (Guanajuato,1998).* pp. 132–142, (2000).
[17] J. P. Hansen, Toric varieties Hirzebruch surfaces and error-correcting codes, *Appl. Algebra Engrg. Comm. Comput.* **13**, 289–300, (2002).
[18] J. P. Hansen, Toric surfaces and codes, techniques and examples, *DMF Preprint: DMF-2004-01-27.* (2004).
[19] S. H. Hansen, Error-correcting codes from higher-dimensional varieties, *Finite Fields Appl.* **7**, 530–552, (2001).
[20] R. Hartshorne, *Algebraic Geometry.* vol. 52, *Graduate Texts in Math.*, (Springer-Verlag, 1977).
[21] J. E. Humphreys, *Linear Algebraic Groups.* vol. 21, *Graduate Texts in Math.*, (Springer-Verlag, 1975).
[22] N. O. Ilten and H. Süß, Ag-codes from polyhedral divisors, *Preprint.* (2008).
[23] D. Joyner, Toric codes over finite fields, *Appl. Algebra Engrg. Comm. Comput.* **15**, 63–79, (2004).
[24] G. Kempf, F. Knudsen, D. Mumford, and B. Saint-Donat, *Toroidal Embeddings I.* vol. 339, *Lectures Notes in Mathematics*, (Springer-Verlag, 1973).
[25] A. G. Khovanskii, Newton polyhedra, a new formula for mixed volume, product of roots of a system equations, *Fields Inst. Commun.* **24**, 325–364, (1999).
[26] S. L. Kleiman, Toward a numerical theory of ampleness, *Ann. of Math. (2).* **84**(3), 293–344, (1966).
[27] J. B. Little. Algebraic geometry codes from higher dimensional varieties. In eds. E. Martínez-Moro, C. Munuera, and D. Ruano, *Advances in Algebraic Geometry Codes*, chapter 7. pp. 257–293. World Scientific, (2008).
[28] J. Little and H. Schenck, Toric surface codes and Minkowski sums, *SIAM J. Discrete Math.* **20**(4), 999–1014 (electronic), (2006).

[29] J. Little and R. Schwarz, On toric codes and multivariate Vandermonde matrices, *Appl. Algebra Engrg. Comm. Comput.* **18**(4), 349–367, (2007).
[30] R. Morelli, Pick's theorem and the Todd class of a toric variety, *Adv. Math.* **100**(2), 183–231, (1993).
[31] T. Oda, *Convex Bodies and Algebraic Geometry. An Introduction to the Theory of Toric Varieties*. Ergebnisse der Mathematik und ihrer Grenzgebiete, 3. Folge, Band 15, (Springer-Verlag, 1984).
[32] D. Ruano, On the parameters of r-dimensional toric codes, *Finite Fields Appl.* **13**(4), 962–976, (2007).
[33] D. Ruano, On the structure of generalized toric codes, *To appear in Journal of Symbolic Computation. Preprint ArXiv:cs.IT/0611010*. (2006).
[34] I. Soprunov and E. Soprunova, Toric surface codes and minkowski length of polygons, *arXiv:0802.2088v1 [math.AG]*. (2008).
[35] B. Sturmfels, *Gröbner Bases and Convex Polytopes*. vol. 8, *AMS University Lecture Series*, (American Mathematical Society, 1996).
[36] H. Verrill and D. Joyner, Computing with toric varieties, *J. Symbolic Comput.* **42**(5), 511–532, (2007).

Chapter 9

Algebraic Geometric Codes over Rings

Katherine G. Bartley and Judy L. Walker*

Department of Mathematics, University of Nebraska, Lincoln, NE 68588–0130,

jwalker@math.unl.edu

In this chapter, algebraic geometric codes over local, Artinian rings are defined and studied. Decoding algorithms for these codes are also presented.

Contents

9.1. Introduction

Whenever information is transmitted across a channel, errors are likely to occur. Since Shannon's groundbreaking paper [36], coding theorists have sought to construct codes which have many codewords, that are easy to encode and decode, and that correct errors. While the main tools used in coding theory have traditionally been those of combinatorics and group theory, this volume is dedicated to codes constructed using algebraic geometry. Such codes were first introduced by Goppa [13] in 1977; see Definition 9.1

*This work was partially supported by the National Science Foundation grant DMS-0602332.

below. Soon after Goppa's original paper, Tsfasman, Vlăduţ and Zink [43] used modular curves to construct a sequence of codes with asymptotically better parameters than any previously known codes. Thus, the study of algebraic geometric codes took on great significance.

The field of coding theory took another major turn with the 1994 paper of Hammons, Kumar, Calderbank, Sloane and Solé [15] that shows that certain nonlinear binary codes are, in fact, nonlinear images of linear codes over the ring $\mathbb{Z}/4\mathbb{Z}$. The study of linear codes over rings has continued to mature into a mathematical field of study in its own right, causing Alexander Barg, Professor of Electrical and Computer Engineering at the University of Maryland, to state at an AMS meeting in October 2006, "We do not have to pretend that what we are doing has anything to do with information transfer any more."

The object of this chapter is to combine these areas of coding theory by introducing and studying algebraic geometric codes over rings. The remainder of this chapter is structured as follows: In Section 9.1.1, we review the required notions from the study of linear codes over rings, and the required background on curves over rings is found in Section 9.1.2. Algebraic geometric codes over rings are defined in Section 9.2 and their basic properties are given there. Section 9.3 considers these codes with respect to weight measures other than the Hamming weight. Three decoding algorithms for algebraic geometric codes over rings are given in Section 9.4: the so-called *basic algorithm* in Section 9.4.1, a list-decoding algorithm in Section 9.4.2, and a variation of the list-decoding algorithm that allows for decoding with respect to weight measures other than the Hamming weight in Section 9.4.3.

For future reference, we end this section with the definition and main theorem on algebraic geometric codes over finite fields.

Definition 9.1 (Goppa, [13]; see also [40], [42]). Let X be a smooth, absolutely irreducible, projective curve over the finite field $\mathbb{F}_q$. Let $\mathcal{P} = \{P_1, \ldots, P_n\} \subseteq X(\mathbb{F}_q)$ be a set of n distinct $\mathbb{F}_q$-rational points on X and let D be a divisor on X such that $\operatorname{Supp} D \cap \mathcal{P} = \emptyset$. The *algebraic geometric codes* associated to X, $\mathcal{P}$ and D are

$$C_L(X, \mathcal{P}, D) = \{(f(P_1), \ldots, f(P_n)) \,|\, f \in L(D)\}$$

and

$$C_\Omega(X, \mathcal{P}, D) = \{(\operatorname{res}_{P_1}(v), \ldots, \operatorname{res}_{P_n}(v)) \,|\, v \in \Omega(D)\}$$

where $L(D)$ is the vector space of rational functions associated to D and $\Omega(D)$ is the vector space of rational differential forms associated to D.

The following theorem gives the basic properties of algebraic geometric codes over finite fields.

Theorem 9.2 (Goppa, [13]; see also [40], [42]).
Let X, $\mathcal{P} = \{P_1, \ldots, P_n\}$ and D be as in Definition 9.1, and assume $2g - 2 < \deg D < n$, where g is the genus of X. Then

(1) $C_L(X, \mathcal{P}, D)$ is a linear code of length n, dimension $\deg D + 1 - g$ and minimum distance at least $\delta_L := n - \deg D$.
(2) $C_\Omega(X, \mathcal{P}, D)$ is a linear code of length n, dimension $n - (\deg D + 1 - g)$ and minimum distance at least $\delta_\Omega := \deg D - 2g + 2$.
(3) $C_L(X, \mathcal{P}, D)^\perp = C_\Omega(X, \mathcal{P}, D)$.
(4) For some canonical divisor K on X, we have $C_\Omega(X, \mathcal{P}, D) = C_L(X, \mathcal{P}, K + \mathcal{P} - D)$, where $\mathcal{P}$ is being used here to mean the Weil divisor $P_1 + \cdots + P_n$.

Definition 9.3. The quantities δ_L and δ_Ω are called the *designed minimum distances* of $C_L(X, \mathcal{P}, D)$ and $C_\Omega(X, \mathcal{P}, D)$, respectively.

9.1.1. *Codes over Rings*

Recall that a code of minimum (Hamming) distance d can correct any error pattern of weight at most $\lfloor \frac{d-1}{2} \rfloor$. Thus one wants to find codes with high minimum distance. On the other hand, the efficiency of a code can be measured in terms of its dimension, which, in the case of a possibly nonlinear code with M codewords over an alphabet of size q, can be interpreted as $\log_q(M)$. Hence one wants to find codes with many codewords. These two goals are at odds with one another, leading one to consider them together: Over a fixed alphabet, what is the largest number of codewords that a code of length n and minimum distance d can have?

This question, of course, is part of what makes algebraic geometry codes so interesting. Soon after the introduction of algebraic geometry codes by Goppa [13], Tsfasman, Vlăduţ and Zink [43] showed that, for $q \geq 49$ a perfect square, there is a sequence of algebraic geometry codes that asymptotically beats the Gilbert-Varshamov bound; this was the first time such a sequence had been shown to exist even though the bound had been known since 1952 [12]. Moreover, this result was a bit of a triumph for mathematicians, who tend to hold dear the belief that structure is good: though many

different methods, both algebraic and random, had been used in an effort to construct codes beating the asymptotic Gilbert-Varshamov bound, no attempt was successful until the powerful tools of algebraic geometry were brought into play.

On the other hand, it has long been known that, for certain lengths and minimum distances, there are nonlinear binary codes with more codewords than any linear code of the same length and minimum distance. One example of this is the Nordstrom-Robinson code [34], a code of length 16 with minimum distance 8 and 256 codewords. The Nordstrom-Robinson code lies at the base of two families of nonlinear codes that have more codewords than all known linear codes with the same length and minimum distance: the Kerdock codes [21] and the Preparata [35] codes. The fact that these good codes are nonlinear was very troubling to many mathematicians. In a breakthrough paper [15], Hammons, Kumar, Calderbank, Sloane and Solé showed that these codes do indeed have a linear structure, when viewed as codes over the ring of integers modulo 4 rather than as binary codes.

We begin with the (standard) definition.

Definition 9.4. Let A be a ring. A *code of length n over A* is a subset C of A^n. If C is a submodule of A^n, we say C is a *linear code* over A. If C is a free A-module, we say that C is *free* and define the *dimension* of C to be $\dim C = \mathrm{rank}_A(C)$.

Although some work on codes over rings was done as early as 1972 [6], interest in these codes did not become widespread until the 1994 paper of Hammons, et al. [15]. In that paper, it was shown that many good nonlinear binary codes, including the Kerdock and Preparata codes (and in particular the Nordstrom-Robinson code) can be obtained as images of linear codes over the ring $\mathbb{Z}/4\mathbb{Z}$ under the *Gray map*. The Gray map $\phi : \mathbb{Z}/4\mathbb{Z} \to \mathbb{F}_2 \times \mathbb{F}_2$ can be defined by

$$\begin{aligned} \phi(0) &= (0,0) \\ \phi(1) &= (0,1) \\ \phi(2) &= (1,1) \\ \phi(3) &= (1,0) \end{aligned}$$

and then extended to a map $\phi : (\mathbb{Z}/4\mathbb{Z})^n \to \mathbb{F}_2^{2n}$ by declaring

$$\phi\left((x_1, x_2, \ldots, x_n)\right) = \left(\phi(x_1), \phi(x_2), \ldots, \phi(x_n)\right).$$

The study of codes over rings has exploded, with many authors exploring many different aspects of the topic. For example, different rings have been considered. Hammons, et al. [15] moved beyond the ring $\mathbb{Z}/4\mathbb{Z}$ and considered codes over *Galois rings*, which are rings formed by adjoining to $\mathbb{Z}/p^e\mathbb{Z}$, for some prime p and some positive integer e, a root of a monic irreducible polynomial that remains irreducible over the residue field $\mathbb{F}_p$. Further generalizations followed, including to *finite chain rings*—finite commutative local rings in which the unique maximal ideal is principal.

The point of the comment of Barg's quoted in the introduction to this chapter was that the questions about codes over rings have become more theoretical, often with the following flavor: "Here's a fundamental property of codes over finite fields. What is the largest class of rings for which this property holds?" A major achievement in this direction came in 1999, when Wood [51] showed that every code C over a fixed ring R satisfies $|C|\left|C^{\perp}\right| = |R^n|$ if and only if R is a finite quasi-Frobenius ring. In the same paper, Wood also showed that the MacWilliams Identities [30] hold for all codes over a fixed ring R if and only if R is a finite Frobenius ring.

The original work [48], [50] of Walker on algebraic geometric codes over rings required that the ring be local[a] and Artinian[b] and, in some instances, Gorenstein[c]. On the other hand, much of the literature on general codes over rings assumes the rings are Frobenius or quasi-Frobenius. In fact, the conditions Gorenstein, Frobenius and quasi-Frobenius coincide for commutative Artinian rings [8]. We note, in particular, that Galois rings and, more generally, finite chain rings are finite, local, Artinian, Gorenstein (hence Frobenius and quasi-Frobenius) rings.

9.1.2. *Curves over Rings*

This section gives some properties of curves over rings that will be needed for the remainder of the chapter. Let A be a local Artinian ring with maximal ideal $\mathfrak{m}$ and finite residue field $\mathbb{F}_q$. Let $\mathbf{X} \subset \mathbb{P}^r_A$ be a *curve* over A, by which we mean that $\mathbf{X}$ is a smooth irreducible projective scheme over $\operatorname{Spec} A$ of relative dimension one. The natural map $f : \mathbf{X} \to \operatorname{Spec}(A)$ is called the *structure morphism* of $\mathbf{X}$ over A. Let

$$X = \mathbf{X} \times_{\operatorname{Spec} A} \operatorname{Spec} \mathbb{F}_q$$

[a]See, e.g., [31] for general algebraic background. A *local* ring is a finite commutative ring with a unique maximal ideal.

[b]A ring is *Artinian* if it satisfies the descending chain condition: any descending chain $I_1 \supseteq I_2 \supseteq \ldots$ of ideals must eventually stabilize.

[c]A local Artinian ring is *Gorenstein* if it is injective as a module over itself.

be the fibre of $\mathbf{X}$ over $\mathfrak{m}$ and let $\phi : X \to \mathbf{X}$ be the canonical embedding. As in [50], we assume that X is absolutely irreducible. Additional information on curves over rings can be found in [16] and [17].

In the field case, the group $\mathrm{Div}(X)$ of Weil divisors modulo linear equivalence, the group $\mathrm{CaCl}(X)$ of Cartier divisors modulo linear equivalence and the Picard group $\mathrm{Pic}(X)$ of isomorphism classes of line bundles are all three isomorphic. In particular, the fact that $\mathrm{Div}(X) \simeq \mathrm{Pic}(X)$ is used implicitly in Definition 9.1 to associate the vector spaces $L(D)$ and $\Omega(D)$, which are really global sections of line bundles, to the Weil divisor D. Although this isomorphism does not hold in the more general setting of curves over local Artinian rings, it is still true that $\mathrm{CaCl}(X) \simeq \mathrm{Pic}(X)$. As the notion of a Cartier divisor is less familiar to many readers than is that of a Weil divisor, we include the definition and essential ingredients of the isomorphism next.

Definition 9.5 (See, e.g., [17]). Let $\mathbf{X}$ be a scheme and let $\mathcal{K}$ denote the sheaf of total quotient rings on $\mathbf{X}$. Denote by $\mathcal{K}^*$ the sheaf of invertible elements in $\mathcal{K}$ and let $\mathcal{O}^*$ be the sheaf of invertible elements of $\mathcal{O}_{\mathbf{X}}$. A *Cartier divisor* D is a global section of the sheaf $\mathcal{K}^*/\mathcal{O}^*$. If a Cartier divisor D is in the image of the natural map $\Gamma(\mathbf{X}, \mathcal{K}^*) \to \Gamma(\mathbf{X}, \mathcal{K}^*/\mathcal{O}^*)$, then D is said to be *principal*. Two Cartier divisors D and D' are *linearly equivalent* if their difference is principal. The *Cartier divisor class group* $\mathrm{CaCl}(\mathbf{X})$ is the group of Cartier divisors modulo linear equivalence.

A Cartier divisor can be represented in the form $D = \{(U_i, f_i)\}$ where $\{U_i\}$ is an open cover of $\mathbf{X}$ and f_i is an element of $\Gamma(U_i, \mathcal{K}^*)$ such that for each i and j,

$$f_i/f_j \in \Gamma(U_i \cap U_j, \mathcal{O}^*).$$

Although the group operation on $\mathcal{K}^*/\mathcal{O}^*$ is multiplication, it is standard (see, e.g., [17]) to use the language of additive groups when talking about Cartier divisors in order to preserve the analogy of Cartier divisors with Weil divisors; we have done this already in Definition 9.5 in using the difference of two divisors to define linear equivalence.

The next definition shows how to associate a line bundle to a Cartier divisor; the subsequent proposition indicates how the isomorphism between $\mathrm{CaCl}(X)$ and $\mathrm{Pic}(X)$ works.

Definition 9.6. Let $D = \{(U_i, f_i)\}$ be a Cartier divisor on $\mathbf{X}$. The sub-

sheaf $\mathcal{O}_{\mathbf{X}}(D)$ of $\mathcal{K}$ is given by

$$\Gamma(U_i, \mathcal{O}_{\mathbf{X}}(D)) = f_i^{-1}\Gamma(U_i, \mathcal{O}_{\mathbf{X}}) = \frac{1}{f_i}\mathcal{O}_{\mathbf{X}}(U_i).$$

Proposition 9.7 (Proposition II.6.13, [17]). *Let* $\mathbf{X}$ *be a curve over a local Artinian ring. Then*

(1) For any Cartier divisor D, $\mathcal{O}_{\mathbf{X}}(D)$ *is an invertible sheaf on* $\mathbf{X}$. *The map* $D \mapsto \mathcal{O}_{\mathbf{X}}(D)$ *gives a 1-1 correspondence between Cartier divisors and invertible subsheaves of* $\mathcal{K}$.
(2) $\mathcal{O}_{\mathbf{X}}(D_1 - D_2) \simeq \mathcal{O}_{\mathbf{X}}(D_1) \otimes \mathcal{O}_{\mathbf{X}}(D_2)^{-1}$
(3) $D_1 \sim D_2$ *if and only if* $\mathcal{O}_{\mathbf{X}}(D_1) \simeq \mathcal{O}_{\mathbf{X}}(D_2)$ *as sheaves.*

As seen in Definition 9.1, an important ingredient in the construction of algebraic geometry codes over finite fields is the notion of $\mathbb{F}_q$-rational points. The analog in the ring case is the notion of A-points, which we define next.

Definition 9.8 (Definition 4.3, [50]). Let $\mathbf{X}$ be a curve over A and let Z be a zero-dimensional closed subscheme of $\mathbf{X}$. Let $i : Z \to \mathbf{X}$ be inclusion and $f : \mathbf{X} \to \operatorname{Spec} A$ be the structure morphism. Then Z is an *A-point* of $\mathbf{X}$ if the composition $f \circ i$ is an isomorphism of schemes.

From Definition 9.8, it follows that $\Gamma(Z, \mathcal{O}_{\mathbf{X}}|_Z) \simeq A$ for any A-point Z of $\mathbf{X}$. It is noted in [50] that every closed point $P \in \mathbf{X}$ that is an $\mathbb{F}_q$-rational point of X is contained in an A-point of $\mathbf{X}$. Furthermore, since $A/\mathfrak{m} \simeq \mathbb{F}_q$, the unique closed point of any A-point Z is an $\mathbb{F}_q$-rational point of X. If Z_1 and Z_2 are A-points containing $\mathbb{F}_q$-rational points P_1 and P_2 respectively, then Z_1 and Z_2 are *disjoint* if $P_1 \neq P_2$.

Given an A-point Z on $\mathbf{X}$, there is a unique, well-defined Cartier divisor (which we will also denote by Z) associated to Z. We can express this divisor explicitly in terms of a *local parameter* for Z in a neighborhood of the unique closed point P contained in Z; see [16] for the original statement (proof omitted) that these local parameters exist and [48] for a detailed proof. Letting $U = \operatorname{Spec} B$ be an affine open neighborhood of P on which the ideal for Z is principal and letting t be a local parameter for Z on U, one can show that $B/(t) \simeq A$ and that t is a unit on the set $U \setminus \{P\}$. Let $V = \mathbf{X} \setminus \{P\}$. Then t is a unit on $U \cap V = U \setminus \{P\}$ and the Cartier divisor for Z can be expressed as $\{(U, t), (V, 1)\}$.

In Section 9.2 below, we will give two constructions of algebraic geometric codes over local Artinian rings. As in the field case of Definition 9.1,

one of these constructions will amount to evaluating rational functions, and the other will amount to taking residues of rational differential forms. To analyze these constructions, versions of the Riemann-Roch Theorem and the Residue Theorem, both of which are well-known in the field case, will be needed. A complete discussion is given in [50]; here we give only the statements of these theorems and the definitions necessary to do so. We treat the Riemann-Roch Theorem first.

Definition 9.9. Let $\mathcal{L}$ be a line bundle on the curve $\mathbf{X}$ defined over the local Artinian ring A with residue field $\mathbb{F}_q$, let $X = \mathbf{X} \times_{\operatorname{Spec} A} \operatorname{Spec} \mathbb{F}_q$, and let $\phi : X \to \mathbf{X}$ be the canonical embedding. Let D' be a Weil divisor on X such that $\phi^*(\mathcal{L}) = \mathcal{O}_X(D')$. Then the *degree* of $\mathcal{L}$ is defined to be

$$\deg\mathcal{L} = \deg D'.$$

Definition 9.10. With $\mathbf{X}$ and X as above, the *genus* of $\mathbf{X}$ is defined to be the genus of X.

Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be line bundles on a curve $\mathbf{X}$ of genus g. Then $\phi^*(\mathcal{L}_1 \otimes \mathcal{L}_2) = \phi^*(\mathcal{L}_1) \otimes \phi^*(\mathcal{L}_2)$ and $\deg(\mathcal{L}_1 \otimes \mathcal{L}_2) = \deg\mathcal{L}_1 + \deg\mathcal{L}_2$. Furthermore, if ω is the canonical line bundle on $\mathbf{X}$ (see Section II.8 of [17] for a definition), then since $\mathbf{X}$ is smooth, $\phi^*(\omega)$ is the canonical line bundle on X. Hence, in this situation, $\deg\omega = \deg\phi^*(\omega) = 2g - 2$.

We can now state the required version of the Riemann-Roch Theorem.

Theorem 9.11 (Theorem 4.7, [50]). *Let $\mathbf{X}$ be a curve of genus g over the local Artinian ring A with residue field $\mathbb{F}_q$ and let $\mathcal{L}$ be a line bundle on $\mathbf{X}$. Let $X = \mathbf{X} \times_{\operatorname{Spec} A} \operatorname{Spec} \mathbb{F}_q$ and let $\mathcal{L}' = \phi^*(\mathcal{L})$, where $\phi : X \to \mathbf{X}$ is the canonical embedding. If $\deg\mathcal{L} > 2g - 2$, then $\Gamma(\mathbf{X}, \mathcal{L})$ is a free A-module of rank $\deg\mathcal{L} + 1 - g$.*

Next, we turn our attention to the Residue Theorem. For more details on the following discussion, see [50]; the treatment there is based on [16]. For the remainder of this section, we assume A is a Frobenius (equivalently, Gorenstein) ring.

Definition 9.12. Let η be the generic point of $\mathbf{X}$, let ω be the canonical sheaf on $\mathbf{X}$, and let ω_η be the stalk of ω at η. Any element of ω_η is called a *rational differential form* on $\mathbf{X}$.

Proposition 9.13. *Let Z be an A-point on the curve $\mathbf{X}$ defined over the ring A, let P be the unique closed point contained in Z and let t be a*

local parameter for Z. For any rational differential form $v \in \omega_\eta$, let $\bar{v}$ be the image of v in ω_η/ω_P, where ω_P is the stalk of ω at P. Then, in a neighborhood of Z, $\bar{v}$ has an expansion of the form

$$\bar{v} = \sum_{j<0} a_j t^j \, dt$$

with each a_j in A. Moreover, if Y is another A-point on $\mathbf{X}$ containing the same closed point P, s is a local parameter for Y and

$$\bar{v} = \sum_{j<0} b_j s^j \, ds$$

is the expansion of $\bar{v}$ in a neighborhood of Y, then $a_{-1} = b_{-1}$.

Definition 9.14. With notation as in Proposition 9.13, the *residue* of v at Z is defined to be $\mathrm{res}_Z(v) = a_{-1}$ and the *residue* of v at P is defined to be $\mathrm{Res}_P(v) = \mathrm{res}_Z(v)$ for any A-point Z containing P.

Theorem 9.15 (Corollary 4.14, [50]). *Let A be a local, Artinian, Frobenius (Gorenstein) ring, let $\mathbf{X}$ be a curve over A, and let S be the set of closed points on $\mathbf{X}$. Then for any rational differential form v on $\mathbf{X}$,*

$$\sum_{P \in S} \mathrm{Res}_P(v) = 0.$$

9.2. Algebraic Geometric Codes over Rings

Algebraic geometric codes over rings were first studied by Walker in [48], [50]. Let A be a local Artinian ring and let $\mathbf{X}$ be a curve over A. Let $\mathcal{Z} = \{Z_1, \ldots, Z_n\}$ be a set of disjoint A-points on $\mathbf{X}$ and let $\mathcal{L}$ be a line bundle on $\mathbf{X}$. For each i, let $\gamma_i : \Gamma(Z_i, \mathcal{L}|_{Z_i}) \to A$ be an isomorphism, and let $\gamma = \{\gamma_1, \ldots, \gamma_n\}$ be the system of these isomorphisms.

Definition 9.16 (Definition 5.1, [50]). Let A, $\mathbf{X}$, $\mathcal{Z}$, $\mathcal{L}$, and γ be as above and let $C_L(\mathbf{X}, \mathcal{Z}, \mathcal{L}, \gamma)$ be the image of the composition α

$$\Gamma(\mathbf{X}, \mathcal{L}) \longrightarrow \oplus \Gamma(Z_i, \mathcal{L}|_{Z_i}) \xrightarrow{\gamma} A^n \qquad (\alpha) \tag{9.1}$$

where the map $\Gamma(\mathbf{X}, \mathcal{L}) \to \oplus \Gamma(Z_i, \mathcal{L}|_{Z_i})$ is given by restriction. Then $C_L(\mathbf{X}, \mathcal{Z}, \mathcal{L}, \gamma)$ is called an *algebraic geometric code over A*.

It should be noted that this definition of algebraic geometric codes over rings is motivated by the "H" construction [42], which gives a generalization of algebraic geometric codes to allow for algebraic varieties of dimension greater than one. In the case that the ring A is a field, the two definitions of algebraic geometric codes coincide. In fact, we can use the fact mentioned in Section 9.1.2 above that $\mathrm{Pic}(X) \simeq \mathrm{CaCl}(X)$ to interpret the system γ of isomorphisms as evaluation, thus making Definition 9.16 a direct generalization of Definition 9.1 from the field case to the ring case. To do this, we first need to define what it means for an A-point to be not in the support of a Cartier divisor.

Definition 9.17 (Definition 5.3, [50]). Let D be a Cartier divisor on $\mathbf{X}$, and let $P \in \mathbf{X}$ be a closed point that is a rational point of $X = \mathbf{X} \times_{\mathrm{Spec}\, A} \mathrm{Spec}\, \mathbb{F}_q$. We say that P *is not in the support of* D if we can write $D = \{(U_i, f_i)\}$, where $f_i \in \mathcal{O}_{\mathbf{X}}(U_i)^{\times}$ for some i such that $P \in U_i$. If Z is an A-point containing P and P is not in the support of D, we say Z *is not in the support of* D as well.

Now let A, $\mathbf{X} \subset \mathbb{P}^r$, $\mathcal{Z}$, $\mathcal{L}$ and γ be as in Definition 9.16 above, and suppose we can find a Cartier divisor D such that $\mathcal{O}_{\mathbf{X}}(D) = \mathcal{L}$ and Z is not in the support of D for every $Z \in \mathcal{Z}$. Let P be the closed point contained in some $Z \in \mathcal{Z}$ and write $D = \{(U_i, f_i)\}$ where $\{U_i = \mathrm{Spec}\, B_i\}$ is an open cover of $\mathbf{X}$ and, for some i, we have that $P \in U_i$ and that f_i is a unit of $\mathcal{O}_{\mathbf{X}}(U_i)$. Write $U = U_i$ and $B = B_i$. Since $P \in U$, we have $Z \subset U$ and so $Z = \mathrm{Spec}\, B/J$ for some ideal J of B such that $B/J \simeq A$. Let $s \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(D))$. Since $\Gamma(U, \mathcal{O}_{\mathbf{X}}(D)) = \frac{1}{f}B$, we have $s|_U = \frac{h}{f}$ for some $h \in B \subset \mathcal{K}(\mathbf{X})$. Suppose Z is given in projective coordinates by $Z = (z_0 : \ldots : z_r)$. Since A is local and $z_0, \ldots, z_r$ generate the unit ideal of A, some z_j is a unit. Without loss of generality, we may assume that $z_0 = 1$ and U is contained in the standard affine open subset of $\mathbb{P}^r$ defined by $z_0 = 1$. Then J is the ideal generated by $z_1 - x_1, \ldots, z_n - x_n$, and the map of (9.1) on the the coordinate corresponding to Z becomes

$$\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(D)) \longrightarrow \Gamma(Z, \mathcal{O}_{\mathbf{X}}(D)|_Z) \xrightarrow{\gamma} A \tag{9.2}$$

and is given by

$$s \mapsto \frac{h(1, z_1, \ldots, z_n)}{f(1, z_1, \ldots, z_n)} \in A.$$

In other words, this map may be thought of as merely evaluating s at Z. When we wish to think of things in this way, we will write $s(Z)$ to

represent the image of $s \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(D))$ under the composite map (9.2), i.e., $s(Z) = \gamma(s|_Z)$.

The relationship between algebraic geometric codes over finite fields and algebraic geometric codes over rings is further strengthened in the next theorem.

Theorem 9.18 (Theorem 5.5, [50]). *Let $\mathbf{X}$, $\mathcal{Z}$ and $\mathcal{P}$ be as before. Let $\mathcal{L}$ be a line bundle on X and let $\mathcal{L}' = \phi^*(\mathcal{L})$, where $\phi : X \to \mathbf{X}$ is the canonical embedding. Let $\gamma = \{\gamma_i : \Gamma(Z_i, \mathcal{L}|_{Z_i}) \to A\}$ be any system of isomorphisms and let $\gamma' = \{\gamma_i'\}$ be the induced system of isomorphisms*

$$\gamma_i' : \Gamma(P_i, \mathcal{L}'|_{P_i}) = \Gamma(Z_i, \mathcal{L}|_{Z_i}) \otimes_A \mathbb{F}_q \to \mathbb{F}_q.$$

Setting $C = C_L(\mathbf{X}, \mathcal{Z}, \mathcal{L}, \gamma)$, $C' = C_{\mathbb{F}_q}(X, \mathcal{P}, \mathcal{L}', \gamma')$ and $\overline{C} = \pi(C)$, where $\pi : A^n \to \mathbb{F}_q^n$ denotes coordinatewise projection, we have $\overline{C} = C'$.

The next theorem summarizes the fundamental properties of the codes $C_L(\mathbf{X}, \mathcal{Z}, \mathcal{L}, \gamma)$. The proof, which we include for completeness, uses Theorem 9.18 above and mimics the proof of the field case.

Theorem 9.19 (Theorem 5.4 and Corollary 5.7, [50]). *Let A, $\mathbf{X}$, $\mathcal{L}$, $\mathcal{Z} = \{Z_1, \ldots, Z_n\}$ and γ be as above. Assume that the residue field of A is finite. Let g denote the genus of $\mathbf{X}$, and suppose $2g - 2 < \deg\mathcal{L} < n$. The $C_L(\mathbf{X}, \mathcal{Z}, \mathcal{L}, \gamma)$ is a free code of length n, dimension $k = \deg\mathcal{L} + 1 - g$ and minimum Hamming distance at least $\delta_L := n - \deg\mathcal{L}$.*

Proof. Set $C = C_L(\mathbf{X}, \mathcal{Z}, \mathcal{L}, \gamma)$. It is clear that C has length n. We will show that it is free of dimension $k = \deg\mathcal{L} + 1 - g$ by showing that the composite map $\alpha : \Gamma(\mathbf{X}, \mathcal{L}) \to A^n$ of Definition 9.16 is injective and then applying Theorem 9.11.

Let $s \in \Gamma(\mathbf{X}, \mathcal{L})$ such that $\alpha(s) = 0$. Let D be any Cartier divisor such that $\mathcal{O}_{\mathbf{X}}(D) \simeq \mathcal{L}$. Write $D = \{(U_j, g_j)\}$ and $Z_i = \{(U_j, g_{ij})\}$ where refinements have been taken if necessary. We may then write the divisor $D - Z_1 - \cdots - Z_n$ as $\{U_j, \frac{g_i}{g_{1j} \ldots g_{nj}}\}$. We first show that $s \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(D - Z_1 - \cdots - Z_n))$, i.e., $s \in \frac{g_i}{g_{1j} \ldots g_{nj}} \mathcal{O}_{\mathbf{X}}(U_j)$ for each j.

Since $s \in \Gamma(\mathbf{X}, \mathcal{L})$, we have $s \in \frac{1}{g_j} \mathcal{O}_{\mathbf{X}}(U_j)$ for each j, i.e., $g_j s \in \mathcal{O}_{\mathbf{X}}(U_j)$. Since $\alpha(s) = 0$ for $i = 1, \ldots, n$, $g_j s \in g_{ij} \mathcal{O}_{\mathbf{X}}(U_j) \subseteq \mathcal{O}_{\mathbf{X}}(U_j)$ for each i and j. Because the Z_i are disjoint, we have

$$\bigcap_i g_{ij} \mathcal{O}_{\mathbf{X}}(U_j) = g_{1j} \ldots g_{nj} \mathcal{O}_{\mathbf{X}}(U_j),$$

which shows $\ker(\alpha) = \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(D - Z_1 - \cdots - Z_n))$. Let $\phi : X \to \mathbf{X}$ be the canonical embedding. Since $\deg\phi^*\mathcal{O}_{\mathbf{X}}(D - Z_1 - \cdots - Z_n) = \deg(D - Z_1 - \cdots - Z_n) < 0$, we have $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(D - Z_1 - \cdots - Z_n)) = 0$ by Nakayama's Lemma [31]. Thus $\alpha : \Gamma(\mathbf{X}, \mathcal{L}) \to A$ is an injection. Since C is the image of α, this means that $C \simeq \Gamma(\mathbf{X}, \mathcal{L})$. By Theorem 9.11, C is free of dimension (rank) $k = \deg\mathcal{L} + 1 - g$ as desired.

Since we now know that C is a free code, Theorem 3.4 of [50] says that the minimum distance of C is precisely that of $\overline{C}$, where $\overline{C}$ is the coordinate-wise projection to $C \subset A^n$ to $\mathbb{F}_q^n$, where $\mathbb{F}_q$ is the residue field of A. By Theorem 9.18, $\overline{C}$ is the algebraic geometric code $C_L(X, \mathcal{P}, \mathcal{L}', \gamma')$ over $\mathbb{F}_q$, where $\mathcal{P}$ is the set of closed points contained in the A-points of $\mathcal{Z}$, $\mathcal{L}' = \phi^*(\mathcal{L})$, and $\gamma' = \{\gamma_i'\}$ is the induced system of isomorphisms

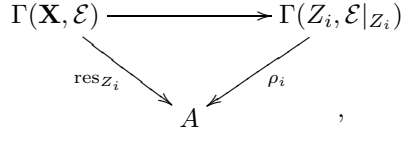
$$\gamma_i' : \Gamma(P_i, \mathcal{L}'|_{P_i}) = \Gamma(Z_i, \mathcal{L}|_{Z_i}) \otimes_A \mathbb{F}_q \to \mathbb{F}_q.$$

The result is now immediate from Theorem 9.2, which says that $C_L(X, \mathcal{P}, \mathcal{L}', \gamma')$ has minimum distance at least $n - \deg\mathcal{L}' = n - \deg\mathcal{L}$. □

Definition 9.20. The quantity $\delta_L := n - \deg\mathcal{L}$ in Theorem 9.19 is called the *designed minimum distance* of $C_L(\mathbf{X}, \mathcal{Z}, \mathcal{L}, \gamma)$.

One very nice property of algebraic geometric codes over finite fields is that they are closed under taking duals. This property also holds for algebraic geometric codes over local Artinian rings that are also Frobenius (or, equivalently, Gorenstein), as we will now show. In an analogous manner to what is done in the field case, we first define *residue codes*.

Let $\mathcal{E} = \omega \otimes \mathcal{O}_{\mathbf{X}}(\mathcal{Z})$, where $\mathcal{Z}$ is the Cartier divisor obtained by taking the sum of the Cartier divisors $Z_1, ..., Z_n$. From Lemma 5.8 of [50], we have $\Gamma(\mathbf{X}, \mathcal{E}) \subset \omega_\eta$. By Lemma 5.9 of [50], for each i, the map res_{Z_i} factors through $\Gamma(Z_i, \mathcal{E}|_{Z_i})$ and there exists an isomorphism $\rho_i : \Gamma(Z_i, \mathcal{E}|_{Z_i}) \to A$ that makes the following diagram commute:



where the map $\Gamma(\mathbf{X}, \mathcal{E}) \to \Gamma(Z_i, \mathcal{E}|_{Z_i})$ is given by restriction.

Definition 9.21. Let A be a local, Artinian, Frobenius ring, let $\mathbf{X}$ be a curve over A, let $\mathcal{Z} = \{Z_1, \ldots, Z_n\}$ be a set of disjoint A-points on $\mathbf{X}$ and

let $\mathcal{L}$ be a line bundle on $\mathbf{X}$. Let ω be the canonical line bundle on $\mathbf{X}$. For each i, let $\gamma_i : \Gamma(Z_i, \mathcal{L}|_{Z_i}) \to A$ be an isomorphism, and let

$$\xi_i : \Gamma(Z_i, \omega \otimes \mathcal{O}_{\mathbf{X}}(\mathcal{Z}) \otimes \mathcal{L}^{-1}|_{Z_i}) \to A$$

be the isomorphism given by the rule

$$\xi_i(v|_{Z_i}) = \rho_i(\gamma_i^{-1}(1)v|_{Z_i}),$$

where $\rho = \{\rho_i\}$ is the system of isomorphisms described above. Let $\gamma = \{\gamma_i\}$ and $\xi = \{\xi_i\}$. The *residue code* $C_\Omega(\mathbf{X}, \mathcal{Z}, \mathcal{L}, \gamma)$ over A is defined to be

$$C_\Omega(\mathbf{X}, \mathcal{Z}, \mathcal{L}, \gamma) = C_L(\mathbf{X}, \mathcal{Z}, \omega \otimes \mathcal{O}_{\mathbf{X}}(\mathcal{Z}) \otimes \mathcal{L}^{-1}, \xi).$$

Remark 9.22. If a Cartier divisor D can be chosen so that $\mathcal{L} \simeq \mathcal{O}_{\mathbf{X}}(D)$ and no $Z \in \mathcal{Z}$ is in the support of D, then γ can be thought of as the evaluation map (as mentioned earlier) and ξ can be thought of as the residue map.

The next theorem gives the basic properties of the codes $C_\Omega(\mathbf{X}, \mathcal{Z}, \mathcal{L}, \gamma)$. Its proof follows immediately from the definitions and Theorem 9.19.

Theorem 9.23. *Let* $\mathbf{X}$, $\mathcal{Z}$, $\mathcal{L}$ *and* γ *be as before. If* $2g - 2 < \deg\mathcal{L} < n$ *then* $C_\Omega(\mathbf{X}, \mathcal{Z}, \mathcal{L}, \gamma)$ *is a free code of dimension* $k_\Omega = n + g - 1 - \deg\mathcal{L}$ *and minimum Hamming distance at least* $\delta_\Omega := \deg\mathcal{L} - 2g + 2$.

Theorem 9.24 (See Theorem 5.12 of [50]). *Let* $\mathbf{X}$, $\mathcal{L}$, $\mathcal{Z}$, γ, *and* ξ *be described as above. Then*

$$C_L(\mathbf{X}, \mathcal{Z}, \mathcal{L}, \gamma)^\perp = C_L(\mathbf{X}, \mathcal{Z}, \omega \otimes \mathcal{O}_{\mathbf{X}}(\mathcal{Z}) \otimes \mathcal{L}^{-1}, \xi).$$

Proof. For $1 \le i \le n$, let $s_i \in \Gamma(Z_i, \mathcal{L}|_{Z_i})$ be the restriction of $s \in \Gamma(\mathbf{X}, \mathcal{L})$, and let $v_i \in \Gamma(Z_i, (\omega \otimes \mathcal{O}_{\mathbf{X}}(\mathcal{Z}) \otimes \mathcal{L}^{-1})|_{Z_i})$ be the image of $v \in \Gamma(\mathbf{X}, \omega \otimes \mathcal{O}_{\mathbf{X}}(\mathcal{Z}) \otimes \mathcal{L}^{-1})$.

As in Theorem 9.15, let S be the set of closed points on $\mathbf{X}$ and let $\mathcal{P} = \{P_1, \ldots, P_n\}$ be the set of closed points contained in $\mathcal{Z} = \{Z_1, \ldots, Z_n\}$. We claim that if $P \in S$ with $\mathrm{Res}_P(sv) \neq 0$, then $P \in \mathcal{P}$. To see this, let $U = \mathbf{X} \setminus \mathcal{P}$ and choose open subsets $V_1, \ldots, V_n$ of $\mathbf{X}$ such that $P_i \in V_i$ and, for $j \neq i$, $P_j \notin V_i$. Then $U \cup V_1 \cup \cdots \cup V_n$ is an open cover of $\mathbf{X}$, and we may express the Cartier divisor $\mathcal{Z}$, i.e., the Cartier divisor that is the sum of the points in $\mathcal{Z}$ as $\{(U, 1), (V_1, t_1), \ldots, (V_n, t_n)\}$, where t_i is a local parameter for Z_i on V_i. If $P \in S \setminus \mathcal{P}$, then $P \in U$ and so $sv|_U \in \Gamma(U, \omega)$, which means that $\mathrm{Res}_P(sv) = 0$.

Using this, we have

$$\begin{aligned}\sum_{i=1}^{n}\gamma_i(s_i)\xi_i(v_i) &= \sum_{i=1}^{n}\gamma_i(s_i)\rho_i(\gamma_i^{-1}(1)v_i) = \sum_{i=1}^{n}\rho_i(\gamma_i(s_i)\gamma_i^{-1}(1)v_i)\\ &= \sum_{i=1}^{n}\rho_i(\gamma_i^{-1}(\gamma_i(s_i)\cdot 1)v_i) = \sum_{i=1}^{n}\rho_i(s_iv_i)\\ &= \sum_{i=1}^{n}\mathrm{res}_{Z_i}(sv) = \sum_{i=1}^{n}\mathrm{Res}_{P_i}(sv) = \sum_{P\in S}\mathrm{Res}_P(sv)\\ &= 0\end{aligned}$$

by Theorem 9.15. □

Remark 9.25. For $\mathbf{X}$, $\mathcal{Z}$, $\mathcal{L}$ and γ as before, $C_L(\mathbf{X},\mathcal{Z},\mathcal{L},\gamma)$ is equivalent to a residue code. To see this, note that

$$\omega\otimes\mathcal{O}_{\mathbf{X}}(\mathcal{Z})\otimes(\mathcal{O}_{\mathbf{X}}(\mathcal{Z})\otimes\omega\otimes\mathcal{L}^{-1})^{-1}\simeq\mathcal{L}.$$

Thus $C_L(\mathbf{X},\mathcal{Z},\mathcal{L},\gamma)$ is equivalent to $C_\Omega(\mathbf{X},\mathcal{Z},\mathcal{O}_{\mathbf{X}}(\mathcal{Z})\otimes\omega\otimes\mathcal{L}^{-1},\psi)$, where

$$\psi=\{\psi_i:\Gamma(\mathbf{X},\mathcal{O}_{\mathbf{X}}(\mathcal{Z})\otimes\omega\otimes\mathcal{L}^{-1}|_{Z_i})\to A\}$$

is any system of isomorphisms.

As an application of the construction of algebraic geometric codes over rings, Walker [49] showed that the Nordstrom-Robinson code, a nonlinear binary code of length 16 with 256 codewords and minimum distance 6 that has more codewords than any linear code of the same length and minimum distance, is the image under the Gray map (see Section 9.1.1) of an algebraic geometric code over the ring $\mathbb{Z}/4\mathbb{Z}$. To do this, she gave explicit equations defining a curve $\mathbf{X}$ over $\mathbb{Z}/4\mathbb{Z}$, an explicit set $\mathcal{Z}$ of $\mathbb{Z}/4\mathbb{Z}$ points on the curve, an explicit Cartier divisor D on $\mathbf{X}$, and a basis for the $\mathbb{Z}/4\mathbb{Z}$-module $\Gamma(\mathbf{X},\mathcal{O}_{\mathbf{X}}(D))$ so that the image under the Gray map of $C_L(\mathbf{X},\mathcal{Z},\mathcal{L},\gamma)$ is the Nordstrom-Robinson code.

9.3. Non-Hamming Weights and Exponential Sums

Recall that one motivation for considering codes over rings other than fields is the paper [15] of Hammons, et al., which showed that certain nonlinear binary codes can be realized as images under the Gray map of linear codes over the ring $\mathbb{Z}/4\mathbb{Z}$. The *Lee weight* on $\mathbb{Z}/4\mathbb{Z}$ is the weight measure that makes the Gray map an isometry between $\mathbb{Z}/4\mathbb{Z}$ and $\mathbb{F}_2^2$: $w_L(0)=0$,

$w_L(1) = w_L(3) = 1$, $w_L(2) = 2$. More generally, we consider the *Euclidean weight* on the ring $\mathbb{Z}/p^l\mathbb{Z}$, which we define as follows.

Definition 9.26 (See [46]). Let $x \in \mathbb{Z}/p^l\mathbb{Z}$. The *Euclidean weight* of x is the distance in the complex plane between $e^{2\pi ix/p^l}$ and the point $(1, 0)$:

$$w_E(x) = \sqrt{\sin^2\left(\frac{2\pi x}{p^l}\right) + \left(1 - \cos\left(\frac{2\pi x}{p^l}\right)\right)^2} = \sqrt{2 - 2\cos\left(\frac{2\pi x}{p^l}\right)}.$$

Notice that the square of the Euclidean weight of $x \in \mathbb{Z}/4\mathbb{Z}$ is precisely twice the Lee weight of x.

For simplicity, we consider the square of the Euclidean weight rather than the Euclidean weight itself. For a vector $\mathbf{x} = (x_1, \ldots, x_n) \in (\mathbb{Z}/p^l\mathbb{Z})^n$, we define the *squared Euclidean weight* of $\mathbf{x}$ to be

$$w_E^2(\mathbf{x}) = \sum_{j=1}^{n} w_E^2(x_j).$$

The observation (see [46]) that allows one to find bounds on the minimum squared Euclidean weight of an algebraic geometric code over $\mathbb{Z}/p^l\mathbb{Z}$ is that there is a close relationship between the squared Euclidean weight and the modulus of a certain exponential sum. More precisely, since $\cos(\frac{2\pi x}{p^l}) = \mathrm{Re}(e^{2\pi ix/p^l})$, we have for $\mathbf{x} \in (\mathbb{Z}/p^l\mathbb{Z})^n$,

$$\begin{aligned} w_E^2(\mathbf{x}) &= \sum_{j=1}^{n} \left(2 - 2\,\mathrm{Re}(e^{2\pi ix_j/p^l})\right) \\ &= 2n - 2\,\mathrm{Re}\sum_{j=1}^{n} e^{2\pi ix_j/p^l} \\ &\geq 2n - 2\left|\sum_{j=1}^{n} e^{2\pi ix_j/p^l}\right|, \end{aligned}$$

and so, to find a lower bound on the minimum Euclidean weight of a linear code over $\mathbb{Z}/p^l\mathbb{Z}$, it is enough to find an upper bound on the modulus of the exponential sum

$$\sum_{j=1}^{n} e^{2\pi ix_j/p^l}$$

over all vectors $\mathbf{x} = (x_1, \ldots, x_n)$ in the code.

Now consider the case where the code in question is $C_L(\mathbf{X}, \mathcal{Z}, \mathcal{L}, \gamma)$, where $\mathbf{X}$ is a curve over the ring $A = \mathbb{Z}/p^l\mathbb{Z}$. Then the exponential sum above becomes

$$\sum_{Z \in \mathcal{Z}} e^{2\pi i f(Z)/p^l}$$

where $f \in \Gamma(\mathbf{X}, \mathcal{L})$.

It is not clear how to get a handle on this exponential sum in the most general case. However, in situations where the A-points are chosen to be algebraic liftings of the closed points, some progress has been made. To explain this progress, we must move to the language of Witt vectors [27]. For a field k of characteristic p, we write $W_l(k)$ for the ring of Witt vectors of length l over k. This ring is local with maximal ideal generated by p, such that $p^l = 0$. In the case that $k = \mathbb{F}_{p^m}$, we have $W_l(k) \simeq GR(p^l, m)$ and, in particular, $W_l(\mathbb{F}_p) \simeq \mathbb{Z}/p^l\mathbb{Z}$.

Definition 9.27. Let $A = W_l(\mathbb{F}_{p^m})$. The *Frobenius map* $F : A \to A$ is given by

$$F((x_0, \ldots, x_{l-1})) = (x_0^p, \ldots, x_{l-1}^p)$$

and the *trace map* $T : A \to W_l(\mathbb{F}_p) \simeq \mathbb{Z}/p^l\mathbb{Z}$ is given by

$$T(x) = x + F(x) + \cdots + F^{m-1}(x).$$

In light of this definition, we can consider a slightly more general class of codes. As usual, let $\mathbf{X}$ be a curve defined over $A = W_l(\mathbb{F}_q) \simeq GR(p^l, m)$ where $q = p^m$ and let $X = \mathbf{X} \times_{\operatorname{Spec} A} \operatorname{Spec} \mathbb{F}_q$, so that X is a curve defined over $\mathbb{F}_q$. Let $\mathcal{Z}$ be a set of disjoint A-points on $\mathbf{X}$, let D be a Cartier divisor on $\mathbf{X}$ such that no $Z \in \mathcal{Z}$ is in the support of D and let $f \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(D))$. As described above, in order to get a lower bound on the squared Euclidean weight of $T(C_L(\mathbf{X}, \mathcal{Z}, \mathcal{L}, \gamma))$, where $\mathcal{L} = \mathcal{O}_{\mathbf{X}}(D)$, it suffices to find an upper bound on the modulus of the sum

$$\sum_{Z \in \mathcal{Z}} e^{2\pi i T(f(Z))/p^l} \tag{9.3}$$

where $f \in \Gamma(\mathbf{X}, \mathcal{L})$. Sums of this type were considered in the case of the projective line by Li [28] and by Kumar, Helleseth and Calderbank [26].

In certain situations, one can transform the sum of (9.3) into a sum of the form

$$\sum_{P \in \mathcal{P}} e^{2\pi i T(\mathbf{f}(P))/p^l}$$

where $\mathcal{P} \subseteq X(\mathbb{F}_q)$ is the set of closed points contained in the A-points of $\mathcal{Z}$ and $\mathbf{f}$ is a Witt vector of rational functions on X associated to f. Then one applies the following theorem, due to Voloch and Walker in [46].

Theorem 9.28 (Theorem 3.1, [46]). *Let $q = p^m$ where p is prime and $m \geq 1$. Let X be a curve of genus g defined over the finite field $\mathbb{F}_q$ with function field $K := \mathbb{F}_q(X)$. Let $Q \in X(\mathbb{F}_q)$ and suppose $f_0, \ldots, f_{l-1}$ have poles only at Q. Consider the Witt vector of rational functions $\mathbf{f} := (f_0, \ldots, f_{l-1}) \in W_l(K)$. Set $X_0 = X \setminus \{Q\}$ and assume that $\mathbf{f}$ is not of the form $F(\mathbf{g}) - \mathbf{g} + c$ for any $\mathbf{g} \in W_l(K)$ and any $c \in W_l(\mathbb{F}_q)$. For $1 \leq i \leq n$, let $\deg f_i = -v_Q(f_i)$ be the order of the pole of f_i at Q. Then*

$$\left| \sum_{P \in X_0(\mathbb{F}_q)} e^{2\pi i T(\mathbf{f}(P))/p^l} \right| \leq \left(2g - 1 + \max\{p^{l-1-i} \deg f_i \,|\, 0 \leq i \leq l-1\}\right) \sqrt{q}.$$

Remark 9.29. This theorem easily extends to the case where the set of poles of $f_0, \ldots, f_{l-1}$ does not consist of a single $\mathbb{F}_q$-rational point on X; the more general version is actually what is given in [46].

The crux of the matter, then, is to be able to transform the sum which involves a rational function on $\mathbf{X}$ and runs over a set of A-points on $\mathbf{X}$ into a sum which involves a Witt vector of rational functions on X and runs over a set of $\mathbb{F}_q$-rational points on X. Voloch and Walker do exactly this in two situations: the case of canonical lifts [29] of ordinary [38] elliptic curves [46], [44]; and the case of plane curves with a unique point at infinity [45]. In each case, the trick is to find an algebraic lifting of points from $X(\mathbb{F}_q)$ to $\mathbf{X}(A)$. This work is extended in [47] to find bounds on the *homogeneous weight* of $T(C)$, where $C = C_L(\mathbf{X}, \mathcal{Z}, \mathcal{L}, \gamma)$ is defined over $GR(p^l, m)$ using either the projective line, the canonical lift of an ordinary elliptic curve, or a plane curve with a unique point at infinity.

Though motivated by the applications to coding theory, the results on exponential sums in [44], [45], [46] and [47] have proven to be of independent interest. Blache [4], [5] extended this work by considering other exponential sums along these and other curves defined over rings. Some of the results were also improved upon by Finotti [9], [10], [11], who used work of Mochizuki [33] to get better bounds on the degrees of the liftings of points.

9.4. Decoding Algebraic Geometric Codes over Rings

Whenever a new class of codes is proposed, it is important to also propose decoding algorithms for the codes in the class. In this section, we present three methods of decoding algebraic geometric codes over rings. First, we give a generalization of the so-called *Basic Algorithm* for decoding algebraic geometric codes over finite fields. This generalization allows for the decoding of algebraic geometric codes over rings, given by the "residue construction" of Definition 9.21 above, with respect to the Hamming distance and can correct any error pattern of weight up to $\lfloor \frac{\delta_\Omega - g - 1}{2} \rfloor$, where δ_Ω is the designed minimum distance of the code. As such codes should really be able to correct any error pattern of weight up to $\lfloor \frac{d-1}{2} \rfloor$ errors, we next provide a generalization of the Guruswami-Sudan [14] list decoding algorithm. Again, this algorithm decodes with respect to the Hamming distance. Since other weight measures, such as the Lee weight, or, more generally, the squared Euclidean weight, are often of interest when codes over rings are considered, we offer a third decoding algorithm — a modification of the second algorithm that, using ideas of Koetter and Vardy [25], works for an arbitrary weight measure.

The results of this section are from the first author's Ph.D. thesis [3] and are previously unpublished.

9.4.1. *The Basic Algorithm for the Hamming Metric*

This section describes a decoding algorithm for a residue code over a finite, local, Artinian, Gorenstein (Frobenius) ring A with respect to the Hamming distance. By Remark 9.25, any algebraic geometric code over A is equivalent to a residue code, so the algorithm decodes all algebraic geometric codes over A. The algorithm is a generalization of the *basic algorithm* for decoding algebraic geometric codes over finite fields. Presentations of the basic algorithm, which itself is a generalization of the Arimoto-Peterson algorithm for decoding Reed-Solomon codes, can be found in [18], [20] and [39].

We begin this section with a proposition that provides motivation for the basic algorithm.

Proposition 9.30 (Compare to Proposition 2.4 of [18]). *Let $C \subset A^n$ be a free code with parity check matrix H and let $\vec{y} \in A^n$. If there exist $\vec{c} \in C$ and $\vec{e} \in A^n$ such that $\vec{y} = \vec{c} + \vec{e}$ with $|\{j \mid e_j \neq 0\}| < d(C)$, then*

$\vec{x} = \vec{e}$ is the unique solution of the system of linear equations given by

$$H\vec{x}^T = H\vec{y}^T$$

and

$$x_j = 0 \text{ for } j \notin J.$$

As before, let A be a finite, local, Artinian, Gorenstein (Frobenius) ring with maximal ideal $\mathfrak{m}$ and finite residue field $\mathbb{F}_q$, and let $\mathbf{X}$ be a curve over A of genus g. Let D be a Cartier divisor on $\mathbf{X}$ such that $2g-2 < \deg\mathcal{O}_{\mathbf{X}}(D) < n$, let $\mathcal{Z} = \{Z_1, \ldots, Z_n\}$ be a set of pairwise disjoint A-points on $\mathbf{X}$, and let $\gamma = \{\gamma_i : \Gamma(Z_i, \mathcal{O}_{\mathbf{X}}(D)|_{Z_i}) \to A\}$ be a system of isomorphisms. For $1 \le i \le n$, let P_i be the closed point contained in Z_i, and let $\mathcal{P} = \{P_1, \ldots, P_n\}$. We will assume that, for each $P_i \in \mathcal{P}$, P_i is not in the support of D. We omit most proofs in this section, as many of the results in this section follow from the definition of a Cartier divisor. The omitted proofs can be found in [3].

Let $C_\Omega = C_\Omega(\mathbf{X}, \mathcal{Z}, \mathcal{O}_{\mathbf{X}}(D), \gamma)$. Since $2g-2 < \deg\mathcal{O}_{\mathbf{X}}(D) < n$, by Theorem 9.23, C_Ω is a free code with minimum distance at least $\delta_\Omega = \deg\mathcal{O}_{\mathbf{X}}(D) - 2g + 2$. Hence, by Proposition 9.30, a received word $\vec{y} = \vec{c} + \vec{e}$, where $\vec{c} \in C_\Omega$ and $\vec{e} \in A^n$, can be correctly decoded if we can find a set $J \subset \{1, \ldots, n\}$ such that $j \in J$ if $e_j \neq 0$ and $|J| < \delta_\Omega$. We shall find this set under that condition that $\mathrm{wt}(\vec{e}) \le \left\lfloor \frac{\delta_\Omega - 1}{2} \right\rfloor$ by finding an *error locator function* for $\vec{y}$.

Definition 9.31 (See also [18]). Let $\vec{y} = \vec{c} + \vec{e}$, where $\vec{c} \in C_\Omega$ and $\mathrm{wt}(\vec{e}) \le \left\lfloor \frac{\delta_\Omega - 1}{2} \right\rfloor$. Set

$$I = \{i \,|\, 1 \le i \le n \text{ and } e_i \neq 0\}.$$

Let F be a Cartier divisor on $\mathbf{X}$ and let $\delta = \{\delta_i : \Gamma(Z_i, \mathcal{O}_{\mathbf{X}}(F)|_{Z_i}) \to A\}$ be a system of isomorphisms. A function

$$s \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F)) \setminus \mathfrak{m}\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F))$$

is an *error locator function* for $\vec{y}$ if $\delta_i(s|_{Z_i}) \in \mathfrak{m}$ for all $i \in I$.

If A is a field, then an error locator function for $\vec{y}$ is simply a rational function that is zero at all error positions. For the rest of this section, unless otherwise stated, let $\vec{y} = \vec{c} + \vec{e}$ be a received word with $\vec{c} \in C_\Omega$ and $\mathrm{wt}(\vec{e}) \le \left\lfloor \frac{\delta_\Omega - 1}{2} \right\rfloor$, let F be Cartier divisor on $\mathbf{X}$ with support disjoint from $\mathcal{P}$, and let $\delta = \{\delta_i : \Gamma(Z_i, \mathcal{O}_{\mathbf{X}}(F)|_{Z_i}) \to A\}$ be a system of isomorphisms.

The following lemma is needed to describe the set of error locator functions for $\vec{y}$.

Lemma 9.32. *Let A, $\mathbf{X}$, $\mathcal{Z}$, F, $\gamma = \{\gamma_i\}$ and $\delta = \{\delta_i\}$ be as before. Then, for each $Z_i \in \mathcal{Z}$, there exists an isomorphism*

$$\tau_i : \Gamma(Z_i, \mathcal{O}_{\mathbf{X}}(D-F)|_{Z_i}) \to A$$

such that, for $s \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F))$ and $v \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(D-F))$,

$$\delta_i(s|_{Z_i})\tau_i(v|_{Z_i}) = \gamma_i(sv|_{Z_i}).$$

Proof. Let $Z_i \in \mathcal{Z}$. Let $s \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F))$ and let $v \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(D-F))$. Write $F = \{(U_j, f_j)\}$ and $D = \{(U_j, g_j)\}$, where $\{U_j\}$ is an open affine cover of $\mathbf{X}$ and refinements have been taken if necessary. For each j, we have $s \in \frac{1}{f_j}\mathcal{O}_{\mathbf{X}}(U_j)$ and $v \in \frac{f_j}{g_j}\mathcal{O}_{\mathbf{X}}(U_j)$, and so $sv \in \frac{1}{g_j}\mathcal{O}_{\mathbf{X}}(U_j)$ for all j. Choose U_j such that $P_i \in U_j$. Then, as discussed in Section 9.2, $U_j = \operatorname{Spec} B$ and $Z_i = \operatorname{Spec} B/J$ for some ideal J of B such that $B/J \simeq A$. Therefore,

$$\Gamma(Z_i, \mathcal{O}_{\mathbf{X}}(D-F)|_{Z_i}) = \frac{\bar{f}_j}{\bar{g}_j}B/J,$$

where $\bar{g}_j$ and $\bar{f}_j$ are the images in B/J of g_j and f_j respectively. Since Z_i is neither in the support of F nor in the support of D, we may assume both g_j and f_j are units of B. Thus $\gamma_i(\frac{1}{\bar{g}_j}) = a_D$ and $\delta_i(\frac{1}{\bar{f}_j}) = a_F$ for some $a_D, a_F \in A^{\times}$. Define the isomorphism $\tau_i : \Gamma(Z_i, \mathcal{O}_{\mathbf{X}}(D-F)|_{Z_i}) \to A$ by the rule

$$\tau_i\left(\frac{\bar{f}_j}{\bar{g}_j}\right) = \frac{a_D}{a_F}.$$

□

Definition 9.33 (See also [18]). Let $\mathbf{X}$, $\mathcal{Z}$, $\mathcal{P}$, $C_\Omega(\mathbf{X}, \mathcal{Z}, \mathcal{O}_{\mathbf{X}}(D), \gamma)$, F and $\delta = \{\delta_i\}$ be as before. Let $\vec{y} \in A^n$ be any received word. The set $K(\vec{y}, F, \delta)$ is defined to be

$$K(\vec{y}, F, \delta) = \big\{ s \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F)) \setminus \mathfrak{m}\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F)) \;\big|\; \textstyle\sum_{i=1}^{n} y_i\delta_i(s|_{Z_i})\tau_i(v|_{Z_i}) = 0 \text{ for all } v \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(D-F)) \big\}$$

where $\{\tau_i\}$ is the system of isomorphisms given in Lemma 9.32.

We shall show that, under certain conditions, the elements of $K(\vec{y}, F, \delta)$ are error locator functions for $\vec{y}$ (see Theorem 9.35 below). Note that since A is finite and both $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F))$ and $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(D-F))$ are finitely generated, the A-modules $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F))$ and $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(D-F))$ contain finitely many elements. Therefore we may calculate $K(\vec{y}, F, \delta)$ by exhaustive search.

If A is a field, then it is shown in [39] that the elements of $K(\vec{y}, F, \delta)$ can be found by solving a system of linear equations. At this time, however, it has not been investigated whether this method of finding elements of $K(\vec{y}, F, \delta)$ holds when A is not a field.

The following lemma is needed to show that $K(\vec{y}, F, \delta)$ is a nonempty set of error locator functions for $\vec{y}$. A proof can be found in [3].

Lemma 9.34. *Let* $\mathbf{X}$, $\mathcal{Z}$, $\mathcal{P}$, $C_\Omega(\mathbf{X}, \mathcal{Z}, \mathcal{O}_\mathbf{X}(D), \gamma)$ *and* $\vec{y} = \vec{c} + \vec{e}$ *be as before. Let* $t = \mathrm{wt}(\vec{e})$, *and let* $I = \{i \mid e_i \neq 0\}$ *be the set of error positions for* $\vec{y}$. *Let* $\mathcal{Q} = \{Z_i \mid i \in I\}$, *and let* Q *be the Cartier divisor obtained by adding up the points of* $\mathcal{Q}$. *Let* F *be a Cartier divisor on* $\mathbf{X}$. *Then*

(1) If $\deg \mathcal{O}_\mathbf{X}(D - F) > t + 2g - 2$, *then* $C_\Omega(\mathbf{X}, \mathcal{Q}, \mathcal{O}_\mathbf{X}(D - F), \tau) = \{\vec{0}\}$, *where* $\tau = \{\tau_i : \Gamma(Z_i, \mathcal{O}_\mathbf{X}(D - F)|_{Z_i}) \to A\}$ *is any system of isomorphisms.*

(2) If $\deg \mathcal{O}_\mathbf{X}(D) > t + g$, *then* $\Gamma(\mathbf{X}, \mathcal{O}_\mathbf{X}(F - Q)) \setminus \mathfrak{m}\Gamma(\mathbf{X}, \mathcal{O}_\mathbf{X}(F - Q)) \neq \emptyset$.

If F *has support disjoint from* $\mathcal{P}$, *then*

(3) For any $s \in \Gamma(\mathbf{X}, \mathcal{O}_\mathbf{X}(F - Q))$ *and* $i \in I$, *we have* $\delta_i(s|_{Z_i}) = 0$ *for any system of isomorphisms* $\{\delta_i : \Gamma(Z_i, \mathcal{O}_\mathbf{X}(F - Q)|_{Z_i}) \to A\}$.

(4) If $s \in \Gamma(\mathbf{X}, \mathcal{O}_\mathbf{X}(F - Q)) \setminus \mathfrak{m}\Gamma(\mathbf{X}, \mathcal{O}_\mathbf{X}(F - Q))$, *then* $s \in K(\vec{y}, F, \delta)$, *where* $\delta = \{\delta_i : \Gamma(Z_i, \mathcal{O}_\mathbf{X}(F)|_{Z_i}) \to A\}$ *is any system of isomorphisms.*

We now show that $K(\vec{y}, F, \delta)$ is a nonempty set of error locator functions for $\vec{y}$.

Theorem 9.35 (Compare to [18]). *With notation as above, assume* $\deg(\mathcal{O}_\mathbf{X}(F)) \geq t + g$ *and* $\deg \mathcal{O}_\mathbf{X}(D - F) > t + 2g - 2$. *Then* $K(\vec{y}, F, \delta)$ *is a nonempty set of error locator functions for* $\vec{y}$.

Proof. Let $s \in K(\vec{y}, F, \delta)$. Then for all $v \in \Gamma(\mathbf{X}, \mathcal{O}_\mathbf{X}(D - F))$, we have

$$\sum_{i=1}^{n} y_i \delta_i(s|_{Z_i}) \tau_i(v|_{Z_i}) = \sum_{i \in I} e_i \delta_i(s|_{Z_i}) \tau_i(v|_{Z_i}) = 0,$$

where $\{\tau_i : \Gamma(Z_i, \mathcal{O}_\mathbf{X}(D - F)|_{Z_i}) \to A\}$ is the system of isomorphisms from Lemma 9.32. Hence the word $\vec{w} \in A^{|I|}$ with entries $w_i = e_i \delta_i(s|_{Z_i})$, $i \in I$, is a codeword of $C_L(\mathbf{X}, \mathcal{Q}, \mathcal{O}_\mathbf{X}(D - F), \tau)^\perp = C_\Omega(\mathbf{X}, \mathcal{Q}, \mathcal{O}_\mathbf{X}(D - F), \tau)$. On the other hand, by part (1) of Lemma 9.34, $C_\Omega(\mathbf{X}, \mathcal{Q}, \mathcal{O}_\mathbf{X}(D - F), \tau) = \{\vec{0}\}$. We conclude $\vec{w} = \vec{0}$, and hence $e_i \delta_i(s|_{Z_i}) = 0$ for all $i \in I$. Note $e_i \delta_i(s|_{Z_i}) = 0$ if either $\delta_i(s|_{Z_i}) = 0$ or $\delta_i(s|_{Z_i})$ and e_i are both nonzero (zero) divisors whose

product is zero. In either case, this implies that $\delta_i(s|_{Z_i})$ is an element of $\mathfrak{m}$, and thus s is an error locator function for $\vec{y}$.

By part (4) of Lemma 9.34,

$$\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F-Q)) \setminus \mathfrak{m}\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F-Q)) \subset K(\vec{y}, F, \delta).$$

Since $\deg\mathcal{O}_{\mathbf{X}}(F) \geq t+g$, we have $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F-Q)) \setminus \mathfrak{m}\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F-Q)) \neq \emptyset$ by part (2) of Lemma 9.34. Therefore $K(\vec{y}, F, \delta)$ is nonempty. □

Let $s \in K(\vec{y}, F, \delta)$. Before we can use Proposition 9.30 to decode $\vec{y}$, we need an upper bound on the number of A-points $Z_i \in \mathcal{Z}$ such that $\delta_i(s|_{Z_i}) \in \mathfrak{m}$.

Proposition 9.36. *With notation as above, let*

$$s \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F)) \setminus \mathfrak{m}\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F)).$$

Then $\delta_i(s|_{Z_i}) \in \mathfrak{m}$ *for at most* $\deg\mathcal{O}_{\mathbf{X}}(F)$ *of the* A*-points in* $\mathcal{Z}$.

Proof. Let $s \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F)) \setminus \mathfrak{m}\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F))$ and consider the word $\vec{s} \in A^n$ defined by $s_i = \delta_i(s|_{Z_i})$. By Theorem 9.18, there exists $\bar{s} \in \Gamma(X, \phi^*(\mathcal{O}_{\mathbf{X}}(F)))$ such that $s_i \equiv \delta_i'(\bar{s}|_{P_i}) \pmod{\mathfrak{m}}$ for all i. Hence $s_i \in \mathfrak{m}$ if and only if $\delta_i'(\bar{s}|_{P_i}) = 0$, if and only if $\bar{s}|_{P_i} = 0$. Because $s \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F)) \setminus \mathfrak{m}\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(F))$, we know $\bar{s} \neq 0$. Since s is a nonzero element of $\Gamma(X, \phi^*(\mathcal{O}_{\mathbf{X}}(F)))$ and F has support disjoint from $\mathcal{P}$, we conclude that $\bar{s}|_{P_i} = 0$ for at most $\deg\phi^*(\mathcal{O}_{\mathbf{X}}(F)) = \deg\mathcal{O}_{\mathbf{X}}(F)$ of the points in $\mathcal{P}$. Hence, $\delta_i(s|_{Z_i}) \in \mathfrak{m}$ for at most $\deg\mathcal{O}_{\mathbf{X}}(F)$ of the A-points in $\mathcal{Z}$. □

Using Proposition 9.36, we immediately obtain the following bounds.

Lemma 9.37 (Compare to [18]). *Suppose that* $\deg\mathcal{O}_{\mathbf{X}}(F) = t+g$ *and* $\deg\mathcal{O}_{\mathbf{X}}(D-F) > 2g-2+t$. *For* $s \in K(\vec{y}, F, \delta)$, *define* $J = \{j \mid \delta_j(s|_{Z_j}) \in \mathfrak{m}\}$. *Then* $t \leq \left\lfloor \frac{\delta_\Omega - g - 1}{2} \right\rfloor$ *and* $|J| < \delta_\Omega$.

We now describe the basic algorithm for decoding the code $C_\Omega(\mathbf{X}, \mathcal{Z}, \mathcal{O}_{\mathbf{X}}(D), \gamma)$ with respect to the Hamming metric. Let F be a Cartier divisor of $\mathbf{X}$ with support disjoint from $\mathcal{P}$ such that $\deg\mathcal{O}_{\mathbf{X}}(F) = t+g$, where $t = \left\lfloor \frac{\delta_\Omega - g - 1}{2} \right\rfloor$. For example, if there exists an $\mathbb{F}_q$-rational point P_0 of X with $P_0 \notin \mathcal{P}$, then we may let $F = \mathcal{O}_{\mathbf{X}}((t+g)Z_0)$ for any A-point Z_0 containing P_0. As before, let $\delta = \{\delta_i : \Gamma(Z_i, \mathcal{O}_{\mathbf{X}}(F)|_{Z_i}) \to A\}$ be a system of isomorphisms, and let H be the parity check matrix for $C_\Omega(\mathbf{X}, \mathcal{Z}, \mathcal{O}_{\mathbf{X}}(D), \gamma)$. The input to the algorithm is a received word $\vec{y}$. The basic algorithm (compare to Algorithm 4.1 of [18]) is then as follows:

Step 1. Compute $K(\vec{y}, F, \delta)$.
Step 2. If $K(\vec{y}, F, \delta) = \emptyset$ stop and output "?".
Step 3. Select an element s of $K(\vec{y}, F, \delta)$ and compute

$$J = \{j \mid \delta_j(s|_{Z_j}) \in \mathfrak{m}\}.$$

Step 4. Solve the system of equations

$$H\vec{x}^T = H\vec{y}^T$$

and

$$x_j = 0 \text{ for } j \notin J.$$

If a unique solution $\vec{e}$ exists output $\vec{y} - \vec{e}$, otherwise output "?".

Putting everything together, we have the following theorem. The proof follows from Theorem 9.35, Lemma 9.37 and Proposition 9.30.

Theorem 9.38 (Compare to Theorem 4.2 of [18]). *The basic algorithm corrects* $\left\lfloor \frac{\delta_\Omega - g - 1}{2} \right\rfloor$ *errors.*

9.4.2. *List Decoding for the Hamming Metric*

Let C be a linear code over a ring A with minimum Hamming distance d. Let $\vec{y}$ be a received vector. It well known that, if at most $\lfloor \frac{d-1}{2} \rfloor$ errors occurred during transmission, then there exists a unique closest codeword to $\vec{y}$ with respect to the Hamming distance. If more than $\lfloor \frac{d-1}{2} \rfloor$ errors occurred, then there may or may not be a unique closest codeword to $\vec{y}$, and many classical decoding algorithms may fail to correctly decode $\vec{y}$ in this case. This is the motivation for the list decoding problem. That is, given a received word $\vec{y}$ and an error bound e, we want to find all codewords $\vec{c} \in C$ such that $d(\vec{c}, \vec{y}) \leq e$, where $d(\vec{c}, \vec{y})$ denotes the Hamming distance between $\vec{c}$ and $\vec{y}$. If a list decoding algorithm finds all codewords within distance e of any received word, then the algorithm is said to be an e *error-correcting algorithm.*

In 1997, Sudan [41] presented a list decoding algorithm for generalized Reed-Solomon codes over finite fields. Shokrollahi and Wasserman [37] showed that Sudan's algorithm could be extended to one-point algebraic geometric codes over finite fields. Guruswami and Sudan [14] then presented an improved algorithm that corrects more errors than the original algorithm and showed that this improved algorithm could also be extended to the case of one-point algebraic geometric codes over finite fields. In [2], Armand

showed that the algorithm holds for Generalized Reed-Solomon codes over commutative rings. In this section, we show that the Guruswami-Sudan algorithm works for *one-point algebraic geometric codes over rings*, which are defined in Definition 9.39 below.

As before, let A be a local Artinian ring with principal maximal ideal $\mathfrak{m}$ and finite residue field $\mathbb{F}_q$. Let $\mathbf{X} \subset \mathbb{P}^r_A$ be a curve over A, and let $X = \mathbf{X} \times_{\mathrm{Spec\ A}} \mathrm{Spec}\ \mathbb{F}_q$ be the fibre of $\mathbf{X}$ over $\mathfrak{m}$. As before, assume that X is absolutely irreducible. Let $\mathcal{Z} = \{Z_1, \ldots, Z_n\}$ be a set of pairwise disjoint A-points of $\mathbf{X}$, and let Z be an A-point of $\mathbf{X}$ such that Z is disjoint from all the points of $\mathcal{Z}$. For $1 \leq i \leq n$, let P_i be the closed point contained in Z_i and let $\mathcal{P} = \{P_1, \ldots, P_n\}$. Let P be the closed point contained in Z.

Definition 9.39. Let $\mathbf{X}$, $\mathcal{Z}$ and Z be as above and let m be a positive integer with $2g - 2 < m < n$, where $n = |\mathcal{Z}|$ and g is the genus of $\mathbf{X}$. Let $\gamma = \{\gamma_i\}$, where $\gamma_i : \Gamma(Z_i, \mathcal{O}_{\mathbf{X}}(mZ)|_{Z_i}) \to A$ is the evaluation map for $1 \leq i \leq n$. Then the algebraic geometric code $C_L(\mathbf{X}, \mathcal{Z}, \mathcal{O}_{\mathbf{X}}(mZ), \gamma)$ is called a *one-point code.*

Remark 9.40. Using Theorem 9.19, we see that the algebraic geometric code $C_L(\mathbf{X}, \mathcal{Z}, \mathcal{O}_{\mathbf{X}}(mZ), \gamma)$ of Definition 9.39 is a free A-code of length $|\mathcal{Z}|$, dimension $m + 1 - g$ and minimum distance at least $n - m$.

Before we can describe the decoding algorithm we must define *valuations* for A-points on $\mathbf{X}$. Let $\mathcal{K}$ be the sheaf of total quotient rings on $\mathbf{X}$, write $\mathcal{K}(\mathbf{X}) = \Gamma(\mathbf{X}, \mathcal{K})$, and let Z be an A-point of $\mathbf{X}$. Set

$$\mathcal{M}_Z = \bigcup_{j=0}^{\infty} \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(jZ)) \subset \mathcal{K}(\mathbf{X}),$$

so that $\mathcal{M}_Z$ is the collection of all functions on $\mathbf{X}$ that have poles only at Z.

Definition 9.41. Let Z be an A-point on $\mathbf{X}$ and $f \in \mathcal{M}_Z$. The *valuation* $\nu_Z : \mathcal{M}_Z \to \{-n \,|\, n \in \mathbb{N}\} \cup \{\infty, 0\}$ is given by

$$\nu_Z(f) = \begin{cases} \infty, & \text{if } f = 0 \\ -m, & \text{if } f \notin \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}((m-1)Z)) \text{ but } f \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(mZ)). \end{cases}$$

If $f \neq 0$ and $\nu_Z(f) = -m$, we say that f has a *pole of order* m at Z. For any A-point W on $\mathbf{X}$ disjoint from Z, define the *valuation* $\nu_{Z,W} : \mathcal{M}_Z \to$

$\mathbb{N} \cup \{\infty, 0\}$ by

$$\nu_{Z,W}(f) = \begin{cases} \infty, & \text{if } f = 0 \\ l, & \text{if } f \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(mZ - lW)) \setminus \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(mZ - (l+1)W)) \end{cases}$$

where $\nu_Z(f) = -m$. If $f \neq 0$ and $\nu_Z(f) = l$, we say that f has a *zero of order* l at W.

If A is a field, then on the set $\mathcal{M}_Z$, the functions defined in Definition 9.41 are equivalent to the discrete valuations defined in Definition I.1.11 of [40].

Remark 9.42. Let Z be an A-point of $\mathbf{X}$, let $f \in \mathcal{M}_Z$ and let W be an A-point on $\mathbf{X}$ disjoint from Z. Then $f \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(jZ - iW))$ if and only if $0 \leq i \leq \nu_{Z,W}(f)$ and $j \geq -\nu_Z(f)$.

The functions ν_Z and $\nu_{Z,W}$ have properties similar to those found in the field case. A proof of the following lemma can be found in [3].

Lemma 9.43. *Let Z be an A-point of $\mathbf{X}$ and let W be an A-point of $\mathbf{X}$ disjoint from Z. Let $r, s \in \mathcal{M}_Z$ and let $a \in A \setminus \{0\}$. Then*

(1) $\nu_Z(a) = 0$ *and* $\nu_{Z,W}(a) = 0$.
(2) $\nu_Z(r+s) \geq \min\{\nu_Z(r), \nu_Z(s)\}$.
(3) $\nu_Z(rs) \geq \nu_Z(r) + \nu_Z(s)$.
(4) $\nu_{Z,W}(r+s) \geq \min\{\nu_{Z,W}(r), \nu_{Z,W}(s)\}$.
(5) $\nu_{Z,W}(rs) \geq \nu_{Z,W}(r) + \nu_{Z,W}(s)$.

The proof of the following proposition follows the spirit of the proof of Theorem 5.4 of [50] and is therefore omitted.

Proposition 9.44. *Let Z be an A-point of $\mathbf{X}$ and let $\{Z_1, \ldots Z_n\}$ be a set of pairwise disjoint A-points of $\mathbf{X}$, all disjoint from Z. Let $h \in \mathcal{M}_Z$. If $\nu_Z(h) \geq -m$ and $\nu_{Z,Z_i}(h) \geq r_i$ for $1 \leq i \leq n$, for some nonnegative integers m, r_1,..., r_n, then*

$$h \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(mZ - r_1 Z_1 - \cdots - r_n Z_n)).$$

For each A-point W disjoint from Z and each nonnegative integer j, we have an evaluation map

$$\gamma_{W,j} : \Gamma(W, \mathcal{O}_{\mathbf{X}}(jZ)|_W) \to A.$$

By the explicit formulation of the evaluation maps given in Section 9.2, the maps $\gamma_{W,j}$ are compatible, i.e. if $h \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(iZ)) \subset \mathcal{M}_Z$, then

$$\gamma_{W,j}(h) = \gamma_{W,i}(h)$$

for all $j \geq i$. For $h \in \mathcal{M}_Z$, we abuse notation and write $h(W)$ to mean $\gamma_{W,j}(h|_W)$ for any nonnegative integer j such that $h \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(jZ))$.

Let $\vec{y} = \vec{c} + \vec{e}$ be a received word, where $\vec{c} \in C_L(\mathbf{X}, \mathcal{Z}, \mathcal{O}_{\mathbf{X}}(mZ), \gamma)$, and let e be a positive integer. To list-decode $\vec{y}$ with error bound e, we want to find all of the functions $h \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(mZ))$ such that $h(Z_i) = y_i$ for at least $n - e$ points of $\mathcal{Z}$. As in the case of algebraic geometric codes over fields [14], the problem of decoding $C_L(\mathbf{X}, \mathcal{Z}, \mathcal{O}_{\mathbf{X}}(mZ), \gamma)$ can be reduced to a reconstruction problem. In this context, the reconstruction problem will be solved by constructing a nonzero polynomial $Q(y) \in \mathcal{M}_Z[y] \subset \mathcal{K}(\mathbf{X})[y]$ such that $Q(h) \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$ for all $h \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(mZ))$, and such that each ordered pair (Z_i, y_i) is a zero of multiplicity at least r (a notion to be defined shortly) of $Q(y)$, where l and r are parameters of the algorithm. In order to find the desired polynomial $Q(y)$, generating sets with specific properties are needed for $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$. The next proposition follows immediately from Nakayama's Lemma [31].

Proposition 9.45 (See page 413 of [19]). *Let R be a commutative ring, M a finitely generated R-module and J the Jacobson radical of R. Let $x_1, \ldots, x_n \in M$. Then $x_1, \ldots, x_n$ generate M if and only if $\overline{x}_1, \ldots, \overline{x}_n$ generate M/JM, where $\overline{x}_i$ is the image in M/JM of x_i.*

Since A is a local Artinian ring, the Jacobson radical of A is equal to $\mathfrak{m}$. Let $l > 2g - 2$ be an integer. We know that $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$ is a free A-module of rank $l + 1 - g$, that $\phi^*(\mathcal{O}_{\mathbf{X}}(lZ)) = \mathcal{O}_X(lP)$ where $\phi : X \to \mathbf{X}$ is the canonical embedding and P is the unique closed point contained in Z, and that $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ)) \otimes_A \mathbb{F}_q \simeq \Gamma(X, \mathcal{O}_X(lP))$. Hence, to find a generating set for $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$ it is sufficient, by Proposition 9.45, to find a generating set for $\Gamma(X, \mathcal{O}_X(lP))$.

In order to define what it means for an ordered pair to be a *zero of multiplicity r* of $Q(y)$, a "shifted" generating set is also needed.

Remark 9.46. For an $\mathbb{F}_q$-rational point P and an integer $l > 2g - 2$ there exist functions $\bar{f}_1, \ldots, \bar{f}_{l+1-g} \in \Gamma(X, \mathcal{O}_X(lP))$ and integers $0 \leq n_1 < n_2 < \ldots < n_{l+1-g} \leq l$ such that $\bar{f}_i$ has a pole of order n_i at P. Thus, for $1 \leq i \leq l + 1 - g$, we have $\nu_P(\bar{f}_i) = -n_i$ and $f_i \in \Gamma(X, \mathcal{O}_X((i + g - 1)P))$.

The next lemma is a restatement of Lemma 21 of [14], using the above remark. It is used to find the "shifted" generating set in Corollary 9.48 below.

Lemma 9.47. *Let P be an $\mathbb{F}_q$-rational point of X and fix an integer $l > 2g-2$. Set $p = \dim\Gamma(X, \mathcal{O}_X(lP)) = l+1-g$. Then for any $\mathbb{F}_q$-rational point $R \neq P$ of X, there is a basis $\{\bar{\psi}_{R,1}, \ldots, \bar{\psi}_{R,p}\}$ of $\Gamma(X, \mathcal{O}_X(lP))$ such that $\nu_R(\bar{\psi}_{R,j}) \geq j-1$ for $1 \leq j \leq p$.*

Corollary 9.48 (Compare to Lemma 21 of [14]). *Let Z be an A-point of $\mathbf{X}$, let $l > 2g-2$ be an integer, and set $p = l+1-g$. Then*

(1) There is a generating set $\{f_1, \ldots, f_p\}$ for $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$ such that, for $1 \leq i \leq p$, $f_i \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}((i+g-1)Z))$.
(2) For any A-point W of $\mathbf{X}$ disjoint from Z, there is a generating set $\{\psi_{W,1}, \ldots, \psi_{W,p}\}$ for $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$ such that $\nu_{Z,W}(\psi_{W,j}) \geq j-1$ for $1 \leq j \leq p$.

Proof. Let P and R be the closed points contained in Z and W respectively. Let $\{\bar{f}_1, \ldots, \bar{f}_p\}$ be as in Remark 9.46.

(1) Since

$$\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}((i+g-1)Z)) \otimes_A \mathbb{F}_q \simeq \Gamma(X, \mathcal{O}_X((i+g-1)P)),$$

for each i with $1 \leq i \leq p$, there exists $f_i \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}((i+g-1)Z))$ such that $f_i \pmod{\mathfrak{m}} = \bar{f}_i$.

(2) By Lemma 9.47, there exists a basis $\{\bar{\psi}_{R,1}, \ldots \bar{\psi}_{R,p}\}$ of $\Gamma(X, \mathcal{O}_X(lP))$ such that $\nu_R(\psi_{R,j}) \geq j-1$ for all j. Hence, $\bar{\psi}_{R,j} \in \Gamma(X, \mathcal{O}_X(lP - (j-1)R)$ for $1 \leq j \leq p$. Choose $\psi_{W,j} \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ - (j-1)W))$ such that $\psi_{W,j}$ reduces modulo $\mathfrak{m}$ to $\bar{\psi}_{R,j}$.

□

Remark 9.49. Let $\{f_1, \ldots, f_p\} \subset \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$ be a generating set for $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$, as given by Corollary 9.48. Then for any choice of $q_{k,j} \in A$, the polynomial

$$Q(y) = \sum_{k=0}^{\lfloor \frac{l-g}{m} \rfloor} \sum_{j=1}^{p-mk} q_{k,j} f_j y^k \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))[y]$$

satisfies $Q(h) \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$ for all $h \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(mZ))$.

Our goal is to find coefficients $q_{k,j}$ such that each ordered pair (Z_i, y_i) is a zero of multiplicity of at least r of $Q(y) = \sum_{k=0}^{\lfloor \frac{l-g}{m} \rfloor} \sum_{j=1}^{p-mk} q_{k,j} f_j y^k$, where r is a parameter of the algorithm; see (9.5) below.

Let W be an A-point of $\mathbf{X}$ disjoint from Z. Let $\{\psi_{W,1}, \ldots, \psi_{W,p}\} \subset \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$ be a generating set for $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$, as described in Corollary 9.48, so that $\nu_{Z,W}(\psi_{W,j}) \geq j-1$ for $1 \leq j \leq p$. Then, for each i with $1 \leq i \leq p$, there are scalars $\Lambda_{W,i,1}, \ldots, \Lambda_{W,i,p} \subset A$ such that

$$f_i = \Lambda_{W,i,1}\psi_{W,1} + \cdots + \Lambda_{W,i,p}\psi_{W,p}.$$

Then

$$Q(y) = \sum_{k_1=0}^{\lfloor \frac{l-g}{m} \rfloor} \sum_{j_2=1}^{p} \sum_{j_1=1}^{p-mk_1} q_{k_1,j_1} \Lambda_{W,j_1,j_2} \psi_{W,j_2} y^{k_1}. \tag{9.4}$$

For any $a \in A$,

$$Q(y+a) = \sum_{k_2=0}^{\lfloor \frac{l-g}{m} \rfloor} \sum_{j_2=1}^{p} q^{(a)}_{k_2,j_2} \psi_{W,j_2} y^{k_2}$$

with coefficients $q^{(a)}_{k_2,j_2}$ given by

$$q^{(a)}_{k_2,j_2} = \sum_{k_1=k_2}^{\lfloor \frac{l-g}{m} \rfloor} \sum_{j_1=1}^{p-mk_1} \binom{k_1}{k_2} a^{k_1-k_2} q_{k_1,j_1} \Lambda_{W,j_1,j_2}.$$

Definition 9.50 (See Section IV of [14]). Let W be an A-point on $\mathbf{X}$ disjoint from Z, let $a \in A$ and let $Q(y)$ be as in Equation 9.4. If $q^{(a)}_{k_2,j_2} = 0$ for all $j_2 + k_2 \leq r$, where $j_2 \geq 1$, $k_2 \geq 0$, and $q_{k_2,j_2} \neq 0$ for some k_2, j_2 such that $k_2 + j_2 = r+1$, then (W, a) is a *zero of multiplicity* r of $Q(y) \in \mathcal{K}(\mathbf{X})[y]$.

We are now ready to describe the algorithm for solving the reconstruction problem. We assume that the algorithm has access to the following information:

(1) The generating set $\{f_1, \ldots, f_p\}$ for $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$ such that, for j with $1 \leq j \leq p$, $f_j \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}((j+g-1)Z))$ (as described in Corollary 9.48).
(2) For each $Z_i \in \mathcal{Z}$, the generating set $\{\psi_{Z_i,1}, \ldots, \psi_{Z_i,p}\}$ for $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$ (as described in Corollary 9.48), and, for $1 \leq j \leq p$, the scalars $\{\Lambda_{Z_i,j,1}, \ldots, \Lambda_{Z_i,j,p}\}$ such that $f_j = \Lambda_{Z_i,j,1}\psi_{Z_i,1} + \cdots + \Lambda_{Z_i,j,p}\psi_{Z_i,p}$.

The input for the algorithm is a set of ordered pairs $\{(Z_1, y_1), \ldots, (Z_n, y_n)\}$, where $Z, Z_1, \ldots Z_n$ are pairwise disjoint A-points of X and $\{y_1, \ldots, y_n\} \subset$

A, and positive integers m and t with $2g-2 < m < n$ and $t < n$. The algorithm is as follows:

Step 0. Set

$$r = 1 + \left\lfloor \frac{2gt + mn + \sqrt{(2gt+mn)^2 - 4(g^2-1)(t^2-mn)}}{2(t^2-mn)} \right\rfloor \tag{9.5}$$

and

$$l = rt - 1. \tag{9.6}$$

Step 1. Find a polynomial

$$Q(y) = \sum_{k_1=0}^{\lfloor \frac{l-g}{m} \rfloor} \sum_{j_1=1}^{p-mk_1} q_{k_1,j_1} f_{j_1} y^{k_1} \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))[y]$$

such that $Q(y) \not\equiv 0 \pmod{\mathfrak{m}}$ and $Q(y)$ has a zero of multiplicity at least r at each ordered pair (Z_i, y_i) for $1 \le i \le n$. (Lemma 9.51 below shows when this is possible.)

Step 2. Find all functions $h \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(mZ))$ such that $y - h$ is a factor of $Q(y)$. For each such h, check to see if $h(Z_i) = y_i$ for at least t values of i with $1 \le i \le n$. If so, output h.

We now show that the algorithm solves the polynomial reconstruction problem. We begin by showing that, under certain conditions, the polynomial $Q(y)$ sought in Step 1 of the algorithm exists.

Lemma 9.51 (Compare to Lemma 25 of [14]).
If $n\binom{r+1}{2} < \frac{(l-g)(l-g+2)}{2m}$, *then the polynomial* $Q(y)$ *sought in Step 1 of the algorithm exists.*

Proof. As in Lemma 25 of [14] we are solving a system of homogeneous linear equations. It is shown in Chapter 4 of [3] that there exists a solution such that $Q(y) \not\equiv 0 \pmod{\mathfrak{m}}$ if the number of unknowns is larger than the number of constraints. In the proof of Lemma 25 of [14], it is shown that there are $n\binom{r+1}{2}$ constraints and at least $\frac{(l-g)(l-g+2)}{2m}$ unknown coefficients. □

Remark 9.52. The proof of Lemma 5 in [2], which uses the McCoy rank of a matrix over A, can be used to show that a non-zero polynomial $Q(y)$ exists, but it is unclear whether this method guarantees that $Q(y) \not\equiv 0 \pmod{\mathfrak{m}}$.

Now that we have proved that $Q(y)$ exists, we will show that, if $h \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(mZ))$ and $h(Z_i) = y_i$ for at least t values of y, then $y - h$ is a factor of $Q(y)$. The following lemma, which follows from the properties of valuations given in Lemma 9.43, is needed. Its proof parallels that of Lemma 23 of [14].

Lemma 9.53 (Compare to Lemma 23 of [14]). *Let $Q(y)$ be the polynomial found in Step 1. For $1 \leq i \leq n$, if $h \in \mathcal{M}_Z$ satisfies $h(Z_i) = y_i$, then $\nu_{Z,Z_i}(Q(h)) \geq r$.*

Lemma 9.54 (Compare to Lemma 24 of [14]).
Let $h \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(mZ))$ and let $Q(y)$ be the polynomial found in Step 1. If $rt > l$ and $h(Z_i) = 0$ for at least t values of i with $1 \leq i \leq n$, then $y - h$ is a factor of $Q(y)$.

Proof. By reindexing if necessary, we may assume that $h(Z_1) = y_1, \ldots, h(Z_t) = y_t$. By Remark 9.49, $Q(h) \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$. By Lemma 9.53, we have $\nu_{Z,Z_i}(Q(h)) \geq r$ for $1 \leq i \leq t$. Since $Z, Z_1, \ldots, Z_t$ are pairwise disjoint,

$$Q(h) \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ - rZ_1 - \cdots - rZ_t))$$

by Proposition 9.44. We know

$$\deg \mathcal{O}_X(lZ - rZ_1 - \cdots - rZ_t) = l - rt,$$

and, since $l < rt$, we have $l - rt < 0$. Hence

$$\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ - rZ_1 - \cdots - rZ_t)) = \{0\},$$

and so $Q(h) = 0$. It follows that $y - h$ is a factor of $Q(y)$. □

The proof of the next lemma is omitted, as it can be found in [14].

Lemma 9.55 (Lemma 26, [14]). *If n, m and t satisfy $t^2 > mn$, then for the choice of r and l made in the algorithm, $\frac{(l-g)(l-g+2)}{2m} > n\binom{r+1}{2}$ and $rt > l$ both hold.*

Finally, putting everything together, we have the following theorem.

Theorem 9.56 (Compare to Theorem 27 of [14]).
The algorithm given above solves the polynomial reconstruction problem for one-point algebraic geometric codes over rings with inputs m, t and points $\{(Z_i, y_i)\}_{i=1}^n$, provided that $t > \sqrt{mn}$.

Proof. If $t > \sqrt{mn}$, then, by Lemma 9.55, $rt > l$ and $\frac{(l-g)(l-g+2)}{2m} > n\binom{r+1}{2}$. Hence, by Lemma 9.51, the polynomial $Q(y)$ sought in Step 1 of the algorithm exists. By Lemma 9.54, if $h \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(mZ))$ satisfies $h(Z_i) = y_i$ for at least t of the points $\{(Z_1, y_1), \ldots (Z_n, y_n)\}$, then $y - h$ is a factor of $Q(y)$. □

Remark 9.57. For generalized Reed-Solomon codes over a Galois ring, Armand [1] proposed a two-stage decoder for C that improves upon the performance of the Guruswami-Sudan algorithm. Armand noted that his two-stage decoding approach could be applied to a one-point algebraic code $C = C_L(\mathbf{X}, \mathcal{Z}, \mathcal{O}_{\mathbf{X}}(mZ), \gamma)$ if

(1) $\overline{C}$ is an algebraic geometric code over $\mathbb{F}_q$ with the same minimum distance as C and
(2) an errors-and-erasure decoding algorithm exists for C.

By Theorem 9.18, $\overline{C}$ is a one-point algebraic geometric code over $\mathbb{F}_q$. Furthermore, since C is a free code, the minimum distances of C and $\overline{C}$ are equal by Theorem 3.4 of [50]. Thus, the first condition is satisfied. Since we can still use the Guruswami-Sudan algorithm when erasures occur [14], there exists an errors-and-erasure decoding algorithm for C.

9.4.3. *The Koetter-Vardy Algorithm for Decoding with Other Metrics*

As discussed above, the Hamming distance is often not the weight measure of interest when studying codes over rings. In this section, we follow the work of Koetter and Vardy in [23] and [25] to show how the Guruswami-Sudan algorithm may be used to decode an algebraic geometric code over a local Artinian ring with respect to any given weight measure on the ring.

Let A be a local Artinian ring with principal maximal ideal $\mathfrak{m}$ and finite residue field $\mathbb{F}_q$. Let $\mathbf{X}$ be a curve over A, and let $X = \mathbf{X} \times_{\operatorname{Spec} A} \operatorname{Spec} \mathbb{F}_q$ be the fibre of $\mathbf{X}$ over $\mathfrak{m}$. As before, assume that X is absolutely irreducible, and let $C = C_L(\mathbf{X}, \mathcal{Z}, \mathcal{O}_{\mathbf{X}}(mZ), \gamma)$ be a one-point algebraic geometric code over A, where $\mathcal{Z} = \{Z_1, \ldots, Z_n\}$ is a set of pairwise disjoint A-points of $\mathbf{X}$, Z is an A-point of $\mathbf{X}$ disjoint from all the points of $\mathcal{Z}$, m is a nonnegative integer such that $2g - 2 < m < n$ and $\gamma = \{\gamma_i : \Gamma(Z_i, \mathcal{O}_{\mathbf{X}}(mZ)|_{Z_i}) \to A\}$ is the system of evaluation maps. For $1 \leq i \leq n$, let P_i be the closed point contained in Z_i, and let $\mathcal{P} = \{P_1, \ldots, P_n\}$. Let P be the closed point contained in Z and fix an ordering on the elements of A, say $a_1, \ldots, a_s$,

where $s = |A|$.

Although the Koetter-Vardy algorithm was originally developed for soft-decision decoding of Reed-Solomon codes over finite fields, Koetter and Vardy also showed that their algorithm could be extended to one-point algebraic geometric codes over finite fields [22]. Due to length constraints, these results were not included in [25]. Many of the results presented in this section are the analogs in the ring case of results of Koetter and Vardy for algebraic geometric codes over finite fields. We begin with the following definitions.

Definition 9.58 (Compare to [23]). Let $R : A \times A \to \mathbb{R}$ be a function such that $R(a, b) \geq 0$ for all a, $b \in A$. Let $\vec{x}$, $\vec{y} \in A^n$. The *cost of* $\vec{x}$ *given* $\vec{y}$, $d_R(\vec{x}, \vec{y})$, is defined by

$$d_R(\vec{x}, \vec{y}) = R(x_1, y_1) + \cdots + R(x_n, y_n).$$

Let e be a nonnegative integer. The set $\mathcal{A}(\vec{y}, e)$ is defined by

$$\mathcal{A}(\vec{y}, e) = \{\vec{x} \in A^n \mid d_R(\vec{x}, \vec{y}) \leq e\}.$$

Definition 9.59 (Compare to Definition 3 of [25]). A *multiplicity matrix* $M = (m_{i,j})$ over the ring A with $|A| = s$ is an $s \times n$ matrix of nonnegative integers. The rows of M are indexed by the elements of A and the columns of M are indexed by the elements of $\mathcal{Z}$. The *cost* $\mathcal{C}(M)$ of M is defined by

$$\mathcal{C}(M) = \frac{1}{2} \sum_{i=1}^{s} \sum_{j=1}^{n} m_{i,j}(m_{i,j} + 1).$$

For a received word $\vec{y}$, the first step of the Koetter-Vardy algorithm is to compute a multiplicity matrix $M_{\vec{y}}$ for $\vec{y}$ and the given cost function d_R. More information about computing $M_{\vec{y}}$ will be given at the end of this section. In the meantime, we will let M denote a given multiplicity matrix with the understanding that when implementing the Koetter-Vardy algorithm, the multiplicity matrix M used must first be computed and depends on the received word $\vec{y}$.

Recall that

$$\mathcal{M}_Z = \bigcup_{j=0}^{\infty} \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(jZ)) \subset \mathcal{K}(\mathbf{X})$$

is the A-module of functions on $\mathbf{X}$ with poles only at the A-point Z. Given a multiplicity matrix M, we will construct a polynomial $Q_M(y) \in \mathcal{M}_Z[y]$

such that $Q_M(y)$ has a zero of multiplicity at least $m_{j,i}$ at each ordered pair $(Z_i, a_j) \in \mathcal{Z} \times A$. The following definitions are useful.

Definition 9.60 (See [22], [25]). Let w_Z and w_y be nonnegative real numbers. For any integer l, define $N_{w_Z,w_y}(l)$ by

$$N_{w_Z,w_y}(l) = |\{j,k \,|\, j \geq 1, k \geq 0, (j+g-1)w_Z + kw_y \leq l\}|$$

and, for any integer δ, define $\Delta_{w_Z,w_y}(\delta)$ by

$$\Delta_{w_Z,w_y}(\delta) = \min\left\{l \in \mathbb{Z} \,|\, N_{w_Z,w_y}(l) > \delta\right\}.$$

Note that if $\{f_1, \ldots, f_{l+1-g}\}$ is a generating set for $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$, as described in Corollary 9.48, then $N_{1,m}$ is the number of (unknown) coefficients $q_{k,j}$ in the polynomial

$$Q(y) = \sum_{k=0}^{\lfloor \frac{l-g}{m} \rfloor} \sum_{j=1}^{l-g+1-mk} q_{k,j} f_j y^k \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))[y],$$

i.e.,

$$N_{1,m}(l) = \sum_{k=0}^{\lfloor \frac{l-g}{m} \rfloor} \sum_{j=1}^{l-g+1-mk} 1 = \left(\left\lfloor \frac{l-g}{m} \right\rfloor + 1\right)(l-g+1-mk).$$

The following bounds on $N_{1,m}(l)$ and $\Delta_{1,m}(\delta)$ are derived in [22]. A proof can be found in [3].

Lemma 9.61 (Koetter and Vardy, [22]; see also [3]). *Let l, $\delta \in \mathbb{Z}$. Then*

$$N_{1,m}(l) > \frac{l^2}{2m} - g\left(\frac{l}{m} + 1\right)$$

and

$$\Delta_{1,m}(\delta) \leq g + 1 + \sqrt{2m(\delta+g) + g^2}.$$

Let M be a multiplicity matrix and let $l = \Delta_{1,m}(\mathcal{C}(M))$. Note that $l > 2g-2$. Let $\{f_1, \ldots, f_{l-g+1}\}$ be a generating set for $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$, as described by Corollary 9.48. As before, our goal is to find coefficients $q_{k,j} \in A$ such that the polynomial

$$Q_M(y) = \sum_{k=0}^{\lfloor \frac{l-g}{m} \rfloor} \sum_{j=1}^{l-g+1-mk} q_{k,j} f_j y^k \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))[y]$$

has a zero of multiplicity at least $m_{j,i}$ at the ordered pair (Z_i, a_j) for $1 \leq i \leq n$, $1 \leq j \leq s$.

For each i, $1 \leq i \leq n$, let $\{\psi_{Z_i,1}, \ldots, \psi_{Z_i,l-g+1}\} \subset \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$ be a generating set of $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))$, as described in Corollary 9.48. Then for each i, $1 \leq i \leq n$, and for each j_1, $1 \leq j_1 \leq l-g+1$, there are coefficients $\Lambda_{Z_i,j_1,1}, \ldots, \Lambda_{Z_i,j_1,l-g+1} \in A$ such that

$$f_{j_1} = \Lambda_{Z_i,j_1,1}\psi_{Z_i,1} + \cdots + \Lambda_{Z_i,j_1,l-g+1}\psi_{Z_i,l-g+1}.$$

Hence, we may write $Q_M(y)$ in the form

$$Q_M(y) = \sum_{k_1=0}^{\lfloor \frac{l-g}{m} \rfloor} \sum_{j_2=1}^{l-g+1} \sum_{j_1=1}^{l-g+1-mk_1} q_{k_1,j_1}\Lambda_{Z_i,j_1,j_2}\psi_{Z_i,j_2}y^{k_1}. \tag{9.7}$$

Let $(Z_i, a_j) \in \mathcal{Z} \times A$. The polynomial $Q_M(y)$ has a zero of multiplicity at least $m_{j,i}$ at (Z_i, a_j) if $q^{(a_j)}_{i,k_2,j_2} = 0$ for all $j_2 \geq 1$, $k_2 \geq 0$ such that $j_2 + k_2 \leq m_{j,i}$, where

$$q^{(a_j)}_{i,k_2,j_2} = \sum_{k_1=k_2}^{\lfloor \frac{l-g}{m} \rfloor} \sum_{j_1=1}^{l-g+1-mk_1} \binom{k_1}{k_2} a_j^{k_1-k_2} q_{k_1,j_1}\Lambda_{Z_i,j_1,j_2}.$$

Therefore, the coefficients of $Q_M(y)$ must satisfy $\frac{m_{i,j}(m_{i,j}+1)}{2}$ linear constraints. Repeating this argument for all ordered pairs of $\mathcal{Z} \times A$, it follows that the coefficients of $Q_M(y)$ must satisfy $\mathcal{C}(M)$ linear constraints.

Lemma 9.62 (See also [22], [25]). *Let M be a multiplicity matrix and let $l = \Delta_{1,m}(\mathcal{C}(M))$. Then there exists a polynomial $Q_M(y) \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(lZ))[y]$ such that*

(1) $Q_M(y) = \sum_{k_1=0}^{\lfloor \frac{l-g}{m} \rfloor} \sum_{j_1=1}^{l-g+1-mk_1} q_{k_1,j_1}f_{j_1}y^{k_1}$ for some $q_{k_1,j_1} \in A$.
(2) $Q_M(y)$ has zero of multiplicity at least $m_{j,i}$ at each ordered pair (Z_i, a_j) for $1 \leq i \leq n$, $1 \leq j \leq s$.
(3) $Q_M(y) \not\equiv 0 \pmod{\mathfrak{m}}$.

Proof. As in the Guruswami-Sudan algorithm, we are solving a system of linear equations. As before, there exists a solution such that $Q_M(y) \not\equiv 0 \pmod{\mathfrak{m}}$ if the number of unknowns is larger than the number of constraints. There are $\mathcal{C}(M)$ constraints and $N_{1,m}(l)$ unknown coefficients. Since $l = \Delta_{1,m}(\mathcal{C}(M))$, it follows from the definition of $\Delta_{1,m}(\mathcal{C}(M))$ that $Q_M(y)$ exists. □

Given a multiplicity matrix M, we shall call a polynomial $Q_M(y)$ that satisfies the conditions described in Lemma 9.62 a *polynomial associated to* M. As in the Guruswami-Sudan algorithm, the next step is to factor $Q_M(y)$. The following definitions are needed in order to describe the roots of $Q_M(y)$ contained in $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(mZ))$.

Definition 9.63 (Compare to [25]). Let $\vec{v} \in A^n$. Define the $s \times n$ matrix $[\vec{v}] = (v_{i,j})$ by

$$v_{i,j} = \begin{cases} 1, & \text{if } a_i = v_j \\ 0, & \text{otherwise.} \end{cases}$$

Let $B = (b_{i,j})$ and $D = (d_{i,j})$ be $s \times n$ matrices over A. Define the inner product $\langle B, D \rangle$ by

$$\langle B, D \rangle = \text{trace}(BD^T) = \sum_{i=1}^{s} \sum_{j=1}^{n} b_{i,j} d_{i,j}.$$

Definition 9.64 (Compare to Definition 4 of [25]). Let $\vec{v} \in A^n$ and let M be a multiplicity matrix. The *score* $\mathcal{S}_M(\vec{v})$ of $\vec{v}$ with respect to M is given by

$$\mathcal{S}_M(\vec{v}) = \langle M, [\vec{v}] \rangle.$$

Using the above definitions, we can describe the roots of $Q_M(y)$ in $\Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(mZ))$. The proof of the next theorem is similar to the proof of Lemma 9.54 and is therefore omitted.

Theorem 9.65 (Compare to Theorem 3 of [25]; see also [22]). *Let M be a multiplicity matrix and let $Q_M(y)$ be a polynomial associated to M. Let $h \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(mZ))$. Then $y - h$ is a factor of $Q_M(y)$ if*

$$\mathcal{S}_M(\vec{h}) > \Delta_{1,m}(\mathcal{C}(M)),$$

where $\vec{h} = (h(Z_1), \ldots, h(Z_n))$.

Let $\vec{y}$ be a received word and let e be an error bound. Recall $\mathcal{A}(\vec{y}, e)$ is the set of codewords $\vec{c} \in C_L(\mathbf{X}, \mathcal{Z}, \mathcal{O}_{\mathbf{X}}(mZ), \gamma)$ such that $d_R(\vec{c}, \vec{y}) \leq e$. We decode $\vec{y}$ as follows.

Step 1. Compute a multiplicity matrix $M = M_{\vec{y}}$ such that $\mathcal{S}_M(\vec{c}) > \Delta_{1,m}(\mathcal{C}(M_{\vec{y}}))$ for all $\vec{c} \in \mathcal{A}(\vec{y}, e)$. (See Remark 9.66 below.)

Step 2. Given $M = M_{\vec{y}}$ from Step 1, find a polynomial

$$Q_M(y) = \sum_{k_1=0}^{\lfloor \frac{l-g}{m} \rfloor} \sum_{j_1=1}^{l-g+1-mk_1} q_{k_1,j_1} f_{j_1} y^{k_1}$$

for some $q_{k_1,j_1} \in A$ such that $Q_M(y)$ has zero of multiplicity at least $m_{j,i}$ at each ordered pair (Z_i, a_j) for $1 \leq i \leq n$, $1 \leq j \leq s$.

Step 3. Find all functions $h \in \Gamma(\mathbf{X}, \mathcal{O}_{\mathbf{X}}(mZ))$ such that $y - h$ is a factor of $Q(y)$. For each such h, check to see if $d_R(\vec{y}, \vec{h}) \leq e$, where $\vec{h} = (h(Z_1), \ldots, h(Z_n))$ as in Theorem 9.65. If $d_R(\vec{y}, \vec{h}) \leq e$, then output h.

Remark 9.66. In Section IV of [24] it is shown how to find $M_{\vec{y}}$ in the case of Reed Solomon codes over fields. Theorem 7 of [24] shows when the computation of $M_{\vec{y}}$ is possible. As this computation does not change significantly for $C_L(\mathbf{X}, \mathcal{Z}, \mathcal{O}_{\mathbf{X}}(mZ), \gamma)$, we omit the details of this step. Information about finding $M_{\vec{y}}$ can be found in Section IV of [24] and in [23]. An example for the case of the Lee weight can be found in [3].

9.5. Conclusion

In this chapter we have introduced and explored algebraic geometric codes over local Artinian rings. In particular, we have shown that algebraic codes over rings have dimension and minimum distance properties similar to those of algebraic geometric codes over fields, and that, as in the field case, the codes are closed under duals. In the second half of the chapter, three decoding algorithms were presented. The basic algorithm and Gurusawmi-Sudan algorithm both decode with respect to the Hamming distance, while the Koetter-Vardy algorithm can be used to decode with respect to other weight measures.

There is still interesting work to be done in the area of algebraic geometric codes over rings. For example, in the basic algorithm the set $K(\vec{y}, F, \delta)$ is computed by exhaustion. For algebraic geometric codes over finite fields, this set can by found by solving a linear system of equations [39]. One may ask if the same is true when working over a local Artinian ring. In the case of the Guruswami-Sudan and the Koetter-Vardy algorithms, we assume that we are able to factor a polynomial $Q(y) \in \mathcal{M}_Z[y]$. As we are no longer factoring $Q(y)$ over a unique factorization domain, it is likely that this step is harder to perform than it is in the field case.

As new propertites and algorithms are discovered or developed for algebraic geometric codes over fields, it is hoped that these concepts and algorithms hold in the ring case. For example, Drake and Matthews [32], [7] have shown how the Guruswami-Sudan algorithm can be used to decode so-called *multi-point codes*, i.e., codes where the divisor in question is of the form $m_1Q_1 + \cdots + m_tQ_t$ rather than simply mQ, over finite fields. It seems plausible that their methods will extend to the case of multi-point codes over local Artinian rings as well, though the details have not yet been worked out.

References

[1] M. A. Armand. Improved list decoding of generalized Reed-Solomon and alternant codes over Galois rings. *IEEE Trans. Inform. Theory*, 51(2):728–733, 2005.

[2] M. A. Armand. List decoding of generalized Reed-Solomon codes over commutative rings. *IEEE Trans. Inform. Theory*, 51(1):411–419, 2005.

[3] K. G. Bartley. *Decoding algorithms for algebraic geometric codes over rings.* PhD thesis, University of Nebraska, 2006.

[4] R. Blache. Majorations de sommes exponentielles sur les anneaux de Galois. *C. R. Acad. Sci. Paris Sér. I Math.*, 332(5):427–430, 2001.

[5] R. Blache. Exponential sums over lifts of points. *J. Number Theory*, 105(2):361–386, 2004.

[6] I. F. Blake. Codes over certain rings. *Information and Control*, 20:396–404, 1972.

[7] N. Drake and G. L. Matthews. Decoding general AG codes up to the minimum distance using lists. Preprint.

[8] D. Eisenbud. *Commutative algebra with a view toward algebraic geometry*, volume 150 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995.

[9] L. R. A. Finotti. Degrees of the elliptic Teichmüller lift. *J. Number Theory*, 95(2):123–141, 2002.

[10] L. R. A. Finotti. Minimal degree liftings of hyperelliptic curves. *J. Math. Sci. Univ. Tokyo*, 11(1):1–47, 2004.

[11] L. R. A. Finotti. Minimal degree liftings in characteristic 2. *J. Pure Appl. Algebra*, 207(3):631–673, 2006.

[12] E. N. Gilbert. A comparison of signalling alphabets. *Bell Syst. Tech. J.*, 31, 1952.

[13] V. D. Goppa. Codes that are associated with divisors. *Problemy Peredači Informacii*, 13(1):33–39, 1977.

[14] V. Guruswami and M. Sudan. Improved decoding of Reed-Solomon and algebraic-geometry codes. *IEEE Trans. Inform. Theory*, 45(6):1757–1767, 1999.

[15] A. R. Hammons, Jr., P. V. Kumar, A. R. Calderbank, N. J. A. Sloane,

and P. Solé. The $\mathbf{Z}_4$-linearity of Kerdock, Preparata, Goethals, and related codes. *IEEE Trans. Inform. Theory*, 40(2):301–319, 1994.

[16] R. Hartshorne. *Residues and duality*. Lecture notes of a seminar on the work of A. Grothendieck, given at Harvard 1963/64. With an appendix by P. Deligne. Lecture Notes in Mathematics, No. 20. Springer-Verlag, Berlin, 1966.

[17] R. Hartshorne. *Algebraic geometry*. Springer-Verlag, New York, 1977. Graduate Texts in Mathematics, No. 52.

[18] T. Høholdt and R. Pellikaan. On the decoding of algebraic-geometric codes. *IEEE Trans. Inform. Theory*, 41(6, part 1):1589–1614, 1995. Special issue on algebraic geometry codes.

[19] N. Jacobson. *Basic algebra. II*. W. H. Freeman and Co., San Francisco, Calif., 1980.

[20] J. Justesen, K. J. Larsen, H. E. Jensen, A. Havemose, and T. Høholdt. Construction and decoding of a class of algebraic geometry codes. *IEEE Trans. Inform. Theory*, 35(4):811–821, 1989.

[21] A. M. Kerdock. A class of low-rate nonlinear binary codes. *Inform. Control*, 20:182–187, 1972.

[22] R. Koetter. Private communication. Extended unpublished preprint of [25], 2003.

[23] R. Koetter and A. Vardy. Decoding of Reed Solomon codes for additive cost functions. In *ISIT '02 (Lausanne, Switzerland)*, page 313, 2002.

[24] R. Koetter and A. Vardy. Soft decoding of Reed Solomon codes and optimal weight assignments. *ITG Fachtagung*, January 2002.

[25] R. Koetter and A. Vardy. Algebraic soft-decision decoding of Reed-Solomon codes. *IEEE Trans. Inform. Theory*, 49(11):2809–2825, 2003.

[26] P. V. Kumar, T. Helleseth, and A. R. Calderbank. An upper bound for Weil exponential sums over Galois rings and applications. *IEEE Trans. Inform. Theory*, 41(2):456–468, 1995.

[27] S. Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2002.

[28] W.-C. W. Li. Character sums over p-adic fields. *J. Number Theory*, 74(2):181–229, 1999.

[29] J. Lubin, J.-P. Serre, and J. Tate. Elliptic curves and formal groups. Proc. of the Woods Hole Summer Institute in Algebraic Geometry, 1964.

[30] F. J. MacWilliams. A theorem on the distribution of weights in a systematic code. *Bell Syst. Tech. J.*, 42:79–94, 1963.

[31] H. Matsumura. *Commutative ring theory*, volume 8 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 1989. Translated from the Japanese by M. Reid.

[32] G. L. Matthews. Viewing multipoint codes as subcodes of one-point codes. Preprint.

[33] S. Mochizuki. A theory of ordinary p-adic curves. *Publ. Res. Inst. Math. Sci.*, 32(6):957–1152, 1996.

[34] A. W. Nordstrom and J. P. Robinson. An optimum nonlinear code. *Inform. Control*, 11:613–616, 1967.

[35] F. P. Preparata. A class of optimum nonlinear double-error correcting codes. *Inform. Control*, 13:378–400, 1968.

[36] C. E. Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 623–656, 1948.

[37] M. A. Shokrollahi and H. Wasserman. Decoding algebraic-geometric codes beyond the error-correction bound. In *STOC '98 (Dallas, TX)*, pages 241–248. ACM, New York, 1999.

[38] Joseph H. Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1986.

[39] A. N. Skorobogatov and S. G. Vlăduţ. On the decoding of algebraic-geometric codes. *IEEE Trans. Inform. Theory*, 36(5):1051–1060, 1990.

[40] H. Stichtenoth. *Algebraic function fields and codes*. Universitext. Springer-Verlag, Berlin, 1993.

[41] M. Sudan. Decoding of Reed Solomon codes beyond the error-correction bound. *J. Complexity*, 13(1):180–193, 1997.

[42] M. Tsfasman, S. Vlăduţ, and D. Nogin. *Algebraic geometric codes: basic notions*, volume 139 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2007.

[43] M. A. Tsfasman, S. G. Vlăduţ, and Th. Zink. Modular curves, Shimura curves, and Goppa codes, better than Varshamov-Gilbert bound. *Math. Nachr.*, 109:21–28, 1982.

[44] J. F. Voloch and J. L. Walker. Lee weights of codes from elliptic curves. In *Codes, curves, and signals (Urbana, IL, 1997)*, volume 485 of *Kluwer Internat. Ser. Engrg. Comput. Sci.*, pages 53–62. Kluwer Acad. Publ., Boston, MA, 1998.

[45] J. F. Voloch and J. L. Walker. Codes over rings from curves of higher genus. *IEEE Trans. Inform. Theory*, 45(6):1768–1776, 1999.

[46] J. F. Voloch and J. L. Walker. Euclidean weights of codes from elliptic curves over rings. *Trans. Amer. Math. Soc.*, 352(11):5063–5076 (electronic), 2000.

[47] J. F. Voloch and J. L. Walker. Homogeneous weights and exponential sums. *Finite Fields Appl.*, 9(3):310–321, 2003.

[48] J. L. Walker. *Algebraic geometric codes over rings*. PhD thesis, University of Illinois, 1996.

[49] J. L. Walker. The Nordstrom-Robinson code is algebraic-geometric. *IEEE Trans. Inform. Theory*, 43(5):1588–1593, 1997.

[50] J. L. Walker. Algebraic geometric codes over rings. *J. Pure Appl. Algebra*, 144(1):91–110, 1999.

[51] J. A. Wood. Duality for modules over finite rings and applications to coding theory. *Amer. J. Math.*, 121(3):555–575, 1999.

This page intentionally left blank

Chapter 10

Generalized Hamming Weights and Trellis Complexity

Carlos Munuera

University of Valladolid, Dept. of Applied Mathematics,
Avda Salamanca SN, 47014 Valladolid
Castilla, Spain
cmunuera@arq.uva.es

Contents

10.1. Introduction

In this chapter we shall study two interesting topics: the generalized Hamming weights and the trellis structure of an AG code. Both can be stated in the more general context of all linear codes, but they have particular properties that make the study more easy for AG codes. At the first look, both problems seem very different, but as we shall see there are connections between them.

The first part is devoted to the generalized Hamming weights. We introduce these weights and explain some of their main properties for linear codes. In section 3 we deal with AG codes, extending the Goppa bound and the order bound on the minimum distance to higher weights. In the second part we study the trellis structure of codes. The problem is stated in section 4, while the link between generalized Hamming weights and trellis complexity is sketched in section 5. Finally, the trellis complexity of AG codes is studied in section 6.

10.2. Generalized Hamming weights

Let $\mathbb{F}_q$ be the finite field with q elements. For a vector $\mathbf{x} \in \mathbb{F}_q^n$ the *support* of $\mathbf{x}$ is the set

$$\mathrm{supp}(\mathbf{x}) = \{i \mid 1 \leq i \leq n, x_i \neq 0\}.$$

We recall that for a linear code $\mathcal{C}$ of length n and dimension k over $\mathbb{F}_q$, the minimum distance (or minimum weight) of $\mathcal{C}$ is $d(\mathcal{C}) = \min\{\#\mathrm{supp}(\mathbf{x}) \mid \mathbf{x} \in \mathcal{C}, \mathbf{x} \neq \mathbf{0}\}$. We can generalize this concept as follows. For a subset $S \subseteq \mathbb{F}_q^n$, we define the support of S as the set of positions where at least one vector in S is not zero,

$$\mathrm{supp}(S) = \bigcup_{\mathbf{x} \in S} \mathrm{supp}(\mathbf{x}).$$

Now, for $1 \leq r \leq k$, we define the r-th *generalized Hamming weight* of $\mathcal{C}$ as

$$d_r(\mathcal{C}) = \min\{\#\mathrm{supp}(S) \mid S \text{ is an } r\text{-dimensional linear subcode of } \mathcal{C}\}$$

and the *weight hierarchy* of $\mathcal{C}$ as the set of its generalized Hamming weights, $\mathrm{GHW}(\mathcal{C}) = \{d_1(\mathcal{C}), \ldots, d_k(\mathcal{C})\}$. As remarked above, $d_1(\mathcal{C})$ is just the minimum distance of $\mathcal{C}$.

The next Proposition summarizes the main properties of these GHWs.

Proposition 10.1. *Let $\mathcal{C}$ be a linear $[n, k]$ code and let $\mathcal{C}^\perp$ be its dual. Then*

(1) (Monotonicity) $1 \leq d_1(\mathcal{C}) < d_2(\mathcal{C}) < \cdots < d_k(\mathcal{C}) \leq n$.
(2) (Generalized Singleton bound) $d_r(\mathcal{C}) \leq n - k + r$.
(3) (Duality) $\{d_r(\mathcal{C}^\perp) \mid 1 \leq r \leq n - k\} \cup \{n + 1 - d_r(\mathcal{C}) \mid 1 \leq r \leq k\} = \{1, \ldots, n\}$.

We leave the proof of this Proposition as an exercise to the reader (or see [2], [38]).

Given a linear $[n, k, d]$ code $\mathcal{C}$, the Singleton bound states that $d \leq n - k + 1$. Codes reaching equality are called *maximum distance separable* (MDS). Property (2) in the above Proposition extends this bound to all GHWs. This leads to the following definition: we say that $\mathcal{C}$ is *r-th rank MDS* if $d_r(\mathcal{C}) = n - k + r$.

If the computation of the minimum distance $d_1(\mathcal{C})$ is usually a difficult problem, the determination of the weight hierarchy in full, seems much more difficult. A more modest goal is to find acceptable bounds on the $d_r(\mathcal{C})$'s. In the next section we accomplish this task for the AG codes.

10.3. Generalized Hamming weights of AG codes

As the GHWs generalize the minimum distance, it seems natural to extend the available bounds on $d_1(\mathcal{C})$ to the remainder weights. In the previous chapters we have found two such bounds: the original Goppa bound and the order (or Feng-Rao) bound (the latter only for the duals of one-point codes).

Let us remember the main definitions concerning AG codes. Let $\mathcal{X}$ be a (projective, geometrically irreducible, non-singular algebraic) curve of genus g defined over the finite field $\mathbb{F}_q$. Let $\mathcal{P} = \{P_1, \ldots, P_n\}$ be a set of n rational distinct points and let us consider the rational divisors $D = P_1 + \cdots + P_n$ and G, with $\mathrm{supp}(G) \cap \mathrm{supp}(D) = \emptyset$ and $1 \leq \deg(G) \leq n + 2g - 1$. The associated code $\mathcal{C} = C(\mathcal{X}, D, G)$ is the image of the evaluation map

$$ev_{\mathcal{P}} : \mathcal{L}(G) \to \mathbb{F}_q^n \ , \ ev_{\mathcal{P}}(f) = (f(P_1), \ldots, f(P_n))$$

where $\mathcal{L}(G)$ is the vector space of rational functions f such that $f = 0$ or $\mathrm{div}(f) + G \geq 0$. If $\ell(G)$ stands for the dimension of $\mathcal{L}(G)$, then the dimension of $\mathcal{C}$ is $\ell(G) - \ell(G - D)$. The number $a = \ell(G - D)$ is the *abundance* of the code. If $a = 0$ then $ev_{\mathcal{P}}$ is injective and the code is called *nonabundant*. If $a > 0$ it is said to be *abundant*.

10.3.1. *The gonality sequence*

The Goppa bound states that the minimum distance of the code $\mathcal{C} = C(\mathcal{X}, D, G)$ verifies $d_1(\mathcal{C}) \geq n - \deg(G)$. In order to extend this bound to higher weights we need a new tool. Recall that the (classical) *gonality*, γ, of the curve $\mathcal{X}$ is defined as the smallest degree of a non constant morphism $\mathcal{X} \to \mathbb{P}^1$ over $\mathbb{F}_q$, where $\mathbb{P}^1$ stands for the projective line. This is equivalent

to saying

$$\gamma = \min\{\deg(A) \mid A \text{ is a rational divisor on } \mathcal{X} \text{ with } \ell(A) \geq 2\}.$$

In the same way, for every positive integer r, we can define the r-th gonality of $\mathcal{X}$ as

$$\gamma_r = \min\{\deg(A) \mid A \text{ is a rational divisor on } \mathcal{X} \text{ with } \ell(A) \geq r\}.$$

The sequence $\mathrm{GS}(\mathcal{X}) = (\gamma_r)_{r=1,2,\ldots}$ is called the *gonality sequence* of $\mathcal{X}$. Some of its main properties are collected in the next Proposition.

Proposition 10.2. *Let $\mathcal{X}$ be a curve having at least one rational point and let (γ_r) be its gonality sequence. Then*

(1) $\gamma_1 = 0$ and the sequence is strictly increasing.
(2) For $r \leq g$ we have $2r - 2 \leq \gamma_r \leq r + g - 1$.
(3) $\gamma_g = 2g - 2$ and $\gamma_r = g + r - 1$ for $r > g$.
(4) Let m be an integer with $0 \leq m \leq 2g-1$. Then $m \in \mathrm{GS}(\mathcal{X})$ if and only if $2g - 1 - m \notin \mathrm{GS}(\mathcal{X})$.

Proof. (1) is clear. The right-hand inequality of (2) and (3) follow from Riemann-Roch Theorem, whereas the left-hand inequality of (2) is a direct consequence of Clifford's Theorem. Let us prove (4). By item (3), there are precisely g gonality numbers in the interval $[0, 2g-1]$. Thus it is enough to show that $2g - 1 - \gamma_i \neq \gamma_j$ for any $i, j = 1, \ldots, g$. Let A be a rational divisor such that $\deg(A) = \gamma_i$ and $\ell(A) \geq i$. Let W be a canonical divisor on $\mathcal{X}$. By the Riemann-Roch Theorem, $\ell(W - A) \geq i + g - \gamma_i - 1$. Suppose that $j \leq i + g - \gamma_i - 1$. Then $\gamma_j \leq \deg(W - A) = 2g - 2 - \gamma_i$ and thus $2g-1-\gamma_i \neq \gamma_j$. Now let $j \geq i+g-\gamma_i$ and suppose by means of contradiction that $2g-1-\gamma_i = \gamma_j$. Let B be a rational divisor such that $\deg(B) = \gamma_j$ and $\ell(B) \geq j$. As above we have $\ell(W - B) \geq j + g - \gamma_j - 1$ so that $\ell(W - B) \geq j + \gamma_i - g \geq i$. This is not possible since $\deg(W - B) = 2g - 2 - \gamma_j = \gamma_i - 1$ and the result is proved. □

The next Corollary states a property that we shall use later.

Corollary 10.3. *For $r = 1, \ldots, g$, we have*

$$\gamma_{g-\gamma_r+r-1} < 2g - 1 - \gamma_r < \gamma_{g-\gamma_r+r}.$$

Proof. Let r be an integer, $1 \leq r \leq g$. The interval $[0, \gamma_r]$ contains exactly $\gamma_r - r + 1$ nongonality numbers hence, according to item (4) in the above Proposition, the interval $[2g - 1 - \gamma_r, 2g - 1]$ contains just the same

amount of gonality numbers. Since $2g-1 \notin \mathrm{GS}(\mathcal{X})$ and $\gamma_g = 2g-2$, we conclude that the first of these gonality numbers is $\gamma_{g-\gamma_r+r}$. Taking into account that $2g-1-\gamma_r \notin \mathrm{GS}(\mathcal{X})$, we get the result. □

The gonality sequence plays an important role in many computations concerning AG codes. However it is usually difficult to determine. For plane curves we have (see [27])

Proposition 10.4. *Let $\mathcal{X}$ be a nonsingular plane curve of degree t over a perfect field $\mathbb{F}$ with at least one rational point. Let r be a positive integer and write $r = \frac{1}{2}(j+1)(j+2) - i$, with $0 \le i \le j$. Then*

$$\gamma_r = \begin{cases} jt - i & \text{if } r \le g; \text{ and} \\ r + g - 1 & \text{if } r > g \end{cases}$$

where g is the genus of $\mathcal{X}$ (that is, $g = (t-1)(t-2)/2$).

To prove this Proposition we need a Theorem due to M. Noether.

Theorem 10.5. *(M. Noether) Let $\mathcal{X}$ be a nonsingular plane curve of degree t and genus g over a perfect field $\mathbb{F}$. Let A be a rational divisor on $\mathcal{X}$.*

(1) If $\deg(A) > t(t-3)$, then $\ell(A) = \deg(A) + 1 - g$.
(2) If $0 \le \deg(A) \le t(t-3)$, write $\deg(A) = jt - i$ with $0 \le j$ and $0 \le i < t$. Then

$$\ell(A) \le \begin{cases} \frac{1}{2}j(j+1) & \text{if } i > j; \text{ and} \\ \frac{1}{2}(j+1)(j+2) - i & \text{if } 0 \le i \le j. \end{cases}$$

We do not include the proof. The interested reader is addressed to [16]. This Theorem holds for perfect fields so, in particular for finite fields, because the dimension $\ell(A)$ of a $\mathbb{F}_q$-rational divisor does not change when we consider A over the algebraic closure of $\mathbb{F}_q$ (see [34] for example). Let us prove now Proposition 10.4.

Proof. If $r > g$ then the Riemann-Roch Theorem implies the result. Let us assume $1 \le r \le g$ and write $r = \frac{1}{2}(j+1)(j+2) - i$, with $0 \le i \le j \le t-3$. Let A be a divisor with $\ell(A) = r$ and set $\deg(A) = j't - i'$, for some $0 \le j' \le t-3$ and $0 \le i' < t$. Let us see that $\gamma_r \ge jt - i$. If $i' \le j'$ then

$$r = \frac{1}{2}(j+1)(j+2) - i \le \frac{1}{2}(j'+1)(j'+2) - i'$$

by Noether's Theorem. So $j < j'$ or $j = j'$ and $i \ge i'$. Thus $jt - i \le j't - i' = \deg(A)$. If $i' > j'$ then

$$\frac{1}{2}j(j+1) < r = \frac{1}{2}(j+1)(j+2) - i \le \frac{1}{2}j'(j'+1)$$

again by Noether's Theorem. So $j < j'$ and then $jt - i < j't - i' = \deg(A)$. Then $\gamma_r \geq jt - i$. To see the equality, take a rational point, say $Q = (0 : 1 : 0)$ (after a change of coordinates if necessary). Let B be the divisor obtained as intersection of $\mathcal{X}$ with the line $Z = 0$. Since $j < t$ then the $\frac{1}{2}(j+1)(j+2)$ functions $x^\alpha y^\beta$ where $x = X/Z$, $y = Y/Z$, $\alpha, \beta \in \mathbb{N}_0$ and $\alpha + \beta \leq j$, are linearly independent and belong to $\mathcal{L}(jB)$. Thus $\ell(jB - iQ) \geq \frac{1}{2}(j+1)(j+2) - i$ and we get the equality. □

10.3.2. *Extending the Goppa bound*

Let us already study the weight hierarchy of an AG code $\mathcal{C}$. As in the case of the minimum distance, the numbers $d_r(\mathcal{C})$ admit and arithmetical interpretation.

Proposition 10.6. *Let $\mathcal{C} = C(\mathcal{X}, D, G)$ be a code of abundance $a \geq 0$. Then*

$$d_r(\mathcal{C}) = \min\{n - \deg(D') \mid 0 \leq D' \leq D, \ell(G - D') \geq r + a\}$$

Proof. If $d_r(\mathcal{C}) = d$ then there exists a subspace $V \subseteq \mathcal{C}$ of dimension r and support size d. Let $V = \langle ev_{\mathcal{P}}(f_1), \ldots, ev_{\mathcal{P}}(f_r)\rangle$. Then $f_1, \ldots, f_r$ are independent functions vanishing at $n - d$ distinct points, say $P_{d+1}, \ldots, P_n$ (up to reordering if necessary). Thus $f_1, \ldots, f_r \in \mathcal{L}(G - P_{d+1} - \cdots - P_n) \setminus \mathcal{L}(G - D)$ and hence $\ell(G - P_{d+1} - \cdots - P_n) \geq r + a$. Conversely, assume that there are $n - d$ distinct points $P_{d+1}, \ldots, P_n$, such that $\ell(G - P_{d+1} - \cdots - P_n) \geq r + a$. Let $\{\phi_1, \ldots, \phi_a\}$ be a basis of $\mathcal{L}(G - D)$ and extend it to a basis $\{\phi_1, \ldots, \phi_a, f_1, f_2 \ldots\}$ of $\mathcal{L}(G - P_{d+1} - \cdots - P_n)$. Let $V = \langle ev_{\mathcal{P}}(f_1), \ldots, ev_{\mathcal{P}}(f_r)\rangle$; thus $\mathrm{supp}(V) \leq d$ and $\dim(V) = r$, hence $d_r(\mathcal{C}) \leq d$. □

Corollary 10.7. *Let $\mathcal{C} = C(\mathcal{X}, D, G)$ be a code of dimension k and abundance a. Then for every r, $1 \leq r \leq k$ we have*

(1) $d_r(\mathcal{C}) \geq n - \deg(G) + \gamma_{r+a}$;
(2) if $r + a > g$ then $d_r(\mathcal{C}) = n - k + r$;
(3) if $r + a = g$ then $d_r(\mathcal{C}) = n - k + r$ or $d_r(\mathcal{C}) = n - k + r - 1$.

Proof. Let $D' \leq D$ be an effective divisor such that $d_r(\mathcal{C}) = n - \deg(D')$. Then $\ell(G - D') \geq r + a$ so $\deg(G) + d_r(\mathcal{C}) - n \geq \gamma_{r+a}$ and (1) is proved. To prove (2) note that $\gamma_{r+a} = r + a + g - 1$ when $r + a > g$, so $n - k + r \leq n - \deg(G) + \gamma_{r+a}$ and the equality follows from the Singleton bound. (3) is proved in the same way. □

Remark 10.8. (1) The statement (1) in the previous Corollary gives $d_1(\mathcal{C}) \geq n - \deg(G) + \gamma_{a+1}$ for $r = 1$. This is known as the *improved Goppa bound*. It is useful for $\deg(G) \geq n$ when the classical Goppa bound does not give any information. (2) The statement (2) shows that a code $C(\mathcal{X}, D, G)$ of abundance a, arising from a curve of genus g, is $(g - a + 1)$-rank MDS.

Let us write $\sigma(\mathcal{C}) = n - k + 1$. According to the Singleton bound we know that $d \leq \sigma(\mathcal{C})$. Actually we can improve this bound as follows.

Corollary 10.9. *Let $\mathcal{C} = C(\mathcal{X}, D, G)$ be a $[n, k, d]$ code of abundance $a \geq 0$. Then $\sigma(\mathcal{C}) - (g - a) \leq d \leq \sigma(\mathcal{C})$.*

Proof. Use the Singleton bound, the improved Goppa bound and the inequality $\gamma_{a+1} \geq 2a$. □

10.3.3. *One-point codes and the Weierstrass semigroup*

A second way to get estimates for the GHWs is to extend the order (or Feng-Rao) bound on $d_1(\mathcal{C})$. In general, this bound gives very good results. However, its field of application is restricted to the duals of one-point codes. In this section we shall recall some basic facts about these codes.

As in the previous sections, let $\mathcal{X}$ be a curve over $\mathbb{F}_q$ of genus g, $\mathcal{P} = \{P_1, \ldots, P_n\}$ be a set of rational points and $D = P_1 + \cdots + P_n$. Assume now that there exist an extra rational point Q, $Q \neq P_i$ for all i. If we take $G = mQ$, then the code $C(\mathcal{X}, D, G)$ is called *one-point*. To study these codes it is useful to consider the algebra

$$\mathcal{L} = \bigcup_{m=0}^{\infty} \mathcal{L}(mQ)$$

and the set $H = \{-v_Q(f) \mid f \in \mathcal{L}\}$, where v_Q is the valuation at Q. Due to the properties of the valuations, H is a semigroup, called the *Weierstrass semigroup* at Q. Recall that this means that $0 \in H$ and $\rho, \sigma \in H$ implies $\rho + \sigma \in H$. Note that $\ell(mQ) > \ell((m-1)Q)$ if and only if there exist a rational function $f \in \mathcal{L}$ such that $-v_Q(f) = m$, hence if $H = \{\rho_1 = 0 < \rho_1 < \cdots\}$, it holds that $\ell(\rho_l Q) = l$, and for a nonnegative integer m, $\ell(mQ) = \max\{l \mid \rho_l \leq m\}$.

The Weierstrass semigroup can be used to give upper bounds on the GHWs. See for example the following result.

Proposition 10.10. *Let $\mathcal{C} = C(\mathcal{X}, D, G)$ be a code of dimension k and abundance a. If there is a rational point $Q \notin \mathcal{P}$, then for every r, $1 \leq r \leq k$,*

such that $C(\mathcal{X}, D, G - \rho_{r+a}Q) \neq 0$ *we have*

$$d_r(C(\mathcal{X}, D, G)) \leq d_1(C(\mathcal{X}, D, G - \rho_{r+a}Q)).$$

Proof. For two effective divisors E_1, E_2 on $\mathcal{X}$, it is known that $\ell(E_1) + \ell(E_2) \leq \ell(E_1 + E_2) + 1$ (see for example [15], Lemma IV.5.5). If $d_1(C(\mathcal{X}, D, G - \rho_{r+a}Q)) = d$ then there exists an effective divisor $D' \leq D$ of degree $n - d$ such that $\ell(G - \rho_{r+a}Q - D') \geq 1 + \ell(G - \rho_{r+a}Q - D)$. Thus

$$\ell(G - \rho_{r+a}Q - D') + \ell(\rho_{r+a}Q) \leq \ell(G - D') + 1$$

and hence $\ell(G - D') \geq r + a$, so $d_r(C(\mathcal{X}, D, G)) \leq d$. □

Let us return to one point codes. The evaluation map can be extended, in the obvious way, to the algebra $\mathcal{L}$,

$$ev_{\mathcal{P}} : \mathcal{L} \to \mathbb{F}_q^n \ , \ ev_{\mathcal{P}}(f) = (f(P_1), \ldots, f(P_n)).$$

Note that this map is now surjective. In fact, from the Goppa's estimates for the dimension of a code, we have that $\dim(C(\mathcal{X}, D, (n + 2g)Q)) = n$. On the other hand, as seen before, for a nonnegative integer m, we have $C(\mathcal{X}, D, mQ) = C(\mathcal{X}, D, \rho_l Q)$, where ρ_l is the largest element in H not exceeding m. Then we can restrict to consider the one-point codes $C(\mathcal{X}, D, \rho_l Q)$, with $0 \leq \rho_l \leq n + 2g$, and their duals $C(\mathcal{X}, D, \rho_l Q)^\perp$. Remark that, as we have seen in a previous chapter, the dual of an AG code is again an AG code: $C(\mathcal{X}, D, G)^\perp = C(\mathcal{X}, D, D + W - G)$, where W is a canonical divisor with simple poles and residue 1 at each point P_i.

For all $i = 0, 1, \ldots$, let $f_i \in \mathcal{L}$ be a function such that $-v_Q(f_i) = \rho_i$. Then $\{f_1, \ldots, f_l\}$ is a basis of $\mathcal{L}(\rho_l)$. In order to simplify the notation, we shall write $\mathbf{f}_i = ev_{\mathcal{P}}(f_i)$ (hence $C(\mathcal{X}, D, \rho_l Q) = \langle \mathbf{f}_1, \ldots \mathbf{f}_l \rangle$) and

$$C(l) = C(\mathcal{X}, D, \rho_l Q)^\perp = \{\mathbf{x} \in \mathbb{F}_q^n \mid \mathbf{x} \cdot \mathbf{f} = 0 \text{ for all } i \leq l\}$$

where $\cdot$ stands for the usual inner product

$$\mathbf{x} \cdot \mathbf{y} = \sum x_i y_i.$$

Example 10.11. The *Hermitian curve* $\mathcal{H}$ is defined over a field $\mathbb{F}_{q^2}$ by the affine equation $y^q + y = x^{q+1}$. It is a nonsingular plane curve of genus $q(q-1)/2$ and it has $n = q^3$ rational affine points plus one point at infinity, Q. This is the maximum possible number of points allowed by the Weyl bound: $\mathcal{H}$ is a maximal curve. One-point codes constructed from this curve and the divisors D, sum of all affine points, and $G = mQ$ are called *Hermitian codes.*

It is easy to check that for all $\alpha, \beta \in \mathbb{F}_{q^2}$ such that $\beta^q + \beta \neq 0$, we have

$$\operatorname{div}(x-\alpha) = \sum_{\beta^q+\beta=\alpha^{q+1}} (\alpha,\beta) - qQ,$$

$$\operatorname{div}(y-\beta) = \sum_{\alpha^{q+1}=\beta^q+\beta} (\alpha,\beta) - (q+1)Q.$$

In particular, if $H = H(Q) = \{\rho_1, \rho_2, \dots\}$ is the Weierstrass semigroup of Q, we have $\langle q, q+1\rangle \subseteq H$. Since both semigroups have equal genus, we conclude that $H = \langle q, q+1\rangle$. Furthermore, according to Proposition 10.4, the gonality sequence of $\mathcal{H}$ verifies $\gamma_i = \rho_i$ for all i. In particular H is symmetric, so $(2g-2)Q$ is a canonical divisor. Let us consider the function

$$f = \sum_{\alpha\in\mathbb{F}_{q^2}} (x-\alpha).$$

Since $\operatorname{div}(f) = D - nQ$, it holds that $D \sim nQ$, hence the dual of $C(\mathcal{H}, D, mQ)$ is $C(\mathcal{H}, D, (n+2g-2-m)Q)$. Finally, according to the properties of valuations, the algebra $\mathcal{L}$ is given by $\mathcal{L} = \langle\{x^i y^j \mid 0 \leq i, 0 \leq j \leq q-1\}\rangle$.

Example 10.12. (Example 10.11 continued). Let us consider the Hermitian code $\mathcal{C} = C(\mathcal{H}, D, \rho Q)$, with $\rho \in H$. Let k be the dimension of this code and be a its abundance. We can say something about the GHWs of $\mathcal{C}$ by using the extended Goppa bound as follows. For a pole number at Q $m \in H$, $m \leq n$, write $m = iq + j(q+1)$ with $0 \leq i$ and $0 \leq j < q$. If either $i < q^2 - q$ or $j = 0$, then there exists an effective divisor $D' \leq D$ such that $mQ \sim D'$. Furthermore note that for a given positive integer t, then either $t \in H$ or $n - t \in H$. According to the extended Goppa bound, for every r, $1 \leq r \leq \min\{k, g-a\}$, we have
a) if $q^2 \leq \rho - \rho_{r+a} \leq n - q^2$, then $d_r(\mathcal{C}) = n - \rho + \rho_{r+a}$.
b) if $a > 0$ then $d_r(\mathcal{C}) \leq \rho_{r+1}$.

The proofs of all these statements are left to the reader as an exercise.

10.3.4. *Extending the order bound*

In this section, following [17], we shall extend the order bound on the minimum distance to all the GHWs of the codes $C(l)$. Let us remember that for given $\mathbf{y} \in \mathbb{F}_q^n$ we define the *syndromes* of $\mathbf{y}$ as $s_i(\mathbf{y}) = \mathbf{f}_i \cdot \mathbf{y}$, $i = 1, 2, \dots$. In the same way we can define the *two-dimensional syndromes* of $\mathbf{y}$ as $s_{ij}(\mathbf{y}) = (\mathbf{f}_i * \mathbf{f}_j) \cdot \mathbf{y}$, where $*$ stands for the coordinate wise product,

$\mathbf{x} * \mathbf{y} = (x_1 y_1, \ldots, x_n y_n)$. Thus $\mathbf{f}_i * \mathbf{f}_j = ev_{\mathcal{P}}(f_i f_j)$. Now write $l(i,j) = \min\{t \mid f_i f_j \in \mathcal{L}(\rho_t Q)\}$, or equivalently $l(i,j) = l$ if $\rho_i + \rho_j = \rho_l$.

Lemma 10.13. *The function l is symmetric and strictly increasing in both arguments.*

Let N be the smallest integer such that $\mathbf{f}_1, \ldots, \mathbf{f}_N$ generate $\mathbb{F}_q^n$. We define the *matrix of syndromes* of $\mathbf{y}$ as

$$S(\mathbf{y}) = (s_{ij}(\mathbf{y}))_{1 \le i,j \le N}.$$

In the next two Lemmas we state some important properties of this matrix.

Lemma 10.14. *Let $\mathbf{y} \in \mathbb{F}_q^n$. Let $S(\mathbf{y})$ the matrix of syndromes of $\mathbf{y}$, $D(\mathbf{y})$ be the diagonal matrix with $\mathbf{y}$ in the diagonal and H be the $N \times n$ matrix whose i-th row is $\mathbf{f}_i$. Then*

$$S(\mathbf{y}) = HD(\mathbf{y})H^t.$$

Lemma 10.15. *Let $\mathbf{y} \in \mathbb{F}_q^n$.*

(1) If $\mathbf{y} \in C(l)$ and $l(i,j) \le l$, then $s_{ij}(\mathbf{y}) = 0$.
(2) If $\mathbf{y} \in C(l)$ and $l(i,j) = l+1$, then $s_{tj}(\mathbf{y}) = 0$ for all $t < i$.
(3) If $\mathbf{y} \in C(l) \setminus C(l+1)$ and $l(i,j) = l+1$, then $s_{ij}(\mathbf{y}) \neq 0$.

Proof. (1) If $l(i,j) \le l$, then $ev_{\mathcal{P}}(f_i f_j) = \mathbf{f}_i * \mathbf{f}_j \in C(l)^{\perp}$, hence $(\mathbf{f}_i * \mathbf{f}_j) \cdot \mathbf{y} = 0$. (2) Since l is strictly increasing in both arguments, we have $l(t,j) < l(i,j)$ and the result follows from (1). (3) If $l(i,j) = l+1$, then $f_i f_j = \lambda f_{l+1} + f$ for some $\lambda \in \mathbb{F}_q^*$ and $f \in \mathcal{L}(\rho_l Q)$. Thus $\mathbf{f}_i * \mathbf{f}_j = \lambda \mathbf{f}_{l+1} + \mathbf{f}$, with $\mathbf{f} = ev_{\mathcal{P}}(f) \in C(l)$, hence $s_{ij}(\mathbf{y}) = (\mathbf{f}_i * \mathbf{f}_j) \cdot \mathbf{y} = \lambda \mathbf{f}_{l+1} \cdot \mathbf{y} + \mathbf{f} \cdot \mathbf{y} = \lambda \mathbf{f}_{l+1} \cdot \mathbf{y} \neq 0$, since $\mathbf{y} \in C(l) \setminus C(l+1)$. □

Lemma 10.15 leads us to consider the set $N(l) = \{(i,j) \in \mathbb{N} \mid l(i,j) = l+1\}$.

Lemma 10.16. *Let $N(l) = \{(i_1, j_1), \ldots, (i_r, j_r)\}$. Then the integers $i_1, \ldots, i_t$ are all different.*

Proof. If $i_t = i_h$ and $j_t < j_h$ then, since l is strictly increasing, we have $l(i_t, j_t) < l(i_h, j_h)$ hence both pairs cannot belong to $N(l)$. □

If $N(l) = \{(i_1, j_1), \ldots, (i_r, j_r)\}$ then, by the symmetry of l, we have $\{i_1, \ldots, i_r\} = \{j_1, \ldots, j_r\}$. We denote this set by $A(l)$, hence $A(l) = \{i \in \mathbb{N} \mid (i,j) \in N(l) \text{ for some } j\}$. From the above Lemma, $\#A(l) = \#N(l)$.

Let $\mathbf{u} \in \mathbb{F}_q^n$ and consider the linear maps given by the matrices $S(\mathbf{u}), D(\mathbf{u}), H$ and H^t, which we can still denote by $S(\mathbf{u}), D(\mathbf{u}), H$ and H^t respectively, that is, $S(\mathbf{u}) : \mathbb{F}_q^N \to \mathbb{F}_q^N, (S(\mathbf{u}))(\mathbf{x}) = \mathbf{x}S(\mathbf{u})$, etc.

Lemma 10.17. *Let $U \subseteq C(l)$ be a subspace. Then $\#supp(U) = \dim\langle \cup_{\mathbf{u}\in U} \mathrm{Im}(S(\mathbf{u}))\rangle$.*

Proof. Let $\mathbf{u} \in U$ and let us consider the diagram

$$\begin{array}{ccc} \mathbb{F}_q^n & \xrightarrow{D(\mathbf{u})} & \mathbb{F}_q^n \\ \downarrow{\scriptstyle H^t} & & \downarrow{\scriptstyle H} \\ \mathbb{F}_q^N & \xrightarrow{S(\mathbf{u})} & \mathbb{F}_q^N \end{array}$$

According to Lemma 10.14 this is a commutative diagram. Furthermore, since the matrix H has full rank, the map H is surjective and the map H^t is injective. Thus $\dim(\mathrm{Im}(S(\mathbf{u}))) = \dim(\mathrm{Im}(D(\mathbf{u})))$, hence $\dim\langle \cup_{\mathbf{u}\in U} \mathrm{Im}(S(\mathbf{u}))\rangle = \dim\langle \cup_{\mathbf{u}\in U} \mathrm{Im}(D(\mathbf{u}))\rangle$. On the other hand, it is clear that $\mathrm{Im}(D(\mathbf{u})$ is the subspace of equations $(X_i = 0, i \notin \mathrm{supp}(\mathbf{u}))$, and thus $\langle \cup_{\mathbf{u}\in U} \mathrm{Im}(D(\mathbf{u}))\rangle$ is the subspace of equations $(X_i = 0, i \notin \mathrm{supp}(U))$. Then $\dim\langle \cup_{\mathbf{u}\in U} \mathrm{Im}(D(\mathbf{u}))\rangle = \#\mathrm{supp}(U)$. □

Lemma 10.18. *Let U be a linear subspace of $C(l)$ of dimension r. Then there exist r integers $l \le l_1 < \cdots < l_r < N$ such that*

$$\dim(C(l_i) \cap U) = \dim(C(l_i + 1) \cap U) + 1.$$

Proof. Let us consider the decreasing chain of subspaces $C(l) \cap U \supseteq C(l+1) \cap U \supseteq \cdots \supseteq C(N) \cap U = (0)$. Since $\dim(C(l) \cap U) = r$, $\dim(C(N) \cap U) = 0$ and at every step t, $\dim(C(t) \cap U) \le \dim(C(t+1) \cap U) + 1$, we conclude that we get equality exactly r times. □

The r-tuple $(l_1, \ldots, l_r)$ is called the *associated r-tuple* to U in $C(l)$. Let $A(l_1, \ldots, l_r) = A(l_1) \cup \cdots \cup A(l_r)$ and $a(l_1, \ldots, l_r) = \#A(l_1, \ldots, l_r)$. The next result gives a bound on the cardinality of the support of a subspace.

Proposition 10.19. *Let U be a linear subspace of $C(l)$ of dimension r and let $(l_1, \ldots, l_r)$ be the associated r-tuple. Then*

$$\#supp(U) \ge a(l_1, \ldots, l_r).$$

Proof. For $i = 1, \ldots, r$, take a vector $\mathbf{u}_i \in C(l_i) \setminus C(l_i + 1)$. Then $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ is a basis of U. Let M_i be the matrix obtained from $S(\mathbf{u}_i)$ by taking the rows of indices in $A(l_1)$. Thus M_i is a $a(l_i) \times N$ matrix

and its rank is $a(l_i)$, according to Lemma 10.15. Now let us consider the $\sum a(l_i) \times N$ matrix

$$M = \begin{pmatrix} M_1 \\ \vdots \\ M_r \end{pmatrix}.$$

Since the rank of this matrix is $\dim\langle \cup_{i=1,\ldots,r} \mathrm{Im}(S(\mathbf{u}_i)) \rangle \leq \dim\langle \cup_{\mathbf{u}\in U} \mathrm{Im}(S(\mathbf{u})) \rangle$, according to Lemma 10.17 it is enough to prove that $\mathrm{rank}(M) \geq a(l_1, \ldots, l_r)$. We shall proceed by induction, by using Gaussian elimination. We begin by applying Gasussian elimination to M with the $a(l_1)$ pivots in the rows of M_1 and the columns of $A(l_1)$. Suppose now that we have found a matrix with pivots in the rows of $M_1, \ldots, M_t$ and the columns of $\cup_{i=1}^{t} A(l_i)$. Apply Gaussian elimination with the pivots in the rows of M_{t+1} and the columns of $A(l_{t+1}) \setminus (\cup_{i=1}^{t} A(l_i))$. Thus, finally we obtain a matrix M' having pivots in the columns of $\cup_{i=1}^{r} A(l_i) = A(l_1, \ldots, l_r)$. Thus its rank, and therefore the rank of M, is at least $a(l_1, \ldots, l_r)$. □

Definition 10.20. The number

$$d_{ORD}(l) = \min\{a(l_1, \ldots, l_r) \mid l \leq l_1 < \cdots < l_r \leq N \text{ and } C(l_i) \neq C(l_i + 1) \text{ for all } i = 1, \ldots, r\}$$

is called the *order bound* on the rth generalized Hamming weight of $C(l)$.

As a consequence of the previous results we can state the following.

Theorem 10.21. $d_r(C(l)) \geq d_{ORD}(l)$.

Example 10.22. (Example 10.12 continued). The computation of the order bound is sometimes difficult. For example, let us consider the Hermitian codes $\mathcal{C} = C(\mathcal{H}, D, \rho Q)$. For simplicity, let us restrict to the second generalized Hamming weight $d_2(\mathcal{C})$. We have $d_2(C(\mathcal{H}, D, \rho)) \geq n - \rho + \rho_2 = n - \rho + q$ and, after the results obtained in Example 10.12, equality holds for $\rho \leq n - q^2 + q + 1$. For $n - q^2 + q + 2 \leq \rho \leq n - 2$, the order bound leads to the true value of d_2 as follows: write $\rho = n - \alpha q + \beta$ with $1 \leq \alpha \leq q - 1$ and $0 \leq \beta \leq q - 1$, then

$$d_2(C(\mathcal{H}, D, \rho)) = \begin{cases} (\alpha + 1)q - \beta \text{ if } \beta = 0 \text{ or } \beta \geq q - \alpha; \text{ and} \\ (\alpha + 1)q - 1 \text{ if } 1 \leq \beta < q - \alpha. \end{cases}$$

The proof is left to the reader. However, it can be found (together with the weight hierarchy in full of all Hermitian codes) in [2].

10.4. Trellis structure of codes

10.4.1. *Trellises and codes*

A *trellis of depth* n is an edge-labeled directed graph $T = (V, E)$ with vertex set V and edge set E satisfying the following conditions:

- V is the union of $(n+1)$ disjoint subsets $V_0, \ldots, V_n$;
- every edge in E that begins at V_i ends at V_{i+1};
- every vertex in V belongs to at least one path from a vertex in V_0 to a vertex in V_n.

Given a trellis T, each vertex in V_i is called a *state at time* i, $i = 0, \ldots, n$. Edges represent transitions between states. Here we consider trellises with V_0 and V_n having just one element (called the *root* and the *toor* respectively) and such that the label alphabet is the finite field $\mathbb{F}_q$ (or a power of $\mathbb{F}_q$). We can associate to each path from V_0 to V_n the ordered n-tuple of edge labels. The set of all such n-tuples is a block code $\mathcal{C}_T$ over $\mathbb{F}_q$. Conversely, given a block code $\mathcal{C} \subseteq \mathbf{F}_q^n$ we say that a trellis T *represents* $\mathcal{C}$ if $\mathcal{C}_T = \mathcal{C}$. Remark that there might exist more than one non-isomorphic trellis representing the same code.

Example 10.23. Figure 1 shows a trellis representation of the $[8, 4, 4]$ binary Reed-Muller code $\mathcal{C}$ with label alphabet $\mathbb{F}_2^2$. We leave as an exercise to see that each codeword in $\mathcal{C} = \langle(10101010), (11001100), (11110000), (11111111)\rangle$ correspond to a unique path though the trellis.

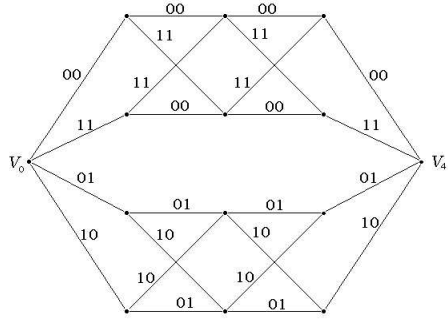


Fig. 10.1. Trellis representation of the $[8, 4, 4]$ binary Reed-Muller code.

The use of trellises in coding theory started with applications to convolutional codes. Later they were employed with block codes mainly for the purpose of soft-decision decoding with the Viterbi algorithm. There are some good reasons to study trellis representations. Two of them are the following:

(1) Trellis based decoding (via the Viterbi algorithm) is one of the most efficient decoding methods for decoding *general* linear codes; and
(2) The complexity of a minimal trellis representation gives a good measure of the complexity of a code (whereas the parameters $[n, k, d]$ do not).

There are several measures of the complexity of a trellis T that represents a code $\mathcal{C} \subseteq \mathbf{F}_q^n$, all of them related to the decoding complexity. The most common is the *state complexity profile*, which is the sequence of state space sizes, $\mathrm{SCP}(T) = \{s_0(T), \ldots, s_n(T)\}$, where $s_i(T) = \log_q(\#V_i)$ with $V_0, \ldots, V_n$ being the underlying partition of the vertex set of T.

If the code $\mathcal{C}$ is linear, and once the order of coordinates of $\mathcal{C}$ is fixed, there exists an unique (up to a graph isomorphism) trellis $T_\mathcal{C}$ that simultaneously minimices all the s_i, that is, such that for each $i = 0, 1, \ldots, n$ and any trellis T that represents $\mathcal{C}$, it holds that $s_i(T_\mathcal{C}) \leq s_i(T)$. We do not give here the proof of this result; the interested reader is addressed to the bibliography (see [37]). The trellis $T_\mathcal{C}$ is called *the minimal trellis* of $\mathcal{C}$. Then the *state complexity profile* of $\mathcal{C}$ is, by definition, $\mathrm{SCP}(T_\mathcal{C})$ and the number

$$s_T(\mathcal{C}) := \max\{s_0(T_\mathcal{C}), s_1(T_\mathcal{C}), \ldots, s_n(T_\mathcal{C})\},$$

is called the *state complexity* of $\mathcal{C}$. An important (although undesirable) property of these numbers is that they may vary when changing the order of coordinates. Let us remember that two codes are *equivalent* if one of them can be obtained from the other by permuting coordinates. We denote by $[\mathcal{C}]$ the set of codes which are equivalent to $\mathcal{C}$. Thus we are lead to consider *the absolute state complexity* of $\mathcal{C}$, namely

$$s[\mathcal{C}] = \min\{s(\mathcal{C}') \mid \mathcal{C}' \in [\mathcal{C}]\}.$$

10.4.2. *Minimal trellises*

Given a $[n, k, d]$ linear code $\mathcal{C}$ over the finite field $\mathbb{F}_q$, its minimal trellis $T = T_\mathcal{C}$ can be constructed in several ways. For our purposes the relevant construction is the one given by Forney. He showed that the state sets $V_0, \ldots, V_n$, are $V_i = \mathcal{C}/(\mathcal{P}_i \oplus \mathcal{F}_i)$, where $\mathcal{P}_i$ and $\mathcal{F}_i$ are respectively the i-th

past and the i-th future subcodes of $\mathcal{C}$; namely, $\mathcal{P}_0 = \mathcal{F}_n = 0$, $\mathcal{P}_n = \mathcal{F}_0 = \mathcal{C}$, and for $i = 1, \ldots, n-1$,

$$\mathcal{P}_i = \{(c_1, \ldots, c_i) \mid (c_1, \ldots, c_i, 0, \ldots, 0) \in \mathcal{C}\},$$
$$\mathcal{F}_i = \{(c_{i+1}, \ldots, c_n) \mid (0, \ldots, 0, c_{i+1}, \ldots, c_n) \in \mathcal{C}\}.$$

(see [37], Theorem 4.13). It follows that $\#V_i$ is a power of q so that $s_i(\mathcal{C}) = k - \Delta_i$, with $\Delta_i = \dim(\mathcal{P}_i) + \dim(\mathcal{F}_i)$, $i = 0, 1, \ldots, n$. A first consequence is the following.

Proposition 10.24. *The state complexity profile of a linear code and that of its dual are identical.*

Proof. Let G and H be, respectively, a generator and a parity check matrix of $\mathcal{C}$. Let H_i (respectively G_i) denote the matrix consisting of the first i columns of H, $H = (\mathbf{h}_1, \cdots, \mathbf{h}_i)$ (resp. columns of G). Consider the sets $W_0 = (\mathbf{0})$ and for $i = 1, \ldots, n$,

$$W_i = \{c_1\mathbf{h}_1 + \cdots + c_i\mathbf{h}_i \mid (c_1, \ldots, c_i, *, \ldots, *) \in \mathcal{C}\} = \langle \text{columns of } H_i G_i^t \rangle$$

where $*$ stands for any element in $\mathbb{F}_q$. Now consider the map $\tau_i = \mathcal{C} \to W_i$ defined by $\tau_i(\mathbf{c}) = c_1\mathbf{h}_1 + \cdots + c_i\mathbf{h}_i$. τ_i is linear, surjective and its kernel is $\mathcal{P}_i \oplus \mathcal{F}_i$. Then $V_i \cong W_i$, hence it is enough to prove the duality property for W_i. Since $W_i(\mathcal{C}^\perp)$ is the column space of $G_i H_i^t$ and the columns of $G_i H_i^t$ are the rows of $H_i G_i^t$, the equality $\dim(W_i) = \dim(W_i(\mathcal{C}^\perp))$ follows from the fact that the row rank of a matrix is equal to its column rank. □

Remark 10.25. The construction of the minimal trellis with $V_i \cong \langle \text{columns of } H_i G_i^t \rangle$ is the so called *BCJR* (Bahl, Cocke, Jelinek and Raviv) *construction.*

Corollary 10.26. *(The Wolf bound)* $s(\mathcal{C}) \leq \min\{k, n-k\}$.

Proof. From the Forney construction we have $s(\mathcal{C}) \leq \dim(\mathcal{C})$ and $s(\mathcal{C}^\perp) \leq \dim(\mathcal{C}^\perp)$. According to Proposition 10.24 we get the result. □

The number $\min\{k, n-k\}$ given by the Wolf bound is usually denoted by $w(\mathcal{C})$. According to Proposition 10.24, to study the state complexity of $\mathcal{C}$ we can restrict ourselves to the case $2k \leq n$. Now by definition

$$s(\mathcal{C}) = k - \Delta,$$

where $\Delta = \Delta(\mathcal{C}) = \min\{\Delta_0, \Delta_1, \ldots, \Delta_n\}$. We set $\Delta[\mathcal{C}] = \max\{\Delta(\mathcal{C}') \mid \mathcal{C}' \in [\mathcal{C}]\}$.

Lemma 10.27. *With the above notation, the following holds:*

(1) $\mathcal{P}_i = 0$ *for* $i = 0, \ldots, d-1$*; in particular* $\min\{\Delta_0, \ldots, \Delta_{d-1}\} = \Delta_{d-1}$.
(2) $\mathcal{F}_i = 0$ *for* $i = n-d+1, \ldots, n$*; in particular* $\min\{\Delta_{n-d+1}, \ldots, \Delta_n\} = \Delta_{n-d+1}$.

Proof. For $i \in \{1, \ldots, d-1\}$ (resp. $i \in \{n-d+1, \ldots, n-1\}$), the effective length of $\mathcal{P}_i$ (resp. $\mathcal{F}_i$) is smaller than the weight of any nonzero codeword in $\mathcal{C}$. Thus $\mathcal{P}_i = 0$ (resp. $\mathcal{F}_i = 0$) and the result follows taking into account the fact that $\dim(\mathcal{F}_0) > \dim(\mathcal{F}_1) > \ldots$ (resp. $\dim(\mathcal{P}_n) > \dim(\mathcal{P}_{n-1}) > \ldots$). □

Proposition 10.28. *For a linear code* $\mathcal{C}$*, we have*

$$s(\mathcal{C}) = \begin{cases} k & \text{whenever } 2d \geq n+2, \\ k - \min\{\Delta_{d-1}, \ldots, \Delta_{n-d+1}\} & \text{otherwise.} \end{cases}$$

Proof. Follows from Lemma 10.27, by taking into account that there exists an integer i with $n-d+1 \leq i \leq d-1$ whenever $2d \geq n+2$. □

10.5. Linking the problems

Up to the moment we have introduced two objects related to a code $\mathcal{C}$: the GHWs and the trellis state complexity. At the first look, both seem very different. However there are close connections relating them. These relations become more clear if we see past and future subcodes $\mathcal{P}_i, \mathcal{F}_i$, of $\mathcal{C}$ as subcodes of support sizes i and $n-i$ respectively. Then their dimensions can be bounded in terms of the GHWs. Thus it is not strange that methods and tools used to study both problems are similar. As an example, we have the following result.

Proposition 10.29. *If there exists an integer* i*,* $1 \leq i \leq k$*, such that either* $d_i(\mathcal{C}) \geq n-d+2$ *or* $d_i(\mathcal{C}^\perp) \geq n-d^\perp+2$*, then* $s(\mathcal{C}) \geq w(\mathcal{C}) - i + 1$.

Proof. Suppose $d_i(\mathcal{C}) \geq n-d+2$. Since the length of $\mathcal{F}_{d-1}$ is $n-d+1$, then $\#\mathrm{supp}(\mathcal{F}_{d-1}) < i$ and hence $\dim(\mathcal{F}_{d-1}) < i$. Since $\mathcal{P}_{d-1} = 0$ we get the result. Similarly if $d_i(\mathcal{C}^\perp) \geq n-d^\perp+2$. □

For completeness, remark that when studying the trellis complexity, instead the GHW (for which we fix the dimension and look for the maximum support size) it is convenient to consider the 'reciprocal' sequence (that is, fix the support size and look for the maximum dimension). Then, given a code $\mathcal{C}$ of length n, for $i = 1, \ldots, n$ we define

$$\kappa_i(\mathcal{C}) = \max\{\dim(S) \mid S \text{ is a linear subcode of } \mathcal{C} \text{ of support size } i\}.$$

The sequence $\{\kappa_1(\mathcal{C}), \ldots, \kappa_n(\mathcal{C})\}$ is called the *dimension/length profile* of $\mathcal{C}$. It is equivalent to the weight hierarchy in the sense that once one of them is known, the other can be determined as follows

$$d_i(\mathcal{C}) = \min\{j \mid \kappa_j(\mathcal{C}) \geq i\} \quad i = 1, \ldots, k,$$
$$\kappa_i(\mathcal{C}) = \max\{j \mid d_j(\mathcal{C}) \leq i\} \quad i = 1, \ldots, n.$$

10.6. Trellis structure of AG codes

Let $\mathcal{X}$ be a curve of genus g and gonality sequence $\mathrm{GS}(\mathcal{X}) = (\gamma_i)$. Let $\mathcal{C} = C(\mathcal{X}, D, G)$ be an AG code of parameters $[n, k, d]$ over $\mathbb{F}_q$ arising from the curve $\mathcal{X}$ and the divisors G and $D = P_1 + \cdots + P_n$. In this section we investigate bounds on the trellis state complexity of $\mathcal{C}$.

10.6.1. *A Goppa-like bound on $s(\mathcal{C})$*

Firstly note that the past and future subcodes of an AG code $\mathcal{C}$ admit a clear interpretation in geometrical terms.

Proposition 10.30. *The past and future subcodes of $\mathcal{C}$ are given by*

$$\mathcal{P}_i \cong \mathcal{C}(\mathcal{X}, D - P_{i+1} - \ldots - P_n, G - P_{i+1} - \ldots - P_n),$$
$$\mathcal{F}_i \cong \mathcal{C}(\mathcal{X}, D - P_1 - \ldots - P_i, G - P_1 - \ldots - P_i).$$

Proof. Let $\mathbf{c} = ev_{\mathcal{P}}(f) \in \mathcal{C}$. Then $\mathbf{c} \in \mathcal{P}_i$ if and only if $f(P_{i+1}) = \cdots = f(P_n) = 0$, that is, if and only if $f \in \mathcal{L}(G - P_{i+1} - \cdots - P_n)$. □

Corollary 10.31. *Let $\mathcal{C} = C(\mathcal{X}, D, G)$ be a code of abundance a. Then*

(1) for $i = 1, \ldots, n-1$ we have

$$s_i(\mathcal{C}) = \ell(G) - \ell(G - P_1 - \cdots - P_i) - \ell(G - P_{i+1} - \cdots - P_n) + a.$$

(2) If $\deg(G) < \lfloor n/2 \rfloor + \gamma_{a+1}$ or $\deg(G) > \lceil n/2 \rceil + 2g - 2 - \gamma_{a+1}$ then $s(\mathcal{C}) = w(\mathcal{C})$.

Proof. (1) follows directly from the above Proposition, by taking dimensions. (2) If $\deg(G) < \lfloor n/2 \rfloor + \gamma_{a+1}$ then apply the improved Goppa bound on the minimum distance to Proposition 10.28. If $\deg(G) > \lceil n/2 \rceil + 2g - 2 - \gamma_{a+1}$ then use the duality property stated in Proposition 10.24. □

Lemma 10.32. *Let i be a positive integer. If $\gamma_{a+i} \geq 2\deg(G) - n + \gamma_{a+1} + 2$ or $\gamma_{a+i} \geq n + 2(2g-2) - 2\deg(G) - \gamma_{a+1} + 2$ then $s(\mathcal{C}) \geq w(\mathcal{C}) - i + 1$.*

Proof. If $i \geq k+1$ then the result is clear. Otherwise it is a consequence of Proposition 10.29 and Corollary 10.7(1). □

Theorem 10.33. $s(\mathcal{C}) \geq w(\mathcal{C}) - (g-a)$.

Proof. Apply the above Lemma to $i = g+1-a$. □

Note that this result is completely analogous to the bound on the minimum distance obtained in Corollary 10.9.

10.6.2. *Another bound on the trellis state complexity*

A more careful analysis gives a new bound on $s(\mathcal{C})$, which is stronger but more difficult to compute. In this section we shall restrict to the case $2k \leq n$.

Lemma 10.34. *If $2k \leq n$ and $n > 2g$, then the code $\mathcal{C} = \mathcal{C}(\mathcal{X}, D, G)$ is non-abundant and $2\deg(G) - n \leq 2g-2$.*

Proof. Suppose that $a = \ell(G-D) \geq 1$. Then the divisor $G-D$ must be special; otherwise $k = \ell(G) - (\deg(G-D)+1-g) \geq (\deg(G)+1-g) - (\deg(G-D)+1-g) = n$ which is a contradiction. By Clifford's Theorem, $\deg(G-D) \leq (\deg(G)-n)/2+1$ and hence

$$k = \ell(G) - a \geq (\deg(G)+1-g) - (\deg(G)-n)/2 - 1 = (\deg(G)+n-2g)/2.$$

From the hypothesis $2k \leq n$ we conclude that $2g \geq \deg(G)$. On the other hand, $\deg(G-D) \geq 0$ as $\ell(G-D) \geq 1$, and thus $2g \geq \deg(G) \geq n$; a contradiction. The second statement follows from the fact that $n/2 \geq k = \ell(G) \geq \deg(G)+1-g$. □

We can consider the gonality sequence $\mathrm{GS}(\mathcal{X})$ as a subset of $\mathbb{N}' = \{-1\} \cup \mathbb{N}_0$. An element in $\mathbb{N}' \setminus \mathrm{GS}(\mathcal{X})$ will be called a *gap* of $\mathcal{X}$. By Proposition 10.2 there are $g+1$ gaps and the biggest one is $2g-1$. Let $\tilde{\ell} : \mathbb{N}' \to \mathbb{N}_0$ be the numerical function defined by $\tilde{\ell}(-1) = 0$ and for $m \in \mathbb{N}_0$

$$\tilde{\ell}(m) = \max\{i \in \mathbb{N} \mid \gamma_i \leq m\}.$$

The function $\tilde{\ell}$ is an increasing step function such that $\tilde{\ell}(2g-2) = g$ and $\tilde{\ell}(2g-1+i) = g+i$ for $i \geq 0$. Moreover, $\tilde{\ell}(a+1) \leq \tilde{\ell}(a)+1$ and equality holds if and only if $a+1 \in \mathrm{GS}(\mathcal{X})$.

Lemma 10.35. *For a rational divisor F with $\deg(F) \geq -1$, it holds that $\ell(F) \leq \tilde{\ell}(\deg(F))$.*

Proof. If $\deg(F) = -1$, then $\ell(F) = 0 = \tilde{\ell}(-1)$. If $\deg(F) \geq 0$, let $i \in \mathbb{N}_0$ be such that $\gamma_i \leq \deg(F) < \gamma_{i+1}$ so that $\tilde{\ell}(\deg(F)) = i$. Thus by definition of γ_{i+1} we must have $\ell(F) \leq i$ and the result follows. □

Let $R : \mathbb{N}' \cap [-1, 2g-2] \to \mathbb{N}$ be the numerical function defined by

$$R(N) = \min\{\tilde{\ell}(a) + \tilde{\ell}(b) : a, b \in \mathbb{N}' \text{ with } a + b = N\}.$$

The main result of this section is the following.

Theorem 10.36. *Let $\mathcal{C} = \mathcal{C}(\mathcal{X}, D, G)$ be an AG code such that $2k \leq n$ and $n > 2g$. Then $\Delta[\mathcal{C}] \leq R(2\deg(G) - n)$ and hence $s[\mathcal{C}] \geq w(\mathcal{C}) - R(2\deg(G) - n)$.*

Proof. Set $m = \deg(G)$. Since the function R depends only on the underlying curve $\mathcal{X}$, it is enough to show that $\Delta(\mathcal{C}) \leq R(2m - n)$. In addition, $w(\mathcal{C}) = \min\{k, n - k\} = k$ and by Proposition 10.28 we can assume that $2d < n + 2$ so that $2m - n \geq -1$ by the Goppa estimate on d. By Lemma 10.34 the code $\mathcal{C}$ is non-abundant and hence the i-th element $s_i = s_i(\mathcal{C})$ in the state complexity profile of $\mathcal{C}$ is given by $s_i = k - \Delta_i$, where $\Delta_i = \ell(G - P_1 - \ldots - P_i) + \ell(G - P_{i+1} - \ldots - P_n)$. Thus, according to Proposition 10.28, and since $d \geq n - m$, we have $s(\mathcal{C}) = w(\mathcal{C}) - \Delta(\mathcal{C})$ where

$$\Delta(\mathcal{C}) = \min\{\Delta_{d-1}, \ldots, \Delta_{n-d+1}\} = \min\{\Delta_{n-m-1}, \ldots, \Delta_{m+1}\}.$$

Let i be an integer with $n - m - 1 \leq i \leq m + 1$ so that $\deg(G - P_1 - \ldots - P_i) \geq -1$ and $\deg(G - P_{i+1} - \ldots - P_n) \geq -1$; then by Lemma 10.35,

$$\Delta_i \leq \tilde{\ell}(\deg(G - P_1 - \ldots - P_i)) + \tilde{\ell}(\deg(G - Pi + 1 - \ldots - P_n)).$$

Now, as $\deg(G - P_1 - \ldots - P_i) + \deg(G - P_{i+1} - \ldots - P_n) = 2m - n$, which is at most $2g - 2$ by Lemma 10.34, the result follows. □

In order to apply the above result we need to know the behavior of the function R. This study is done in the rest of this section. In particular, we shall compute $R(2g - 2)$ whenever $g > 0$ and also explicitly describe R for the case of plane curves.

Lemma 10.37. *Let $N \in \mathbb{N}' \cap [-1, 2g - 2]$.*

(1) *R is an increasing function such that $R(N) \leq R(N + 1) \leq R(N) + 1$.*
(2) *If $N < \gamma_i - 1$, then $1 \leq R(N) \leq i - 1$;*
(3) *$R(N) \leq \lfloor (N + 1)/2 \rfloor + 1$;*
(4) *There exists a gap $a = a(N)$ of $\mathcal{X}$ with $a \leq N/2$ such that $R(N) = \tilde{\ell}(a) + \tilde{\ell}(N - a)$.*

Proof. From the definition of R it is clear that $R(N) \geq 1$ and that $R(-1) = 1$. (1) Let $R(N+1) = \tilde{\ell}(a) + \tilde{\ell}(b)$ with $a + b = N + 1$ and $a \leq b$. From $a+(b-1) = N$ we have $R(N) \leq \tilde{\ell}(a)+\tilde{\ell}(b-1) \leq \tilde{\ell}(a)+\tilde{\ell}(b) = R(N+1)$ since $\tilde{\ell}$ is an increasing function. Now suppose that $R(N) < R(N+1)$ and let $R(N) = \tilde{\ell}(a') + \tilde{\ell}(b')$ with $a' + b' = N$. Then from $(a'+1)+b' = N+1$, $R(N+1) \leq \tilde{\ell}(a'+1)+\tilde{\ell}(b')$ and thus $\tilde{\ell}(a'+1) > \tilde{\ell}(a')$. Therefore $R(N+1) = R(N)+1$ since $\tilde{\ell}(a'+1) = \tilde{\ell}(a')+1$. (2) From $N = -1+(N+1)$ it follows that $R(N) \leq \tilde{\ell}(N+1)$; the latter number is at most $i-1$ by hypothesis and (2) follows. (3) There exists $i \in \{1, \ldots, g\}$ such that $\gamma_i \leq N+1 < \gamma_{i+1}$. Then by (2), $R(N) \leq i$, and the latter number is at most $(N+3)/2$ by Lemma 10.35(2). (4) Let $R(N) = \tilde{\ell}(a) + \tilde{\ell}(b)$ with $a \leq b = N - a$ and suppose that $a \in GS(\mathcal{X})$. We have $\tilde{\ell}(a-1) = \tilde{\ell}(a) - 1$, and $\tilde{\ell}(b+1) \leq \tilde{\ell}(b) + 1$. Then $\tilde{\ell}(a-1)+\tilde{\ell}(b+1) \leq \tilde{\ell}(a)+\tilde{\ell}(b) = R(N)$ and thus $R(N) = \tilde{\ell}(a-1)+\tilde{\ell}(b+1)$. If $a-1$ is a gap of $\mathcal{X}$, then we are done; otherwise we repeat the above argument. □

Remark 10.38. From Lemma 10.37(1)(3), we have that $R(N) \leq R(2g-2) \leq g$ whenever $N \in \mathbb{N}' \cap [-1, 2g-2]$. Then Theorem 10.36 yields the main result in the previous section, namely $s[\mathcal{C}] \geq w(\mathcal{C}) - g$, provided that $2k \leq n$ and $n > 2g$. We are going to improve this result via Proposition 10.41 and Theorem 10.42 below.

Lemma 10.39. *Let $i \in \mathbb{N}', N \in \mathbb{N}' \cap [-1, 2g-2]$ and $r \in \mathbb{N}$ with $i + r \leq N+1$. If $A = \{i, i+1, \ldots, i+r\} \subseteq \mathbb{N}'$ is a set of $r+1$ consecutive integers such that $i+1, \ldots, i+r$ are gaps of $\mathcal{X}$, then* $\min\{\tilde{\ell}(a) + \tilde{\ell}(N-a) : a \in A\} = \tilde{\ell}(i+r) + \tilde{\ell}(N-i-r)$.

Proof. Let $a = i + j$ with $1 \leq j \leq r$. By hypothesis a is a gap of $\mathcal{X}$, hence $\tilde{\ell}(a) = \tilde{\ell}(i)$. Then $\tilde{\ell}(a) + \tilde{\ell}(N-a)$ is minimum when $\tilde{\ell}(N-a)$ is so. Since $\tilde{\ell}$ is an increasing function, this happens when a is the largest element in A. □

Proposition 10.40. *For $N \in \mathbb{N}' \cap [-1, 2g-2]$, we have*

$$R(N) = \min\{\tilde{\ell}(a) + \tilde{\ell}(N-a) : -1 \leq a \leq N/2, a = \lfloor N/2 \rfloor \text{ or } a \in \mathbb{N}' \setminus GS(\mathcal{X}) \text{ with } a+1 \in GS(\mathcal{X})\}.$$

Proof. By Lemma 10.37(4), $R(N) = \tilde{\ell}(a) + \tilde{\ell}(N-a)$ for some gap a of $\mathcal{X}$ such that $a \leq N/2$. Suppose that $a < \lfloor N/2 \rfloor$. If each integer a' with $a < a' \leq \lfloor N/2 \rfloor$ is a gap of $\mathcal{X}$, then from Lemma 10.39 $R(N) = \tilde{\ell}(\lfloor N/2 \rfloor) + \tilde{\ell}(\lceil N/2 \rceil)$; otherwise, by Lemma 10.39 again, we can assume $a+1 \in G(\mathcal{X})$ and the result follows. □

According to Remark 10.38, it is useful to compute $R(2g-2)$. The result is the following.

Proposition 10.41. *Let* $\mathrm{GS}(\mathcal{X}) = (\gamma_i)$ *be the gonality sequence of* $\mathcal{X}$. *Then* $R(2g-2) = g - \max\{\gamma_i - 2(i-2) | i = 1, \ldots, g\}$.

Proof. By definition of R, $R(2g-2) = \min\{\tilde{\ell}(a) + \tilde{\ell}(2g-2-a) \mid -1 \leq a \leq 2g-1, a+1 \in \mathrm{GS}(\mathcal{X})\}$. Write $a = \gamma_i - 1$, $1 \leq i \leq g$. Then, since $\tilde{\ell}(\gamma_i - 1) = i-1$ and, according to Corollary 10.3, $\tilde{\ell}(2g-\gamma_i-1) = g-\gamma_i+i-1$, we obtain the result. □

Now Theorem 10.36, Remark 10.38 and the above computation of $R(2g-2)$ imply the following.

Theorem 10.42. *Let* $\mathcal{C} = \mathcal{C}(\mathcal{X}, D, G)$ *be an AG code such that* $2k \leq n$ *and* $n > 2g$, *where* g *is the genus of* $\mathcal{X}$. *Let* γ_2 *be the gonality of* $\mathcal{X}$ *over* $\mathbb{F}_q$. *Then* $s[\mathcal{C}] \geq w(\mathcal{C}) - g + \gamma_2 - 2$.

In the remaining part of this section we study the function R on a plane curve $\mathcal{X}$ of degree $r+1$. In this case the genus of $\mathcal{X}$ is $g = r(r-1)/2$ and its gonality sequence $\mathrm{GS}(\mathcal{X})$ is obtained from the semigroup generated by r and $r+1$ (cf. Proposition 10.4). For an integer $a \in \mathbb{N}_0$, let α and β be the non-negative integers defined by

$$a = \alpha r + \beta\,, \qquad 0 \leq \beta < r\,.$$

It is clear that $a \in \mathrm{GS}(\mathcal{X})$ if and only $\beta \leq \alpha$.

Lemma 10.43.

$$\tilde{\ell}(a) = \frac{\alpha(\alpha+1)}{2} + \min\{\alpha, \beta\} + 1\,.$$

Proof. If $a = 0$, the formula is true so let $a > 0$. Suppose first that $a \in \mathrm{GS}(\mathcal{X})$ so that $\min\{\alpha, \beta\} = \beta$. Then $\tilde{\ell}(a) = 1 + 2 + \ldots + \alpha + \beta + 1$ and we the claimed formula follows. Now let a be a gap of $\mathcal{X}$ so that $\beta > \alpha$. We have $\tilde{\ell}(a) = \tilde{\ell}(\alpha r + \alpha)$ and the result follows by applying the above computation to $\alpha r + \alpha \in \mathrm{GS}(\mathcal{X})$. □

Lemma 10.44. *Let* $N \in \mathbb{N}' \cap [-1, 2g-2]$ *and* $a = \alpha r + \beta$ *a gap of* $\mathcal{X}$ *with* $\alpha \geq 1$ *such that* $a \leq N/2$. *Then* $\tilde{\ell}(a) + \tilde{\ell}(N-a) \leq \tilde{\ell}(a-r) + \tilde{\ell}(N-(a-r))$.

Proof. Set $a' = a - r$, that is, $a' = (\alpha - 1)r + \beta$. Let $b = N - a = \delta r + \epsilon$, with $0 \leq \epsilon < r$ so that $b' = N - a' = (\delta + 1)r + \epsilon$. From Lemma 10.43 we have $\tilde{\ell}(a) - \tilde{\ell}(a') = \alpha + 1$ and $\tilde{\ell}(b) - \tilde{\ell}(b') \leq -\delta - 1$. Now the result follows since $a \leq b = N - a$ implies $\delta \geq \alpha$. □

Thus Proposition 10.40 for the case of a plane curve becomes as follows.

Proposition 10.45. *For $N \in \mathbb{N}_0 \cap [0, 2g-2]$ let α and β be the integers defined by $\lfloor N/2 \rfloor = \alpha r + \beta$ with $0 \leq \beta < r$. Assume that $\alpha \geq 1$.*

(1) *If $\lfloor N/2 \rfloor$ is a gap of $\mathcal{X}$, then*

$$R(N) = \min\{\tilde{\ell}(\lfloor \frac{N}{2} \rfloor) + \tilde{\ell}(\lceil \frac{N}{2} \rceil), \tilde{\ell}(\alpha r - 1) + \tilde{\ell}(N - \alpha r + 1)\}.$$

(2) *If $\lfloor N/2 \rfloor \in GS(\mathcal{X})$, then $R(N) = \tilde{\ell}(\alpha r - 1) + \tilde{\ell}(N - \alpha r + 1)$.*

Proof. It follows from Proposition 10.40 and Lemma 10.44. □

To improve this result we shall introduce the notion of "jump". An integer N with $0 \leq N \leq 2g-2$ is called *a jump of $\mathcal{X}$* whenever $R(N) > R(N-1)$ (so, if $R(N) = R(N-1) + 1$ by Lemma 10.37(1)). We denote by $U(\mathcal{X})$ the set of jumps of $\mathcal{X}$. Clearly $\#U(\mathcal{X}) = R(2g-2)$ and this number can be computed via the above Proposition. More precisely the following holds.

Lemma 10.46. *Let $\mathcal{X}$ be a (non-singular) plane curve of degree $r+1$. Then*

(1) $\#U(\mathcal{X}) = \begin{cases} r^2/4 & \text{if } r \text{ is even,} \\ (r^2-1)/4 & \text{if } r \text{ is odd;} \end{cases}$

(2) $U(\mathcal{X}) = \{\alpha r + \beta \mid -1 \leq \alpha \leq r-1,\ 0 \leq \beta \leq r-1, \text{ and } 2\beta + 2 \leq \alpha, \text{ or } \beta = r-1\} \setminus \{2g-1\}$.

Proof. (1) Let us compute $R(2g-2)$. If r is even, then $g - 1 = (r-2)(r+1)/2 = (r-2)r/2 + (r-2)/2$ and thus it belongs to $\mathrm{GS}(\mathcal{X})$. By Proposition 10.45, $R(2g-2) = \tilde{\ell}(\alpha r + 1) + \tilde{\ell}(2g - 2 - \alpha r + 1)$ with $\alpha = (r-2)/2$. Now the result follows by applying Lemma 10.43. The case r odd is similar. (2) Let us denote by T the set of the right-hand side in the equality in Item (2). We claim that $\#T = R(2g-2)$. Indeed $\#T = \sum_{\beta=0}^{\lfloor (r-4)/2 \rfloor} (r - 2\beta - 3) + r - 1 = R(2g-2)$. Therefore it is enough to show that $T \subseteq U(\mathcal{X})$. From Proposition 10.45 and Lemma 10.43 it is easily seen that all elements in $U(\mathcal{X})$ are jumps. Then the proof is complete. □

Finally the promised improved description of $R(N)$ in the case of plane curves is as follows.

Proposition 10.47. *Let $\mathcal{X}$ be a (non-singular) plane curve of degree $r+1$ and $N \in \mathbb{N}' \cap [-1, 2g-2]$. Let α and β be the integers defined by $N = \alpha r + \beta$ with $0 \leq \alpha \leq r-2$ and $-1 \leq \beta \leq r-2$.*

(1) *If* $\beta > \lfloor \alpha/2 \rfloor - 1$, *then* $R(N) = R(\alpha r + \lfloor \alpha/2 \rfloor - 1)$;
(2) *If* $\beta \leq \lfloor \alpha/2 \rfloor - 1$, *then*

$$R(N) = \begin{cases} \alpha(\alpha+2)/4 + \beta + 1 & \text{if } \alpha \text{ is even,} \\ (\alpha+1)^2/4 + \beta + 1 & \text{if } \alpha \text{ is odd.} \end{cases}$$

Proof. (1) In this case $\alpha r + \lfloor \alpha/2 \rfloor - 1$ is the largest jump of $\mathcal{X}$ not exceeding N and (1) follows. (2) Here place all the integers from -1 to $2g-2$ in an array according to the corresponding values of α and β. The j-th row of the array contains $\lfloor (j+2)/2 \rfloor$ jumps of $\mathcal{X}$ which are precisely the ones in the first $\lfloor (j+2)/2 \rfloor$ columns of the array. The number of jumps from -1 to N is: $\beta + 1$ in the row α plus $2(\sum_{i=1}^{(\alpha-1)/2} i) = \alpha(\alpha+2)/2$ if α is even, and $\beta + 1$ in row α plus $2(\sum_{i=1}^{(\alpha-2)/2} i) = (\alpha+1)^2/4$ if α is odd. □

Example 10.48. Let us consider the Hermitian curve $\mathcal{H}$, introduced in example 10.11, and the Hermitian code $\mathcal{C} = C(\mathcal{H}, D, \rho Q)$, $\rho \in H(Q)$. Let us consider the particular case $q = 5$. Here $n = 125$ and $g = 10$. In order to apply Theorem 10.36, we have to choose ρ such that $-1 \leq 2m - 125 \leq 18$, that is $62 \leq \rho \leq 71$. One easily checks that $2k \leq n$ and $n > 2g$. The next table contains the results given by Theorem 10.36 and Proposition 10.47. The row 'Wolf' contains the Wolf upper bound on $s(\mathcal{C})$; the row 'true' contains the true values of $s(\mathcal{C})$, which have been computed by Blackmore and Norton in [4]. The row 'LB' contains the results obtained by using Theorem 10.36.

Table 10.1. Trellis state complexity of Hermitian codes over $\mathbb{F}_{25}$.

ρ	62	63	64	65	66	67	68	69	70	71
Wolf	53	54	55	56	57	58	59	60	61	62
LB	52	53	54	54	55	55	55	56	55	56
true	53	53	54	54	55	56	56	56	57	56

10.7. Bibliographical notes

The generalized Hamming weights were introduced by Helleseth, Kløve and Mykkeltveit in 1977, [18], in connexion with weight distributions. They were later rediscovered by Wei [38], motivated by their cryptographical applications, following the work of Ozarov and Wyner [26] on codes for wire-tap channels. More recent applications connect the GHWs with the trellis structure of codes (as seen in this chapter) or with the problem of list decoding from erasures, see [14].

Since the work of Wei the study of the GHWs has attracted considerable interest, and now the full hierarchy is known, or at least partial results are available, for many codes, including Hamming, cyclic, BCH, RS, RM, product codes, etc. See [7, 8, 12, 19, 39] (and many others).

The relevance of the gonality sequence in the study AG codes was first noticed by Pellikaan [27], who also found the gonality sequence of plane curves (Proposition 10.4) in [28]. This sequence was introduced by Yang, Kumar and Stichtenoth in [41], which is also the first article devoted to study the GHWs of AG codes (mainly coming from Hermitian curves). The results of that work were later improved and generalized by Munuera in [22]. Our exposition of the extension of the Goppa bound to all the GHWs is based in these two papers, [22, 41], see also [23]. Making use of these results and ideas, some authors have found the complete weight hierarchy of some AG codes, see for example [6] for hyperelliptic curves and [2] for Hermitian curves. The extension of the order bound to all GHWs was given by Heijnen and Pellikaan, [17]. In this work, the complete weight hierarchy of Reed-Muller codes was also determined. This line of research has been continued by several authors. For example, in a series of papers, Shibuya, Sakaniva and others, derived bounds to all linear codes, defined by their parity check matrices, see [30–32]. These ideas have been completed by Geil, [13]. In particular, the study of t-rank MDS codes is treated in [31].

A global geometric approach to the GHWs was given by Tsfasman and Vladut [35]. For a generalization of the weight hierarchy to the called *support weight distribution*, see [33]

Trellises were introduced by Forney in 1967, [9], in order to study the Viterbi algorithm [10] for decoding convolutional codes. After these articles, trellises became ubiquitous in the theory of convolutional codes. The use with block codes started with an article of Bahl, Cocke, Jelinek and Raviv, [1], following an unpublished work of Forney. This paper also presented a way to construct the trellis of a linear code (the BCJR construction). However the interest in trellis representations started later, with the observation, made by Wolf [40], that they can be used for maximum-likelihood decoding of linear codes with the Viterbi algorithm. The relation between the trellis complexity and the generalized Hamming weights was discovered in [21] and later studied in [11]. More details about the history, the bibliography and the theory of trellises, can be obtained in the excellent survey [37]

The study of the trellis structure of AG codes begun with the works of Shany and Be'ery [29], and mainly Blackmore and Norton, who in a series of

papers, [3–5], found the translation of past and future subcodes to algebraic geometric terms and stated some important properties. In particular, in [4] the state complexity of Hermitian codes is obtained. Our presentation in this chapter, follows the structure of [24] and [25].

Finally, for a general reference on AG codes, the reader is addressed to [20, 34, 36].

References

[1] L.R. Bahl, J. Cocke, F. Jelinek and J. Raviv, Optimal decoding of linear codes for minimizing symbol error rate, *IEEE Trans. Information Theory* **20**, 284-287 (1974).

[2] A. Barbero and C. Munuera, The weight hierarchy of Hermitian codes, *SIAM J. Discrete Math.* **13**, 79-104 (2000).

[3] T. Blackmore and G. Norton, On trellis structures for Reed-Muller codes, *Finite Fields and Appl.* **6**, 37-90 (2000).

[4] T. Blackmore and G. Norton, Determining when the absolute sate complexity of a Hermitian code achieves the DLP bound, *SIAM J. Discrete Math.* **15**, 14-40 (2001).

[5] T. Blackmore and G. Norton, Lower bounds on the state complexity of Geometric Goppa codes, *Codes, Des. and Cryptography* **25**, 95-115 (2002).

[6] M. de Boer, The generalized Hamming weights of some hyperelliptic codes, *J. Pure Appl. Algebra* **123**, 153-163 (1998).

[7] G. Cohen, S. Lytsin and G. Zémor, Upper bounds on the generalized distances, *IEEE Trans. Inform. Theory* **40**, 2090-2092 (1994).

[8] G.L. Feng, K.K. Tzeng and V.K. Wei, On the generalized Hamming weights for several classes of cyclic codes, *IEEE Trans. Information Theory* **38**, 1125-1130 (1992).

[9] D. Forney, Final report on a coding system design for advanced solar missions. Contract NAS2-3637, NASA Ames Research Center (1967).

[10] D. Forney, The Viterbi algorithm, *Proc. IEEE* **61**, 268-278 (1973).

[11] D. Forney, Dimension/Length profiles and trellis complexity of linear block codes, *IEEE Trans. Inform. Theory* **40**, 1741-1752 (1994).

[12] G. van der Geer and M. van der Vlugt, On generalized Hamming weights of BCH codes, *IEEE Trans. Inform. Theory* **40**, 913-920 (1994).

[13] O. Geil and C. Thommesen, *On the Feng-Rao bound for Generalized Hamming weights*. In ed. M. Fossorier, *AAECC 2006*, pp. 295-3006 (Springer-Verlag, LNCS 3857, Berlin 2006).

[14] V. Guruswami, List decoding from erasures: bounds and code constructions, *IEEE Trans. Inform. Theory* **49**, 2826–2833 (2003).

[15] R. Hartshorne, *Algebraic Geometry*. (Springer-Verlag, GTM-52, New York, 1987).

[16] R. Hartshorne, Generalized divisors on Gorenstein curves and a Theorem of Noether, *J. Math Kyoto Univ.* **26**, 375-386 (1986).

[17] P. Heijnen and R. Pellikkaan, Generalized Hamming weights of q-ary Reed-Muller codes, *IEEE Trans. Inform. Theory* **44**, 181-196 (1998).

[18] T. Helleseth, T. Kløve and J. Mykkeltveit, The weight distribution of irreducible cyclic codes with block lengths $n_1(q^l - 1/N)^n$, *Discrete Math.* **18**, 179-211 (1977).

[19] T. Helleseth, T. Kløve and Ø. Ytrehus, Generalized Hamming weights of linear codes, *IEEE Trans. Inform. Theory* **38**, 1133-1140 (1992).

[20] T. Høholdt, J.H. van Lint and R. Pellikaan, *Algebraic Geometry codes.* In eds V. Pless, W.C. Huffman and R.A Brualdi, *Handbook of Coding Theory*, pp. 871-961 (Elsevier, Amsterdam, 1998).

[21] T. Kasami, T. Takata, T. Fujiwara and S. Lin, On complexity trellis structure of linear block codes, *IEEE Trans. Inform. Theory* **39**, 1057-1064 (1993).

[22] C. Munuera, On the generalized Hamming weights of Geometric Goppa codes, *IEEE Trans. Inform. Theory* **40**, 2092-2099 (1994).

[23] C. Munuera and J.I. Farrán, Goppa-like bounds for the generalized Hamming distances, *Discrete Applied Math.* **128**, 145-156 (2003).

[24] C. Munuera and F. Torres, A Goppa-like bound on the trellis state complexity of algebraic geometric codes, *IEEE Trans. Inform. Theory* **49**, 733-737 (2003).

[25] C. Munuera and F. Torres, Bounding the trellis state complexity of algebraic geometric codes, *Applicable Algebra Eng. Comm. Computing* **15**, 81-100 (2004).

[26] L. Ozarow and A.D. Wyner, Wire-tap channel II, *Bell Labs Tech. J.* **63**, 2135-2157 (1984).

[27] R. Pellikaan, *On the gonality of curves, abundant codes and decoding.* In eds. H. Stichtenoth and M. Tsfasman, *Coding Theory and Algebraic Geometry*, pp. 132-144 (Springer-Verlag, LNM-1518, Berlin 1992).

[28] R. Pellikaan, *On special divisors and the two-variable zeta function of algebraic curves over finite fields.* In eds. R. Pellikaan, M. Perret and S. Vladut, *Arithmetic, Geometry and Coding Theory*, pp. 175-184 (de Gruyter, Berlin 1996).

[29] Y. Shany and Y. Be'ery, Bounds on the state complexity of codes from the Hermitian function field, *IEEE Trans. Inform. Theory* **46**, 1523-1527 (2000).

[30] T. Shibuya, J. Mizutani and K. Sakaniwa, On generalized Hamming weights of codes constructed from affine algebraic sets, Proc. AAECC-12, Springer-Verlag, Lecture Notes in Computer Science 1255 (1993), pp. 231-253.

[31] T. Shibuya, R. Hasagawa and K. Sakaniwa, A lower bound for the generalized Hamming weights and a condition for t-rank MDS, *IEICE Trans. Fundamentals* **E-82-A**, 1979-1989 (1999).

[32] T. Shibuya and K. Sakaniwa, A note on a lower bound for generalized Hamming weights, *IEICE Trans. Fundamentals* **E-84-A**, 3138-3145 (2001).

[33] J. Simonis, The effective length of subcodes, *Appl. Algebra Eng. Commun. Comput.* **5**, 371-377 (1994).

[34] H. Stichtenoth, *Algebraic function fields and codes.* (Springer-Verlag, Berlin, 1993).

[35] M. Tsfasman and S. Vladut, Geometric approach to higher weights, *IEEE Trans. Inform. Theory* **41**, 1564-1588 (1995).

[36] M. Tsfasman and S. Vladut, *Algebraic-Geometric codes.* (Kluwer Academic Publishers, Dordrecht-Boston-London, 1991).

[37] A. Vardy, *Trellis structure of codes.* In eds. V.S.Pless, W.C. Huffman and R.A. Brualdi, *Handbook of Coding Theory*, pp. 1981-2117 (Elsevier, Amsterdam, 1998).

[38] V.K. Wei, Generalized Hamming weights for linear codes, *IEEE Trans. Inform. Theory* **40**, 1412-1418 (1995).

[39] V.K. Wei and K. Yang, On the generalized Hamming weights of product codes, *IEEE Trans. Information Theory* **39**, 1709-1713 (1993).

[40] J.K. Wolf, Efficient maximum-likelihood decoding of linear block codes using a trellis, *IEEE Trans. Information Theory* **24**, 76-80 (1978).

[41] K. Yang, P.V. Kumar and H. Stichtenoth, On the weight hierarchy of geometric Goppa codes, *IEEE Trans. Inform. Theory* **24**, 913-920 (1994).

This page intentionally left blank

Chapter 11

Algebraic Geometry Constructions of Convolutional Codes

J.A. Domínguez Pérez, J.M. Muñoz Porras and G. Serrano Sotelo

Department of Mathematics
University of Salamanca
Plaza de la Merced 1-4, 37008 Salamanca, Spain
{jadoming,jmp,laina}@usal.es

Algebraic-geometric techniques to construct linear codes can be applied to construct convolutional codes, using algebraic curves over function fields. In this way we construct convolutional Goppa codes and provide a systematic way for constructing convolutional codes with prescribed properties. We study convolutional Goppa codes defined by the projective line and elliptic curves in detail.

Contents

11.1. Introduction

The notion of convolutional codes was introduced by Peter Elias [4] in 1955, considering the codification as a time-dependent process: the *codified word* at some instant depends not only on the *information word* at that instant, but also on the previous words; the number of the previous words

on which the codifications depends is called the *memory* of the codifier; in this scheme the codification of a word can be interpreted as a certain *convolution* with other words.

The use of convolutional codes was very important after the discovery by Andrew Viterbi [19] in 1967 of the decodification algorithm known by his name. The compatibility of this algorithm depends exponentially on the memory of the codifier.

One of the main applications of convolutional codes is the transmission of information through the *deep-space*, where there are strong limitations of potency, but in general no restrictions to the wide band. The communication systems used in artificial satellites transmit telemetric information: orders from earth stations to satellites and tracking. In the telemetric channel, where the rate of the code is relatively high, convolutional and block codes are used.

For instance, the space missions *Pioneer* 10 and 11 to Jupiter and Saturn in 1972-73 used a convolutional code of rate 1/2 and memory 31. After Viterbi, a *planetary standard* as a convolutional code of rate 1/2 and memory 6 was implemented. This code was used for the first time in the *Voyager 1* mission (1980-81), concatenated with a Reed-Solomon code, and in the *Galileo* (1986) and *Voyager 2* missions (1989), concatenated with other convolutional and block codes.

Convolutional codes are also used in the construction of *turbo-codes*, introduced in 1983 by Berrou, Glavieux and Thitimajshima [1]. These are the codes currently used in wireless communications.

There are different approaches to the study of convolutional codes: they can be studied as sequential circuits, as discrete linear systems, etc. From an algebraic point of view, the fundamental reference is the work of G. David Forney Jr. [5] in 1970. Then came the works of Robert J. McEliece [11] in 1977 and Philippe Piret in 1988 [14]. More recently, McEliece [12] in 1998 again gave an introduction to the algebraic theory of convolutional codes, which clarifies the previous approaches.

The recent work of Joachim Rosenthal, Roxana Smarandache [16] and Heide Gluesing-Luerssen [17] in 1999 and 2001, Vakhtang Lomadze [9] in 2001, and the authors of this chapter [2] in 2004, has shown that the use of techniques of Algebraic Geometry are very useful in the study of convolutional codes.

Our contribution continues with five sections. In §2 we give an introduction to the general theory of convolutional codes. Section §3 is devoted to constructing convolutional Goppa codes in terms of algebraic curves de-

fined over the field of rational functions in one variable over a finite field (see [13]).

In §4 we analyze the notion of weight and free distance for convolutional codes, and their possible geometric interpretation in the case of convolutional Goppa codes.

Finally, §5 and §6 are devoted to studying some explicit examples of convolutional Goppa codes defined by the projective line and elliptic curves.

11.2. Convolutional Codes

Given a finite field $\mathbb{F}_q$, representing the symbols in which an information word $u \in \mathbb{F}_q^k$ is written, a *linear block encoder* is essentially an injective linear map

$$\begin{aligned} G\colon \mathbb{F}_q^k &\hookrightarrow \mathbb{F}_q^n \\ u &\mapsto x = u \cdot G \end{aligned} \quad ,$$

whose image subspace is the *linear code* $\mathcal{C}_k = \operatorname{Im} G \subset \mathbb{F}_q^n$.

The map G is a $k \times n$ matrix with entries in $\mathbb{F}_q$, which is called a *generator matrix* of the code (their rows are a basis of $\mathcal{C}_k$); k is the *dimension* of the code, and n is its *length*. Alternatively, one can define the code $\mathcal{C}_k$ by its implicit equations

$$\mathcal{C}_k = \{x \in \mathbb{F}_q^n \,/\, H \cdot x = 0\}\,,$$

where H is an $(n-k) \times n$ matrix with entries in $\mathbb{F}_q$, called a *parity-check (control) matrix* of the code.

In practical applications, the codification process is not limited to a single word, but a sequence of information words depending on time $u(t) \in \mathbb{F}_q^k$, $t = 0, 1, 2, \ldots$, which after the codification are transformed in the sequence of codified words

$$x(t) = u(t) \cdot G \in \mathcal{C}_k \in \mathbb{F}_q^n\,, \quad t = 0, 1, 2, \ldots$$

The codified word $x(t)$ at the instant t depends only on the information word $u(t)$ at the same instant t.

The basic idea of convolutional codification is to allow $x(t)$ to depend not only on $u(t)$ but also on $u(t-1), \ldots, u(t-m)$ for some positive integer m, which is the *memory* of the code. One can then consider a linear block code as a convolutional code with zero memory.

Let us explain this with greater precision. One can write a sequence of words as a polynomial in one variable z whose coefficients are the sequence,

$$U(z) = \sum_t u(t) z^t \in \mathbb{F}_q[z]^k \,,$$

and the product by z^i can be considered as a *delay operator*

$$z^i U(z) = \sum_t u(t) z^{t+i} = \sum_t u(t-i) z^t \in \mathbb{F}_q(z)^k \,.$$

One can now define a convolutional (polynomial) encoder as an injective homomorphism of $\mathbb{F}_q[z]$-modules

$$\begin{aligned} G(z) \colon \mathbb{F}_q[z]^k &\hookrightarrow \mathbb{F}_q[z]^n \\ U(z) &\mapsto X(z) = U(z) \cdot G(z) \end{aligned} \,,$$

where $G(z)$ is a $k \times n$ matrix with entries in $\mathbb{F}_q[z]$.

If we allow the possibility of performing *feedback*, this means that we can reverse the delay, and define convolutional codification more generally over $\mathbb{F}_q(z)$, the field of fractions of $\mathbb{F}_q[z]$, i.e., the localization of the ring $\mathbb{F}_q[z]$ with respect to the multiplicative system $S = \{Q(z) \in \mathbb{F}_q[z] / Q(z) \neq 0\}$. Thus, a convolutional encoder is an injective $\mathbb{F}_q(z)$-linear map

$$G(z) \colon \mathbb{F}_q(z)^k \hookrightarrow \mathbb{F}_q(z)^n \,,$$

with the entries of $G(z)$ also in $\mathbb{F}_q(z)$.

Definition 11.1. An (n, k) *convolutional code* $\mathcal{C}_k$ over $\mathbb{F}_q$ is a linear subspace of dimension k over $\mathbb{F}_q(z)$ of $\mathbb{F}_q(z)^n$.

The integers (n, k) are called, respectively, the *length* and *dimension* of the convolutional code, and n/k is the *ratio* of $\mathcal{C}_k \subseteq \mathbb{F}_q(z)^n$.

Example 11.2. The subspace of dimension 1 in $\mathbb{F}_2(z)^3$ defined by

$$\mathcal{C}_1 = \langle 1 + z, 1 + z^2, z + z^3 \rangle \,,$$

is a $(3, 1)$-convolutional code over $\mathbb{F}_2$, in which the code word $x(t)$ at the instant t is obtained in terms of the information words $u(t)$, $u(t-1)$, $u(t-2)$ and $u(t-3)$ in the following way:

$$x(t) = (u(t) + u(t-1), u(t) + u(t-2), u(t-1) + u(t-3)) \,.$$

11.2.1. *Convolutional encoders. Linear systems and circuits*

Definition 11.3. A *convolutional encoder*, or *codifier*, for a convolutional code $\mathcal{C}_k \subseteq \mathbb{F}_q(z)^n$ is a $k \times n$ matrix

$$G(z) = \begin{pmatrix} G_1(z) \\ \vdots \\ G_k(z) \end{pmatrix},$$

where $G_i(z) = (G_{i1}(z), \ldots, G_{in}(z)) \in F(z)^n$ are a basis of $\mathcal{C}_k$.

Equivalently, $G(z)$ is a *generator matrix* that defines an injective linear *encoding map*:

$$\begin{aligned} G(z)\colon F(z)^k &\hookrightarrow F(z)^n \\ U(z) &\mapsto X(z) = U(z) \cdot G(z), \end{aligned}$$

such that $\operatorname{Im} G(z) = \mathcal{C}_k \subset F(z)^n$.

Given two codifiers $G(z)$ and $G'(z)$ of the same convolutional code $\mathcal{C}_k$, there exists an element (the *base change*) $B(z) \in GL(k, \mathbb{F}_q(z))$ of the linear group of dimension k over $\mathbb{F}_q(z)$ such that:

$$G'(z) = B(z) \cdot G(z).$$

Example 11.4. The convolutional code $\mathcal{C}_1$ defined in example 11.2 has the following different encoders

$$\begin{aligned} G(z) &= (\ 1+z, 1+z^2,\ z+z^3\ ) \\ G'(z) &= (\ z,\ z+z^2, z^2+z^3\ ) \\ G''(z) &= (\ 1,\ 1+z,\ z+z^2\ ) \\ G'''(z) &= (\ \tfrac{1}{1+z},\ 1,\ z\ ) \end{aligned}$$

and the following identities are satisfied

$$G(z) = \frac{1+z}{z} \cdot G'(z) = (1+z) \cdot G''(z) = (1+z^2) \cdot G'''(z).$$

Definition 11.5. A *polynomial encoder* is a convolutional encoder $G(z)$ whose entries are in $\mathbb{F}_q[z]$.

If $G_i(z)$ is the i-th row of a polynomial encoder $G(z)$, the *degree* of $G_i(z)$, denoted $e_i = \operatorname{Degree} G_i(z)$, is the highest of the degrees of its components

$$e_i = \max_{1 \le j \le n} (\operatorname{Degree} G_{ij}(z)).$$

The *memory* of a polynomial encoder $G(z)$, denoted m_G, is the maximun degree of its rows

$$m_G = \max_{1 \leq i \leq k} e_i \,.$$

The *degree* of a polynomial encoder $G(z)$, denoted δ_G, is the sum of the degrees of its rows

$$\delta_G = \sum_{i=1}^{k} e_i \,.$$

Remark 11.6. By reordering the rows of $G(z)$, one can henceforth assume that

$$e_1 \leq \cdots \leq e_k = m_G \,.$$

Every convolutional code has polynomial encoders: if $G(z)$ is an arbitrary encoder and $\mu(z)$ the least common multiple of the denominators of its coefficients, $\mu(z)\, G(z)$ is a polynomial encoder of the same code. Thus, for any encoder $G(z)$ one can consider the *degree* δ_G as the degree of the polynomial encoder $\mu(z)\, G(z)$. In particular, *linear block codes* are convolutional codes for which there are *zero degree* encoders.

Definition 11.7. A convolutional encoder $G(z)$ is called *realizable* if the denominators of its entries are not multiples of z.

The notion of realizable encoder is related to the possibility of describing the encoder as a *linear system* and therefore the possibility of constructing a *physical circuit* which performs the encoding process. Given a $k \times n$ realizable encoder $G(z)$ of degree δ_G, it can be decomposed as

$$G(z) = D + E(z) \cdot C \,,$$

where $D = G(0)$ is a $k \times n$ matrix with entries in $\mathbb{F}_q$; C is a $\delta_G \times n$ matrix with entries in $\mathbb{F}_q$, and $E(z)$ is a $k \times \delta_G$ matrix defining the *morphism of states* with values in the *space of state-variables* $\mathbb{F}_q[z]^{\delta_G}$

$$\begin{aligned} E(z) \colon \mathbb{F}_q[z]^k &\to \mathbb{F}_q[z]^{\delta_G} \\ U(z) &\mapsto S(z) = U(z) \cdot E(z) \,. \end{aligned}$$

With these notations, the encoding morphism can be expressed in terms of the *state-variables* $S(z)$ as:

$$\begin{aligned} X(z) &= U(z) \cdot G(z) = S(z) \cdot C + U(z) \cdot D \,. \\ z^{-1} S(z) &= U(z) \cdot z^{-1} E(z) \,. \end{aligned}$$

Denoting $B = (z^{-1}E(z))|_{z=0}$ and decomposing $z^{-1}E(z) = B + E(z) \cdot A$, where A is a $\delta_G \times \delta_G$ matrix with entries in $\mathbb{F}_q$, one obtains the following identity:

$$z^{-1}S(z) = U(z) \cdot (E(z) \cdot A + B) = S(z) \cdot A + U(z) \cdot B\,.$$

Thus, a convolutional realizable encoder is equivalent to a *time-invariant linear system with a finite number of state variables.* The *state space description* of the encoder is given by

$$\left.\begin{aligned} s(t+1) =& s(t) \cdot A_{\delta_G \times \delta_G} + u(t) \cdot B_{k \times \delta_G} \\ x(t) =& s(t) \cdot C_{\delta_G \times n} + u(t) \cdot D_{k \times n} \end{aligned}\right\}. \tag{11.1}$$

By solving the states in the first equation, one obtains $x(t)$ as a function of a set $u(t), u(t-1), \ldots$. If this set is infinite, we say that the encoder has *infinite memory*, whereas if it is finite $u(t), u(t-1), \ldots, u(t-m_G)$, we call m_G the *memory* of the realizable encoder $G(z)$.

Example 11.8. The encoder $G'''(z) = (\frac{1}{1+z}, 1, z)$ of example 11.4 is realizable and its degree is $\delta_{G'''} = 2$. Applying the above description, one obtains:

$$\begin{aligned} D =& G'''(z)|_{z=0} = (1,1,0) \\ G'''(z) - D =& \left(\frac{z}{1+z}, 0, z\right) = \frac{1}{1+z}\left(z, 0, z+z^2\right) \\ =& \underbrace{\frac{1}{1+z}\left(z, z^2\right)}_{E(z)} \cdot \underbrace{\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}}_{C} \\ B =& (z^{-1}E(z))|_{z=0} = (1,0) \\ z^{-1}E(z) - B =& \frac{1}{1+z}(z,z) = \underbrace{\frac{1}{1+z}(z,z^2)}_{E(z)} \cdot \underbrace{\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}}_{A} \end{aligned}$$

and finally the state equations are:

$$\left.\begin{aligned} (s_1(t+1), s_2(t+1)) =& (s_1(t), s_2(t)) \cdot \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + u_1(t) \cdot (1,0)\,, \\ (x_1(t), x_2(t), x_3(t)) =& (s_1(t), s_2(t)) \cdot \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} + u_1(t) \cdot (1,1,0) \end{aligned}\right\}.$$

Remark 11.9. Let us consider the equations of the encoder of the above example 11.8

$$\left.\begin{aligned}(s_1(t+1), s_2(t+1)) =& (u_1(t) + s_1(t), s_1(t)) \,, \\ (x_1(t), x_2(t), x_3(t)) =& (u_1(t) + s_1(t), u_1(t), s_1(t) + s_2(t)) \,.\end{aligned}\right\}$$

The state s_1 depends recurrently on itself; that is, there is *feed-back*, and by substituting the first equation in the second one obtains

$$\begin{aligned}x_1(t) = u_1(t) + s_1(t) =& u_1(t) + u_1(t-1) + s_1(t-1) = \\ =& u_1(t) + u_1(t-1) + u_1(t-2) + s_1(t-2) = \\ =& u_1(t) + u_1(t-1) + u_1(t-2) + u_1(t-3) + \dots\end{aligned}$$

Accordingly, the code word depends indefinitely on the information word, which means that the encoder has infinite memory, although its degree is finite.

The generator matrix of an encoder can be recovered from its expression as a linear system: writing equations (11.1) as

$$\left.\begin{aligned}z^{-1}S(z) =& S(z) \cdot A + U(z) \cdot B \\ X(z) =& S(z) \cdot C + U(z) \cdot D\end{aligned}\right\} \,,$$

and eliminating the state variables, one concludes:

$$G(z) = B \cdot (z^{-1} \operatorname{Id}_{\delta_G} - A)^{-1} \cdot C + D \,.$$

Example 11.10. Let us consider the matrices of the linear system defined in example 11.8:

$$A = \begin{pmatrix}1 & 1 \\ 0 & 0\end{pmatrix} \,, \quad B = \begin{pmatrix}1 & 0\end{pmatrix} \,, \quad C = \begin{pmatrix}1 & 0 & 1 \\ 0 & 0 & 1\end{pmatrix} \,, \quad D = \begin{pmatrix}1 & 1 & 0\end{pmatrix} \,.$$

One has $\left(z^{-1} \operatorname{Id}_2 - A\right)^{-1} = \frac{1}{1+z} \begin{pmatrix}z & z^2 \\ 0 & z + z^2\end{pmatrix}$, and hence:

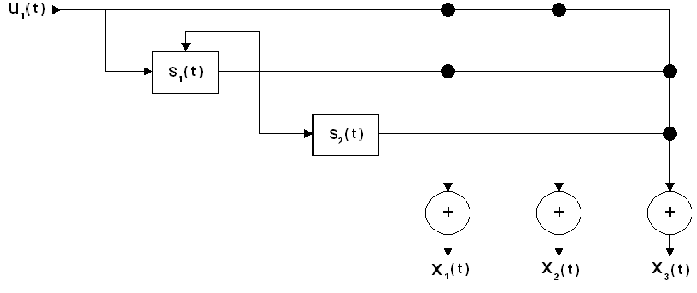
$$\underbrace{\frac{1}{1+z} \begin{pmatrix}1 & 0\end{pmatrix} \cdot \begin{pmatrix}z & z^2 \\ 0 & z + z^2\end{pmatrix}}_{E(z)} \cdot \begin{pmatrix}1 & 0 & 1 \\ 0 & 0 & 1\end{pmatrix} + \begin{pmatrix}1 & 1 & 0\end{pmatrix} = \left(\tfrac{1}{1+z}, 1, z\right) = G'''(z) \,.$$

The description of realizable encoders as linear systems allows us to express them as physical devices called *sequential circuits*, composed of *memory boxes* for delay operations (as many as the memory of the encoder) and *sum boxes* for addition operations.

Example 11.11. Let us consider the encoder of example 11.8. Its expression as linear systems is given by the equations:

$$\left.\begin{aligned} s_1(t+1) &= u_1(t) + s_1(t) \\ s_2(t+1) &= s_1(t) \end{aligned}\right\}, \qquad \left.\begin{aligned} x_1(t) &= u_1(t) + s_1(t) \\ x_2(t) &= u_1(t) \\ x_3(t) &= s_1(t) + s_2(t) \end{aligned}\right\},$$

which correspond to the circuit:



11.2.2. *Basic encoders. Degree of a convolutional code*

Let $\mathcal{C}_k \subseteq \mathbb{F}_q(z)^n$ be a (n,k)-convolutional code. Any polynomial encoder $G(z)$ for $\mathcal{C}_k$ induces a morphism of $\mathbb{F}_q[z]$-modules:

$$\phi_G \colon \mathbb{F}_q[z]^k \hookrightarrow \mathbb{F}_q[z]^n \,,$$

whose localization with respect to the multiplicative system $S = \{Q(z) \in \mathbb{F}_q[z] / Q(z) \neq 0\}$ is the encoding map:

$$G(z) = (\phi_G)_S \colon \mathbb{F}_q(z)^k \hookrightarrow \mathbb{F}_q(z)^n \,.$$

The existence of a *decoding map* $G^{-1}(z) \colon \mathbb{F}_q(z)^n \to \mathbb{F}_q(z)^k$ retract of $G(z)$ (i.e. $G(z) \cdot G^{-1}(z) = \mathrm{Id}_k$) can be deduced from the exact sequence:

$$0 \to \mathbb{F}_q[z]^k \xrightarrow{\phi_G} \mathbb{F}_q[z]^n \to \mathbb{F}_q[z]^n / \operatorname{Im} \phi_G \to 0 \,.$$

For each multiplicative system $S \subset \mathbb{F}_q[z]$, one has the exact sequence:

$$0 \to \mathbb{F}_q[z]_S^k \xrightarrow{(\phi_G)_S} \mathbb{F}_q[z]_S^n \to (\mathbb{F}_q[z]^n / \operatorname{Im} \phi_G)_S \to 0 \,.$$

The existence of a retract of $(\phi_G)_S$ is equivalent to saying that $(\mathbb{F}_q[z]^n/\operatorname{Im}\phi_G)_S$ is a free module isomorphic to $\mathbb{F}_q[z]_S^{n-k}$. Let us consider the decomposition of $\mathbb{F}_q[z]^n/\operatorname{Im}\phi_G$ as:

$$\mathbb{F}_q[z]^n/\operatorname{Im}\phi_G \simeq \mathbb{F}_q[z]^{n-k} \oplus T\,,$$

where T is the torsion submodule of $\mathbb{F}_q[z]^n/\operatorname{Im}\phi_G$. Then, $(\phi_G)_S$ is a retract if and only if $T_S = 0$. On the other hand,

$$T \simeq \mathbb{F}_q[z]/\langle\gamma_i(z)\rangle \oplus \cdots \oplus \mathbb{F}_q[z]/\langle\gamma_k(z)\rangle\,,$$

where $\gamma_i(z) \in \mathbb{F}_q[z]$ are the *invariant factors* of ϕ_G,

$$\gamma_i(z) = \Delta_i(z)/\Delta_{i-1}(z)\,,$$

where $\Delta_i(z)$ is the highest common divisor of the minors of order i of $G(z)$. Thus, the vanishing of T_S is equivalent to the condition of $\Delta_k(z)$ being invertible in $\mathbb{F}_q[z]_S$.

Definition 11.12. A polynomial encoder $G(z)$ is *non-catastrophic* if any of the following equivalent conditions are satisfied:

(1) $G(z)$ has a right inverse in $\mathbb{F}_q[z]_S$, where $S = \{z^l, l \geq 0\}$.
(2) $\Delta_k(z) = z^l$, $l \geq 0$.

The word *catastrophic* is used by Massey and Sain [10] to refer to encoders in which a code word $X(z)$ of finite length can be obtained from an information word $U(z)$ of infinite length, and hence the code word may contain infinite errors (*catastrophic errors*).

Example 11.13. Let us consider the encoder $G'(z) = (z, z + z^2, z^2 + z^3)$ of example 11.4. One has that $\Delta_1(z) = z$, and therefore $G'(z)$ is a non-catastrophic encoder, which has a (non-unique) right inverse, such as for instance:

$$(G')^{-1}(z) = \begin{pmatrix} \frac{1}{z} + P(z)(1+z) + Q(z)(1+z^2) \\ P(z) \\ Q(z) \end{pmatrix}\,,$$

where $P(z), Q(z)$ are arbitrary polynomials in $\mathbb{F}_q[z]$.

Definition 11.14. A polynomial encoder $G(z)$ is *basic* if any of the following equivalent conditions are satisfied:

(1) $G(z)$ has a right inverse in $\mathbb{F}_q[z]$.
(2) $\Delta_k(z) = 1$.

(3) $\gamma_1(z) = \cdots = \gamma_k(z) = 1$.

In particular, any basic encoder is non-catastrophic.

Example 11.15. The encoder $G''(z) = (1, 1+z, z+z^2)$ of example 11.4 satisfies the condition $\gamma_1(z) = 1$, and is therefore basic. A right inverse of $G''(z)$ is:

$$G^{-1}(z) = \begin{pmatrix} 1 + P(z)(1+z) + Q(z)(1+z^2) \\ P(z) \\ Q(z) \end{pmatrix} .$$

The existence of basic encoders for all convolutional codes was proved in a constructive way by Forney [5], using the Smith algorithm for the computation of invariant factors.

Theorem 11.16. *All convolutional codes admit basic encoders.*

Proof. Let $G(z)$ be a polynomial encoder for an (n,k)-convolutional code. The Smith algorithm allows us to compute the invariant factors $\gamma_1(z), \ldots, \gamma_k(z)$ of ϕ_G and two unimodular matrices $B(z) \in GL(k, \mathbb{F}_q[z])$, $C(z) \in GL(n, \mathbb{F}_q[z])$ such that:

$$B(z) \cdot G(z) \cdot C(z) = \big(\Gamma(z) \mid 0_{k\times(n-k)}\big) ,$$

where

$$\Gamma(z) = \begin{pmatrix} \gamma_1(z) & & \\ & \ddots & \\ & & \gamma_k(z) \end{pmatrix} .$$

From the above identity, one obtains:

$$\Gamma(z)^{-1} \cdot B(z) \cdot G(z) = (\mathrm{Id}_k \,|0) \cdot C(z)^{-1} ,$$

which has invariant factors equal to 1 and is therefore a basic encoder whose right inverse is the polynomial matrix determined by the first k columns of $C(z)$. □

Example 11.17. If we apply the above algorithm to the polynomial encoder $G(z) = (1+z, 1+z^2, z+z^3)$ introduced in example 11.2, we obtain:

$$\Gamma(z) = (1+z) \quad B(z) = (1) , \quad C(z) = \begin{pmatrix} 1 & 1+z & z+z^2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} ,$$

and the corresponding basic encoder is

$$\Gamma(z)^{-1} \cdot B(z) \cdot G(z) = \frac{1}{1+z} \cdot (1) \cdot (1+z, 1+z^2, z+z^3) = (1, 1+z, z+z^2)\,.$$

The basic encoders are a *maximal class* in the set of polynomials encoders, in the following sense:

Theorem 11.18. *Let $G(z)$ and $G'(z)$ be two polynomial encoders of an (n,k) convolutional code. One has:*

(1) If $G(z)$ is basic and $\operatorname{Im}\phi_G \subseteq \operatorname{Im}\phi_{G'}$, then $\operatorname{Im}\phi_G = \operatorname{Im}\phi_{G'}$.
(2) If $G(z)$ and $G'(z)$ are basic, then $\operatorname{Im}\phi_G = \operatorname{Im}\phi_{G'}$.

Proof. (1) One has a conmutative diagram:

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \mathbb{F}_q[z]^k & \xrightarrow{\phi_G} & \operatorname{Im}\phi_G & \hookrightarrow & \mathbb{F}_q[z]^n & \longrightarrow & \mathbb{F}_q[z]^n/\operatorname{Im}\phi_G \longrightarrow 0 \\
 & & \| & & \downarrow f & & \| & & \downarrow h \\
0 & \longrightarrow & \mathbb{F}_q[z]^k & \xrightarrow{\phi_{G'}} & \operatorname{Im}\phi_{G'} & \hookrightarrow & \mathbb{F}_q[z]^n & \longrightarrow & \mathbb{F}_q[z]^n/\operatorname{Im}\phi_{G'} \longrightarrow 0
\end{array}$$

and $\ker h = \operatorname{Coker} f$. Since f is injective, one has that $\operatorname{Coker} f = \operatorname{Im}\phi_{G'}/\operatorname{Im}\phi_G$, and $\mathbb{F}_q[z]^n/\operatorname{Im}\phi_G$ is free, since $G(z)$ is basic. Thus, $\ker h$ is torsion-free and hence $\operatorname{Im}\phi_{G'}/\operatorname{Im}\phi_G = 0$.

(2) The submodule $\operatorname{Im}\phi_G + \operatorname{Im}\phi_{G'}$ generates the convolutional code and contains $\operatorname{Im}\phi_G$ and $\operatorname{Im}\phi_{G'}$. Thus, applying (1), one has $\operatorname{Im}\phi_G = \operatorname{Im}\phi_G + \operatorname{Im}\phi_{G'} = \operatorname{Im}\phi_{G'}$. □

Remark 11.19. This theorem implies that basic encoders are *invariant with respect the action of the unimodular group* $GL(k, \mathbb{F}_q[z]) = \operatorname{Aut}_{\mathbb{F}_q[z]} \mathbb{F}_q[z]^k$.

Given two basic encoders $G(z)$ and $G'(z)$ of an (n,k) convolutional code, there exists a $B(z) \in GL(k, \mathbb{F}_q[z])$ such that:

$$G'(z) = B(z) \cdot G(z)\,.$$

Corollary 11.20. *Let $G(z)$ and $G'(z)$ be two polynomial encoders of an (n,k) convolutional code, and let us assume that $G(z)$ is basic. If one denotes*

$$\bar{\delta}_G = \textit{maximum degree of the minors of order } k \textit{ of } G(z)\,,$$

then:

$$\bar{\delta}_G \leq \bar{\delta}_{G'}\,.$$

In particular, if $G'(z)$ is also basic, then $\bar{\delta}_G = \bar{\delta}_{G'}$.

Proof. $G(z)$ and $G'(z)$ define the same convolutional code, and hence there exists a $B(z) \in GL(k, \mathbb{F}_q(z))$ such that $G'(z) = B(z) \cdot G(z)$. Since $G(z)$ is basic, from theorem 11.18 $\operatorname{Im} \phi_{G'} \subseteq \operatorname{Im} \phi_G$,

$$\begin{array}{ccc} \mathbb{F}_q[z]^k & \xrightarrow{\phi_G} & \operatorname{Im} \phi_G \\ {\scriptstyle B(z)}\uparrow & & \uparrow \\ \mathbb{F}_q[z]^k & \xrightarrow{\phi_{G'}} & \operatorname{Im} \phi_{G'} \end{array},$$

and one deduces that $B(z) \in GL(k, \mathbb{F}_q[z])$.

Because the minors of $G'(z)$ are the minors of $G(z)$ multiplied by the determinant of $B(z)$, one concludes that $\bar{\delta}_{G'} \geq \bar{\delta}_G$. If $G'(z)$ is also basic, one deduces the equality. □

In the sense of corollary 11.20, the degree of any basic encoder is an *invariant* of the convolutional code.

Definition 11.21. The *degree δ of a convolutional code $\mathcal{C}_k \subseteq \mathbb{F}_q(z)^n$* is

$$\delta = \bar{\delta}_G \,,$$

where $G(z)$ is any basic encoder of $\mathcal{C}_k$.

Sometimes the degree $\bar{\delta}_G$ of a polynomial encoder is also called *internal degree* (see McEliece [12]) to distinguish it from the so-called *external degree*, used to refer to the degree δ_G (see definition 11.5).

11.2.3. *Minimal basic encoders. Canonical encoders*

In the implementation of convolutional codes as physical devices it is convenient to find *minimal encoders*, in the sense that the corresponding circuit had a minimum quantity of memory boxes. The formalization of the concept of minimality can be expressed in terms of the degree δ of the code.

Theorem 11.22. *(Forney [5]) For each (n, k)-convolutional code of degree δ there exists at least one basic encoder $G(z)$ such that*

$$\delta = \delta_G \leq \delta_{G'} \,,$$

for all realizable encoders $G'(z)$ of the convolutional code. These basic encoders $G(z)$ are called minimal basic encoders.

Definition 11.23. (McEliece [12]) Given an (n,k)-convolutional code $\mathcal{C}_k \subseteq \mathbb{F}_q(z)^n$, a polynomial encoder $G(z)$ is called

- *canonical*, if $\delta_G \leq \delta_{G'}$, for every polynomial encoder $G'(z)$ of $\mathcal{C}_k$.
- *reduced*, if $\delta_G = \bar{\delta}_G$.

Theorem 11.24. *(McEliece [12])A polynomial encoder is canonical if and only if it is basic and reduced. Moreover, the set of row degrees is the same for every canonical encoder,*

$$e_1 \leq \cdots \leq e_k = m_G\,.$$

These invariants *of the convolutional code are called the* Forney indices. *The maximum degree $e_k = m_G$ is called the* memory *of the convolutional code.*

Remark 11.25. A polynomial encoder $G(z) = (G_{ij}(z))$ is reduced if and only if the matrix $\bar{G} = (\bar{g}_{ij})$, $\bar{g}_{ij} \in \mathbb{F}_q$, defined by the coefficients of the terms of highest degree in each row, has rank k (McEliece [12]).

Thus, one has the following method for constructing reduced encoders: if the matrix $\bar{G}$ does not have rank k, there exists a zero linear combination between its rows

$$\sum_{i=1}^{k} \lambda_i \bar{g}_{ij} = 0\,, \quad \text{with } \lambda_i \in \mathbb{F}_q\,, \quad 1 \leq j \leq n\,,$$

from which one can construct a linear combination between the rows of $G(z)$ by eliminating the terms of highest degree,

$$\sum_{i=1}^{k} \lambda_i z^{e_k - e_i} G_{ij}(z) = 0\,,$$

and this allows us to replace a row of $G(z)$ by a new one to obtain a new encoder with the lowest degree in each row. Applying this process several times, we finally obtain a reduced encoder.

Example 11.26. For the encoders of example 11.4 we have:

$G(z)$	Δ_k	Basic	δ_G	$\bar{\delta}_G$	Reduced
$(1+z, 1+z^2, z+z^3)$	$1+z$	No	3	3	Yes
$(z, z+z^2, z^2+z^3)$	z	No	3	3	Yes
$(1, 1+z, z+z^2)$	1	Yes	2	2	Yes

In particular, one deduces that $\mathcal{C}_1 \subset \mathbb{F}_2(z)^3$ is a convolutional code with memory equal to its degree, $\delta = 2$.

11.2.4. *Dual code. Parity-check (control) matrix*

Let us consider, over the $\mathbb{F}_q(z)$-vectorial space $\mathbb{F}_q(z)^n$, the pairing $\langle\ ,\ \rangle$

$$\begin{aligned} &\mathbb{F}_q(z)^n \times \mathbb{F}_q(z)^n \to \mathbb{F}_q(z) \\ &(X(z), Y(z)) \mapsto \langle X(z), Y(z)\rangle = \sum_{i=1}^{n} X_i(z) Y_i(z)\,, \end{aligned} \tag{11.2}$$

where $X(z) = (X_1(z), \ldots, X_n(z)), Y(z) = (X_1(z), \ldots, X_n(z)) \in \mathbb{F}_q(z)^n$.

Definition 11.27. Given an (n,k)-convolutional code $\mathcal{C}_k \subseteq \mathbb{F}_q(z)^n$, the dual code is the $\mathbb{F}_q(z)$-subspace defined by

$$\mathcal{C}_k^{\perp} = \{Y(z) \in \mathbb{F}_q(z)^n \ / \ \langle X(z), Y(z)\rangle = 0 \text{ for every } X(z) \in \mathcal{C}\}\,.$$

Theorem 11.28. *$\mathcal{C}_k^{\perp}$ is an $(n-k,k)$ convolutional code of degree equal to the degree of $\mathcal{C}_k$.*

Proof. Let $G(z)$ be a basic encoder $\mathcal{C}_k$. Since $G(z)$ is basic, then $\mathbb{F}_q[z]^n / \operatorname{Im}\phi_G \simeq \mathbb{F}_q[z]^{n-k}$ is free, and one has an exact sequence

$$0 \to \mathbb{F}_q[z]^k \xrightarrow{\phi_G} \mathbb{F}_q[z]^n \xrightarrow{\pi_G} \mathbb{F}_q[z]^{n-k} \to 0\,,$$

and taking $\operatorname{Hom}_{\mathbb{F}_q[z]}(\quad, \mathbb{F}_q)$ one obtains the exact sequence of free $\mathbb{F}_q[z]$-modules

$$0 \to \mathbb{F}_q[z]^{n-k} \xrightarrow{\pi^*} \mathbb{F}_q[z]^n \xrightarrow{\phi_G^*} \mathbb{F}_q[z]^k \to 0\,.$$

By construction,

$$\mathcal{C}_k^{\perp} \simeq \operatorname{Im}\pi_G^*\,,$$

from which one concludes that the matrix $H(z)$ defining π_G^* is a basic encoder of $\mathcal{C}_k^{\perp}$ and $\delta_G = \delta_H$. Moreover, one has:

$$H(z) \cdot G(z)^T = 0\,, \tag{11.3}$$

an equality that allows us to compute $H(z)$ from $G(z)$ or viceversa. □

Definition 11.29. A *parity-check (control) matrix* for an (n,k)-convolutional code $\mathcal{C}_k$ is every $(n, n-k)$-generator matrix $H(z)$ of its dual code $\mathcal{C}_k^{\perp}$.

We can easily compute a parity-check matrix $H(z)$ from equation (11.3) when we have a generator matrix $G(z)$ in which the first k columns have rank k, making a base change to turn these columns into the identity matrix of order k.

Definition 11.30. An encoder $G(z)$ of an (n,k)-convolutional code $\mathcal{C}_k \subseteq \mathbb{F}_q(z)^n$ is *systematic* if it is takes the form

$$G(z) = (\mathrm{Id}_{k\times k} \mid \bar{G}(z)_{k\times(n-k)}) .$$

Let $G(z)$ be a systematic encoder and let us decompose the parity-check matrix $H(z)$ as:

$$H(z) = (\bar{H}(z)_{(n-k)\times k} \mid \bar{H}'(z)_{(n-k)\times(n-k)}) .$$

From equation (11.3), one deduces that $\bar{H}(z)+\bar{H}'(z)\cdot\bar{G}(z)^T = 0$. Therefore, $H(z) = \bar{H}'(z)\cdot(-\bar{G}(z)^T \mid \mathrm{Id}_{(n-k)\times(n-k)})$. Thus, we can take simply as a parity-check matrix for $\mathcal{C}_k$

$$H(z) = (-\bar{G}(z)^T \mid \mathrm{Id}) .$$

Example 11.31. In example 11.4, the encoder $G''(z) = (1, 1+z, z+z^2)$ is systematic. A parity check control matrix is:

$$H(z) = \begin{pmatrix} 1+z & 1 & 0 \\ z+z^2 & 0 & 1 \end{pmatrix} .$$

Remark 11.32. There is another possible notion of dual codes (see [15]), according to a *time reversal* $\mathbb{F}_q$-linear pairing

$$\begin{aligned} [\ ,\ ] : \mathbb{F}_q((z))^n \times \mathbb{F}_q((z))^n &\to \mathbb{F}_q \\ (X(z), Y(z)) &\mapsto \sum_t \langle x(t), y(-t)\rangle , \end{aligned}$$

where $X(z) = \sum_t x(t)z^t, Y = \sum_i y(t)z^t \in \mathbb{F}_q((z))^n$, and $\langle\ ,\ \rangle$ is the standard $\mathbb{F}_q$-bilinear form on $\mathbb{F}_q^n$. The duality with respect to this pairing $[\ ,\ ]$ is therefore a duality *over the base field* $\mathbb{F}_q$, whereas the duality with respect to the pairing (11.2) is over the field $\mathbb{F}_q(z)$.

11.3. Convolutional Goppa codes

(For an overview of linear Goppa codes, see [7] or [8], and also section 1.4 in this book [3])

Let $(X, \mathcal{O}_X)$ be a smooth projective curve over $\mathbb{F}_q(z)$ of genus g. Let us denote by Σ_X the field of rational functions of X, and let us assume that $\mathbb{F}_q(z)$ is algebraically closed in Σ_X. Both the Riemann-Roch and the Residue theorems (see for instance [6]) still hold under this hypothesis.

Given a set $p_1, \ldots, p_n$ of n different $\mathbb{F}_q(z)$-rational points of X, if $\mathcal{O}_{p_i}$ denotes the local ring at the point p_i, with maximal ideal $\mathfrak{m}_{p_i}$, and t_i is a local parameter at p_i, one has the exact sequences:

$$\begin{aligned} 0 \to \mathfrak{m}_{p_i} \to \mathcal{O}_{p_i} \to \mathcal{O}_{p_i}/\mathfrak{m}_{p_i} \simeq \mathbb{F}_q(z) \to 0\,, \\ s(t_i) \mapsto s(p_i)\,. \end{aligned} \tag{11.4}$$

Let us consider the divisor $D = p_1 + \cdots + p_n$, with its associated invertible sheaf $\mathcal{O}_X(D)$. One then has an exact sequence of sheaves

$$0 \to \mathcal{O}_X(-D) \to \mathcal{O}_X \to Q \to 0\,, \tag{11.5}$$

where the quotient Q is a sheaf with support at the points p_i.

Let G be a divisor on X of degree r, with support disjoint from D. Tensoring the exact sequence (11.5) by the associated invertible sheaf $\mathcal{O}_X(G)$, one obtains:

$$0 \to \mathcal{O}_X(G-D) \to \mathcal{O}_X(G) \to Q \to 0\,. \tag{11.6}$$

For each divisor F over X, let us denote its $\mathbb{F}_q(z)$-vector space of global sections by

$$L(F) \equiv \Gamma(X, \mathcal{O}_X(F)) = \{s \in \Sigma_X \mid (s) + F \geq 0\}\,,$$

where (s) is the divisor defined by $s \in \Sigma_X$. Taking global sections in (11.6), one obtains

$$\begin{aligned} 0 \to L(G-D) \to L(G) \stackrel{\alpha}{\to} \mathbb{F}_q(z) \times \overset{n}{\cdots} \times \mathbb{F}_q(z) \to \ldots \\ s \mapsto (s(p_1), \ldots, s(p_n))\,. \end{aligned}$$

Definition 11.33. The convolutional Goppa code $\mathcal{C}(D,G)$ associated with the pair (D,G) is the image of the $\mathbb{F}_q(z)$-linear map $\alpha\colon L(G) \to \mathbb{F}_q(z)^n$.

Analogously, given a subspace $\Gamma \subseteq L(G)$, one defines the convolutional Goppa code $\mathcal{C}(D,\Gamma)$ as the image of $\alpha_{|\Gamma}$.

Remark 11.34. The above definition is more general than the one given in [2] in terms of families of curves $X \to \mathbb{A}^1_{\mathbb{F}_q}$. In fact, given such a family, the fibre X_η, over the generic point $\eta \in \mathbb{A}^1_{\mathbb{F}_q}$, is a curve over $\mathbb{F}_q(z)$. However, not every curve over $\mathbb{F}_q(z)$ extends to a family over $\mathbb{A}^1_{\mathbb{F}_q}$.

By construction, $\mathcal{C}(D,G)$ is a convolutional code of length n and dimension

$$k \equiv \dim L(G) - \dim L(G-D)\,.$$

Prop 11.35. Let us assume that $2g-2 < r < n$. Accordingly, the evaluation map $\alpha\colon L(G) \hookrightarrow \mathbb{F}_q(z)^n$ is injective, and the dimension of $\mathcal{C}(D,G)$ is

$$k = r + 1 - g\,.$$

Proof. If $r < n$, $\dim L(G-D) = 0$, the map α is injective and $k = \dim L(G)$. If $2g-2 < r$, $\dim L(G) = 1-g+r$ by the Riemann-Roch theorem. □

11.3.1. *Dual convolutional Goppa codes*

Given a convolutional Goppa code $\mathcal{C}(D,G)$, let $\mathcal{C}(D,G)^\perp$ be its dual convolutional code, in the sense of definition 11.27.

Theorem 11.36. *$\mathcal{C}^\perp(D,G)$ is also a convolutional Goppa code, in the following sense: If K denotes the canonical divisor of rational differential forms over X, then $\mathcal{C}^\perp(D,G)$ is the image of the $\mathbb{F}_q(z)$-linear map $\beta\colon L(K+D-G) \to \mathbb{F}_q(z)^n$, given by*

$$\beta(\eta) = (\mathrm{Res}_{p_1}(\eta), \ldots, \mathrm{Res}_{p_n}(\eta))\,.$$

Proof. Following the construction of $\mathcal{C}(D,G)$, we start by tensoring the exact sequence (11.4) by $\mathfrak{m}_{p_i}^* = \mathrm{Hom}_{\mathcal{O}_{p_i}}(\mathfrak{m}_{p_i}, \mathcal{O}_{p_i})$, and we obtain:

$$\begin{aligned} 0 \to \mathcal{O}_{p_i} \to \mathfrak{m}_{p_i}^* \to \mathcal{O}_{p_i}/\mathfrak{m}_{p_i} \otimes_{\mathcal{O}_{p_i}} \mathfrak{m}_{p_i}^* \simeq \mathbb{F}_q(z) \to 0 \\ t_i^{-1}s(t_i) \mapsto s(p_i)\,. \end{aligned} \tag{11.7}$$

Again tensoring (11.7) by $\mathfrak{m}_{p_i}/\mathfrak{m}_{p_i}^2$, the tangent space of differentials at the point p_i, one obtains:

$$\begin{aligned} 0 \to \mathfrak{m}_{p_i}/\mathfrak{m}_{p_i}^2 \to \mathfrak{m}_{p_i}^* \otimes_{\mathcal{O}_{p_i}} \mathfrak{m}_{p_i}/\mathfrak{m}_{p_i}^2 \to \mathbb{F}_q(z) \to 0 \\ t_i^{-1}s(t_i)dt_i \mapsto s(p_i)\,, \end{aligned} \tag{11.8}$$

where $s(p_i) = \mathrm{Res}_{p_i}(t_i^{-1}s(t_i)dt_i)$.

This allows us to define a new convolutional Goppa code associated with the pair of divisors $D = p_1 + \cdots + p_n$ and G; tensoring (11.5) by the line sheaf $\mathcal{O}_X(K+D-G)$, one has:

$$0 \to \mathcal{O}_X(K-G) \to \mathcal{O}_X(K+D-G) \to Q \to 0\,. \tag{11.9}$$

Taking global sections, one has

$$\begin{aligned} 0 \to L(K-G) \to L(K+D-G) \xrightarrow{\beta} \mathbb{F}_q(z) \times \overset{n}{\cdots} \times \mathbb{F}_q(z) \to \ldots \\ \eta \mapsto (\mathrm{Res}_{p_1}(\eta), \ldots, \mathrm{Res}_{p_n}(\eta))\,. \end{aligned}$$

The image of β is a subspace of $\mathbb{F}_q(z)^n$, whose dimension can be calculated by the Riemann-Roch theorem:

$$\begin{aligned}&\dim L(K+D-G) - \dim L(K-G)\\ &= \dim L(G-D) - (r-n) - 1 + g - (\dim L(G) - r - 1 + g)\\ &= n-k\,.\end{aligned}$$

Moreover, $\operatorname{Im}\beta$ is the subspace $\mathcal{C}(D,G)^\perp \subset \mathbb{F}_q(z)^n$, because they have the same dimension, and by the Residue theorem for every $\eta \in L(K+D-G)$ and every $s \in L(G)$ one has

$$\langle \beta(\eta), \alpha(s)\rangle = \sum_{i=1}^{n} s(p_i)\operatorname{Res}_{p_i}(\eta) = \sum_{i=1}^{n} \operatorname{Res}_{p_i}(s\eta) = 0\,.$$

□

Under the hypothesis $2g-2 < r < n$, the map β is injective, and $\mathcal{C}^\perp(D,G)$ is a convolutional code of length n and dimension

$$\dim L(K+D-G) = n - (1-g+r)\,.$$

Remark 11.37. In the context of duality in the sense of the pairing $[\ ,\ ]$ of remark 11.32,

$$[X(z), Y(z)] = \operatorname{Res}_{z=0}\left(\langle X(z), Y(z)\rangle \frac{dz}{z}\right) = \sum_{i=1}^{n} \operatorname{Res}_{z=0}\left(X_i(z)Y_i(z)\frac{dz}{z}\right)\,.$$

Thus, the duality for convolutional Goppa codes in the sense of Definition 11.27 is related to the residues at the points of X, and the duality with respect to the pairing $[\ ,\ ]$ is related to the residues in the variable of the base field.

11.4. Weights and (free)distance

For vectors $x = (x_1, \ldots, x_n) \in \mathbb{F}_q^n$, the *(Hamming) weight* is defined as $hwt(x) = \#\{i \mid x_i \neq 0\}$ and the *(Hamming) distance* between $x, y \in \mathbb{F}_q^n$ can be defined as the weight $hwt(y-x)$. In the setting of linear coding theory, the corresponding notion of *minimum weight (distance)* of the words in the code is one of the most important parameters of the code.

For convolutional codes, one needs an analogous notion for polynomial vectors $X(z) = (X(z)_1, \ldots, X_n(z)) \in \mathbb{F}_q[z]^n$. However, it is possible to define two kinds of weights. First, one can simply define the *Hamming weight of* $X(z)$ as

$$hwt(X(z)) = \#\{i \mid X_i(z) \neq 0\}\,.$$

Thus, the concept of *minimum Hamming weight of a convolutional code* does not reflect the performance of convolutional codes over noisy channels in convolutional coding theory. Of course the minimum Hamming weight of a convolutional Goppa code $\mathcal{C}(D,G)$ can be bounded using the Riemann-Roch theorem, as in the usual Goppa codes.

However, when one considers $X(z) \in \mathbb{F}_q[z]^n$ as a polynomial with vector coefficients

$$X(z) = \sum_t x(t) z^t \,, \text{ where } x(t) = (x_1(t), \ldots, x_n(t) \in \mathbb{F}_q^n \,,$$

one can define a more natural notion of *weight in convolutional coding theory*:

Definition 11.38. The *weight* of $X(z) \in \mathbb{F}_q[z]^n$ is

$$wt(X(z)) = \sum_t hwt(x(t)) \,.$$

Definition 11.39. The *(free) distance* of a (n,k)-convolutional code $\mathcal{C}_k \subseteq \mathbb{F}_q(z)^n$ is

$$d = \text{Min}\{wt(X(z)) \mid X(z) \in \mathcal{C}_k \cap \mathbb{F}_q[z]^n \,,\ X(z) \neq 0\} \,.$$

In particular, if the degree of the code is zero, $\mathcal{C}_k$ is a linear code and the (free) distance is the (minimum) distance as linear code.

As in the case of linear codes, the distance d is one of the most important parameters in convolutional coding theory.

In particular, an interesting problem is to find upper bounds for d. A possible solution is to link convolutional codes with linear codes, fixing the degree of the words in $\mathcal{C}_k$:

Theorem 11.40. *(McEliece [12]) For an (n,k)-convolutional code $\mathcal{C}_k \subseteq \mathbb{F}_q(z)^n$ of Forney indices $e_1 \leq \cdots \leq e_k$, let*

$$\mathcal{C}_L = \{X(z) \in \mathcal{C} \ / \ \text{Degree}(X(z)) \leq L\} \,.$$

Identifying the set of all possible n-dimensional polynomial vectors of degree $\leq L$ over $\mathbb{F}_q$ with $\mathbb{F}_q^{n(L+1)}$, one can see $\mathcal{C}_L$ as an $\mathbb{F}_q$-linear code of length $n(L+1)$ and a certain dimension k_L. Then,

$$k_L = \sum_{i=1}^{k} \text{Max}(L+1-e_i, 0) \,,$$

and

$$d \leq \text{Min}_{L \geq 0} \left(\text{Max}\{\textit{distance of possible } (n(L+1), k_L) \textit{ linear codes}\}\right) \,.$$

This result can be used to calculate the distance d of a convolutional code $\mathcal{C}_k$ from the distances d_L of its linear (sub)codes $\mathcal{C}_L$.

Also it is possible for convolutional codes to find a bound of d analogous to the Singleton bound for linear codes.

Theorem 11.41. *(Rosenthal, Smarandache [16]) If $\mathcal{C}_k \subseteq \mathbb{F}_q(z)^n$ is an (n,k)-convolutional code of degree δ, its distance is bounded by:*

$$d \leq (n-k)\big(\lfloor \frac{\delta}{k} \rfloor + 1\big) + \delta + 1\,.$$

If d achieves this bound, then $\mathcal{C}_k$ will be called a Maximum Distance Separable (MDS) *convolutional code.*

For linear codes, there is a geometric interpretation of distance, in terms of *balls* in the words of the code. Moreover, for linear Goppa codes the distance can be viewed in terms of the number of zeroes of certain meromorphic functions, which allows us to use the Riemann-Roch theorem to make very precise computations.

In the case of convolutional Goppa codes $\mathcal{C}(D,G)$ of length n over a curve X defined over $\mathbb{F}_q(z)$, the interpretation of the notion of weight in geometric terms is much more difficult. Let us assume that X can be extended to a family of curves X_U over $U = \operatorname{Spec} \mathbb{F}_q[z] = \mathbb{A}^1_{\mathbb{F}_q}$ (as in [2]). Let X_0 be the fibre of X_U over the origin of U. The points $p_1, \ldots, p_n$ of the divisor D define sections $p_i(z)\colon \mathbb{A}^1_{\mathbb{F}_q} \to X_U$ and the polynomial words of the code $\mathcal{C}(D,G)$ are defined by evaluating the sections $s \in L(G)$ along the sections $p_i(z)$.

Let p be one of the points defined by D, C_p the curve of X_U defined as the image of the section $p(z)$, and q_0 the intersection of C_p with X_0; that is, $q_0 = p(0)$. Let us assume that $L(G)$ is a very ample linear series [6], and let us assume that X_U is immersed in $\mathbb{P}^N_{\mathbb{F}_q} \times \mathbb{P}^1_{\mathbb{F}_q}$ using the linear series $L(G)$. Let us denote by $\pi_r(q_0)$ the r-th osculating plane to the curve C_p at the point q_0. One has a sequence of strict inclusions:

$$\pi_0(q_0) = q_0 \subset \pi_1(q_0) \subset \pi_2(q_0) \subset \cdots \subset \pi_r(q_0) \subset \cdots\,.$$

The evaluation of s at p, $s(p)$, can be expressed by:

$$s(p) = s_0 + s_1 z + \cdots + s_n z^n\,,$$

where $s_0 = s(0)$ and s_r, the r-th coefficient, can be interpreted as the r-th jet of $s(z)$ at the point q_0.

With this interpretation in mind, one has that $s_r = 0$ if and only if

$$H_s \cap \pi_r(q_0) \neq \emptyset \text{ and } H_s \cap \pi_{r-1}(q_0) \subsetneqq H_s \cap \pi_r(q_0)\,,$$

where H_s is the hyperplane defined by the section s.

Accordingly, the problem of computing the number $\#\{r \mid s_r = 0\}$ can be translated into a problem of enumerative geometry over finite fields.

The main problem here is to develop the classical theory of osculating planes and all the classical computations in the case of finite base fields. This is not an easy problem, but its solution would allow one to give a geometric interpretation of the distance of convolutional Goppa codes.

11.5. Convolutional Goppa Codes over the projective line

Let $X = \mathbb{P}^1_{\mathbb{F}_q(z)} = \operatorname{Proj} \mathbb{F}_q(z)[x_0, x_1]$ be the projective line over the field $\mathbb{F}_q(z)$, and let us denote by $t = x_1/x_0$ the affine coordinate.

Let $p_0 = (1, 0)$ be the origin point, $p_\infty = (0, 1)$ the point at infinity, and let $p_1, \ldots, p_n$ be different rational points of $\mathbb{P}^1$, $p_i \neq p_0, p_\infty$. Let us define the divisors $D = p_1 + \cdots + p_n$ and $G = rp_\infty - sp_0$, with

$$0 \leq s \leq r < n\,.$$

Since $g = 0$, the evaluation map $\alpha\colon L(G) \to \mathbb{F}_q(z)^n$ is injective, and $\operatorname{Im}\alpha$ defines a convolutional Goppa code $\mathcal{C}(D, G)$ of length n and dimension $k = r - s + 1$.

Let us choose the functions $t^s, t^{s+1}, \ldots, t^r$ as a basis of $L(G)$. If $\alpha_i \in \mathbb{F}_q(z)$ is the local coordinate of the point p_i, $i = 1, \ldots, n$, the matrix of the evaluation map α is the following generator matrix for the code $\mathcal{C}(D, G)$:

$$G = \begin{pmatrix} \alpha_1^s & \alpha_2^s & \ldots & \alpha_n^s \\ \alpha_1^{s+1} & \alpha_2^{s+1} & \ldots & \alpha_n^{s+1} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^r & \alpha_2^r & \ldots & \alpha_n^r \end{pmatrix}. \tag{11.10}$$

The dual convolutional Goppa code $\mathcal{C}^\perp(D, G)$ also has length n, and dimension $n - k = n - r + s - 1$.

To construct $\mathcal{C}^\perp(D, G)$, let us choose in $L(K + D - G)$ the basis of rational differential forms:

$$\left\langle \frac{dt}{t^s \prod_{i=1}^n (t - \alpha_i)}, \frac{t\ dt}{t^s \prod_{i=1}^n (t - \alpha_i)}, \ldots, \frac{t^{n-r+s-2} dt}{t^s \prod_{i=1}^n (t - \alpha_i)} \right\rangle,$$

and let us calculate the residues:

$$\begin{aligned}
&\mathrm{Res}_{p_j}\left(\frac{t^m dt}{t^s \prod_{i=1}^{n}(t-\alpha_i)}\right)\\
&= \mathrm{Res}_{p_j}\left(\frac{(t-\alpha_j+\alpha_j)^m d(t-\alpha_j)}{(t-\alpha_j)(t-\alpha_j+\alpha_j)^s \prod_{\substack{i=1\\ i\neq j}}^{n}(t-\alpha_j+\alpha_j-\alpha_i)}\right)\\
&= \frac{\alpha_j^m}{\alpha_j^s \prod_{\substack{i=1\\ i\neq j}}^{n}(\alpha_j-\alpha_i)}\,.
\end{aligned}$$

If one sets $h_j = \frac{1}{\alpha_j^s \prod_{\substack{i=1\\ i\neq j}}^{n}(\alpha_j-\alpha_i)}$, then the matrix H of

$\beta\colon L(K+D-G) \to \mathbb{F}_q(z)^n$,

$$H = \begin{pmatrix} h_1 & h_2 & \dots & h_n \\ h_1\alpha_1 & h_2\alpha_2 & \dots & h_n\alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ h_1\alpha_1^{n-r+s-2} & h_2\alpha_2^{n-r+s-2} & \dots & h_n\alpha_n^{n-r+s-2} \end{pmatrix}, \tag{11.11}$$

is a generator matrix for the dual code $\mathcal{C}^{\perp}(D,G)$, and therefore a parity-check matrix for $\mathcal{C}(D,G)$. In fact, one has $H \cdot G^T = 0$.

Remark 11.42. The matrix in (11.11) suggests that $\mathcal{C}^{\perp}(D,G)$ is an alternating code over the field $\mathbb{F}_q(z)$, and we can thus apply to $\mathcal{C}(D,G)$ some kind of Berlekamp-Massey decoding algorithm as a linear code over $\mathbb{F}_q(z)$.

Example 11.43. Let $a, b \in \mathbb{F}_q$ be two different non-zero elements, and

$$\alpha_i = a^{i-1}z + b^{i-1}\,, i = 1, \dots, n\,, \text{ with } n < q\,.$$

We present some examples of convolutional Goppa codes with canonical generator matrices, whose distance d attains the generalized Singleton bound 11.41 (i.e., they are MDS convolutional codes), and we include their encoding equations as linear systems.

- Field $\mathbb{F}_3(z)$, $\mathbb{F}_3 = \{0,1,2\}$:

$$\begin{aligned}
&G = \begin{pmatrix} z+1 & z+2 \end{pmatrix}\\
&H = \begin{pmatrix} \frac{1}{2(z+1)} & \frac{1}{z+2} \end{pmatrix}\\
&A = \begin{pmatrix} 0 \end{pmatrix},\ B = \begin{pmatrix} 1 \end{pmatrix},\ C = \begin{pmatrix} 1 & 1 \end{pmatrix},\ D = \begin{pmatrix} 1 & 2 \end{pmatrix}\\
&(n,k,\delta,d) = (2,1,1,4)\,.
\end{aligned}$$

- Field $\mathbb{F}_4(z)$, $\mathbb{F}_4 = \{0, 1, \alpha, \alpha^2\}$ where $\alpha^2 + \alpha + 1 = 0$:

$$G = \begin{pmatrix} 1 & 1 & 1 \\ z+1 & \alpha z + \alpha^2 & \alpha^2 z + \alpha \end{pmatrix}$$
$$H = \left(\frac{1}{(\alpha^2 z+\alpha)(\alpha z+\alpha^2)} \ \frac{1}{(\alpha^2 z+\alpha)(z+1)} \ \frac{1}{(\alpha z+\alpha^2)(z+1)} \right)$$
$$A = \begin{pmatrix} 0 \end{pmatrix}, \ B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \ C = \begin{pmatrix} 1 & \alpha & \alpha^2 \end{pmatrix}, \ D = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \alpha^2 & \alpha \end{pmatrix}$$
$$(n, k, \delta, d) = (3, 2, 1, 3).$$

- Field $\mathbb{F}_4(z)$:

$$G = \begin{pmatrix} z+1 & z+\alpha & z+\alpha^2 \end{pmatrix}$$
$$H = \begin{pmatrix} \frac{1}{z+1} & \frac{\alpha}{z+\alpha} & \frac{\alpha^2}{z+\alpha^2} \\ 1 & \alpha & \alpha^2 \end{pmatrix}$$
$$A = \begin{pmatrix} 0 \end{pmatrix}, \ B = \begin{pmatrix} 1 \end{pmatrix}, \ C = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}, \ D = \begin{pmatrix} 1 & \alpha & \alpha^2 \end{pmatrix}$$
$$(n, k, \delta, d) = (3, 1, 1, 6).$$

- Field $\mathbb{F}_5(z)$, $\mathbb{F}_5 = \{0, 1, 2, 3, 4\}$:

$$G = \begin{pmatrix} (z+1)^2 & (z+2)^2 & (z+4)^2 \end{pmatrix}$$
$$H = \begin{pmatrix} \frac{2}{(z+1)^2} & \frac{2}{(z+2)^2} & \frac{1}{(z+4)^2} \\ \frac{2}{z+1} & \frac{2}{z+2} & \frac{1}{z+4} \end{pmatrix}$$
$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ B = \begin{pmatrix} 1 & 0 \end{pmatrix}, \ C = \begin{pmatrix} 2 & 4 & 3 \\ 1 & 1 & 1 \end{pmatrix}, \ D = \begin{pmatrix} 1 & 4 & 1 \end{pmatrix}$$
$$(n, k, \delta, d) = (3, 1, 2, 9).$$

- Field $\mathbb{F}_5(z)$:

$$G = \begin{pmatrix} z+1 & 2z+3 & 4z+4 & 3z+2 \\ (z+1)^2 & (2z+3)^2 & (4z+4)^2 & (3z+2)^2 \end{pmatrix}$$
$$H = \begin{pmatrix} \frac{4}{a^2 bc} & \frac{4}{bcd^2} & \frac{4}{a^2 bc} & \frac{4}{bcd^2} \\ \frac{4}{abc} & \frac{3}{bcd} & \frac{1}{abc} & \frac{2}{bcd} \end{pmatrix}$$
where $a = z+1$, $b = z+2$, $c = z+3$ and $d = z+4$,
$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \ B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \ C = \begin{pmatrix} 1 & 2 & 4 & 3 \\ 2 & 2 & 2 & 2 \\ 1 & 4 & 1 & 4 \end{pmatrix}, \ D = \begin{pmatrix} 1 & 3 & 4 & 2 \\ 1 & 4 & 1 & 4 \end{pmatrix}$$
$$(n, k, \delta, d) = (4, 2, 3, 8).$$

11.6. Convolutional Goppa Codes over elliptic curves

We can obtain convolutional codes from elliptic curves in the same way. Let $X \subset \mathbb{P}^2_{\mathbb{F}_q(z)}$ be a plane elliptic curve over $\mathbb{F}_q(z)$, and let us denote by (x, y) the affine coordinates in $\mathbb{P}^2_{\mathbb{F}_q(z)}$. Let p_∞ be the infinity point, and $p_1, \ldots, p_n$ rational points of X, with $p_i = (x_i(z), y_i(z))$. Let us define $D = p_1 + \cdots + p_n$ and $G = rp_\infty$.

The *canonical* basis of $L(G)$ is $\{1, x, y, \ldots, x^a y^b\}$, with $2a+3b = r$ (and $b = 0, 1$ so that there are no linear combinations). Thus, the evaluation map $\alpha\colon L(G) \to \mathbb{F}_q(z)^n$ is:

$$\alpha(x^i y^j) = (x_1^i(z) y_1^j(z), \ldots, x_n^i(z) y_n^j(z)).$$

The image of a subspace $\Gamma \subseteq L(G)$ under the map α provides a Goppa convolutional code.

We present a couple of examples obtained from elliptic curves that, although not MDS, have distance approaching that bound.

Example 11.44. We consider the curve over $\mathbb{F}_2(z)$

$$y^2 + (1+z)xy + (z+z^2)y = x^3 + (z+z^2)x^2\,,$$

and the points

$$\begin{aligned} p_1 &= (z^2+z, z^3+z^2) \\ p_2 &= (0, z^2+z) \\ p_3 &= (z, z^2)\,. \end{aligned}$$

$L(G)$ is the subspace generated by $\{1, x\}$. Thus, the valuation map α is defined by the matrix

$$\begin{pmatrix} 1 & 1\,1 \\ z^2+z & 0\ z \end{pmatrix}.$$

This code has distance $d = 2$. The maximum distance for its parameters is 3.

Example 11.45. Let us now consider the curve over $\mathbb{F}_2(z)$

$$y^2 + (1+z+z^2)xy + (z^2+z^3)y = x^3 + (z^2+z^3)x^2\,,$$

and the points

$$\begin{aligned} p_1 &= (z^3+z^2, 0) \\ p_2 &= (0, z^3+z^2) \\ p_3 &= (z^3+z^2, z^5+z^3) \\ p_4 &= (z^2+z, z^3+z) \\ p_5 &= (z^2+z, z^4+z^2)\,. \end{aligned}$$

Again we take $L(G)$ as the subspace generated by $\{1, x\}$. Therefore, the valuation map α is defined by the matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ z^3+z^2 & 0 & z^3+z^2 & z^2+z & z^2+z \end{pmatrix}.$$

This code has distance $d = 4$. The maximum distance for its parameters is 5.

Remark 11.46. Every elliptic curve X over $\mathbb{F}_q(z)$ can be considered the generic fibre of a fibration $\mathcal{X} \to U = \operatorname{Spec} \mathbb{F}_q[z]$, with some fibres singular curves of genus 1. The global structure of this fibration is related to the singular fibres (see [18]); the translation into the language of coding theory of the arithmetic and geometric properties of the fibration is the first step in the program of applying the general construction to the effective construction of good convolutional Goppa codes of genus 1.

References

[1] C. Berrou, A. Glavieux and P. Thitimajshima, Near Shannon limit error-correcting coding and decoding: Turbo codes (1), *IEEE Int. Conf. on Communications (ICC'93)*, 1064–1070, (1993).

[2] J.A. Domínguez Pérez, J.M. Muñoz Porras and G. Serrano Sotelo, Convolutional Codes of Goppa type, *Algebra Engrg. Comm. Comput.*, **15** (1), 51–61, (2004).

[3] I.M. Duursma Algebraic geometry codes: general theory. In eds. E. Martínez-Moro, C. Munuera, and D. Ruano, *Advances in Algebraic Geometry Codes*, chapter 1. pp. 1–48. World Scientific, (2008).

[4] P. Elias, Coding for noisy channels, *I.R.E. Nat. Conv. Record* **3**, 34–45, (1955).

[5] G.D. Forney Jr, Convolutional codes I: Algebraic structure, *IEEE Trans. Inform. Theory*, **16** (3), 720–738, (1970).

[6] R. Hartshorne, Algebraic geometry, (Graduate Texts in Mathematics, vol. 52), Springer-Verlag, New York, (1977).

[7] T. Høholdt, J.H. van Lint and R. Pellikaan, Algebraic Geometric Codes, in *Handbook of coding theory, Vol. I*, 871–962, North-Holland, Amsterdam, (1998).

[8] J.H. van Lint and G. van der Geer, Introduction to Coding Theory and Algebraic Geometry (DMV Seminar, vol. 12), Birkhäuser, Basel, (1998)

[9] V. Lomadze, Convolutional Codes and Coherent Sheaves, *Algebra Engrg. Comm. Comput.*, **12** (4), 273–326, (2001).

[10] J.L. Massey and M.K. Sain, Inverses of linear sequential circuits, *IEEE Trans. Comp.* **19** (5), 330–337, (1968).

[11] R.J. McEliece, The theory of information and coding, (vol. 3 of the Encylcopedia of Mathematics and Its Applications). Addison-Wesley, Reading, MA, (1977).

[12] R.J. McEliece, The algebraic theory of convolutional codes, in *Handbook of coding theory, Vol. I*, 1065–1138, North-Holland, Amsterdam, (1998).

[13] J.M. Muñoz Porras, J.A. Domínguez Pérez, J.I. Iglesias Curto and G. Serrano Sotelo, Convolutional Goppa Codes, *IEEE Trans. Inform. Theory*, **52** (1), 340–344, (2006).

[14] P. Piret, Convolutional codes: An Algebraic Approach, MIT Press, Cambridge, MA, (1988).

[15] J.Rosenthal, Connections between linear systems and convolutional codes, in *Codes, systems, and graphical models (Minneapolis, MN, 1999) IMA Vol. Math. Appl.* 39–66, Springer, New York, (2001)

[16] J. Rosenthal and R.Smarandache, Maximum distance separable convolutional codes, *Appl. Algebra Engrg. Comm. Comput.*, **10** (1), 15–32, (1999).

[17] R.Smarandache, H. Gluesing-Luersen and J. Rosenthal, Constructions of MDS-Convolutional Codes, *IEEE Trans. Inform. Theory*, **47** (5), 2045–2049, (2001).

[18] J. Tate, Algorithm for determining the type of a singular fiber in an elliptic pencil, in *Modular functions of one variable, IV (Proc. Internat. Summer School, Univ. Antwerp, Antwerp,1972)*, (Lecture Notes in Math., vol. 476), 33–52, Springer-Verlag, Berlin, (1975).

[19] A.J. Viterbi, Error bounds for convolutional codes and an asymptotically optimum decoding algorithm, *IEEE Trans. Inf. Theory* **13** (2), 260–269, (1967).

This page intentionally left blank

Chapter 12

Quantum Error-Correcting Codes from Algebraic Curves

Jon-Lark Kim and Gretchen L. Matthews

Department of Mathematics, University of Louisville, Louisville, KY 40292 USA
jl.kim@louisville.edu
Department of Mathematical Sciences, Clemson University, Clemson, SC 29634 USA
gmatthe@clemson.edu

This chapter discusses quantum error-correcting codes constructed from algebraic curves. We give an introduction to quantum coding theory including bounds on quantum codes. We describe stabilizer codes which are the quantum analog of classical linear codes and discuss the binary and q-ary CSS construction. Then we focus on quantum codes from algebraic curves including the projective line, Hermitian curves, and hyperelliptic curves. In addition, we describe the asymptotic behaviors of quantum codes from the Garcia-Stichtenoth tower attaining the Drinfeld-Vlăduţ bound.

Contents

12.1. Introduction

One of the applications of algebraic geometry (AG) codes is their use in the construction of quantum error-correcting codes. Quantum error-correction was developed by Shor [31] and has become one of key ingredients in quantum computation and quantum information theory. Calderbank and Shor [6] and Steane [34] independently showed that quantum error-correcting codes can be constructed via classical linear codes over finite fields, known as the CSS construction. At the same time, Gottesman developed the stabilizer formalism [13]. Shortly thereafter, nonbinary quantum codes were studied by Rains [28] and Ashikhmin and Knill [2].

In this chapter, we start with a brief introduction to quantum information and quantum correction (Section 1.2). Interested readers can refer to the book [27]. Then in Section 1.3, we describe how to construct quantum error-correcting codes (in particular, stabilizer codes) from classical codes via the CSS construction. Finally Section 1.3 explains quantum codes from algebraic geometry codes. We consider quantum Reed-Solomon codes, quantum Hermitian codes, quantum codes from hyperelliptic curves, and quantum codes from multipoint AG codes. We also discuss asymptotic behaviors of quantum codes from AG codes.

12.2. Quantum information and error correction

12.2.1. *Background and terminology*

The classical unit of information is the bit, which is either 0 or 1. The quantum analog of the classical $0-1$ bit is the qubit, which is short for quantum bit. A qubit is of the form

$$\alpha|0\rangle + \beta|1\rangle \text{ where } \alpha, \beta \in \mathbb{C}.$$

Often, the normalization condition that $\mid \alpha \mid^2 + \mid \beta \mid^2 = 1$ is assumed to reflect that upon observation the qubit collapses to 0 with probability $\mid \alpha \mid^2$ and to 1 with probability $\mid \beta \mid^2$. Notice that the qubit may be viewed as a vector in $\mathbb{C}^2$. As in classical coding theory, one may consider larger alphabets such as $\mathbb{F}_q$ where $q = p^m$ and p is prime. Here, the units of information are quantum digits, called qudits. To describe a qudit, fix a basis $\{|a\rangle : a \in \mathbb{F}_q\}$ for the complex vector space $\mathbb{C}^q$. Then a qudit (also

called a q-ary quantum state) is of the form

$$\sum_{a\in\mathbb{F}_q} \alpha_a|a\rangle \text{ where } \alpha_a \in \mathbb{C}.$$

Now the state of an n-qubit system may be viewed as a vector in the n-fold tensor product

$$(\mathbb{C}^q)^{\otimes n} = \underbrace{\mathbb{C}^q \otimes \cdots \otimes \mathbb{C}^q}_{n} \cong \mathbb{C}^{q^n}.$$

In this setting, we now define a quantum code.

Definition 12.1. Given a prime power q, a q-ary quantum code of length n is a complex subspace of $(\mathbb{C}^q)^{\otimes n}$.

Throughout this chapter, q denotes a power of a prime p.

We next discuss how quantum codes guard against errors. Unlike the classical case, it is not immediately obvious that this is even possible. More pointedly, classical codes protect information by adding redundancy with the most elementary example of this being a repetition code. However, quantum information cannot be duplicated in the same sense due to the following observation, called the No Cloning Theorem.

Theorem 12.2. *(No Cloning Theorem) There is no quantum operation that takes the state $|\psi\rangle$ to $|\psi\rangle \otimes |\psi\rangle$ for all states $|\psi\rangle$.*

Proof. Suppose there is such an operation. Then given $|\psi\rangle \neq |\phi\rangle$,

$$|\psi\rangle + |\phi\rangle \mapsto |\psi\rangle|\psi\rangle + |\phi\rangle|\phi\rangle$$

since $|\psi\rangle \mapsto |\psi\rangle$ and $|\phi\rangle \mapsto |\phi\rangle$. However,

$$|\psi\rangle + |\phi\rangle \mapsto (|\psi\rangle + |\phi\rangle)\,(|\psi\rangle + |\phi\rangle)$$

which is a contradiction since

$$|\psi\rangle|\psi\rangle + |\phi\rangle|\phi\rangle \neq (|\psi\rangle + |\phi\rangle)\,(|\psi\rangle + |\phi\rangle)\,.$$

□

Despite the inability to copy quantum information, quantum codes do exist. Peter Shor produced the first example in 1995 [31] which was followed by a larger family found by Shor and Calderbank in 1996 [6]. To better understand the errors in a quantum system, it is helpful to consider the following (albeit oversimplified) analogy as in [15]: Given a linear code C of length n and dimension k over $\mathbb{F}_q$, C partitions $\mathbb{F}_q^n$ into cosets

$$\mathbb{F}_q^n = C \cup (C + e_1) \cup (C + e_2) \cup \cdots \cup \left(C + e_{q^{n-k}-1}\right)$$

and errors act on C by translation whereas a q-ary quantum code Q of length n and dimension k gives rise to an orthogonal decomposition

$$\mathbb{C}^{q^n} = Q \oplus E_1 Q \oplus E_2 Q \oplus \cdots \oplus E_{q^{n-k}-1} Q$$

and errors act on Q as unitary transformations. To be more precise, we next describe the types of errors encountered by an n-qudit q-ary system.

When dealing with qubits, there are three types of errors that may occur: a bit flip, phase flip, and a combination of bit and phase flips. These errors on a single qubit may be represented by 2×2 matrices:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \text{ and } Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = iXZ.$$

Indeed,

$$X|a\rangle = |a \oplus 1\rangle, Z|a\rangle = (-1)^a |a\rangle, \text{ and } Y|a\rangle = i(-1)^a |a \oplus 1\rangle.$$

The matrices X, Y, and Z are called Pauli matrices.

More generally, let $q = p^m$ where p is prime. Given $a, b \in \mathbb{F}_q$ we have dit flip and phase flip errors acting on a single qudit as

$$T_a|u\rangle = |u + a\rangle$$

and

$$R_b|u\rangle = \xi^{Tr(bu)}|u\rangle$$

where $\xi = e^{\frac{2\pi i}{p}}$ is a p^{th} root of unity and $Tr : \mathbb{F}_q \to \mathbb{F}_p$ is the trace function. These operators may be expressed by matrices as follows. Suppose that $\{\gamma_1, \gamma_2, \ldots, \gamma_m\}$ is a basis for $\mathbb{F}_q$ as an $\mathbb{F}_p$-vector space. Given $a, b \in \mathbb{F}_q$, $a = \sum_{i=1}^m a_i \gamma_i$ and $b = \sum_{i=1}^m b_i \gamma_i$ for some $a_i, b_i \in \mathbb{F}_p$. Let $T, R \in \mathbb{C}^{p \times p}$ be the matrices

$$T = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \text{ and } R = \begin{bmatrix} \xi & & & & \\ & \xi^2 & & & \\ & & \xi^3 & & \\ & & & \ddots & \\ & & & & \xi^{p-1} \end{bmatrix};$$

that is,

$$[T]_{i,j} = \delta_{i,j-1 \bmod p} \text{ and } [R]_{i,j} = \xi^i \delta_{i,j}$$

where

$$\delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

and the rows and columns are indexed $0, \ldots, p-1$. Now, matrices corresponding to the dit flip and phase flip errors described above are

$$T_a := T^{a_1} \otimes T^{a_2} \otimes \cdots \otimes T^{a_m} \text{ and } R_b := R^{b_1} \otimes R^{b_2} \otimes \cdots \otimes R^{b_m}.$$

Clearly,

$$\begin{aligned} T_a R_b &= \left(T^{a_1} \otimes T^{a_2} \otimes \cdots \otimes T^{a_m}\right)\left(R^{b_1} \otimes R^{b_2} \otimes \cdots \otimes R^{b_m}\right) \\ &= T^{a_1} R^{b_1} \otimes T^{a_2} R^{b_2} \otimes \cdots \otimes T^{a_m} R^{b_m}. \end{aligned}$$

Note that $\{T_a R_b : a, b \in \mathbb{F}_q\}$ is an orthogonal basis for $\mathbb{C}^q$ under the trace inner product $\langle A, B\rangle := Tr\left(A^\dagger B\right)$ where $A^\dagger$ denotes the Hermitian transpose of A. Thus, the span of $\{T_a R_b : a, b \in \mathbb{F}_q\}$ is the set of errors on a single qudit.

Next, we consider errors on an n-state system, that is, a system of n qudits. Given $a = (a_1, \ldots, a_n), b = (b_1, \ldots, b_n) \in \mathbb{F}_q^n$, define

$$T_a := T_{a_1} \otimes T_{a_2} \otimes \cdots \otimes T_{a_n} \text{ and } R_b := R_{b_1} \otimes R_{b_2} \otimes \cdots \otimes R_{b_n}.$$

Then

$$\begin{aligned} T_a R_b &= \left(T_{a_1} \otimes T_{a_2} \otimes \cdots \otimes T_{a_n}\right)\left(R_{b_1} \otimes R_{b_2} \otimes \cdots \otimes R_{b_n}\right) \\ &= T_{a_1} R_{b_1} \otimes T_{a_2} R_{b_2} \otimes \cdots \otimes T_{a_n} R_{b_n}. \end{aligned}$$

Given $a, b \in \mathbb{F}_q^n$, set $E_{a,b} := T_a R_b$. We sometimes write E_{ab} to mean $E_{a,b}$. Then the set

$$\mathcal{E}_n := \left\{E_{a,b} : a, b \in \mathbb{F}_q^n\right\}$$

is an error basis for $\mathbb{C}^{q^n}$. Hence, the error group for an n-state q-ary system is

$$G_n = \left\{\xi^i E_{a,b} : a, b \in \mathbb{F}_q^n, 0 \le i \le p-1\right\},$$

a group of order pq^{2n} with center $Z(G_n) = \langle \xi I\rangle$.

We now discuss when errors are correctable by a quantum code C. Let $\{|\psi_j\rangle : 1 \le j \le k\}$ be a basis for C. In order for errors E and F to be correctable, $E|\psi_i\rangle$ and $F|\psi_j\rangle$ must be distinguishable (meaning orthogonal) for all $i \ne j$; that is,

$$\langle \psi_i | E^\dagger F | \psi_j \rangle = 0.$$

Because measurement disturbs the state, error correction cannot be done by measurement; that is, an operation that causes measurement is not allowed. This includes anything that gives information about the state. For example, if $\langle \psi_i | E^\dagger F | \psi_i \rangle \ne \langle \psi_j | E^\dagger F | \psi_j \rangle$ for some $1 \le i, j \le k$, then

this measurement gives information about the state. Hence, an additional requirement for E and F to be correctable errors is that

$$\langle\psi_i|E^\dagger F|\psi_i\rangle = \langle\psi_j|E^\dagger F|\psi_j\rangle$$

for all $1 \leq i, j \leq k$. This discussion is summarized in the following result due to Knill and Laflamme [24] and Bennett, DiVincenzo, Smolin, and Wootters [4].

Theorem 12.3. *A set $\mathcal{A}$ of errors is correctable by a code C with basis $\{|\psi_j\rangle : 1 \leq j \leq k\}$ if and only if*

$$\langle\psi_i|E_a^\dagger E_b|\psi_j\rangle = C_{ab}\delta_{ij}$$

where E_a and E_b run over all possible errors in $\mathcal{A}$ and C_{ab} depends only on a and b (not on i and j).

The weight of an error $\xi^i E_{\mathbf{a},\mathbf{b}} \in G_n$ is the number of its nonidentity components, meaning

$$wt\left(\xi^i E_{\mathbf{a},\mathbf{b}}\right) = n - |\{i : a_i = b_i = 0\}|.$$

Given this notion of weight, we can now define the minimum distance of a q-ary quantum code C of length n to be

$$d = \max\left\{d : \begin{array}{r} \langle u|v\rangle = 0 \text{ and } wt(E) \leq d-1 \Rightarrow \langle u|E|v\rangle = 0 \\ \forall |u\rangle, |v\rangle \in C \text{ and } \forall E \in G_n \end{array}\right\}.$$

Definition 12.4. An $[[n, k, d]]_q$ code is a q-ary quantum code of length n, dimension k, and minimum distance d.

We will write $[[n, k, \geq d]]_q$ code to mean an q-ary quantum code of length n, dimension k, and minimum distance at least d. As is standard, a classical linear code of length n, dimension k, and minimum distance d (resp. at least d) is called an $[n, k, d]$ (resp. $[n, k, \geq d]$) code.

An $[[n, k, d]]_q$ code C is pure if and only if

$$wt(E) \leq d - 1 \Rightarrow \langle u|E|v\rangle = 0$$

for all $|u\rangle, |v\rangle \in C$ and all $E \in G_n$. Notice that the words u and v are not required to be orthogonal here. A weaker condition is that of nondegeneracy. An $[[n, k, d]]_q$ code C is nondegenerate if and only if

$$wt(E) \leq d - 1 \Rightarrow |u\rangle \text{ and } E|v\rangle \text{ are linearly independent}$$

for all $|u\rangle, |v\rangle \in C$ and all $E \in G_n$; otherwise C is said to be degenerate. While the term degenerate has seemingly negative connotations, we will see in the next subsection that this is not necessarily the case.

12.2.2. *Bounds on quantum codes*

Many of the classical coding theory bounds have analogs that apply to quantum codes.

Theorem 12.5. *(Quantum Hamming Bound) Given any* $[[n,k,d]]_q$ *nondegenerate quantum code,*

$$\sum_{j=0}^{\lfloor \frac{d-1}{2} \rfloor} \binom{n}{j} \left(q^2-1\right)^j q^k \leq q^n.$$

Proof. Suppose that C is a $[[n,k,d]]_q$ nondegenerate quantum code. Since C is nondegenerate, any two linearly independent correctable errors produce orthogonal q^k-dimensional subspaces of $\mathbb{C}^{q^n}$. Given $0 \leq j \leq \frac{d-1}{2}$, any j errors are correctable and there are

$$\binom{n}{j} \left(q^2-1\right)^j$$

such errors. From this, the bound follows. □

Notice that this bound only applies to nondegenerate codes. This suggests that it might be possible for a degenerate code to have parameters exceeding this bound. In 1997, Gottesman [13] proved that degenerate single- and double-error-correcting binary codes satisfy the bound given in Theorem 12.5. Nearly a decade later, it was shown for degenerate q-ary stabilizer codes of minimum distance 3 [22] and minimum distance 5 [1]. A major open problem in quantum coding theory is to determine if there is a Hamming bound that applies to degenerate codes.

Quantum codes also satisfy MacWilliams identities [32]. Using these, one can prove a quantum analog of the classical singleton bound.

Theorem 12.6. *(Quantum Singleton Bound) If* C *is a* $[[n,k,d]]_q$ *code with* $k > 1$*, then*

$$k + 2d \leq n + 2.$$

A quantum maximum distance separable (MDS) code is a quantum code which attains the Singleton bound. Rains [28, Theorem 2] showed that all quantum MDS codes are pure. There is an interesting relationship betweeen quantum MDS codes and classical MDS codes. If Q is a quantum MDS stabilizer code with $n-2d+2 > 0$, then it gives rise to classical MDS codes [22, Lemma 61]. Recall that the MDS conjecture for classical codes

says: "If there is a nontrivial $[n,k,d]_q$ MDS code, then $n \leq q+1$ unless q is even and $k=3$ or $k=q-1$ in which case $n \leq q+2$." The classical MDS conjecture implies that there are no nontrivial MDS stabilizer codes of lengths greater than q^2+1, except when q is even and $d=4$ or $d=q^2$ in which case $n \leq q^2+2$ [22, Corollary 65]. Therefore, the discovery of certain quantum MDS codes could provide a route to disproving the classical MDS conjecture. This is an active area of research in quantum error-correcting codes.

Theorem 12.7. *(Quantum Gilbert-Varshamov Bound) Suppose that* $2 \leq k < n$, $d \geq 2$ *and* $n \equiv k \bmod 2$. *If*

$$\sum_{j=1}^{d-1} \binom{n}{j} (q^2-1)^{j-1} < \frac{q^{2n}-1}{q^{n+k}-q^{n-k}},$$

then there exists a $[[n,k,d]]_q$ *code.*

Recently, Feng and Ma proved a Gilbert-Varshamov type bound which guarantees the existence of pure codes.

Theorem 12.8. *(Gilbert-Varshamov Bound for pure stabilizer codes) [9, Theorem 1.4] Suppose that* $2 \leq k < n$, $d \geq 2$ *and* $n \equiv k \bmod 2$. *If*

$$\sum_{j=1}^{d-1} \binom{n}{j} (q^2-1)^{j-1} < \frac{q^{n-k+2}-1}{q^2-1},$$

then there exists a $[[n,k,d]]_q$ *pure code.*

Asymptotically, these two bounds coincide. We will consider the asymptotic version in Section 12.4. The statements in Theorems 12.7 and 12.8 may be made a bit stronger. Under the given hypotheses, there exists a stabilizer code with the given parameters. Stabilizer codes are discussed in the next section.

12.3. Relating quantum codes and classical codes

While classical linear codes may be compactly described in terms of a basis, this may not be the most concise expression for a quantum code (see Gottesman's thesis [13] for some examples illustrating this). In fact, for a large class of quantum codes called stabilizer codes, another algebraic structure is more useful. Stabilizer codes over $\mathbb{F}_2$ were introduced by Gottesman in his thesis [13], and many of the same ideas were discovered independently

by Calderbank, Rains, Shor, and Sloane [5] and used in the famous CSS constuction.

12.3.1. *Stabilizer codes*

Some believe stabilizer codes to be the quantum analog of linear codes. The stabilizer can be thought of as the quantum analog of a classical parity check matrix. While not every code is a stabilizer code, the following is true: Given a quantum code C, there is a stabilizer code C' such that $C \subseteq C'$ [22]; hence, knowledge of (lower bounds on) the error-correcting capability of stabilizer codes provides information about the capabilities of arbitrary quantum codes.

Definition 12.9. A q-ary quantum stabilizer code C of length n is a joint eigenspace of operators of an abelian subgroup S of G_n; that is,

$$C = \left\{ u \in \mathbb{C}^{q^n} : Mu = u \ \forall M \in S \right\}.$$

The fact that S is abelian guarantees that the code is nontrivial. To see this, suppose $M, N \in S$. Then

$$MN|\psi\rangle = M|\psi\rangle = |\psi\rangle \text{ and } NM|\psi\rangle = N|\psi\rangle = |\psi\rangle$$

which imply

$$(MN - NM)\,|\psi\rangle = MN|\psi\rangle - NM|\psi\rangle = 0.$$

It follows that $MN = NM$ or $|\psi\rangle = 0$. As a result, S must be abelian or $C = \{|0\rangle\}$. (If S is nonabelian, it is standard to extend S by $Z(G_n)$.)

We do not have space to prove or even mention all of the facts on stabilizer codes. Instead, we point the reader to the excellent references [2], [13], and [22]. There one can find the following result.

Proposition 12.10. *A stabilizer code C with stabilizer $S \subseteq G_n$ has $\frac{q^n}{|S|}$ codewords and minimum distance* $\min wt\{M \in N(S) \setminus S\}$ *where $N(S)$ denotes the normalizer of S.*

In the next subsection, we consider some stabilizer codes constructed from classical linear codes.

12.3.2. *CSS construction*

In this section, we describe a large class of quantum stabilizer codes based on classical linear codes.

Recall that an additive code of length n over $\mathbb{F}_4$ is an additive subgroup of $\mathbb{F}_4^n$. Write $\mathbb{F}_4 = \{0, \omega, \omega^2, 1\}$ where $\omega^2 = \omega + 1$ so that $\overline{\omega} = \omega^2$. Then $\{\omega, \overline{\omega}\}$ is a basis for $\mathbb{F}_4$ as an $\mathbb{F}_2$-vector space. Hence, given $v \in \mathbb{F}_4^n$,

$$v = \omega a + \overline{\omega} b$$

for some $a, b \in \mathbb{F}_2^n$. This defines a bijection

$$\begin{array}{rccc} f: & \mathbb{F}_4^n & \to & \mathbb{F}_2^{2n} \\ & \omega a + \overline{\omega} b & \mapsto & (a|b)\,. \end{array}$$

This bijection may be composed with

$$\begin{array}{rccc} g: & \mathbb{F}_2^{2n} & \to & G_n \\ & (a|b) & \mapsto & E_{ab} \end{array}$$

to produce

$$\begin{array}{rccc} \phi: & \mathbb{F}_4^n & \to & G_n \\ & \omega a + \overline{\omega} b & \mapsto & E_{ab}. \end{array}$$

In [5], additive codes over $\mathbb{F}_4$ are used to construct quantum codes via the following major result. Here, the inner product employed is the trace inner product defined by

$$u * v := Tr\,(u \cdot v)$$

for all $u, v \in \mathbb{F}_4^n$, where the trace map is

$$\begin{array}{rcl} Tr: \mathbb{F}_4 \to & \mathbb{F}_2 \\ x & \mapsto x + \overline{x} \end{array}$$

and $u \cdot v := \sum_{i=1}^{n} u_i v_i$ is the usual inner product. Recall that a code C is self-orthogonal (or weakly self-dual) provided $C \subseteq C^{\perp}$.

Theorem 12.11. *[5, Theorem 2] Suppose that $D \subseteq \mathbb{F}_4^n$ is an additive self-orthogonal code such that $|D| = 2^{n-k}$ and $D^{\perp} \setminus D$ has no vectors of weight less than d. Then any eigenspace of $\phi(D)$ is a $[[n, k, d]]_2$ code.*

Classical binary linear codes may be employed in Theorem 12.11 as follows. Suppose that C_1 is a $[n, k_1, d_1]_2$ code and C_2 is $[n, k_2, d_2]_2$ code where $C_1 \subseteq C_2$. Then

$$D = \omega C_1 + \overline{\omega} C_2^{\perp} \subseteq \mathbb{F}_4^n$$

is an additive code over $\mathbb{F}_4$. Moreover, D is self-orthogonal with respect to the trace inner product. To see this, note that

$$\begin{aligned} Tr\left((\omega a+\overline{\omega}b)\cdot\overline{(\omega a'+\overline{\omega}b')}\right) &= \textstyle\sum_{i=1}^{n} a_i b_i' Tr\left(\overline{\omega}\right)+a_i' b_i Tr\left(\omega\right) \\ &= (a|b)\cdot(a'|b') \\ &= 0 \end{aligned}$$

for all $(\omega a+\overline{\omega}b),(\omega a'+\overline{\omega}b')\in D$ as $a,a'\in C_1\subseteq C_2$ and $b,b'\in C_2^{\perp}$. Applying Theorem 12.11 to D as above produces what is commonly called the CSS construction for binary quantum codes, one of the most important constructions of quantum codes. This turns out to be a special case of a q-ary construction which is given in Corollary 12.15.

Theorem 12.12. *(Binary CSS construction) Suppose that C_1 and C_2 are binary linear codes of length n and dimensions k_1 and k_2 respectively with $C_1\subseteq C_2$. Then there exists a $\left[\left[n,k_2-k_1,min\{d(C_2\setminus C_1),d\left(C_1^{\perp}\setminus C_2\right)\}\right]\right]_2$ code.*

Following Rains' work on nonbinary quantum codes [28], Ashikhmin and Knill developed a q-ary analog to Theorem 12.11. Notice that a code C of length n over $\mathbb{F}_4$ is additive if and only if C is an $\mathbb{F}_2$-subspace of $\mathbb{F}_4^n$. Hence, in the q-ary case where $q=p^m$, the notion of an additive code is replaced with that of an $\mathbb{F}_p$-subspace. Such a code is said to be $\mathbb{F}_p$-linear. More precisely, we have the following definition.

Definition 12.13. Suppose $q=p^m$ where p is prime. An $\mathbb{F}_p$-linear code of length n over $\mathbb{F}_q$ is an $\mathbb{F}_p$-subspace of $\mathbb{F}_q^n$.

Consider

$$\begin{array}{rccc} g: & \mathbb{F}_q^{2n} & \rightarrow & G_n \\ & (a|b) & \mapsto & E_{ab}. \end{array}$$

To generalize Theorem 12.11 to the q-ary case, one may use a generalization of the trace inner product defined about. Given $(a|b),(a'|b')\in\mathbb{F}_q^{2n}$, set

$$(a|b)*(a'|b')=Tr\left(a\cdot b'-a'\cdot b\right)$$

where $Tr:\mathbb{F}_q\rightarrow\mathbb{F}_p$ is the usual trace map.

Theorem 12.14. *[2, p. 3069] Suppose that $D\subseteq\mathbb{F}_q^{2n}$ is an $\mathbb{F}_p$-linear code which is self-orthogonal with respect to $*$ such that $|D|=p^r$. Then any eigenspace of $g(D)$ is a $\left[\left[n,n-\frac{r}{m},d\left(D^{\perp *}\setminus D\right)\right]\right]_q$ code.*

Classical q-ary codes may be employed in Theorem 12.14. To do so, consider a degree two extension $\mathbb{F}_{q^2}$ of $\mathbb{F}_q$. Suppose that ω is a primitive element of $\mathbb{F}_{q^2}$ so that $\{\omega, \overline{\omega}\}$ is a basis for $\mathbb{F}_{q^2}$ over $\mathbb{F}_q$. Define

$$\begin{array}{rccc} f: & \mathbb{F}_{q^2}^n & \to & F_q^{2n} \\ & \omega a + \overline{\omega} b & \mapsto & (a|b)\,. \end{array}$$

This results in a q-ary version of the CSS construction.

Corollary 12.15. *(q-ary CSS construction) [16, 22, 23] Suppose that C_1 and C_2 are linear codes over $\mathbb{F}_q$ of length n and dimensions k_1 and k_2 respectively with $C_1 \subseteq C_2$. Then there exists a $\left[[n, k_2 - k_1, min\{d(C_2 \setminus C_1), d\left(C_1^{\perp} \setminus C_2\right)\}]\right]_q$ code.*

Proof. Set $C = \omega C_1 + \overline{\omega} C_2^{\perp} \subseteq \mathbb{F}_{q^2}^n$ and $D = f(C) \subseteq \mathbb{F}_q^{2n}$. Then D is self-orthogonal with respect to $*$ (see [23, Lemma 2.5, Proposition 2.6]). Now Theorem 12.14 gives the desired result. □

Next, we see how another inner product on $\mathbb{F}_{q^2}^n$ may be utilized to construct quantum codes over $\mathbb{F}_q$. Recall that the Hermitian inner product on $\mathbb{F}_{q^2}^n$ is given by

$$u *_h v := \sum_{i=1}^{n} u_i v_i^q.$$

In [2, Theorem 4], it is shown that a code which is self-orthogonal with respect to the Hermitian inner product is also self-orthogonal with respect to $*$. This idea can be used to construct q-ary quantum codes.

Corollary 12.16. *[2, Corollary 1] Suppose that D is a $[n, k, d]_{q^2}$ code which is self-orthogonal with respect to the Hermitian inner product. Let $D^{\perp_h}$ denote the Hermitian dual of D. Then there exists a $\left[[n, n-2k, min\{wt\left(D^{\perp_h} \setminus D\right)\}]\right]_q$ code.*

An $[n, k, d]_q$ code is pure if its dual contains no nonzero vectors of weight less than d. For example, a self-dual code is pure. Suppose a quantum code Q is constructed from a classical code C in the CSS construction (taking $C_1 = C_2 = C$ in Corollary 12.15). Then Q is pure if and only if C is pure.

12.4. Quantum codes constructed from algebraic geometry codes

In this section we employ algebraic geometry codes in the construction of quantum codes. We consider several families of such codes as well as

asymptotic results. To begin, we review the notation used in this section.

Let X be a smooth, projective, absolutely irreducible curve of genus g over a finite field $\mathbb{F}_q$. Let $\mathbb{F}_q(X)$ denote the field of rational functions on X defined over $\mathbb{F}_q$, and let $\Omega(X)$ denote the set of differentials on X defined over $\mathbb{F}_q$. The divisor of a rational function f (resp. differential η) will be denoted by (f) (resp. (η)). Given a divisor A on X defined over $\mathbb{F}_q$, let

$$\mathcal{L}(A) = \{f \in \mathbb{F}_q(X) : (f) \geq -A\} \cup \{0\}$$

and

$$\Omega(A) = \{\eta \in \Omega(X) : (\eta) \geq A\} \cup \{0\}.$$

Let $\ell(A)$ denote the dimension of $\mathcal{L}(A)$ as an $\mathbb{F}_q$-vector space. The support of a divisor D is denoted by $suppD$.

Algebraic geometry codes $C_{\mathcal{L}}(D, G)$ and $C_{\Omega}(D, G)$ can be constructed using divisors $D = \sum_{i=1}^{n} P_i$ and $G = \sum_{i=1}^{m} \alpha_i Q_i$ on X where $P_1, \ldots, P_n, Q_1, \ldots, Q_m$ are pairwise distinct $\mathbb{F}_q$-rational points and $\alpha_i \in \mathbb{N}$ for all i, $1 \leq i \leq m$. In particular,

$$C_{\mathcal{L}}(D, G) := \{(f(P_1), \ldots, f(P_n)) : f \in \mathcal{L}(G)\}$$

and

$$C_{\Omega}(D, G) := \{(res_{P_1}(\eta), \ldots, res_{P_n}(\eta)) : \eta \in \Omega(G - D)\}.$$

These codes are sometimes called m-point codes since the divisor G has m distinct $\mathbb{F}_q$-rational points in its support. Typically, an m-point code is constructed by taking the divisor D to be the sum of all $\mathbb{F}_q$-rational points not in the support of G, and we will keep this convention. We will use the term multipoint code to mean an m-point code with $m \geq 2$.

The two algebraic geometry codes above are related in that

$$C_{\mathcal{L}}(D, G)^{\perp} = C_{\Omega}(D, G).$$

If $\deg G < n$, then $C_{\mathcal{L}}(D, G)$ has length n, dimension $\ell(G)$, and designed distance $n - \deg G$. If $\deg G > 2g - 2$, then $C_{\Omega}(D, G)$ has dimension $\ell(K + D - G)$, where K is a canonical divisor, and designed distance $\deg G - (2g - 2)$. The minimum distance of each of the codes $C_{\mathcal{L}}(D, G)$ and $C_{\Omega}(D, G)$ is at least its designed distance.

For more background on AG codes, the reader may consult [12], [35], or [39].

12.4.1. *Families of quantum codes from one-point AG codes*

12.4.1.1. *Quantum Reed-Solomon codes*

Perhaps the most popular family of AG codes is the class of Reed-Solomon codes which are one-point AG codes on the projective line. Prior to the work on nonbinary quantum codes [2], Grassl, Geiselmann, and Beth [17] generalized some of the ideas in [5] from $\mathbb{F}_4$ to higher degree extensions of $\mathbb{F}_2$. Specifically, they considered Reed-Solomon codes over $\mathbb{F}_{2^t}$ and their binary expansions. Let $\{b_1, \ldots b_t\}$ be a basis for $\mathbb{F}_{2^t}$ as an $\mathbb{F}_2$-vector space. Define

$$\begin{array}{rccc} \mathcal{B}: & \mathbb{F}_{2^t} & \to & \mathbb{F}_2^t \\ & \sum_{i=1}^t a_i b_i & \mapsto & (a_1, \ldots, a_t)\,. \end{array}$$

Given a $[n,k,d]_{2^t}$ code C, $\mathcal{B}(C)$ is a $[tn, tk, \geq d]_2$ code. By [17, Theorem 1],

$$\mathcal{B}(C)^\perp = \mathcal{B}^\perp\left(C^\perp\right).$$

Hence, if the basis is chosen to be self-dual (which it can be according to [30, Theorem 4]) and the code C is self-orthogonal, then

$$\mathcal{B}(C) \subseteq \mathcal{B}\left(C^\perp\right) = \mathcal{B}\left(C\right)^\perp.$$

Recall that an $[n,k,d]_{2^t}$ Reed-Solomon code is self-dual provided $2k < n$. Using this fact together with their precursor to Corollary 12.15, Grassl et. al. obtain the following.

Proposition 12.17. *[17] Given $\delta > \frac{2^t-1}{2} + 1$, there is a quantum Reed-Solomon code with parameters $[[t\left(2^t-1\right), t\left(2\delta - 2^t - 1\right), \geq 2^t - \delta + 1]]_2$.*

Proof. Let C be an $[2^t - 1, 2^t - \delta, \delta]_{2^t}$ Reed-Solomon code where $\delta > \frac{2^t-1}{2} + 1$. Then C is self-orthogonal. Now apply Corollary 12.15 with $C_1 = C_2 = \mathcal{B}(C)$ where $\mathcal{B}$ is a self-dual basis for $\mathbb{F}_{2^t}$ over $\mathbb{F}_2$. The result follows immediately. □

See [14] for applications of other cyclic codes to the construction of quantum codes.

Quantum Reed-Solomon codes over fields of odd characteristic may be constructed too. We do not provide the details here as this construction is a special case of a result in Subsection 12.4.2. Extended Reed-Solomon codes have also been used to construct quantum MDS codes as in [16].

Table 12.1. Parameters of the Hermitian code $C_{\mathcal{L}}(P_1 + \cdots + P_{q^3}, \alpha P_\infty)$

α	$k(\alpha)$	$d(\alpha)$
$0 \leq \alpha \leq q^2 - q - s$ $\underline{\alpha} = sq + t$ $0 \leq b \leq a \leq q-1$	$\frac{a(a+1)}{2} + b + 1$	$q^3 - \underline{\alpha}$
$q^2 - q - 2 < \alpha < q^3 - q^2 + q$	$\alpha + 1 - \frac{q(q-1)}{2}$	$n - \alpha$
$q^3 - q^2 + q \leq \alpha < q^3$ $\alpha = q^3 - q^2 + aq + b$ $0 \leq a, b \leq q-1$	$\alpha + 1 - \frac{q(q-1)}{2}$	$q^3 - \alpha$ if $a < b$ $q^3 - \alpha + b$ if $a \geq b$
$q^3 \leq \alpha \leq q^3 + q^2 - q - 2$ $\underline{q^3 + q^2 - q - 2 - \alpha} = aq + b$ $0 \leq b \leq a \leq q-1$	$q^3 - \frac{a(a+1)}{2} - b - 1$	$a + 2$ if $b = a$ $a + 1$ if $b < a$

12.4.1.2. *Quantum Hermitian codes*

Next to Reed-Solomon codes, Hermitian codes are certainly the most studied algebraic geometry codes. Recall that the exact parameters of one-point Hermitian codes are known due to [41]. For reference, Table 12.1. gives the dimension $k(\alpha)$ and minimum distance $d(\alpha)$ of the code $C_{\mathcal{L}}(P_1 + \cdots + P_{q^3}, \alpha P_\infty)$ where $P_1, \ldots, P_{q^3}, P_\infty$ are all of the $\mathbb{F}_{q^2}$-rational points of the Hermitian curve defined by $y^q + y = x^{q+1}$. Here $\underline{\alpha} = max\{a \in H(P_\infty) : a \leq \alpha\}$ is the largest element of the Weierstrass semigroup at the point P_∞ that is no bigger than α.

If $\alpha_1 < \alpha_2$, then

$$C_{\mathcal{L}}(D, \alpha_1 P_\infty) \subseteq C_{\mathcal{L}}(D, \alpha_2 P_\infty).$$

Applying Corollary 12.15 with $C_1 = C_{\mathcal{L}}(D, \alpha_1 P_\infty)$ and $C_2 = C_{\mathcal{L}}(D, \alpha_2 P_\infty)$ yields the following fact.

Theorem 12.18. *[29, Theorem 3] For $0 \leq \alpha_1 < \alpha_2 \leq q^3 + q^2 - q - 2$, there exists a $\left[\left[q^3, k(\alpha_2) - k(\alpha_1), \geq min\{d(\alpha_2), d\left(q^3 + q^2 - q - 2 - \alpha_1\right)\}\right]\right]_{q^2}$ code where $k(\alpha)$ and $d(\alpha)$ are given in Table 12.1.*

Quantum Hermitian codes can also be constructed using Hermitian codes which are self-orthogonal with respect to the Hermitian inner product. Recall that the dual of the one-point Hermitian code $C_{\mathcal{L}}(D, \alpha P_\infty)$ over $\mathbb{F}_{q^2}$ is given by

$$C_{\mathcal{L}}(D, \alpha P_\infty)^\perp = C_{\mathcal{L}}\left(D, \left(q^3 + q^2 - q - 2 - \alpha\right) P_\infty\right)$$

as shown in [36, 38]. It follows that $C_{\mathcal{L}}(D, \alpha P_\infty)$ is self-orthogonal if $2\alpha \leq q^3 + q^2 - q - 2 - \alpha$. Using this, one can prove that $C_{\mathcal{L}}(D, \alpha P_\infty)$ is self-orthogonal with respect to the Hermitian inner product for $0 \leq \alpha \leq q^2 - 2$

(see [29, Lemma 7] for details). Now Corollary 12.16 gives another family of quantum Hermitian codes.

Theorem 12.19. *[29, Theorem 8] If $0 < \alpha \leq q^2 - 2$, then there exists a $\left[\left[q^3, q^3 - 2k(\alpha), \geq d\left(q^3 + q^2 - q - 2 - \alpha\right)\right]\right]_q$ code where $k(\alpha)$ and $d(\alpha)$ are given in Table 12.1.*

12.4.2. *More general AG constructions*

The quantum Reed-Solomon and quantum Hermitian codes defined earlier in this section are special cases of a more general construction for quantum codes from AG codes detailed in this section.

Let X be a smooth, projective, absolutely irreducible curve of genus g over a finite field $\mathbb{F}_q$. Suppose that A and B are divisors on X such that $A \leq B$, and let $D = P_1 + \cdots + P_n$ be another divisor on X whose support consists of n distinct $\mathbb{F}_q$-rational points none of which are in the support of A or B. Then

$$\mathcal{L}(A) \subseteq \mathcal{L}(B)$$

and so

$$C_{\mathcal{L}}(D, A) \subseteq C_{\mathcal{L}}(D, B).$$

Applying Corollary 12.15, we find a large family of quantum codes from AG codes.

Theorem 12.20. *Let A, B, and $D = P_1 + \cdots + P_n$ be divisors on a smooth, projective, absolutely irreducible curve X of genus g over $\mathbb{F}_q$. Assume that $A \leq B$ and $(suppA \cup suppB) \cap suppD = \emptyset$ and $\deg B < n$. Then there exists a $[[n, \ell(B) - \ell(A), d]]_q$ code where*

$$\begin{aligned} d &\geq \min\{d\left(C_{\mathcal{L}}(D, B) \setminus C_{\mathcal{L}}(D, A)\right), d\left(C_{\Omega}(D, A) \setminus C_{\Omega}(D, B)\right)\} \\ &\geq \min\{n - \deg B, \deg A - (2g - 2)\}. \end{aligned}$$

Proof. This follows immediately from Corollary 12.15 (taking $C_1 = C_{\mathcal{L}}(D, A)$ and $C_2 = C_{\mathcal{L}}(D, B)$) and the fact that $\deg A \leq \deg B < n$ implies $dim C_{\mathcal{L}}(D, B) = \ell(B)$ and $dim C_{\mathcal{L}}(D, A) = \ell(A)$. □

In the next example, we see how one may apply Theorem 12.20 to a multipoint code.

Example 12.21. Let X be a smooth, projective, absolutely irreducible curve of genus g over $\mathbb{F}_q$. Consider the m-point code $C_{\mathcal{L}}(D, \sum_{i=1}^{m} a_i Q_i)$

on X over $\mathbb{F}_q$. Since $\mathbb{F}_q$ is finite, the class number of the function field of X over $\mathbb{F}_q$ is finite [35, Proposition V.1.3]. Hence, there exists a rational function f with divisor

$$(f) = \sum_{i=2}^{m} b_i Q_i - b_1 Q_1$$

where $b_i \geq a_i$ for all i, $2 \leq i \leq m$, and $b_1 := \sum_{i=2}^{m} b_i$. Multiplication by f gives rise to a vector space isomorphism

$$\begin{array}{ccc} \phi : \mathcal{L}\left(\sum_{i=1}^{m} a_i Q_i\right) & \to & \mathcal{L}\left((a_1+b_1)\,Q_1 - \sum_{i=2}^{m}(b_i-a_i)\,Q_i\right) \\ h & \mapsto & fh \end{array}$$

which in turn induces an isometry ϕ^* of codes

$$C_{\mathcal{L}}\left(D, \sum_{i=1}^{m} a_i Q_i\right) \cong C_{\mathcal{L}}\left(D, (a_1+b_1)\,Q_1 - \sum_{i=2}^{m}(b_i-a_i)\,Q_i\right).$$

Since $(a_1+b_1)\,Q_1 - \sum_{i=2}^{m}(b_i-a_i)\,Q_i \leq (a_1+b_1)\,Q_1$,

$$C_{\mathcal{L}}\left(D, (a_1+b_1)\,Q_1 - \sum_{i=2}^{m}(b_i-a_i)\,Q_i\right) \subseteq C_{\mathcal{L}}\left(D, (a_1+b_1)\,Q_1\right).$$

Therefore, if $a_1 + b_1 < |suppD|$ then Theorem 12.20 yields a quantum code over $\mathbb{F}_q$ of length $|suppD|$ and dimension

$$\ell\left((a_1+b_1)\,Q_1\right) - \ell\left((a_1+b_1)\,Q_1 - \sum_{i=2}^{m}(b_i-a_i)\,Q_i\right).$$

A bound on the minimum distance is given by the theorem also. However, the weights of words in multipoint codes are not typically known. As a result, determining the minimum distance of the quantum code may be challenging. A notable exception to this is family of two-point Hermitian codes whose exact minimum distance has been determined in the extensive recent work of Homma and Kim [18], [19], [20], [21].

Of course, one may also apply Theorem 12.20 to nested multipoint codes. While this construction provides a great deal of flexibility, it produces codes whose minimum distances may be hard to determine. For this reason, we will not elaborate on this idea here.

Next, we consider how Corollary 12.16 may be applied to AG codes. The idea is a generalization of Theorem 12.19.

Lemma 12.22. *The algebraic geometry code $C_{\mathcal{L}}(D,G)$ is self-orthogonal with respect to the Hermitian inner product if there exists a differential η such that $v_{P_i}(\eta) = -1$, $\eta_{P_i}(1) = 1$ for $1 \le i \le n$, and*

$$D + (\eta) \ge (q+1)\, G. \tag{12.1}$$

Proof. Let $D = P_1 + \cdots + P_n$ and G be divisors on a smooth, projective, absolutely irreducible curve X over $\mathbb{F}_q$ where $P_1, \ldots, P_n$ are distinct $\mathbb{F}_q$-rational points not in the support of G. Recall that the dual of $C_{\mathcal{L}}(D,G)$ may be expressed as

$$C_{\mathcal{L}}(D,G)^{\perp} = C_{\mathcal{L}}(D, D - G + (\eta))$$

where η is a differential on X such that $v_{P_i}(\eta) = -1$ and $\eta_{P_i}(1) = 1$ for $1 \le i \le n$. Notice that for $h \in \mathcal{L}(G)$,

$$\begin{array}{rll} ev(f) *_h ev(h) = 0 & \text{iff} & \sum_{i=1}^{n} f(P_i) h^q(P_i) = 0 \quad \forall f \in \mathcal{L}(G) \\ & \text{iff} & h^q \in \mathcal{L}(D - G + (\eta)) \\ & \text{iff} & q(h) \ge G - D - (\eta) \\ & \text{if} & -qG \ge G - D - (\eta) \end{array}$$

where $ev(f) := (f(P_1), \ldots, f(P_n))$ It follows that given $h \in \mathcal{L}(G)$, $ev(f) *_h ev(h) = 0$ for all $f \in \mathcal{L}(G)$ if

$$D + (\eta) \ge (q+1)\, G.$$

□

The next result is a consequence of the lemma above. Here, P_{00} denotes the common zero of the functions x and y on the Hermitian curve over $\mathbb{F}_{q^2}$.

Proposition 12.23. *Suppose that $0 \le a + b < q^2 - 2$. Then the two-point code $C_{\mathcal{L}}(D, aP_\infty + bP_{00})$ on the Hermitian curve defined by $y^q + y = x^{q+1}$ over $\mathbb{F}_{q^2}$ is self-orthogonal with respect to the Hermitian inner product.*

Proof. Take $\eta = \frac{y^{b+1}}{z} dz$ where $z = x^{q^2} - x$. Then

$$(\eta) = \left(q^3 + q^2 - q - (b+1)(q+1)\right) P_\infty - \left((b+1)(q+1) + 1\right) P_{00} - D$$

and the conditions of Lemma 12.22 are satisfied. □

Proposition 12.24. *Let $0 \le a + b < q^2 - 2$. Then there exists a $\left[\left[q^3 - 1, q^3 - 2\ell\,(aP_\infty + bP_{00}) - l, d\right]\right]_q$ code where*

$$d = min\{wt\left(C_{\mathcal{L}}(D, aP_\infty + bP_{00})^{\perp_h} \setminus C_{\mathcal{L}}(D, aP_\infty + bP_{00})\right)\}.$$

12.4.3. *Quantum codes from hyperelliptic curves*

In this subsection, we review Niehage's construction of quantum codes using hyperelliptic curves over finite fields [26]. This approaches uses ideas of Matsumoto [25].

Given $a_1, \ldots, a_n \in \mathbb{F}_q \setminus \{0\}$, define a weighted symplectic inner product on $\mathbb{F}_q^{2n}$ by

$$u *_a v := \sum_{i=1}^{n} a_i \left(u_i v_{i+n} - u_{i+n} v_i\right).$$

The weighted symplectic inner product gives more flexibility in the construction of quantum codes. However, a code C which is self-orthogonal with respect to $*_a$ may not be self-orthogonal with respect to the standard symplectic inner product $*$. To correct for this, the codewords of C are multiplied by $(a_1, \ldots, a_n, 1, \ldots, 1)$. This is detailed in the following lemma.

Lemma 12.25. *[26, Lemma 1] Let C be a linear code of length $2n$ over $\mathbb{F}_q$ that is self-orthogonal with respect to $*_a$. Let M denote the generator matrix for the quantum code defined by C. Then the code C' with generator matrix*

$$M' := M \cdot diag\,(a_1, \ldots, a_n, 1, \ldots, 1)$$

is a stabilizer code (with respect to the standard symplectic inner product) with the same parameters as C.

Proof. Suppose that $C \subseteq \mathbb{F}_q^{2n}$ is self-orthogonal with respect to $*_a$. Then

$$0 = u *_a v = \sum_{i=1}^{n} a_i \left(u_i v_{i+n} - u_{i+n} v_i\right) = \sum_{i=1}^{n} \left(\left(a_i u_i\right) v_{i+n} - u_{i+n} \left(a_i v_i\right)\right)$$

for all $u, v \in C$. This proves that

$$C' := \{(a_1 c_1, \ldots, a_n c_n, c_{n+1}, \ldots, c_{2n}) : c \in C\}$$

is self-orthogonal with respect to $*$. □

Next, we describe how to use $*_a$ and a hyperelliptic curve X over $\mathbb{F}_q$ to produce quantum codes. Let X be a smooth, projective, absolutely irreducible curve over $\mathbb{F}_q$ with an automorphism σ of order two that fixes the elements of $\mathbb{F}_q$. Set

$$D = P_1 + \cdots + P_n + \sigma P_1 + \cdots + \sigma P_n$$

where $P_1, \ldots, P_n, \sigma P_1, \ldots, \sigma P_n$ are distinct $\mathbb{F}_q$-rational points on X, and take G to be a divisor on X defined over $\mathbb{F}_q$ that is fixed by σ and $suppG \cap suppD = \emptyset$. Suppose η is a differential on X satisfying

$$v_{P_i}(\eta) = v_{\sigma P_i}(\eta) = -1$$

and

$$res_{P_i}(\eta) = -res_{\sigma P_i}(\eta)$$

for all $1 \leq i \leq n$. Then it can be shown (as in [26, Proposition 3] and [25, Proposition 1]) that

$$C_{\mathcal{L}}(D,G)^{\perp_a} = C_{\mathcal{L}}(D, D - G + (\eta)).$$

By an argument similar to that of Lemma 12.22, if $G \leq D - G + (\eta)$ then $C_{\mathcal{L}}(D,G)$ is self-orthogonal with respect to $*_a$. Now Lemma 12.25 implies that $C_{\mathcal{L}}(D,G)'$ is self-orthogonal with respect to $*$. This construction gives rise to quantum AG codes from hyperelliptic curves as discussed in [26].

12.4.4. *Asymptotic results*

Since their introduction by Goppa [11], algebraic geometry codes have been a tool for obtaining asymptotic results [40]. In this section, we describe families of asymptotically good quantum codes from AG codes.

Given a family of quantum $[[n_i, k_i, d_i]]$ codes, let $R = \lim_{n\to\infty} \frac{k_i}{n_i}$ and $\delta = \lim_{n\to\infty} \frac{d_i}{n_i}$. If $R > 0$ and $\delta > 0$, then the family is called *good*.

In [3], Ashikhmin, Litsyn, and Tsfasman proved that there exist asymptotically good families of binary quantum codes as follows.

Theorem 12.26. *[3] For any $\delta \in (0, \frac{1}{18}]$ and R lying on the broken line given by the piecewise linear function*

$$R(\delta) = 1 - \frac{1}{2^{m-1}-1} - \frac{10}{3}m\delta \text{ for } \delta \in [\delta_m, \delta_{m-1}],$$

where $m = 3, 4, 5, \ldots, \delta_2 = \frac{1}{18}, \delta_3 = \frac{3}{56}$, and

$$\delta_m = \frac{3}{5}\frac{2^{m-2}}{(2^{m-1}-1)(2^m-1)} \text{ for } m = 4, 5, 6, \ldots,$$

there exist polynomially constructible families of binary quantum codes with $n \to \infty$ and asymptotic parameters greater than or equal to (δ, R).

Later, Chen, Ling, and Xing improved the above theorem on certain intervals.

Theorem 12.27. *[8] Let*

$$\delta_t = \frac{2}{3}\frac{2^t - 3}{(2t+1)(2^t-1)}.$$

For $t \geq 3$ and $\delta \in (0, \delta_t)$, there exist polynomially constructible families of binary quantum codes with $n \to \infty$ and asymptotic parameters $(\delta, R_1(\delta))$, where $R_1(\delta) = 3t(\delta_t - \delta)$.

Remark 12.28. When $t = 3$, the above theorem gives the line given by $R_1 + 9\delta = \frac{30}{49}$ in $(0, \frac{10}{147})$. This line exceeds the Ashikhmin-Litsyn-Tsfasman bound in the interval $(\frac{8}{147}, \frac{1}{18})$.

Kim and Walker [23] generalized the ideas of Chen-Ling-Xing's construction to non-binary quantum codes and obtained the following.

Theorem 12.29. *[23] Let p be any prime number. If p is odd, choose integers $t \geq 1$ and $r \geq 0$ such that $2t + r \leq p + 1$. If $p = 2$, then choose integers $t \geq 3$ and $r = 1$. Let*

$$\delta(p, r, t) = \frac{(r+1)(p^t - 3)}{(r+2)(2t+r)(p^t-1)}.$$

Then for any δ with $0 < \delta < \delta(p, r, t) < \frac{1}{4}$, there exist polynomially constructible families of p-ary quantum codes with $n \to \infty$ and asymptotic parameters at least $(\delta, R_p(\delta))$, where

$$R_p(\delta) = \frac{2t(r+2)}{r+1}(\delta(p, r, t) - \delta).$$

Note that when $p = 2$, this theorem implies Theorem 12.27.

Proof. (Sketch of proof) We follow [23]. Let X be a smooth, projective, absolutely irreducible curve over $\mathbb{F}_q$ of genus g. Let G be a divisor, which is a multiple of a fixed $\mathbb{F}_q$-rational point P_0, and let D be the sum of all the other N $\mathbb{F}_q$-rational points on X. We pick any integers m_1 and m_2 such that $2g - 2 < m_1 < m_2 < N$. Then we consider the codes $T_j := C_{\mathcal{L}}(D, m_j P_0)$ for $j = 1, 2$. Then $T_1 \subset T_2$ and T_j $(j = 1, 2)$ is an $[N, m_j - g + 1, \geq N - m_j]$ code over $\mathbb{F}_q$ and its dual $T_j^{\perp}$ is an $[N, N - m_j + g - 1, \geq m_j - 2g + 2]$ code over $\mathbb{F}_q$.

From now on, we assume that the ground field is $\mathbb{F}_{q^2}$, where $q = p^t$ with p a prime. We want to obtain linear codes C_j over $\mathbb{F}_p$ from T_j over $\mathbb{F}_{q^2}$ for $j = 1, 2$ via concatenation defined as follows. Consider an $\mathbb{F}_p$-linear map $\sigma : \mathbb{F}_{q^2} \to \mathbb{F}_p^{2t+r}$ such that the image of σ is a $[2t + r, 2t, r + 1]$ Reed-Solomon code over $\mathbb{F}_p$ for some nonnegative integer r. If p is 2, we can choose $t \geq 1$ and $r = 1$. If p is odd, we choose t and r such that $2t + r \leq p + 1$ or $0 \leq r \leq p - 2t + 1$ due to the fact that Reed-Solomon codes over $\mathbb{F}_p$ exist only for lengths at most $p + 1$. We map T_j via σ componentwisely to get $C_j := \sigma(T_j)$. Then C_j $(j = 1, 2)$ is an $\mathbb{F}_p$-linear $[(2t + r)N,\ 2t(m_j - g + 1),\ \geq (r + 1)(N - m_j)]$ code. Further it can be shown [8] that for any vector $\mathbf{x} \in C_1^{\perp} \backslash C_2^{\perp}$, we have the weight of $\mathbf{x}$ is $\geq m_1 - 2g + 2$.

Hence using the CSS construction (Corollary 12.15), we obtain a quantum $[[n, k, d]]_p$ code $B = B(X)$ with parameters $n = (2t + r)N,\ k = 2t(m_2 - m_1),\ d \geq \min\{(r + 1)(N - m_2),\ m_1 - 2g + 2\}$. Furthermore, by letting $l = m_2 - m_1$, one can show that for any integers l and r with $0 < l \leq N - 2g$ and $0 \leq r \leq p + 1 - 2t$, there is a quantum $[[n, k, d]]_p$ code $B = B(X)$ with parameters $n = (2t+r)N,\ k = 2tl,\ d \geq \frac{r+1}{r+2}(N-2g-l+1)$.

Let $\mathbf{X} = \{X\}$ be a Garcia-Stichtenoth tower of polynomially constructible curves over $\mathbb{F}_{q^2}$ where $q = p^t$ with increasing genus $g = g(X)$ [10]. We know that $\mathbf{X}$ attains the Drinfeld-Vlăduţ bound, i.e., $\limsup_{X \in \mathbf{X}} \frac{\#X(\mathbb{F}_{q^2})}{g} = q - 1$. Then for any sequence of integers $\{l = l(X) \mid X \in \mathbf{X}\}$ with $0 < l \leq N-2g$ for each X, we have $0 < \limsup_{x \in \mathbf{X}} \frac{l}{N} \leq 1 - \frac{2}{q-1}$. As in [8], for a fixed $\lambda \in (0, 1 - \frac{2}{q-1})$, we let $\lambda := \limsup_{x \in \mathbf{X}} \frac{l}{N}$. Then

$$R := \limsup_{x \in \mathbf{X}} \frac{2tl}{(2t + r)N} = \frac{2t}{2t + r}\lambda,$$

and

$$\delta := \limsup_{x \in \mathbf{X}} \frac{\frac{r+1}{r+2}(N - 2g - l + 1)}{(2t + r)N)} = \frac{r + 1}{(r + 2)(2t + r)} \left(1 - \frac{2}{q - 1} - \lambda\right).$$

Solving for λ in terms of δ, we get the following.

$$R_p(\delta) := R = \frac{2t}{2t + r} \left(1 - \frac{2}{q - 1}\right) - \frac{2t(r + 2)}{r + 1}\delta.$$

Using $\delta(p, r, t)$ defined in Theorem 12.29, we finally get

$$R_p(\delta) = \frac{2t(r + 2)}{r + 1}(\delta(p, r, t) - \delta).$$

□

Another approach to finding asymptotically good quantum codes uses the construction of Subsection 12.4.3 and the tower of function fields in [37, Theorem 1.7]. We refer the reader to [26] for these results.

12.5. Bibliographical notes

The literature on quantum error-correcting codes is massive. The first paper on quantum error-correcting codes is by Shor (Scheme for reducing decoherence in quantum memory Phys. Rev. A 52 (1995)). Calderbank, Rains, Shor, and Sloane (Quantum error correction via codes over $GF(4)$, IEEE Trans. Inform. Theory, vol. 44, (1998)) described the correspondence between binary additive quantum codes and additive self-orthogonal codes over $\mathbb{F}_4$. Nielsen and Chuang (Quantum Computation and Quantum Information, Cambridge University Press, Cambridge, 2000) is a widely used textbook in both quantum computation and quantum information theory.

Motivated by the fact that there exist good families of algebraic geometry codes meeting the Tsfasman-Vladut-Zink bound, which is better than the Gilbert-Varshamov bound, Ashikhmin, Litsyn, and Tsfasman (Asymptotically good quantum codes, Phys. Rev. A 63 (2001)) showed that asymptotically good binary quantum codes can be obtained from algebraic geometry codes in a polynomial construction. Some improvements in this direction have been made by Chen (Some good quantum error-correcting codes from algebraic-geometric codes, IEEE Trans. Inform. Theory, vol. 47, 2001), Chen, Ling, and Xing (Asymptotically good quantum codes exceeding the Ashikhmin-Litsyn-Tsfasman bound, IEEE Trans. Inform. Theory, vol. 47, 2001), Kim and Walker (Nonbinary quantum error-correcting codes from algebraic curves, Discrete Math. (2007)), Sarvepalli, Klappenecker (Nonbinary quantum codes from Hermitian curves, Applied algebra, algebraic algorithms and error-correcting codes, 136–143, Lecture Notes in Comput. Sci., 3857, Springer, Berlin, 2006), Niehage (Nonbinary quantum Goppa codes exceeding the quantum Gilbert-Varshamov bound, Quantum Inf. Process. 6 (2007)), and others.

References

[1] S. A. Aly, A note on the quantum Hamming bound, arXiv:0711.4603v1 [quant-ph].

[2] A. Ashikhmin and E. Knill, Nonbinary quantum stabilizer codes. IEEE Trans. Inform. Theory 47 (2001), no. 7, 3065–3072.

[3] A. Ashikhmin, S. Litsyn, and M. A. Tsfasman, Asymptotically good quantum codes, Phys. Rev. A 63 (2001), 032311.

[4] C. Bennett, D. DiVincenzo, J. Smolin, and W. Wootters, Mixed state entanglement and quantum error correction, Phys. Rev. A 54 (1996), 3824.

[5] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, Quantum error correction via codes over $GF(4)$, IEEE Trans. Inform. Theory, vol. 44, pp. 1369–1387, (1998).

[6] A. R. Calderbank and P. W. Shor, Good quantum error-correcting codes exist, Phys. Rev. A 54 (1996), 1098–1105.

[7] H. Chen, Some good quantum error-correcting codes from algebraic-geometric codes, IEEE Trans. Inform. Theory, vol. 47, pp. 2059–2061, (2001).

[8] H. Chen, S. Ling, and C. Xing, Asymptotically good quantum codes exceeding the Ashikhmin-Litsyn-Tsfasman bound, IEEE Trans. Inform. Theory, vol. 47, pp. 2055–2058, (2001).

[9] K. Feng and Z. Ma, A finite Gilbert-Varshamov bound for pure stabilizer quantum codes, IEEE Trans. Inform. Theory, vol. 50 (2004), no. 12, 3323–3325.

[10] A. Garcia and H. Stichtenoth, A tower of Artin-Schreier extensions of function fields attaining the Drinfeld-Vlăduţ bound, Invent. Math. 121 (1995), no. 1, 211–222.

[11] V. D. Goppa, *Algebraico-geometric codes*, Math. USSR-Izv. **21** (1983), 75–91.

[12] V. D. Goppa, *Geometry and Codes*, Kluwer, 1988.

[13] D. Gottesman, Stabilizer Codes and Quantum Error Correction, Ph.D. dissertation, California Inst. of Technol., Pasadena, CA, 1997.

[14] M. Grassl and T. Beth, Cyclic quantum error-correcting codes and quantum shift registers, R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. 456 (2000), no. 2003, 2689–2706.

[15] M. Grassl and T. Beth, Relations between classical and quantum error-correcting codes, in Proceedings Workshop "Physik und Informatik", DPG-Frhjahrstagung, Heidelberg, Mrz 1999, 45–57.

[16] M. Grassl, T. Beth, and M. Rötteler, On optimal quantum codes, Intl. J. Quantum Information 2 (2004) 55-64.

[17] M. Grassl, W. Geiselmann, and Th. Beth, Quantum Reed-Solomon codes, Applied algebra, algebraic algorithms and error-correcting codes (Honolulu, HI, 1999), 231–244, Lecture Notes in Comput. Sci., 1719, Springer, Berlin, 1999.

[18] M. Homma and S. J. Kim, The complete determination of the minimum distance of two-point codes on a Hermitian curve, Des. Codes Cryptogr. 40

(2006), no. 1, 5–24.
[19] M. Homma and S. J. Kim, Toward the determination of the minimum distance of two-point codes on a Hermitian curve, Des. Codes Cryptogr. 37 (2005), no. 1, 111–132.
[20] M. Homma and S. J. Kim, The two-point codes on a Hermitian curve with the designed minimum distance, Des. Codes Cryptogr. 38 (2006), no. 1, 55–81.
[21] M. Homma and S. J. Kim, The two-point codes with the designed distance on a Hermitian curve in even characteristic, Des. Codes Cryptogr. 39 (2006), no. 3, 375–386.
[22] A. Ketkar, A. Klappenecker, S. Kumar, P. K. Sarvepalli, Nonbinary Stabilizer Codes over Finite Fields IEEE Transactions on Information Theory, Volume 52, Issue 11, pages 4892 - 4914, (2006).
[23] J.-L. Kim and J. L. Walker, Nonbinary quantum error-correcting codes from algebraic curves, Discrete Math. (2007), doi:10.1016/j.disc.2007.08.038.
[24] E. Knill and R. Laflamme, A theory of quantum error-correcting codes, Phys. Rev. A, vol. 55, no. 2, pp. 900–911, (1997).
[25] R. Matsumoto, Improvement of Ashikhmin-Litsyn-Tsfasman bound for quantum codes, IEEE Trans. Inform. Theory 48 (2002), no. 7, 2122–2124.
[26] A. Niehage, Nonbinary quantum Goppa codes exceeding the quantum Gilbert-Varshamov bound, Quantum Inf. Process. 6 (2007), no. 3, 143–158.
[27] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information, Cambridge University Press, Cambridge, 2000.
[28] E. M. Rains, Nonbinary quantum codes, IEEE Trans. Inform. Theory 45 (1999), 1827–1832.
[29] P. K. Sarvepalli and A. Klappenecker, Nonbinary quantum codes from Hermitian curves, Applied algebra, algebraic algorithms and error-correcting codes, 136–143, Lecture Notes in Comput. Sci., 3857, Springer, Berlin, 2006.
[30] G. Seroussi and A. Lempel, Factorization of symmetric matrices and trace-orthogonal bases in finite fields, SIAM J. Comput. 9 (1980), no. 4, 758–767.
[31] P. W. Shor, Scheme for reducing decoherence in quantum memory Phys. Rev. A 52 (1995), 2493.
[32] P. Shor and R. Laflamme, Quantum analog of the MacWilliams identities for classical coding theory, Phys. Rev. Lett 78 (1997), 1600-1602.
[33] A. M. Steane, Enlargement of Calderbank-Shor-Steane quantum codes. IEEE Trans. Inform. Theory 45 (1999), no. 7, 2492–2495.
[34] A. M. Steane, Multiple-particle interference and quantum error correction. Proc. Roy. Soc. London Ser. A 452 (1996), no. 1954, 2551–2577.
[35] H. Stichtenoth, Algebraic Function Fields and Codes, Springer-Verlag, 1993.
[36] H. Stichtenoth, Self-dual Goppa codes, J. Pure Appl. Algebra 55 (1988), no. 1-2, 199–211.
[37] H. Stichtenoth, Transitive and self-dual codes attaining the Tsfasman-Vlăduţ-Zink bound. IEEE Trans. Inform. Theory 52 (2006), no. 5, 2218–2224.
[38] H. J. Tiersma, Remarks on codes from Hermitian curves, IEEE Trans. Inform. Theory 33 (1987), no. 4, 605–609.

[39] M. A. Tsfasman and S. G. Vlădut, Geometric approach to higher weights, IEEE Trans. Inform. Theory **41** (1995), no. 6, 1564–1588.
[40] M. A. Tsfasman, S. G. Vlădut, and T. Zink, Modular curves, Shimura curves, and Goppa codes better than the Varshamov-Gilbert bound, Math. Nachrichtentech., **109** (1982), 21–28.
[41] K. Yang and P. V. Kumar, On the true minimum distance of Hermitian codes, Coding Theory and Algebraic Geometry, Proceedings, Luminy, 1991, Lecture Notes in Mathematics **1518**, Springer-Verlag, 1992, 99–107.