# The Covering Radius of Long Primitive Ternary BCH Codes

by

Ralf Franken

A thesis submitted to
the Faculty of Information and Mathematical Sciences
at the University of Glasgow
for the degree of Doctor of Philosophy

March 2005

# Statement

This thesis is submitted in accordance with the regulations for the degree of Doctor of Philosophy in the University of Glasgow. It is the record of research carried out in the Department of Mathematics between June 2001 and December 2004.

Chapter 1 is a summary of widely known facts. Unless otherwise stated, all parts of Chapters 2–6 are original work by the author carried out under the guidance of his supervisor Prof. S. D. Cohen.

No part of this thesis has been previously submitted for a degree at the University of Glasgow or any other university.

# Acknowledgements

I would like to express my warm gratitude to my supervisor, Prof. Stephen D. Cohen, for his constant stimulation and encouragement. He was always ready to offer his advice, I learned a lot from him, and it was a pleasure working with him.

Special thanks also to Steve Pride, Lynne Alexander and Caroline Finlayson.

My research course was funded jointly by the EPSRC (fees) and the Faculty of Information and Mathematical Sciences, University of Glasgow (living allowance). I am grateful to both institutions for their financial support, which enabled me to do this course.

Finally, my stay in Glasgow was not only a valuable academic experience, but I also had a really great time! I thank my fellow PhD students, all scientific and non-scientific staff in the Maths Department, and all the friends I made during activities in- and outside the University, for being part of it.

Vaterstetten/Glasgow, March 2005

RALF FRANKEN

Dedicated to my father,

without whom it would have been impossible.

# Summary

This thesis is about the generalisation of a method to determine an asymptotic upper bound for the covering radius of primitive BCH codes. The method was introduced by S. D. Cohen in the mid-1990s for binary codes. It reduces the coding-theoretical problem to the complete splitting of a single polynomial $F(x)$ over a finite field, which is then established using results that have their roots in ramification theory of function fields.

The opening chapter introduces the covering radius problem for BCH codes along with its full coding-theoretical background and some history.

As a first result, the transformation from the covering radius problem to a polynomial splitting problem is extended to primitive $p$-ary BCH codes, where $p$ is an arbitrary prime. The process, during which an explicit "ready-to-use" form of the general $F$ is derived, is summarised in one theorem (Theorem 6).

The foundations for arranging the splitting of $F$ (via certain adjustable coefficients) were laid in previous work by Cohen, which is presented in extracts. By combining the key strategy of this with new ideas to meet the special requirements of the non-binary case, sufficient criteria for the splitting are obtained; these come in the form of conditions on polynomials $f_0$ and $f_1$, where $F$ has been parameterised as $f_0 + uf_1$ ($u$ an indeterminate). Several other lemmas are proved to deal with the establishing of the conditions. All these results are valid for arbitrary primes $p \geq 3$, so that with this the desired general version of the method has been made available.

The second half of the thesis is an in-depth study of the application of the method to ternary codes whose designed distance is of the form $\delta = 3t + 2$ ($t \in \mathbb{N}$). It is shown that for all $t \equiv 0 \pmod 4$ the covering radius $\rho$ takes its minimal value $\delta - 1$ whenever the length of the code exceeds $[(\delta - 1)!\,(\delta - 3)]^2$; for $t \equiv 2 \pmod 4$ the same holds when the length $3^m - 1$ of the code has even $m$. The case $t = 1$ is used as an example to illustrate how the method can fail to yield a result.

Subsequently, an improvement of the method for $p \geq 3$ is achieved by exploiting a certain factor $\varepsilon$ (not visible for $p = 2$ because in this case it is always 1). Coming back to the ternary case with $\delta = 3t + 2$, the example $t = 1$ becomes instructive in two ways: for odd $m$, the refined version of the method finally allows to prove $\rho = \delta - 1$ (showing that the improvement really is one), while for even $m$ this turns out to be the only case in the text where $\rho$ does demonstrably *not* attain its lower bound $\delta - 1$, but $\delta$ instead. Further modifications of the method are explored, and finally evidence is gathered that a result similar to that for $t = 1$ may hold for all odd $t$.

The final chapter discusses briefly possible directions for a continuation of the research.

# Contents

**Statement** ii

**Acknowledgements** iii

**Summary** v

**Preface** viii

**1 The Covering Radius Problem for BCH Codes** 1
- 1.1 Some basic coding theory . . . 1
- 1.2 BCH codes . . . 3
- 1.3 The covering radius problem for BCH codes . . . 6
- 1.4 A lower bound for $\rho(C_{P,n,\delta})$ . . . 7

**2 From the Covering Radius of a BCH Code to the Splitting of a Polynomial** 9
- 2.1 Starting point: Helleseth systems . . . 9
- 2.2 Linear transformation of variables . . . 11
- 2.3 Conversion into the problem of the splitting of a polynomial over $\mathbb{F}_q$ . . . 12
- 2.4 Expressing the $s_k$ $(p \nmid k)$ in terms of the $s_k$ $(p|k)$ . . . 15
- 2.5 Re-structuring of $F(x)$ and summary . . . 17
- 2.6 An example . . . 19

**3 How to Achieve the Splitting** 22
- 3.1 The splitting criterion . . . 22
- 3.2 Practical conditions . . . 23
- 3.3 Co-primality is never a problem . . . 27
- 3.4 About decomposition of rational functions . . . 28
- 3.5 About the discriminant $\Delta_x$ . . . 32
- 3.6 Another preliminary lemma . . . 35

**4 The Primitive Ternary Case with Designed Distance $\equiv 2 \pmod 3$** 36
- 4.1 The central result and set-up of its proof . . . 36
- 4.2 The "Standard Case" . . . 37
  - Co-primality . . . 38
  - Indecomposability . . . 38
  - Simplicity—Step 1 . . . 40
  - Simplicity—Step 2 . . . 41

Simplicity—Step 3: reduction . . . . . 42
Simplicity—Step 3: solving the decomposition problem . . . . . 43
4.3 The "Exceptional Case" . . . . . 48
Co-primality . . . . . 48
Indecomposability . . . . . 48
Condition (iii) . . . . . 54
Discriminants . . . . . 55
Discriminants—the remaining open cases . . . . . 56
4.4 The "Degenerate Case" . . . . . 61
4.5 Failure of the approach for $t = 1$ . . . . . 64

**5 Refinement of the Approach** **65**
5.1 Why there should be room for improvement . . . . . 65
5.2 A refined approach . . . . . 67
5.3 The case $t = 1$ ($\delta = 3t + 2$) re-visited . . . . . 68
5.4 No new results for even $t$ . . . . . 74
5.5 Working with one $\varepsilon_i$ different from the rest . . . . . 75
5.6 Economical change of $\varepsilon$ . . . . . 78
5.7 Notes on the case of odd $t$ greater than 1 . . . . . 79

**6 And next?** **85**
6.1 Open questions for $p = 3$ and $\delta \equiv 2 \pmod 3$ . . . . . 85
6.2 The ternary case with $\delta \equiv 0 \pmod 3$ . . . . . 86
6.3 Primes $p$ greater than 3 . . . . . 87
6.4 Non-primitive codes . . . . . 89

**Bibliography** **90**

# Preface

In mathematics, generalisations come with varying degrees of excitement. Where some computation is more or less done again with different numbers, this is usually considered a technical detail and not a subject of an academic discussion. Then there are those generalisations which are essentially straightforward, but due to a substantial growth in complexity require great discipline in carrying out, and when completed help to see the original matter clearer. Certainly the most interesting ones are those where unexpected things happen, completely new aspects come to light, and nobody knows where the way really goes until the end is reached.

In this PhD thesis we generalise a method of solving a coding-theoretical problem. "Coding theory" is the technical term for the mathematical theory of error-correcting codes, a discipline with both an applied and a pure side to it. We will not go too deep into this theory here; all that is needed to understand the origin of the problem will be explained in the first chapter. (Indeed, once the problem has been transformed into the context of polynomials over finite fields, coding theory will play a minor role in this study. Even without mentioning codes at all, the remaining part of the problem—finding solutions to certain equation systems by making a suitable polynomial split into linear factors over a finite field—would make an interesting topic in its own right.)

The method at the centre of our attention is due to S. D. Cohen and was published in 1997. It brings together a classic coding-theoretical observation by T. Helleseth, auxiliary results from ramification theory of function fields (an application of Galois theory), and a certain amount of hands-on manipulation of polynomials. The method was originally developed for binary primitive BCH codes and will be generalised here to such codes over arbitrary prime alphabets ("$p$-ary" codes). This happens in Chapters 2 and 3, and for a first orientation of the reader a quick overview is given now. The skeleton of the method is as follows.

First Part:
- covering radius problem for BCH code
- $\Updownarrow$
- solvability problem for Helleseth systems
- $\Downarrow$
- splitting problem for polynomial $F$

Second Part:
- uses results from Galois theory to show that certain coefficients of $F$ can be chosen such that $F$ splits completely over the relevant finite field

First, following Helleseth's idea, equivalence is shown between the coding-theoretical

problem and the solvability of certain equation systems (Section 2.1). Then, following Cohen's idea, a polynomial $F$ is determined whose splitting over the finite field of order $q$, where $q-1$ is the length of the code in question, implies a solution to the system (Sections 2.3–2.5). Between the two steps there can be a normalisation procedure (Section 2.2). Together this may be called the First Part of the method, and extending it from 2 to $p$ belongs, in the opinion of the author, to the second type of generalisation mentioned at the beginning.

Proving the splitting of $F$ is a different thing—the Second Part. To this end, we take up and develop further (in Sections 3.1 and 3.2) the relevant theory of Cohen, which derives conditions for the splitting from a Galois-theoretic criterion. The conditions concern elementary properties (functional decomposability, multiplicities of factors, etc.) of polynomials $f_0$ and $f_1$, after $F$ has been brought into the form $f_0 + uf_1$ (with $u$ incorporating adjustable coefficients of $F$). The remaining sections of Chapter 3 are devoted to proving auxiliary results for the practical application.

This gives the general method, but before results are obtained, the hardest part remains: $F$ takes a different form depending on the alphabet size $p$, the designed distance $\delta$ and the assumed upper bound $r$ for the covering radius of the code in question, and the conditions have to be checked individually in each case.

Chapters 4 and 5 test the power of the method in one particular case ($p = 3$, $\delta \equiv 2 \pmod 3$, $r = \delta - 1$). Results are found, but also obstacles. The latter lead to various attempts to improve the method. Not all of these are successful; nevertheless, some of the "dead-end streets" have been included as genuine parts of the knowledge gained about the limits of the method (and so that no one else has to try them again). In all approaches it was a principle to be as general as possible.

In the categories of the first paragraph, Chapters 3–5 belong without doubt to the most interesting type of generalisation. The most prominent features of which the binary case gives no indication at all are the occurrence of a certain "Situation A" (defined in Lemma 8) and the consequences of the *real* choice of $\varepsilon$ in (2.4), the latter being more or less the foundation of the whole of Chapter 5.

Since successful application of the method requires a considerable amount of work and original ideas, it turned out to be beyond the scope of this thesis to carry it out for more than one case. Some remarks about the various others which wait to be attacked are collected in the final chapter, and it is hoped that research on this will carry on beyond the present project.

The thesis is written on a level that should be accessible for a beginning postgraduate student of any mathematical direction. All prerequisites from coding theory are provided in the first chapter, but some familiarity with finite fields will be assumed (roughly to the extent of the first two chapters of [LiNi]). For a full understanding of parts of Chapter 3, some knowledge of algebraic function fields is necessary. Appropriate references to textbooks are made.

We list some symbols which may be uncommon or give rise to uncertainties:

| | |
|---|---|
| $\mathbb{N}$ | $\{1, 2, 3, \ldots\}$ |
| $\mathbb{N}_0$ | $\mathbb{N} \cup \{0\}$ |
| $\lvert A \rvert$ | cardinality of the set $A$ |
| $A \backslash B$ | set of elements in $A$ which are not in $B$ |
| $A \subset B$ | $A \subseteq B \ \wedge \ A \neq B$ |
| $A^\top$ | transpose of the matrix $A$ |
| $\langle a_1, a_2, \ldots \rangle$ | ideal generated by ring elements $a_1, a_2, \ldots$ |
| $\operatorname{ord}_b a$ | order of $a$ modulo $b$ $(a, b \in \mathbb{N})$ |
| $\mathbb{A}^*$ | multiplicative group of the field $\mathbb{A}$ |

All other notation is either completely standard or explained in the main text when used for the first time. We write LHS and RHS for the left-hand side and right-hand side of an equation. Polynomials are stated without their arguments where possible; however, we will avoid mixed expressions, i.e. write $xf(x)$, not $xf$—this is to avoid confusion when we have to deal with expressions like $xf(x^p)$. The statement $f = 0$ means that $f$ is the zero polynomial, as opposed to the algebraic equation $f(x) = 0$. The term "splitting" will always mean complete splitting into linear factors.

Finally, the author thinks that a PhD thesis is a good place to put down certain details for which in other publications there is not enough space. In this spirit, elaborate proofs have been included on two occasions for results stated in the literature without further justification. Readers who are only interested in the main thread will no doubt wish to skip these tedious technical passages, and for this purpose they have been marked with

on the margin. (Apologies to the examiners, who will have to read everything.)

# Chapter 1

# The Covering Radius Problem for BCH Codes

Our investigation starts with a problem from coding theory. This is introduced here, along with all necessary coding-theoretical preliminaries for this thesis.

## 1.1 Some basic coding theory

To begin with, we list the basic notions and facts from coding theory relevant to our problem. The reader who wishes a more comprehensive introduction to algebraic coding theory and the ideas behind it is referred to the literature, e.g. Chapters 2 and 3 of [vLi].

Let $P$ be a prime power and $\mathbb{F}_P$ the finite field with $P$ elements.

A ***P*-ary linear $(n, k)$ code** $\mathcal{C}$ is a $k$-dimensional subspace of the $\mathbb{F}_P$-vector space $\mathbb{F}_P^n$. The integer $n$ is called the **length** of $\mathcal{C}$. The elements of $\mathbb{F}_P^n$ are called **words**, those in $\mathcal{C}$ **codewords**.

The **(Hamming) weight** of a word $a = (a_0, \ldots, a_{n-1})$ is defined by

$$w(a) \;:=\; \big|\{i \in \{0, \ldots, n-1\} :\; a_i \neq 0\}\big|$$

and gives rise to the **(Hamming) distance**

$$d(a, b) \;:=\; w(a - b) \qquad (a, b \in \mathbb{F}_P^n)\,,$$

which is easily seen to be a metric on $\mathbb{F}_P^n$.

Two important parameters of every code $\mathcal{C} \neq \{0\}$ are its **minimal distance** $d(\mathcal{C})$ and its **covering radius** $\rho(\mathcal{C})$. The first is defined as

$$d(\mathcal{C}) \;:=\; \min_{\substack{a,b \in \mathcal{C} \\ a \neq b}} d(a, b)\,;$$

for the second let

$$B_r(a) := \{b \in \mathbb{F}_P^n : d(a,b) \leq r\} \qquad (r \in \mathbb{N}_0),$$

then

$$\rho(\mathcal{C}) := \min\{r \in \mathbb{N} : \bigcup_{c \in \mathcal{C}} B_r(c) = \mathbb{F}_P^n\},$$

i.e. the covering radius of $\mathcal{C}$ is the smallest radius for which the balls around the codewords cover the whole space $\mathbb{F}_P^n$.

One way of describing a linear $(n,k)$ code $\mathcal{C}$ is to give a matrix $H$ with $n$ columns and entries in $\mathbb{F}_P$ such that

$$\forall\, a \in \mathbb{F}_P^n : \quad a \in \mathcal{C} \iff Ha^\top = 0^\top.$$

Let $l$ be the number of rows of $H$. If $l = n - k$, then $H$ is called a **parity check matrix** for $\mathcal{C}$, and $Ha^\top$ is called the **syndrome** of $a$ (with respect to $H$). The map

$$\begin{array}{rccc} \varphi_H : & \mathbb{F}_P^n/\mathcal{C} & \longrightarrow & \mathbb{F}_P^l \\ & a + \mathcal{C} & \longmapsto & Ha^\top \end{array}$$

is then an isomorphism of $\mathbb{F}_P$-vector spaces.

A first step towards proving results about the covering radius of linear codes is the following little lemma.

**Lemma 1** *Let $\mathcal{C}$ be a linear $(n,k)$ code with parity check matrix $H$.*

(i) *A word $a \in \mathbb{F}_P^n$ has distance $r$ from a codeword if and only if its syndrome $Ha^\top$ can be written as a linear combination of the columns of $H$ with exactly $r$ non-zero terms.*

(ii) *The covering radius of $\mathcal{C}$ is equal to the smallest integer $r$ such that every column vector $v^\top \in \mathbb{F}_P^l$ can be written as a linear combination of at most $r$ columns of $H$.*

*Proof.*

(i) Let $c \in \mathcal{C}$ with $d(a,c) = r$. Then $w(a-c) = r$ and $Ha^\top = Ha^\top - Hc^\top = H(a-c)^\top$ is a linear combination of exactly $r$ columns of $H$.
Conversely, take $a \in \mathbb{F}_P^n$ and consider its syndrome $Ha^\top \in \mathbb{F}_P^l$. By assumption, there is some $b \in \mathbb{F}_P^n$ with $w(b) = r$ and $Ha^\top = Hb^\top$. Then $H(a-b)^\top = Ha^\top - Hb^\top = 0$, i.e. $a - b \in \mathcal{C}$, and $a$ has distance $r$ from the codeword $a - b$.

(ii) By surjectivity of $\varphi_H$, every $v^\top \in \mathbb{F}_P^l$ is the syndrome of some $a \in \mathbb{F}_P^n$. Hence, by (i), a radius $r \in \mathbb{N}$ covers the space $\mathbb{F}_P^n$ iff every $v^\top \in \mathbb{F}_P^l$ can be written as a linear combination of at most $r$ columns of $H$. Taking the minimal such $r$ yields the result. □

Observe that the statement of Lemma 1(ii) makes no sense if $H$ is a matrix with more than $n - k$ rows: then the map $\varphi_H$ is no longer surjective, and some elements of $\mathbb{F}_P^l$ cannot be represented as linear combinations of columns of $H$ at all. (Hence our insistence in the definition that a parity check matrix have exactly $n - k$ rows.)

## 1.2 BCH codes

Some codes have a very useful property: a code $\mathcal{C}$ is called **cyclic**, if it is closed under cyclic shifts, that is

$$\forall\ (c_0, \ldots, c_{n-1}) \in \mathcal{C} :\ (c_{n-1}, c_0, \ldots, c_{n-2}) \in \mathcal{C}\,.$$

To study cyclic codes of length $n$, one works with advantage in the principal ideal domain $\mathbb{F}_P[x]/\langle x^n - 1\rangle =: \mathcal{R}_n$ by identifying a word $(a_0, \ldots, a_{n-1})$ with the polynomial $a_0 + a_1 x + a_2 x^2 + \ldots + a_{n-1} x^{n-1}$. Then the cyclic linear codes of length $n$ are exactly the ideals in $\mathcal{R}_n$, and the code generated by a polynomial $g(x) \in \mathcal{R}_n$ has dimension $n - \deg g$. (See for example [vLi] §6.1.)

A famous class of cyclic linear codes, of great importance for both theory and practice, are the so-called BCH codes. The "BCH" stands for the initials of their discoverers: R. C. Bose and D. K. Ray-Chaudhuri (1960), and independently A. Hocquenghem (1959).

Every (narrow sense) BCH code is determined by three parameters:

| | | |
|---|---|---|
| alphabet size | $P$ | a prime power, |
| length | $n$ | an integer $\geq 2$ co-prime to $P$, |
| designed distance | $\delta$ | $\in \{2, \ldots, n\}$. |

The ***P*-ary BCH code of length *n* with designed distance *δ***, which we will denote by $\mathcal{C}_{P,n,\delta}$, is defined as follows. Let $m := \operatorname{ord}_n P$, and let $\eta$ be a primitive $n$-th root of unity in $\mathbb{F}_{P^m}$ (the smallest extension of $\mathbb{F}_P$ that contains such a root). Denote by $m_i(x)$ the minimal polynomial of $\eta^i$ over $\mathbb{F}_P$. Then $\mathcal{C}_{P,n,\delta}$ is the ideal in $\mathcal{R}_n$ generated by the least common multiple of $m_1(x), \ldots, m_{\delta-1}(x)$. Equivalently,

$$\mathcal{C}_{P,n,\delta} \ :=\ \{c(x) \in \mathcal{R}_n :\ c(\eta^i) = 0 \ \text{ for } i = 1, \ldots, \delta - 1\}\,.$$

For a given BCH code $\mathcal{C}_{P,n,\delta}$ we fix the notation

$$m \ =\ \operatorname{ord}_n P\,, \qquad q \ :=\ P^m \qquad \text{and} \qquad N \ :=\ \frac{q-1}{n}\,.$$

The integer $N$ is called the **degree of primitivity** of $\mathcal{C}_{P,n,\delta}$; if $N = 1$, the code is called **primitive**. In the context of BCH codes, the letters $P$, $n$, $\delta$, $m$, $q$ and $N$ will be used with these meanings throughout this work without further explanation. (Later, $p$ will replace $P$ in the case of a prime number.)

To find all possible BCH codes with a fixed alphabet size $P$, one may also start from $M \in \mathbb{N}$. Then there exists a BCH code $\mathcal{C}_{P,n,\delta}$ as defined above (with $m = M$) for every divisor $n \neq 1$ of $P^M - 1$ with $\operatorname{ord}_n P = M$ (and every designed distance in the appropriate range). Note that in general we have only $\operatorname{ord}_n P \mid M$. But equality holds always trivially for $n := P^M - 1$, so that the primitive $P$-ary BCH code of length $P^M - 1$ exists for every $M \in \mathbb{N}$ (and every $\delta \in \{2, \ldots, P^M - 1\}$).

Since $f(x^P) = f^P(x)$ for every $f(x) \in \mathbb{F}_P[x]$, it is obvious that in the definition of $\mathcal{C}_{P,n,\delta}$ all $i \in \{1, \ldots, \delta - 1\}$ which are multiples of $P$ can be omitted. In particular, for $\delta \equiv 1 \pmod{P}$ we have $\mathcal{C}_{P,n,\delta} = \mathcal{C}_{P,n,\delta-1}$.

The next (rather technical) lemma shows that there is no further such redundancy if the code is long enough with respect to its designed distance. We will need this in Section 2.1 to find parity check matrices for BCH codes, so that we can apply Lemma 1 to such codes and establish Theorem 4. This will pave the way for our approach. With Lemma 2 and Theorem 4 we make precise and prove, for arbitrary characteristic, statements indicated by T. Helleseth in [Hel], pp. 158–163.

The proof of Lemma 2 uses the fact that the minimal polynomial of two elements of $\mathbb{F}_{P^m}$ over $\mathbb{F}_P$ are either equal—in which case the elements are called **conjugates**—or co-prime. The conjugates of $\alpha \in \mathbb{F}_P$ in this situation are given by $\{\alpha, \alpha^P, \alpha^{P^2}, \ldots, \alpha^{P^{m-1}}\}$, a set of $d$ distinct elements, each repeated $\frac{m}{d}$ times, where $d$ is the degree of their common minimal polynomial.

**Lemma 2 (Helleseth)** *Suppose the BCH code $\mathcal{C}_{P,n,\delta}$ satisfies $\delta - 1 < \sqrt{q}/N$. Then the code has generator polynomial*

$$g(x) = \prod_{\substack{i=1 \\ P \nmid i}}^{\delta-1} m_i(x) ,$$

*or equivalently (with $\eta$ a primitive $n$-th root of unity in $\mathbb{F}_q$)*

$$\mathcal{C}_{P,n,\delta} = \{c(x) \in \mathcal{R}_n : \quad c(\eta^i) = 0 \ \text{ for all } i = 1, \ldots, \delta - 1 \ \text{ with } \ P \nmid i \} ,$$

*and its dimension is $n - ms$, where $s := \delta - 1 - \left\lfloor \frac{\delta - 1}{P} \right\rfloor$.*

*Proof.* Write $S := \{i \in \{1, \ldots, \delta - 1\} : \ P \nmid i\}$. Then $s = |S|$.

We will show that the condition $\delta - 1 < \dfrac{P^{\frac{m}{2}}}{N}$ implies

$$\forall i, j \in S : \ \forall k \in \{1, \ldots, m-1\} : \ i \not\equiv P^k j \pmod{n} . \tag{1.1}$$

This has two consequences:

1. Suppose $m_i(x) = m_j(x)$ for $i, j \in S$, $i \neq j$.
   Then $\eta^i$ and $\eta^j$ are conjugates, i.e. $\eta^i \in \{\eta^{Pj}, \ldots, \eta^{P^{m-1}j}\}$, so that $i \equiv P^k j \pmod{n}$ for some $k \in \{1, \ldots, m-1\}$. Hence (1.1) implies that all minimal polynomials $m_i(x)$ with indices in $S$ are distinct.

2. Suppose $\deg m_i < m$ for some $i \in S$.
   Then some elements of $\{\eta^i, \eta^{Pi}, \ldots, \eta^{P^{m-1}i}\}$ coincide, i.e. there are $k_1, k_2 \in \{0, \ldots, m-1\}$ with $P^{k_1} i \equiv P^{k_2} i \pmod{n}$ and $k_1 \neq k_2$, or equivalently $k \in \{1, \ldots, m-1\}$ with $i \equiv P^k i \pmod{n}$. Hence (1.1) implies also that $\deg m_i = m$ for all $i \in S$.

Therefore it follows from (1.1) that $\mathcal{C}_{P,n,\delta}$ has generator polynomial

$$g(x) \;=\; \operatorname*{lcm}_{i=1,\ldots,\delta-1} m_i(x) \;=\; \prod_{i\in S} m_i(x)$$

and dimension $n - \deg g = n - ms$.

It remains to prove (1.1). Assume first that $1 \le k \le \frac{m}{2}$. Then

$$0 \;<\; \underbrace{|P^k j - i|}_{\not\equiv 0 \pmod P} \;\le\; (\delta-1)P^k - 1 \;<\; \frac{P^{\frac{m}{2}+k}}{N} - 1 \;\le\; \frac{P^m}{N} - 1 \;\le\; \frac{P^m - 1}{N} \;=\; n\,,$$

so $P^k j - i$ cannot be a multiple of $n$.

Now consider $\frac{m}{2} < k \le m-1$. Suppose $P^k j - i = zn$ for some $z \in \mathbb{Z}$. Then, because

$$P^k j - i \;\ge\; P^k - (\delta-1) \;>\; P^k - \frac{P^{\frac{m}{2}}}{N} \;\ge\; P^k - P^{\frac{m}{2}} \;>\; 0\,,$$

we must have $z \ge 1$. $\quad(*)$

On the other hand, because

$$P^k j - i \;\le\; P^k(\delta-1) - 1 \;<\; \frac{P^{\frac{m}{2}+k}}{N} - 1 \;\le\; \frac{P^{\frac{3m}{2}-1}}{N} - 1 \qquad (1.2)$$

and

$$(P^{\frac{m}{2}} + N)\cdot n \;=\; \frac{P^{\frac{3m}{2}}}{N} + P^m - \frac{P^{\frac{m}{2}}}{N} - 1 \;>\; \frac{P^{\frac{3m}{2}-1}}{N} - 1 \overset{(1.2)}{>} P^k j - i\,,$$

we must have $z < P^{\frac{m}{2}} + N$. $\quad(**)$

Now

$$P^k j - i = zn \quad\Rightarrow\quad NP^k j = z(P^m - 1) + Ni \quad\Rightarrow\quad Nj = zP^{m-k} + \frac{Ni - z}{P^k}\,,$$

hence $(Ni - z)/P^k$ must be an integer; however,

$$Ni - z \overset{(*)}{\le} N(\delta-1) - 1 \;<\; P^{\frac{m}{2}} - 1 \;<\; P^k$$

and

$$Ni - z \overset{(**)}{>} N - (P^{\frac{m}{2}} + N) \;>\; -P^k\,,$$

so this is only possible for $Ni - z = 0$. But then we have

$$\underbrace{Nj}_{\not\equiv 0 \pmod P} \;=\; \underbrace{zP^{m-k}}_{\equiv 0 \pmod P}\,,$$

a contradiction. So again $P^k j - i$ cannot be a multiple of $n$. □

We conclude this section about BCH codes with a well-known fact that justifies the term "designed distance". The proof is in most textbooks on coding theory, see for instance Theorem 6.6.2 in [vLi].

**Theorem 1**

*The true minimal distance of a BCH code $\mathcal{C}_{P,n,\delta}$ is at least its designed distance $\delta$.*

It is possible to have $d(\mathcal{C}_{P,n,\delta}) > \delta$ (a family of examples is mentioned in [Hel], proof of Lemma 3.4). However, this does not occur with codes which are sufficiently long compared to their designed distance (see Theorem 3.5 of [Hel] and the references there).

## 1.3 The covering radius problem for BCH codes

With Theorem 1 it is easy to construct, for example, codes with arbitrarily large minimal distance over a given alphabet. Obviously the BCH construction allows good control of the minimal distance. By contrast, it is not at all clear what the covering radius of such codes should be.

Indeed, determining the covering radius of a general BCH code is hard ([Hel], p. 158). In view of Lemma 2 and the remarks at the end of the previous section, we will not concern ourselves with numerical values for specific codes, but rather be interested in asymptotic bounds. More precisely, by the **covering radius problem for BCH codes** we mean:

> *Find expressions $B_*, B^*$ and conditions on $P$, $n$ and $\delta$ such that $B_* \leq \rho(\mathcal{C}_{P,n,\delta}) \leq B^*$ whenever the conditions are satisfied.*

The first general results of this kind were proved by T. Helleseth in his pioneering paper [Hel], 1985. He found for binary BCH codes (with $\delta$ odd):

$$\delta - 2 \leq \rho(\mathcal{C}_{2,n,\delta}) \leq \delta \quad \text{whenever} \quad 2^m \geq [(\delta-2)N - 1]^{2\delta}\,. \tag{1.3}$$

The true challenge here lies in the upper bound. The lower bound, as the next section will show, is easy to obtain and generalises also to non-binary codes as

$$\delta - 1 \leq \rho(\mathcal{C}_{P,n,\delta}) \quad \text{whenever} \quad \sqrt{P^m}/N > \delta - 1 \quad (\delta \not\equiv 1 \pmod P)^\dagger\,.$$

For binary BCH codes, the upper bound and the corresponding required size of $2^m$ in (1.3) were then gradually improved by various authors ([Shp], [Tie], [SkVl]), until in 1993 it was known that for primitive binary codes ($\delta$ odd)

$$\rho(\mathcal{C}_{2,2^m-1,\delta}) = \delta - 2 \quad \text{whenever} \quad 2^m > 10^4(\delta-2)d^5\mathfrak{p}^3(\lfloor 4\log d\rfloor)\,, \tag{1.4}$$

where $d$ denotes the product of all odd numbers from 1 to $\delta - 2$, and $\mathfrak{p}(k)$ the $k$-th prime ([MoMo]). Finally, in his paper [Coh97], S. D. Cohen proved again $\rho(\mathcal{C}_{2,2^m-1,\delta}) \leq \delta - 2$, using a new method which yielded a much lower and nicer bound for $2^m$ than that in (1.4), namely $2^m > [(\delta-2)!\,(\delta-4)]^2$. It is this method that we will generalise to the non-binary case in this thesis.

---

$^\dagger$All binary BCH codes can be (and usually are) viewed as having odd designed distance, whereas for $P > 2$ we will prefer to assume $\delta \not\equiv 1 \pmod P$. This explains the apparent discrepancy that in the binary case we have a lower bound of $\delta - 2$ while in all other cases it is $\delta - 1$. See Theorem 3 for a unified statement.

The premier source of results for non-binary BCH codes is the dissertation [Kai] of Y. Kaipainen, 1995. We quote his main results (Theorems 3.0.1, 4.0.4 and 5.3.8 in the original work).

**Theorem 2 (Kaipainen)**

(i) Primitive ternary codes.—*The covering radius of* $\mathcal{C}_{3,3^m-1,\delta}$ *is at most*

$$\begin{cases} \delta+1 & \text{if } \delta \equiv 0 \pmod 3, \\ \delta & \text{if } \delta \equiv 2 \pmod 3, \end{cases}$$

*whenever* $3^m > q_0$, *where* $q_0$ *is a constant depending only on* $\delta$.

(ii) Codes with large characteristic.—*Let* $p \geq 5$ *be a prime,* $P$ *a power of* $p$, $\delta \leq p$. *Then*

$$\rho(\mathcal{C}_{P,n,\delta}) \leq \delta \quad \text{whenever } P > P_0 \text{ and } P^m > q_0,$$

*where* $q_0$ *is a constant depending on* $\delta$ *and* $N$, *and* $P_0$ *is a constant depending only on* $\delta$.

(iii) Primitive codes with small characteristic and large alphabet.—*Let* $p < \delta$ *and* $P$ *be a power of* $p$. *Then*

$$\rho(\mathcal{C}_{P,P^m-1,\delta}) \leq \delta \quad \text{whenever } P > P_0 \text{ and } P^m > q_0,$$

*where* $P_0$ *and* $q_0$ *are constants depending only on* $\delta$.

In [Kai] it is also suggested explicitly (on p. 49) that a generalisation of Cohen's method may provide a concrete value for $q_0$ in Theorem 2(i).

## 1.4 A lower bound for $\rho(\mathcal{C}_{P,n,\delta})$

If $\mathcal{C}, \mathcal{C}' \subseteq \mathbb{F}_P^n$ are two codes with $\mathcal{C} \subseteq \mathcal{C}'$, then $\mathcal{C}$ is called a **subcode** of $\mathcal{C}'$ and $\mathcal{C}'$ a **supercode** of $\mathcal{C}$. For example, it is clear from their definition that the $P$-ary BCH codes of fixed length $n$ are **nested**, i.e.

$$\mathcal{C}_{P,n,\delta} \subseteq \mathcal{C}_{P,n,\delta'} \quad \text{whenever} \quad \delta \geq \delta'.$$

More precisely, from Lemma 2 we have (for sufficiently small $\delta$)

$$\ldots \subseteq \mathcal{C}_{P,n,\delta} \subseteq \ldots \subseteq \mathcal{C}_{P,n,Pt+2} \subset \mathcal{C}_{P,n,Pt+1} = \mathcal{C}_{P,n,Pt} \subset \mathcal{C}_{P,n,Pt-1} \subseteq \ldots \subseteq \mathcal{C}_{P,n,2} \subset \mathbb{F}_P^n;$$

in particular for binary and ternary codes:

$$\ldots \subseteq \mathcal{C}_{2,n,\delta} \subseteq \ldots \subset \mathcal{C}_{2,n,7} = \mathcal{C}_{2,n,6} \subset \mathcal{C}_{2,n,5} = \mathcal{C}_{2,n,4} \subset \mathcal{C}_{2,n,3} = \mathcal{C}_{2,n,2} \subset \mathbb{F}_2^n,$$

$$\ldots \subseteq \mathcal{C}_{3,n,\delta} \subseteq \ldots \subset \mathcal{C}_{3,n,7} = \mathcal{C}_{3,n,6} \subset \mathcal{C}_{3,n,5} \subset \mathcal{C}_{3,n,4} = \mathcal{C}_{3,n,3} \subset \mathcal{C}_{3,n,2} \subset \mathbb{F}_3^n.$$

An easy but highly useful observation, due to [GPZ], is the following (more about it can be found in Section I.H of [Surv]).

**Lemma 3 (Supercode Lemma)**

*Let $\mathcal{C}$ and $\mathcal{C}'$ be codes in $\mathbb{F}_P^n$ with $\mathcal{C} \subset \mathcal{C}'$. Then the covering radius of $\mathcal{C}$ is at least the minimal distance of $\mathcal{C}'$.*

*Proof.* There exists $c_0 \in \mathcal{C}'\backslash\mathcal{C}$. This has distance at least $d(\mathcal{C}')$ from all other codewords in $\mathcal{C}'$, in particular all codewords in $\mathcal{C}$. Hence for $r < d(\mathcal{C}')$ we have

$$c_0 \notin \bigcup_{c\in\mathcal{C}} B_r(c)\,,$$

therefore $\rho(\mathcal{C}) \geq d(\mathcal{C}')$. □

This allows to derive quickly the lower bound for the covering radius of long BCH codes mentioned in the previous section.

**Theorem 3 (Helleseth)** *Suppose the BCH code $\mathcal{C}_{P,n,\delta}$ satisfies $\delta - 1 < \sqrt{q}/N$. Then its covering radius is at least*

$$\begin{cases} \delta - 1 & \text{if } \delta \not\equiv 1 \pmod{P}\,, \\ \delta - 2 & \text{if } \delta \equiv 1 \pmod{P}\,. \end{cases}$$

*Proof.* Assume first that $\delta \not\equiv 1 \pmod{P}$. Then we deduce from Lemma 2 that $\mathcal{C}_{P,n,\delta-1}$ is a proper supercode of $\mathcal{C}_{P,n,\delta}$. Hence $\rho(\mathcal{C}_{P,n,\delta}) \geq d(\mathcal{C}_{q,n,\delta-1}) \geq \delta - 1$ by the Supercode Lemma and Theorem 1.

In the case $\delta \equiv 1 \pmod{P}$ we have $\mathcal{C}_{P,n,\delta-1} = \mathcal{C}_{P,n,\delta}$, but $\mathcal{C}_{P,n,\delta-2}$ is a proper supercode of $\mathcal{C}_{P,n,\delta}$. The rest of the argument is as above. □

Having settled this, we can now turn to the upper bound and the generalisation of the method from [Coh97].

# Chapter 2

# From the Covering Radius of a BCH Code to the Splitting of a Polynomial

## 2.1 Starting point: Helleseth systems

As announced in the remarks leading up to Lemma 2, we complete in this section the spelling out of details of the idea by Helleseth which will form the basis for our attack (and did so also for [Kai]). It says that the covering radius problem for a long BCH code is equivalent to the question whether a certain type of system of non-linear equations over a finite field has a solution.

**Theorem 4 (Helleseth)** *Let $C_{P,n,\delta}$ be a BCH code with $\delta - 1 < \sqrt{q}/N$. Then its covering radius is equal to the smallest integer $r$ such that for any choice of*

$$a_k \in \mathbb{F}_q \qquad (k = 1, \ldots, \delta - 1; \;\; P \nmid k)$$

*the system*

$$\begin{array}{ccccccccc}
\varepsilon_1 x_1^N & + & \varepsilon_2 x_2^N & + & \ldots & + & \varepsilon_r x_r^N & = & a_1 \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
\varepsilon_1 x_1^{kN} & + & \varepsilon_2 x_2^{kN} & + & \ldots & + & \varepsilon_r x_r^{kN} & = & a_k \quad (P \nmid k) \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
\varepsilon_1 x_1^{(\delta-1)N} & + & \varepsilon_2 x_2^{(\delta-1)N} & + & \ldots & + & \varepsilon_r x_r^{(\delta-1)N} & = & a_{\delta-1}
\end{array} \tag{2.1}$$

*has a solution* $\varepsilon_1, \ldots, \varepsilon_r \in \mathbb{F}_P$, $x_1, \ldots, x_r \in \mathbb{F}_q$.

*Proof.* Let $S$ and $s$ be as in Section 1.2, i.e.

$$S = \{i \in \{1, \ldots, \delta - 1\} : \; P \nmid i\}, \qquad s = |S|,$$

and let $\zeta$ be a primitive element of $\mathbb{F}_q = \mathbb{F}_{P^m}$. Then $\eta := \zeta^N$ is a primitive $n$-th root of unity in $\mathbb{F}_q$, and by Lemma 2

$$C_{P,n,\delta} \;=\; \{c(x) \in \mathcal{R}_n : \; c(\eta^i) = 0 \;\text{ for all } i \in S\}.$$

Therefore the matrix

$$\tilde{H} := \begin{pmatrix} 1 & \eta & \eta^2 & \eta^3 & \cdots & \eta^{n-1} & \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \\ 1 & \eta^i & \eta^{2i} & \eta^{3i} & \cdots & \eta^{(n-1)i} & (i \in S) \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \\ 1 & \eta^{\delta-1} & \eta^{2(\delta-1)} & \eta^{3(\delta-1)} & \cdots & \eta^{(n-1)(\delta-1)} & \end{pmatrix}$$

clearly satisfies

$$\forall\, a \in \mathbb{F}_P^n : \quad a \in \mathcal{C}_{P,n,\delta} \iff \tilde{H}a^\top = 0^\top \ (\in \mathbb{F}_P^s)\,.$$

If each entry of $\tilde{H}$ is replaced by a column of its $m$ coordinates with respect to a fixed ordered basis of $\mathbb{F}_q$ over $\mathbb{F}_P$, one checks easily that the resulting $ms \times n$ matrix $H$ with entries in $\mathbb{F}_P$ also satisfies

$$\forall\, a \in \mathbb{F}_P^n : \quad a \in \mathcal{C}_{P,n,\delta} \iff Ha^\top = 0^\top \ (\in \mathbb{F}_P^{ms})\,.$$

Since $\mathcal{C}_{P,n,\delta}$ has dimension $n - ms$ by Lemma 2, we conclude that $H$ is a parity check matrix for $\mathcal{C}_{P,n,\delta}$. Therefore we can apply Lemma 1(ii) to see that $\rho(\mathcal{C}_{P,n,\delta})$ is equal to the smallest integer $r$ such that every column $v \in \mathbb{F}_P^{ms}$ can be written as a linear combination with coefficients in $\mathbb{F}_P$ of at most $r$ columns of $H$. The result now follows by re-interpreting $v$ and $H$ over $\mathbb{F}_q$ and observing that, for a given $(a_i)_{i\in S}$, a linear combination

$$\begin{pmatrix} a_1 \\ \vdots \\ a_{\delta-1} \end{pmatrix} = \sum_{j=1}^{r} \varepsilon_j \cdot (k_j\text{-th column of } \tilde{H})$$

is the same as a solution $\varepsilon_1, \ldots, \varepsilon_r \in \mathbb{F}_P$, $x_1, \ldots, x_r \in \mathbb{F}_q$ to the system (2.1) with $x_i = \zeta^{k_i - 1}$ $(i = 1, \ldots, r)$. □

We can assume that the system (2.1) contains also for $P|k$ equations of the form $\varepsilon_1 x_1^{kN} + \ldots + \varepsilon_r x_r^{kN} = a_k$, namely with $a_{jP} := a_j^P$. Because $a^P = a$ for all $a \in \mathbb{F}_P$ and $(a+b)^P = a^P + b^P$ for all $a, b \in \mathbb{F}_q$, these equations are redundant and affect neither the solvability nor the solutions of the system. The system is still completely determined by an $S$-tuple of values in $\mathbb{F}_q$.

We will refer to (2.1) briefly as the **$N$-Helleseth system of size $\delta - 1 \times r$ with value vector $(a_k)_{k\in S}$**, and a solution with $\varepsilon_1, \ldots, \varepsilon_r \in \mathbb{F}_P$ and $x_1, \ldots, x_r \in \mathbb{F}_q$ will be called a **$(P, q)$-solution** to the system.

Theorem 4 says that to prove $\rho(\mathcal{C}_{P,n,\delta}) \le r$ for a code $\mathcal{C}_{P,n,\delta}$ with $\delta - 1 < \sqrt{q}/N$ it suffices to show that the $N$-Helleseth system of size $\delta - 1 \times r$ has a $(P, q)$-solution for every value vector $(a_k) \in \mathbb{F}_q^S$. By the results of Section 1.4, we only need to consider $\delta \not\equiv 1 \pmod{p}$ and $r \ge \delta - 1$.

From now on we restrict our attention to a much more special case. Our aim for the rest of this thesis is as follows:

$$\begin{array}{l} \textit{Let } C_{p,q-1,\delta} \textit{ be a } p\textit{-ary primitive BCH code, } p \textit{ a prime.} \\ \textit{We want to prove an upper bound } r \textit{ for its covering radius} \\ \textit{(assuming } \delta - 1 < \sqrt{q}\,,\ \delta \not\equiv 1 \pmod{p} \textit{ and } r \geq \delta - 1\,\textit{).} \end{array} \tag{2.2}$$

To achieve (2.2) we must show that the 1-Helleseth system of size $\delta - 1 \times r$ has a $(p,q)$-solution for every $(a_k) \in \mathbb{F}_q^S$. In other words, we want to find a solution to

$$\begin{array}{ccccccccc} \varepsilon_1 x_1 & + & \varepsilon_2 x_2 & + & \ldots & + & \varepsilon_r x_r & = & a_1 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ \varepsilon_1 x_1^{\delta-1} & + & \varepsilon_2 x_2^{\delta-1} & + & \ldots & + & \varepsilon_r x_r^{\delta-1} & = & a_{\delta-1} \end{array} \tag{2.3}$$

with $\varepsilon_1, \ldots, \varepsilon_r \in \mathbb{F}_p$ and $x_1, \ldots, x_r \in \mathbb{F}_q$.

Later, when we want to apply Newton's identities, we will be forced to specialise this to $\varepsilon_1 = \ldots = \varepsilon_r =: \varepsilon \in \mathbb{F}_p^*$. With $\sigma_k := \sigma_k(x_1, \ldots, x_r) := x_1^k + \ldots + x_r^k$ for $k \in \mathbb{N}$, the system (2.3) can then conveniently be written as

$$\varepsilon \sigma_k = a_k \qquad (k \in S)\,. \tag{2.4}$$

## 2.2 Linear transformation of variables

Before making the transition from the Helleseth system (2.3) to a polynomial, it is useful to study the effect of a linear transformation of variables on the system.

**Lemma 4** *Let a system of the form (2.3) be given, and let $c \in \mathbb{F}_q$. Put*

$$y_i := x_i - c \qquad \textit{for } i = 1, \ldots, r\,. \tag{2.5}$$

*Then the original system has a solution if and only if the system*

$$\begin{array}{ccccccccc} \varepsilon_1 y_1 & + & \varepsilon_2 y_2 & + & \ldots & + & \varepsilon_r y_r & = & b_1 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ \varepsilon_1 y_1^{\delta-1} & + & \varepsilon_2 y_2^{\delta-1} & + & \ldots & + & \varepsilon_r y_r^{\delta-1} & = & b_{\delta-1} \end{array}$$

*with*

$$b_k := a_k - c^k(\varepsilon_1 + \ldots + \varepsilon_r) - \sum_{\mu=1}^{k-1} \binom{k}{\mu} c^{k-\mu} b_\mu \in \mathbb{F}_q \tag{2.6}$$

*has one, and solutions of the two systems are linked by (2.5).*

*Proof.*

With $x_i = y_i + c$, the LHS of the $k$-th equation of the original system becomes

$$\begin{aligned}
&\varepsilon_1(y_1+c)^k \ + \ \ldots \ + \ \varepsilon_r(y_r+c)^k \ = \\
&= \ \varepsilon_1 \sum_{\mu=0}^{k} \binom{k}{\mu} y_1^\mu c^{k-\mu} \ + \ \ldots \ + \ \varepsilon_r \sum_{\mu=0}^{k} \binom{k}{\mu} y_r^\mu c^{k-\mu} \ = \\
&= \ c^k(\varepsilon_1 + \ldots + \varepsilon_r) \ + \ (\varepsilon_1 y_1^k + \ldots + \varepsilon_r y_r^k) \ + \\
&\qquad + \left( \varepsilon_1 \sum_{\mu=1}^{k-1} \binom{k}{\mu} y_1^\mu c^{k-\mu} \ + \ \ldots \ + \ \varepsilon_r \sum_{\mu=1}^{k-1} \binom{k}{\mu} y_r^\mu c^{k-\mu} \right) ,
\end{aligned}$$

so that the equation is equivalent to

$$\varepsilon_1 y_1^k \ + \ \ldots \ + \ \varepsilon_r y_r^k \ = \ a_k \ - \ c^k(\varepsilon_1 + \ldots + \varepsilon_r) \ - \ \sum_{\mu=1}^{k-1} \binom{k}{\mu} c^{k-\mu}(\varepsilon_1 y_1^\mu + \ldots + \varepsilon_r y_r^\mu) \, .$$

For $k = 1$, the RHS of this is equal to $a_1 - c(\varepsilon_1 + \ldots + \varepsilon_r) = b_1$. And inductively, for $k > 1$ it is equal to

$$a_k \ - \ c^k(\varepsilon_1 + \ldots + \varepsilon_r) \ - \ \sum_{\mu=1}^{k-1} \binom{k}{\mu} c^{k-\mu} b_\mu \ = \ b_k \, . \qquad \square$$

Note: there is no generalisation of this to translating the variables $x_i$ by different individual values $c_i$, as the induction step relies on the fact that for $\mu = 1, \ldots, k-1$ the common factor $c^{k-\mu}$ can be isolated from $\varepsilon_1 y_1^\mu + \ldots + \varepsilon_r y_r^\mu$.

What can we gain from this? Coming back to the situation (2.4), where all the $\varepsilon_i$ are equal to one $\varepsilon \in \mathbb{F}_p^*$, we have in particular

$$b_1 \ = \ a_1 - cr\varepsilon \, .$$

Thus, for $r \not\equiv 0 \pmod{p}$ it is possible to obtain any $b_1 \in \mathbb{F}_q$ via the above transformation by taking $c := (a_1 - b_1)/r\varepsilon$. As a consequence, rather than solving (2.4) for all $(a_k) \in \mathbb{F}_q^S$ we may assume that $a_1$ is an arbitrary fixed value in $\mathbb{F}_q$, for instance zero.

For $r \equiv 0 \pmod{p}$ this assumption cannot be made, as the transformation then leaves $a_1$ unchanged for all $c$. (Assumptions, of a more complex nature, about $a_k$ with higher indices are possible, but will not be used in this study.)

## 2.3 Conversion into the problem of the splitting of a polynomial over $\mathbb{F}_q$

The link between a Helleseth system with constant coefficients (2.4) and the roots of a polynomial is provided be the Newton identities (cf. [LiNi], Theorem 1.75).

**Theorem 5 (Newton identities)** *Let $\sigma_k$ be as defined in Section 2.1, and for $i = 0, \ldots, r$ let $S_i := S_i(x_1, \ldots, x_r)$ be the $i$-th elementary symmetric polynomial in $x_1, \ldots, x_r$, i.e.*

$$\begin{aligned} S_0 &= 1\,, \\ S_1 &= x_1 + x_2 + \ldots + x_r\,, \\ &\vdots \\ S_r &= x_1 x_2 \cdots x_r\,. \end{aligned}$$

*Then the following formula holds for all $k = 1, \ldots, r$:*

$$\sum_{l=0}^{k-1} (-1)^l \sigma_{k-l} S_l \;=\; (-1)^{k+1} k S_k\,.$$

*And for all $k > r$:*

$$\sum_{l=0}^{r} (-1)^l \sigma_{k-l} S_l \;=\; 0\,.$$

Now suppose a fixed system of the form (2.4) is given. Express $r$ (the number of variables $x_i$) as $r =: pt + M$ with $t = \lfloor \frac{r}{p} \rfloor$ and $M \in \{0, \ldots, p-1\}$. Then apply the following procedure.

1. Choose arbitrary elements $s_p, s_{2p}, \ldots, s_{tp} \in \mathbb{F}_q$.
   If $\delta - 1 < r$, choose also arbitrary elements $a_k \in \mathbb{F}_q$ for $k = \delta, \ldots, r$ with $p \nmid k$, and fill possible gaps for $p|k$ with $a_{jp} := a_j^p$.

2. Define $s_0 := 1$ and

$$s_k \;:=\; \frac{(-1)^{k+1}}{\varepsilon k} \sum_{l=0}^{k-1} (-1)^l a_{k-l} s_l \;\in \mathbb{F}_q \qquad \text{for } k = 1, \ldots, r \text{ with } p \nmid k\,.$$

3. Form the polynomial $F(x) := \displaystyle\sum_{i=0}^{r} (-1)^i s_i x^{r-i} \;\in \mathbb{F}_q[x]$.

4. Factorize $F(x)$ over $\overline{\mathbb{F}}_q$, the algebraic closure of $\mathbb{F}_q$, into

$$F(x) \;=\; \prod_{i=1}^{r} (x - \gamma_i)\,.$$

**Lemma 5**

*The roots $\gamma_1, \ldots, \gamma_r$ of $F(x)$ form a solution to the system $\varepsilon\sigma_k = a_k$ $(k = 1, \ldots, r)$, and in particular to the subsystem $\varepsilon\sigma_k = a_k$ $(k = 1, \ldots, \delta - 1)$.*
*If $F(x)$ splits completely over $\mathbb{F}_q$, we obtain a $(p, q)$-solution to both systems.*

*Proof.* For the purpose of this proof, let $\sigma_k$ stand for $\sigma_k(\gamma_1, \ldots, \gamma_r)$. We must show that for all $k = 1, \ldots, r$

$$\varepsilon\sigma_k \;=\; a_k\,. \tag{2.7}$$

Expressing $F(x)$ in terms of the elementary symmetric polynomials in its roots,

$$F(x) \;=\; \prod_{i=1}^{r}(x - \gamma_i) \;=\; \sum_{i=0}^{r}(-1)^i S_i(\gamma_1, \ldots, \gamma_r)x^{r-i}\,,$$

and comparing coefficients with the definition of $F(x)$ shows that $S_i(\gamma_1, \ldots, \gamma_r) = s_i$ for all $i = 1, \ldots, r$. Hence it follows from the Newton identities that

$$(-1)^{k+1}ks_k \;=\; \sum_{l=0}^{k-1}(-1)^l\sigma_{k-l}s_l \qquad \text{for all } k = 1, \ldots, r\,. \tag{2.8}$$

Now use induction.

$\underline{k = 1}$: (2.8) says that $s_1 = \sigma_1 s_0 = \sigma_1$. By definition, $s_1 = \varepsilon^{-1}a_1$. Hence $\varepsilon\sigma_1 = a_1$.

$\underline{k = 2, \ldots, r}$: Suppose (2.7) has already been shown for all indices up to $k - 1$.

If $k \equiv 0 \pmod{p}$, write $k = jp$. From the induction hypothesis we know $\varepsilon\sigma_j = a_j$, and $\varepsilon\sigma_k = a_k$ follows by taking $p$-th powers on both sides.

If $k \not\equiv 0 \pmod{p}$, we have

$$\begin{aligned}
\sigma_k s_0 \;&\overset{(2.8)}{=}\; -\sum_{l=1}^{k-1}(-1)^l\sigma_{k-l}s_l \;+\; (-1)^{k+1}ks_k \;=^{\text{definition of } s_k}\\
&=\; -\sum_{l=1}^{k-1}(-1)^l\sigma_{k-l}s_l \;+\; \frac{1}{\varepsilon}\sum_{l=0}^{k-1}(-1)^l a_{k-l}s_l \;=^{\text{induction hypothesis}}\\
&=\; -\sum_{l=1}^{k-1}(-1)^l\sigma_{k-l}s_l \;+\; \sum_{l=1}^{k-1}(-1)^l\sigma_{k-l}s_l \;+\; \varepsilon^{-1}a_k s_0\,,
\end{aligned}$$

so that again $\varepsilon\sigma_k \;=\; a_k$. □

With this the problem (2.2) about the covering radius of $\mathcal{C}_{p,q-1,\delta}$ has been transformed into the problem of making a polynomial split completely over $\mathbb{F}_q$.

Note that we do not get the "converse": if for some $(a_k) \in \mathbb{F}_q^S$ there exists no choice of $s_p, \ldots, s_{pt} \in \mathbb{F}_q$ for which $F(x)$ splits, this does not imply a lower bound for $\rho(\mathcal{C}_{p,q-1,\delta})$. This is because we had to choose all $\varepsilon_i$ equal in order to accommodate the Newton identities—but there may still exist a solution to one of the Helleseth systems with $\varepsilon_i \neq \varepsilon_j$ for some $i, j \in \{1, \ldots, r\}$.

## 2.4 Expressing the $s_k$ $(p \nmid k)$ in terms of the $s_k$ $(p|k)$

A look back at the procedure on page 13 shows that the elements to be varied in order to achieve the splitting of $F(x)$ are exactly the $s_p, s_{2p}, \ldots, s_{tp}$ and the $a_k$ with $\delta \le k \le r$ and $p \nmid k$ (if any of the latter are present). Everything else is then completely determined. Therefore our next wish is to express $F(x)$ in a form that uses only the $s_k$ with indices divisible by $p$.

**Lemma 6** *Let* $s_p, s_{2p}, \ldots, s_{tp}$, $a_k$ $(k = 1, \ldots, r)$, $s_0 = 1$ *and*

$$s_k = \frac{1}{k\varepsilon} \sum_{l=0}^{k-1} (-1)^{k+l+1} a_{k-l} s_l \qquad (k = 1, \ldots, r;\ p \nmid k)$$

*be the elements chosen or defined in the first two steps of the procedure on page 13. Define* $A_0 := 1$ *and, for* $k = 1, \ldots, r$,

$$A_k := \begin{cases} 0, & \text{if } p|k, \\ -\dfrac{1}{k\varepsilon} \displaystyle\sum_{\mu=0}^{k-1} a_{k-\mu} A_\mu, & \text{if } p \nmid k. \end{cases} \tag{2.9}$$

*Then*

$$s_k = \sum_{\nu=0}^{\lfloor \frac{k}{p} \rfloor} (-1)^{k-\nu p} A_{k-\nu p} s_{\nu p} \qquad \text{for all } k = 1, \ldots, r.$$

*Proof.* Induction on $k$. For $p|k$ we have the trivial statement $s_k = s_k$, so that only for $p \nmid k$ there is something to prove.

$\underline{k = 1}$: We have $s_1 = \varepsilon^{-1} a_1 s_0$ and $A_1 = -\varepsilon^{-1} a_1$, hence $s_1 = -A_1 s_0$.

$\underline{k \ge 2\ (p \nmid k)}$:

$$s_k \overset{\text{def}}{=} \frac{1}{k\varepsilon} \sum_{l=0}^{k-1} (-1)^{k+l+1} a_{k-l} s_l =$$

$$= \sum_{\substack{l=0 \\ p|l}}^{k-1} \frac{1}{k\varepsilon} (-1)^{k-l+1} a_{k-l} A_0 s_l + \sum_{\substack{l=0 \\ p \nmid l}}^{k-1} \frac{1}{k\varepsilon} (-1)^{k+l+1} a_{k-l} \underbrace{\sum_{\nu=0}^{\lfloor \frac{l}{p} \rfloor} (-1)^{l-\nu p} A_{l-\nu p} s_{\nu p}}_{= s_l \text{ by induction hypothesis}} =$$

$$= \underbrace{\sum_{\substack{l=0 \\ p|l}}^{k-1} \sum_{\nu=\lfloor \frac{l}{p} \rfloor}^{\lfloor \frac{l}{p} \rfloor} \frac{1}{k\varepsilon} (-1)^{k-\nu p+1} a_{k-l} A_{l-\nu p} s_{\nu p}}_{\text{complicated way of writing the first sum with } l = \nu p} + \underbrace{\sum_{\substack{l=0 \\ p|l}}^{k-1} \sum_{\nu=0}^{\lfloor \frac{l}{p} \rfloor - 1} \frac{1}{k\varepsilon} (-1)^{k-\nu p+1} a_{k-l} \underbrace{A_{l-\nu p}}_{=0} s_{\nu p}}_{\text{complicated way of writing an additional zero term}} +$$

$$+ \sum_{\substack{l=0 \\ p \nmid l}}^{k-1} \sum_{\nu=0}^{\lfloor \frac{l}{p} \rfloor} \frac{1}{k\varepsilon} (-1)^{k-\nu p+1} a_{k-l} A_{l-\nu p} s_{\nu p},$$

and with $\frac{1}{k\varepsilon}(-1)^{k-\nu p+1}a_{k-l}A_{l-\nu p}s_{\nu p} =: B(l,\nu)$ for $l = 0,\ldots,k-1$ and $\nu = 0,\ldots,\lfloor\frac{l}{p}\rfloor$, this whole expression becomes

$$\sum_{l=0}^{k-1}\sum_{\nu=0}^{\lfloor\frac{l}{p}\rfloor} B(l,\nu)\ .$$

If further $B(l,\nu) := 0$ for $\nu > \lfloor\frac{l}{p}\rfloor$, then

$$s_k \;=\; \sum_{l=0}^{k-1}\sum_{\nu=0}^{\lfloor\frac{k}{p}\rfloor} B(l,\nu) \;=\; \sum_{\nu=0}^{\lfloor\frac{k}{p}\rfloor}\sum_{l=0}^{k-1} B(l,\nu) \;=\; \sum_{\nu=0}^{\lfloor\frac{k}{p}\rfloor}\sum_{l=\nu p}^{k-1} B(l,\nu)\,,$$

where the last step uses that for $l = 0,\ldots,\nu p-1$ we have $\nu > \frac{l}{p} \geq \lfloor\frac{l}{p}\rfloor$ and therefore $B(l,\nu) = 0$.

Finally,

$$\begin{aligned}
s_k \;&=\; \sum_{\nu=0}^{\lfloor\frac{k}{p}\rfloor}\sum_{l=\nu p}^{k-1} \frac{1}{k\varepsilon}(-1)^{k-\nu p+1}a_{k-l}A_{l-\nu p}s_{\nu p} \;\overset{l =: \mu+\nu p}{=}\\
&=\; \sum_{\nu=0}^{\lfloor\frac{k}{p}\rfloor}\sum_{\mu=0}^{k-1-\nu p} \frac{1}{k\varepsilon}(-1)^{k-\nu p+1}a_{k-\mu-\nu p}A_{\mu}s_{\nu p} \;\overset{\text{characteristic } p}{=}\\
&=\; \sum_{\nu=0}^{\lfloor\frac{k}{p}\rfloor}(-1)^{k-\nu p}\left(-\frac{1}{(k-\nu p)\varepsilon}\sum_{\mu=0}^{(k-\nu p)-1} a_{(k-\nu p)-\mu}A_{\mu}\right)s_{\nu p} \;=\\
&=\; \sum_{\nu=0}^{\lfloor\frac{k}{p}\rfloor}(-1)^{k-\nu p}A_{k-\nu p}s_{\nu p}\,, \quad \text{as claimed.} \qquad \square
\end{aligned}$$

Substituting the expression of Lemma 6 for the $s_k$, we obtain a version of $F(x)$ which uses only the $s_{\nu p}$ and $A_k$. Definition (2.9) establishes a one-to-one correspondence between the $(a_k) \in \mathbb{F}_q^S$ and the $(A_k) \in \mathbb{F}_q^S$. It is reversed via

$$a_k \;=\; \begin{cases} a_j^p\,, & \text{if } k = jp\,,\\ -k\varepsilon A_k \;-\; \sum_{\mu=1}^{k-1} a_{k-\mu}A_{\mu}\,, & \text{if } p \nmid k\,. \end{cases} \tag{2.10}$$

In these definitions, the $A_k$ and $a_k$ with indices $k = 1,\ldots,\delta-1$ are independent of those with indices $k = \delta,\ldots,r$. Moreover, $A_1 = -\varepsilon^{-1}a_1$, in particular $A_1 = 0$ if and only if $a_1 = 0$. Therefore, in order to achieve (2.2), it suffices to make $F(x)$ split over $\mathbb{F}_q$ for any $(A_k) \in \mathbb{F}_q^S$, where the same assumptions can be made about the $A_k$ as were made about the $a_k$ at the end of Section 2.2, and the $A_k$ for $k = \delta,\ldots,r$ with $p \nmid k$ can be chosen arbitrarily.

## 2.5 Re-structuring of $F(x)$ and summary

Finally, some manipulations are needed to bring $F(x)$ into a shape that makes it suitable to be subjected to our splitting criteria in Chapter 3. We have

$$F(x) \stackrel{\text{def}}{=} \sum_{i=0}^{r}(-1)^i s_i x^{r-i} = \sum_{\substack{i=0 \\ p|i}}^{r}(-1)^i s_i x^{r-i} + \sum_{\substack{i=0 \\ p\nmid i}}^{r}(-1)^i s_i x^{r-i} =$$

$$= \sum_{\nu=0}^{t}(-1)^{\nu p} s_{\nu p} x^{r-\nu p} + \sum_{\alpha=1}^{M}\sum_{\mu=0}^{t}(-1)^{\mu p+\alpha} s_{\mu p+\alpha} x^{r-(\mu p+\alpha)} + \tag{2.11}$$

$$+ \sum_{\alpha=M+1}^{p-1}\sum_{\mu=0}^{t-1}(-1)^{\mu p+\alpha} s_{\mu p+\alpha} x^{r-(\mu p+\alpha)} .$$

The second of the three expressions in (2.11) is equal to

$$\sum_{\alpha=1}^{M}\sum_{\mu=0}^{t}(-1)^{\mu p+\alpha}\underbrace{\left(\sum_{\nu=0}^{\mu}(-1)^{(\mu p+\alpha)-\nu p} A_{(\mu p+\alpha)-\nu p} s_{\nu p}\right)}_{= s_{\mu p+\alpha} \text{ by Lemma 6}} x^{r-(\mu p+\alpha)} =$$

$$= \sum_{\mu=0}^{t}\sum_{\nu=0}^{\mu}\sum_{\alpha=1}^{M}\underbrace{(-1)^{\nu p} A_{(\mu-\nu)p+\alpha} s_{\nu p} x^{r-(\mu p+\alpha)}}_{=: B(\mu,\nu,\alpha) \text{ for short}} =$$

$$= \sum_{\substack{\mu,\nu=0 \\ \nu\le\mu}}^{t}\sum_{\alpha=1}^{M} B(\mu,\nu,\alpha) = \sum_{\nu=0}^{t}\sum_{\mu=\nu}^{t}\sum_{\alpha=1}^{M} B(\mu,\nu,\alpha) =$$

$$= \sum_{\nu=0}^{t}\sum_{\alpha=1}^{M}\sum_{\mu=\nu}^{t}(-1)^{\nu p} A_{(\mu-\nu)p+\alpha} s_{\nu p} x^{pt+M-(\mu p+\alpha)} =$$

$$= \sum_{\nu=0}^{t}\sum_{\alpha=1}^{M}(-1)^{\nu p} x^{M-\alpha}\, g_{t-1-\nu}^{(M-\alpha)}(x^p)\, s_{\nu p}$$

with $g_{t-1-\nu}^{(M-\alpha)}(x^p) := \sum_{\mu=\nu}^{t} A_{(\mu-\nu)p+\alpha}(x^p)^{t-\mu}$ for $\nu = 0,\ldots,t$ and $\alpha = 1,\ldots,M$.

Similarly, the third expression in (2.11) is equal to

$$\sum_{\alpha=M+1}^{p-1}\sum_{\mu=0}^{t-1}(-1)^{\mu p+\alpha}\left(\sum_{\nu=0}^{\mu}(-1)^{(\mu p+\alpha)-\nu p} A_{(\mu p+\alpha)-\nu p} s_{\nu p}\right) x^{r-(\mu p+\alpha)} = \ldots =$$

$$= \sum_{\nu=0}^{t-1} \sum_{\alpha=M+1}^{p-1} \sum_{\mu=\nu}^{t-1} (-1)^{\nu p} A_{(\mu-\nu)p+\alpha} s_{\nu p} x^{pt+M-(\mu p+\alpha)+p-p} \quad =$$

$$= \sum_{\nu=0}^{t} \sum_{\alpha=M+1}^{p-1} (-1)^{\nu p} x^{p+M-\alpha} \, g_{t-1-\nu}^{(p+M-\alpha)}(x^p) \, s_{\nu p}$$

with $g_{t-1-\nu}^{(p+M-\alpha)}(x^p) := \sum_{\mu=\nu}^{t-1} A_{(\mu-\nu)p+\alpha}(x^p)^{t-\mu-1} \quad (\nu = 0, \ldots, t; \alpha = M+1, \ldots, p-1)$

(for $\nu = t$ this definition means in particular $g_{-1}^{(p+M-\alpha)} = 0$).

With all this put together, $F(x)$ equals

$$\sum_{\nu=0}^{t} (-1)^{\nu p} s_{\nu p} \left[ x^{r-\nu p} + \sum_{\alpha=1}^{M} x^{M-\alpha} \, g_{t-1-\nu}^{(M-\alpha)}(x^p) + \sum_{\alpha=M+1}^{p-1} x^{p+M-\alpha} \, g_{t-1-\nu}^{(p+M-\alpha)}(x^p) \right].$$

Finally, by replacing $s_{\nu p}$ for odd $\nu p$ with its negative, the factor $(-1)^{\nu p}$ can be dropped.

Thus the results of this chapter can be summed up as follows.

**Theorem 6**

*Let $p$ be a prime and $q$ a power of $p$.*

*In order to show that a primitive $p$-ary BCH code $C_{p,q-1,\delta}$ with $\delta - 1 < \sqrt{q}$ has covering radius at most $r = pt + M$ $(M \in \{0, \ldots, p-1\})$, it suffices to show that for every choice of*

$$A_k \in \mathbb{F}_q \qquad (k = 1, \ldots, \delta - 1; \; p \nmid k) \tag{2.12}$$

*there exist elements*

$$s_p, s_{2p}, \ldots, s_{tp} \in \mathbb{F}_q \qquad \text{and} \qquad A_k \in \mathbb{F}_q \quad (k = \delta, \ldots, r; \; p \nmid k)$$

*such that the polynomial*

$$F(x) \; = \; \sum_{\nu=0}^{t} s_{\nu p} \left[ x^{r-\nu p} + \sum_{\alpha=1}^{M} x^{M-\alpha} \, g_{t-1-\nu}^{(M-\alpha)}(x^p) + \sum_{\alpha=M+1}^{p-1} x^{p+M-\alpha} \, g_{t-1-\nu}^{(p+M-\alpha)}(x^p) \right]$$

*with*

$$\begin{aligned} s_0 &= 1, \\ g_{t-1-\nu}^{(M-\alpha)}(x^p) &= \sum_{\mu=\nu}^{t} A_{(\mu-\nu)p+\alpha} x^{p(t-\mu)}, \\ g_{t-1-\nu}^{(p+M-\alpha)}(x^p) &= \sum_{\mu=\nu}^{t-1} A_{(\mu-\nu)p+\alpha} x^{p(t-1-\mu)}, \end{aligned}$$

*splits completely over* $\mathbb{F}_q$ .

*Moreover, if* $r \not\equiv 0 \pmod{p}$, *one can take* $k \geq 2$ *in (2.12) and assume that* $A_1$ *is any fixed element (zero or other).*

The definition of the $g$-polynomials above reflects the way they were derived. However, a more convenient form to work with is

$$g_{t-1-\nu}^{(M-\alpha)}(x^p) = \sum_{i=0}^{t-\nu} A_{(t-\nu-i)p+\alpha} x^{pi},$$
$$g_{t-1-\nu}^{(p+M-\alpha)}(x^p) = \sum_{i=0}^{t-1-\nu} A_{(t-1-\nu-i)p+\alpha} x^{pi}.$$

## 2.6 An example

The re-structured polynomial $F(x)$ in Theorem 6 is at the centre of our approach. Its building principle is best illustrated with a specific example. We do it here for $p = 3$ and $r = 3t + 1$, as it arises e.g. from $\delta \equiv 2 \pmod{3}$ and $r = \delta - 1$. This may not be the most instructive example—the reader is encouraged to work out one with $p \geq 5$ and $M \in \{2, \ldots, p-2\}$—but it is the situation that will prevail in the rest of this thesis.

With $p$ and $r$ as said, one finds that there are $g^{(2)}$- and $g^{(0)}$-polynomials, whose expansions are as follows:

$$\begin{aligned}
g_{t-1}^{(2)}(x^3) &= A_2 x^{3t-3} + A_5 x^{3t-6} + \ldots + A_{3t-7} x^6 + A_{3t-4} x^3 + A_{3t-1}, \\
g_{t-2}^{(2)}(x^3) &= A_2 x^{3t-6} + A_5 x^{3t-9} + \ldots + A_{3t-7} x^3 + A_{3t-4}, \\
&\vdots \\
g_2^{(2)}(x^3) &= A_2 x^6 + A_5 x^3 + A_8, \\
g_1^{(2)}(x^3) &= A_2 x^3 + A_5, \\
g_0^{(2)}(x^3) &= A_2, \\
g_{-1}^{(2)}(x^3) &= 0,
\end{aligned}$$

and

$$\begin{aligned}
g_{t-1}^{(0)}(x^3) &= A_1 x^{3t} + A_4 x^{3t-3} + A_7 x^{3t-6} + \ldots + A_{3t-5} x^6 + A_{3t-2} x^3 + A_{3t+1}, \\
g_{t-2}^{(0)}(x^3) &= A_1 x^{3t-3} + A_4 x^{3t-6} + A_7 x^{3t-9} + \ldots + A_{3t-5} x^3 + A_{3t-2}, \\
&\vdots \\
g_2^{(0)}(x^3) &= A_1 x^9 + A_4 x^6 + A_7 x^3 + A_{10}, \\
g_1^{(0)}(x^3) &= A_1 x^6 + A_4 x^3 + A_7, \\
g_0^{(0)}(x^3) &= A_1 x^3 + A_4, \\
g_{-1}^{(0)}(x^3) &= A_1;
\end{aligned}$$

if we assume $A_1 = 0$, the $g^{(0)}$-polynomials are shortened to

$$\begin{aligned}
g^{(0)}_{t-1}(x^3) &= A_4x^{3t-3} + A_7x^{3t-6} + \ldots + A_{3t-5}x^6 + A_{3t-2}x^3 + A_{3t+1}\,,\\
g^{(0)}_{t-2}(x^3) &= A_4x^{3t-6} + A_7x^{3t-9} + \ldots + A_{3t-5}x^3 + A_{3t-2}\,,\\
&\vdots\\
g^{(0)}_{2}(x^3) &= A_4x^6 + A_7x^3 + A_{10}\,,\\
g^{(0)}_{1}(x^3) &= A_4x^3 + A_7\,,\\
g^{(0)}_{0}(x^3) &= A_4\,,\\
g^{(0)}_{-1}(x^3) &= 0\,.
\end{aligned}$$

Together these form

$$\begin{aligned}
F(x) \;=\; & \Big[x^{3t+1} + x^2\, g^{(2)}_{t-1}(x^3) + g^{(0)}_{t-1}(x^3)\Big] \;+\\
& +\; s_3 \;\cdot\Big[x^{3t-2} + x^2\, g^{(2)}_{t-2}(x^3) + g^{(0)}_{t-2}(x^3)\Big] \;+\\
& +\; \ldots \;+\\
& +\; s_{3t-6}\cdot\Big[x^7 + x^2\, g^{(2)}_{1}(x^3) + g^{(0)}_{1}(x^3)\Big] \;+\\
& +\; s_{3t-3}\cdot\Big[x^4 + x^2\, g^{(2)}_{0}(x^3) + g^{(0)}_{0}(x^3)\Big] \;+\\
& +\; s_{3t}\;\cdot\Big[x + g^{(0)}_{-1}(x^3)\Big]\,.
\end{aligned}$$

Here it may also be of interest to look at the expansions of the polynomials in square brackets: the top one is

$$\begin{aligned}
&x^{3t+1} + A_1x^{3t} + A_2x^{3t-1} + A_4x^{3t-3} + \ldots\\
&\qquad\qquad \ldots + A_{3(t-i)-1}x^{3i+2} + A_{3(t-i)+1}x^{3i} + \ldots + A_{3t-1}x^2 + A_{3t+1}\,,
\end{aligned}$$

the penultimate one is $x^4 + A_1x^3 + A_2x^2 + A_4$ ($= x^4 + A_2x^2 + A_4$ if $A_1 = 0$), and the last one is $x + A_1$ (or simply $x$, if $A_1 = 0$).

One final word about our choice of notation. The reader will have noticed that in the symbol $g^{(i)}_j$

(1) the $i$ stands for the power of $x$ together with which the polynomial $g^{(i)}_j(x^p)$ appears in $F(x)$,

(2) for $i$ of the form $p + M - \alpha$ ($\alpha = M + 1, \ldots, p - 1$), the $j$ corresponds to the maximal degree that $g^{(i)}_j$ can have as a polynomial in $x$.

However, (2) does not hold for $i = M - \alpha$ ($\alpha = 1, \ldots, M$), except for $\alpha = 1$ when we assume $A_1 = 0$. Otherwise the index $j$ would have to be one higher.

This is somewhat inconsequent, but for a good reason: the overwhelming part of the situations that we will deal with in this thesis will have $M = 1$ and $A_1 = 0$, in

which case (1) and (2) cover everything and our notation is more natural. Moreover, apart from avoiding much unnecessary awkwardness in Chapters 4–6, the chosen notation is more[†] consistent with [FrCoh]. The only concession we have to make is to keep in mind that for $i = M - \alpha$ ($\alpha = 1, \ldots, M$) the degree of $g_j^{(i)}(x^p)$ is bounded by $p(j+1)$ rather than $pj$.

[†]There is, however, some inconsistency as far as $t$ is concerned: in [FrCoh] we wrote $\delta = 3t - 1$, whereas now we prefer $\delta = pt + M$. For fixed $\delta$, this makes odd $t$ even and vice versa.

# Chapter 3

# How to Achieve the Splitting

With the First Part of our method complete, what we now need is a means to ensure that the polynomial $F(x)$ in the situation of Theorem 6 splits completely over $\mathbb{F}_q$. This Second Part of the method is independent of the First Part inasmuch as *any* result to this effect (known or to still be found) could be used. We stick to the criterion which was also used in [Coh97], and which stems from [Coh98].

Theorem 7 and Lemma 7 below are extracts from the work of Cohen. They are taken from a wider context, namely an extensive (and sometimes rather intricate) study of ramification in function fields, for which Sections 3 and 4 of [Coh98] are the primary reference (besides [Coh99] and others). The proofs are too extensive to be repeated here, but some key steps will be mentioned with the intention to convey to the reader a feeling for the underlying ideas and some understanding how the generalisations in the second part of Section 3.2 work. This requires some basic knowledge of Algebraic Function Field Theory; an excellent introduction to this area is [Sti].

Sections 3.3–3.6 gather some auxiliary material to deal with the practical conditions for the splitting obtained in Section 3.2. Some of the statements may seem almost too trivial to be called a lemma, but the purpose in doing so is twofold: to lift some of the burden of Chapter 4, thereby making the proof of Theorem 8 more transparent, and to supply certain arguments which tend to be used frequently in connection with our method in a compact and widely transferable form.

Notation: $\mathcal{S}_n$ denotes the symmetric group, $\mathcal{A}_n$ the alternating group of degree $n$. From Section 3.2 on, $\mathbb{K}$ stands for a general field with an extension field $\mathbb{L}$ and algebraic closure $\overline{\mathbb{K}}$. All notation from previous chapters is retained.

## 3.1 The splitting criterion

Here is the result around which the whole method is built.

**Theorem 7 (Cohen)**

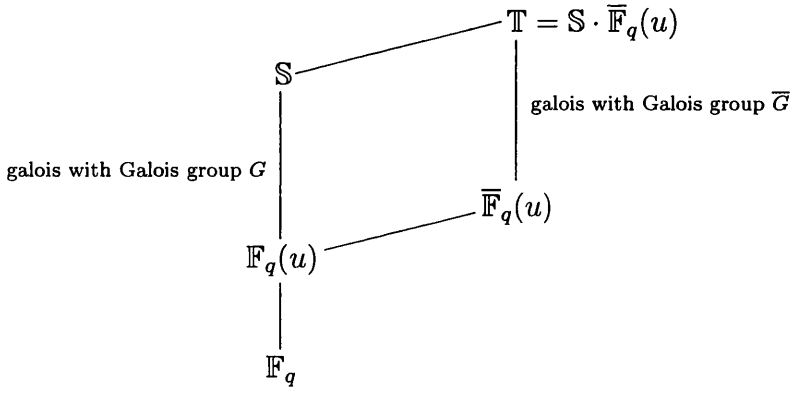
*Let $F_u(x) := f_0(x) + uf_1(x)$, where $u$ is an indeterminate and $f_0, f_1 \in \mathbb{F}_q[x]$ are monic polynomials with $r := \deg f_0 > \deg f_1 \geq 0$ and $f_0/f_1 \notin \mathbb{F}_q(x^p)$.*

*Write $G$ for the Galois group of $F_u$ over $\mathbb{F}_q(u)$ and $\overline{G}$ for that of $F_u$ over $\overline{\mathbb{F}}_q(u)$.*

*Then there exists an $\alpha \in \mathbb{F}_q$ such that $F_\alpha(x)$ splits completely into (distinct) linear factors over $\mathbb{F}_q$, provided $f_0$ and $f_1$ are co-prime, $q > [r! \cdot (r-2)]^2$ and $\overline{G} = G$.*

This is a shortened version of Lemma 5.1 of [Coh98]. (There deg $f_1 \geq 2$, but this is for other reasons and not essential; cf. [Coh99], which is without this restriction.)

We explain briefly the significance of the condition $\overline{G} = G$. Let $\mathbb{S}$ and $\mathbb{T}$ be the respective splitting fields of $F_u$ over the rational function fields $\mathbb{F}_q(u)$ and $\overline{\mathbb{F}}_q(u)$:




The original statement of the lemma in [Coh98] is an estimate which shows that for sufficiently large $q$ the number of $\alpha \in \mathbb{F}_q$ as in the theorem is positive. This is derived from the Hasse-Weil bound for $\mathbb{S}$ and the Hurwitz genus formula applied to $\mathbb{S}/\mathbb{F}_q(u)$. For the former to be valid, it is necessary that $\mathbb{F}_q$ is the full constant field of $\mathbb{S}$, meaning that $\mathbb{F}_q$ is algebraically closed in $\mathbb{S}$. This is equivalent to $\overline{G} = G$.

$G$ and $\overline{G}$ are sometimes called **arithmetic** and **geometric Galois group** of $F_u$, respectively. They always satisfy $\overline{G} \subseteq G \subseteq \mathcal{S}_r$ (viewed as groups of permutations of the roots of $F_u$), so it suffices to concentrate on $G \subseteq \overline{G}$.

By setting the $s_p, s_{2p}, \ldots, s_{tp}$ in Theorem 6 equal to polynomial expressions in an indeterminate $u$, it is easy to bring $F(x)$ into the form $F_u(x)$ of Theorem 7. Usually we will choose all but few of the $s_p, s_{2p}, \ldots, s_{tp}$ equal to zero.

In the same way, parameters $v_1, \ldots, v_k$ (usually one $v$ will suffice) can be brought in, such that one or both of $f_0, f_1$ depend on them. Then it remains to show that these can be chosen in $\mathbb{F}_q$ in such a way that the conditions for the splitting in Theorem 7 are satisfied.

Of these, the co-primality of $f_0$ and $f_1$ is an obvious necessity which will never be problematic to arrange. The condition $q > [r!\,(r-2)]^2$ will appear in our theorems as a lower bound for the validity of our statements about the covering radius of codes. Our efforts go now into establishing practical criteria (i.e. elementary properties the polynomials $f_0$ and $f_1$ must have) for $\overline{G} = G$.

## 3.2 Practical conditions

First we need some terminology (valid in an arbitrary field extension $\mathbb{L}/\mathbb{K}$).

Let $g, h \in \mathbb{K}[x]$, $h \neq 0$. The rational function $f := g/h$ is called **decomposable over** $\mathbb{L}$ if there exist $Q_1, Q_2, R_1, R_2 \in \mathbb{L}[x]$ with $\gcd(Q_1, Q_2) = \gcd(R_1, R_2) = 1$ such that, with $Q := Q_1/Q_2$ and $R := R_1/R_2$,

(i) $f(x) = Q(R(x))$, and

(ii) neither $Q$ nor $R$ is a fraction of linear polynomials.

We call the quadruple $(Q_1, Q_2, R_1, R_2)$ a **non-trivial decomposition** of $f$ over $\mathbb{L}$. If no such non-trivial decomposition exists, $f$ is said to be **indecomposable** over $\mathbb{L}$.

Further, a polynomial $g \in \mathbb{K}[x]$ will be called **simple** if exactly one of its linear factors in $\overline{\mathbb{K}}[x]$ has multiplicity two and all others have multiplicity one.

Now we can formulate:

**Lemma 7 (Cohen)** *Suppose that, in the situation of Theorem 7, the polynomial $F_u$ satisfies all of the following conditions:*

(i) *$f_0$ and $f_1$ are co-prime,*

(ii) *$f_0/f_1$ is indecomposable over $\mathbb{F}_q$,*

(iii) *$F_\beta$ is simple for some $\beta \in \overline{\mathbb{F}}_q$.*

*Then $\overline{G} = G = \mathcal{S}_r$.*

This is Lemma 3.1 of [Coh98] with alternative (a). Condition (i) makes $G$ a transitive subgroup of $\mathcal{S}_r$, and then, as detailed in [Fri], Lemma 2, or [Coh91], Lemma 3.1, condition (ii) implies that $G$ is primitive (see e.g. [Wie] for definitions). If now $\overline{G}$ (the smaller group) contains a transposition, then it follows (see [Coh98], p. 326/327) that the subgroup of $\overline{G}$ generated by transpositions equals $\mathcal{S}_r$, hence the result. Securing the existence of a transposition in $\overline{G}$ is the purpose of condition (iii). This is a key point and deserves some elaboration.

For a root $\alpha$ of $F_u$ we have $\overline{\mathbb{F}}_q(u) \subseteq \overline{\mathbb{F}}_q(\alpha)$. Write $\hat{\mathbb{F}}_q := \overline{\mathbb{F}}_q(u) \cup \{\infty\}$.

The set of places of the rational function field $\overline{\mathbb{F}}_q(u)$ is $\{P_\beta\colon\ \beta \in \hat{\mathbb{F}}_q\}$, where $P_\beta$ for $\beta \in \overline{\mathbb{F}}_q$ corresponds to the $(u-\beta)$-adic and $P_\infty$ to the $\frac{1}{u}$-adic valuation. Similarly, the places of $\overline{\mathbb{F}}_q(\alpha)$ are given by $\{\mathfrak{p}_\beta\colon\ \beta \in \hat{\mathbb{F}}_q\}$. ([Sti], I.2.)

For $\beta \in \overline{\mathbb{F}}_q$, factorise $F_\beta(x)$ as

$$\prod_{i=1}^{r(\beta)} \big(x - \gamma_i(\beta)\big)^{e_i(\beta)} \qquad \text{with distinct } \gamma_i(\beta) \in \overline{\mathbb{F}}_q\,,$$

i.e. the $e_i(\beta)$ are the multiplicities of the roots of $F_\beta$ in $\overline{\mathbb{F}}_q$. Further, factorise $f_1$ as

$$\prod_{i=1}^{r(\infty)-1} \big(x - \gamma_i(\infty)\big)^{e_i(\infty)} \qquad \text{with distinct } \gamma_i(\infty) \in \overline{\mathbb{F}}_q\,,$$

i.e. the $e_i(\infty)$ for $i = 1, \ldots, r(\infty) - 1$ are the multiplicities of the roots of $f_1$ in $\overline{\mathbb{F}}_q$. Set $\gamma_{r(\infty)}(\infty) := \infty$ and $e_{r(\infty)}(\infty) := \deg f_0 - \deg f_1$. Then for all $\beta \in \hat{\mathbb{F}}_q$ the ramification index of $\mathfrak{p}_{\gamma_i(\beta)}$ over $P_\beta$ is $e_i(\beta)$. ([Coh98], p. 327/328.)

Now fix $\beta \in \hat{\mathbb{F}}_q$ (subsequently we suppress the argument $\beta$ after $r$, $e_i$ etc.). For a place $\mathfrak{P}$ of $\mathbb{T}$ lying over $P_\beta$, let $\mathbb{F}^{(0)}$ denote the inertia field and $\mathbb{F}^{(1)}$ the first ramification field of $\mathfrak{P}$ over $P_\beta$. We have the inclusions

$$\overline{\mathbb{F}}_q(u) \subseteq \mathbb{F}^{(0)} \subseteq \mathbb{F}^{(1)} \subseteq \mathbb{T},$$

and in Galois correspondence to this

$$\overline{G} \supseteq G^{(0)} \supseteq G^{(1)} \supseteq \{\mathrm{id}_{\mathbb{T}}\},$$

where $G^{(0)} = \mathrm{Gal}(\mathbb{T}, \mathbb{F}^{(0)})$ is the inertia group and $G^{(1)} = \mathrm{Gal}(\mathbb{T}, \mathbb{F}^{(1)})$ the first ramification group of $\mathfrak{P}$ over $P_\beta$. ([Sti], III.8.)

Suppose $P_\beta$ is "tamely ramified", i.e. $p \nmid e_1, \ldots, e_r$. Then it is shown in [Coh98]: $F_u(x)$ factorises over $\mathbb{F}^{(0)}$ as $H_1(x) \cdots H_r(x)$, where the $H_i$ are co-prime irreducible polynomials with $\deg H_i = e_i$, and the cyclic group $G^{(0)}$ is generated by a product of disjoint cycles of lengths $e_1, \ldots, e_r$. Some examples illustrate how conclusions can be drawn from this:

(1) $p \neq 2$, $F_\beta$ simple $\Rightarrow$ $G^{(0)}$ is generated by a transposition.
(This proves Lemma 7 for odd $p$.)

(2) $p \neq 2$, $\deg f_0 - \deg f_1 = 2$, $f_1$ square-free $\Rightarrow$
$\Rightarrow$ same conclusion (now with ramification "at infinity": $\beta = \infty$).

(3) $p \neq 2$, $e_j = 2$ for exactly one $j \in \{1, \ldots, r\}$, $e_i$ odd and $p \nmid e_i$ for all $i \neq j$ $\Rightarrow$
$\Rightarrow$ $\sigma^\ell$ ($\sigma$ a generator of $G^{(0)}$, $\ell := \mathrm{lcm}(e_i \colon i \neq j)$) is a transposition.

(4) $p_0$ prime, $p \neq p_0$, $e_j = p_0$ for exactly one $j \in \{1, \ldots, r\}$,
$e_i$ not divisible by $p_0$ or $p$ for all $i \neq j$ $\Rightarrow$
$\Rightarrow$ as in the previous example, $G^{(0)}$ contains the $p_0$-cycle $\sigma^\ell$.

Now suppose $P_\beta$ is "fairly tamely", but not tamely ramified, i.e. $p^2 \nmid e_1, \ldots, e_r$ but $p | e_i$ for at least one $i \in \{1, \ldots, r\}$. Then, by [Coh98], $F_u = H_1 \cdots H_r$ as before; in addition, if $p \nmid e_i$, then $H_i$ splits completely over $\mathbb{F}^{(1)}$, and if $e_i = p\tilde{e}_i$, then $H_i$ factorises over $\mathbb{F}^{(1)}$ as a product of $\tilde{e}_i$ irreducible polynomials of degree $p$. Example of a possible conclusion:

(5) $e_j = p$ for exactly one $j \in \{1, \ldots, r\}$, $p \nmid e_i$ for all $i \neq j$ $\Rightarrow$
$\Rightarrow$ $F_u = H_1 \cdots H_r$, where $H_j$ is irreducible of degree $p$ and the $H_i$ for $i \neq j$ split completely over $\mathbb{F}^{(1)}$ $\Rightarrow$
$\Rightarrow$ $G^{(1)}$ is generated by a $p$-cycle.

With $p = 2$ and $F_\beta$ simple, this completes the proof of Lemma 7.

In certain situations it is impossible to establish condition (iii) of Lemma 7. For instance, if $p \geq 3$ and $f_0'' = f_1'' = 0$ identically, then $F_\beta$ cannot have a root of multiplicity exactly 2 for any $\beta$. This makes it necessary to have an alternative.

In the binary case, one option is to replace condition (iii) with

$$\deg f_0 - \deg f_1 = 2, \text{ and all factors of } f_1 \text{ have odd multiplicity.}$$

(This works by example (5) with $\beta = \infty$.) In [Coh97], the resulting modification of Lemma 7 was enough to settle the binary case completely. Though results where a transposition is obtained from $\deg f_0 - \deg f_1 = 2$ can be generalised to arbitrary prime characteristic, they are useless for our specific purpose when $p \geq 3$, because the structure of $F(x)$ (the degrees of the expressions in square brackets decrease in steps of $p$) causes the difference of degrees between $f_0$ and $f_1$ to be always a multiple of $p$.

With the ternary case in view, the next reasonable thing to do is to replace condition (iii) instead with

> $\deg f_0 - \deg f_1 = 3$, and all factors of $f_1$ have multiplicity not divisible by 3 or $p$.

This yields a 3-cycle in $\overline{G}$ (example (5) with $\beta = \infty$ for $p = 3$; example (4) with $p_0 = 3$ and $\beta = \infty$ for $p > 3$). If we further manage to show that $f_0/f_1$ is indecomposable over $\overline{\mathbb{F}}_q$, it follows that $\overline{G}$ is a primitive group with a 3-cycle in it, hence, by Theorem 13.3 of [Wie], it must contain the alternating group $\mathcal{A}_n$. This is almost what we want.

In order to fill the last gap, the **discriminant** proves to be the tool of choice. A well-known criterion says that, if $\mathrm{char}\ \mathbb{K} \neq 2$, then for a polynomial $g \in \mathbb{K}[x]$ of degree $n$ the Galois group $\mathrm{Gal}(g, \mathbb{K})$ is contained in the alternating group $\mathcal{A}_n$ if and only if the discriminant of $g$ is a square in $\mathbb{K}$. This can be helpful in two ways. Let $\Delta_x := \mathrm{discrim}(F_u(x), x)$.

- If $\Delta_x$ is a square in $\mathbb{F}_q(u)$, then $G$ is contained in $\mathcal{A}_r$, and equality follows from $\mathcal{A}_r \subseteq \overline{G} \subseteq G \subseteq \mathcal{A}_r$.
- If $\Delta_x$ is a non-square in $\overline{\mathbb{F}}_q(u)$, then $\overline{G}$ is not contained in $\mathcal{A}_r$. As it contains $\mathcal{A}_r$ itself, it can only be equal to $\mathcal{S}_r$, and equality follows from $\mathcal{S}_r = \overline{G} \subseteq G \subseteq \mathcal{S}_r$.

(Of course, neither situation may be the case.) We have proved:

**Lemma 8** *Let $p \geq 3$ (prime). Suppose that, in the situation of Theorem 7, the polynomial $F_u$ satisfies all of the following conditions:*

(i) *$f_0$ and $f_1$ are co-prime,*

(ii) *$f_0/f_1$ is indecomposable over $\overline{\mathbb{F}}_q$,*

(iii) $\deg f_0 - \deg f_1 = 3$ *and $f_1$ has no factor of multiplicity divisible by 3 or $p$,*

(iv) *$\Delta_x$ is either a non-square in $\overline{\mathbb{F}}_q(u)$ ("Situation S") or a square in $\mathbb{F}_q(u)$ ("Situation A").*

*Then* $\overline{G} = G = \begin{cases} \mathcal{S}_r & \text{in "Situation S",} \\ \mathcal{A}_r & \text{in "Situation A".} \end{cases}$

And this is the strategy for our application of Lemma 7 or 8, or other results of this type. Suppose $f_0, f_1$ depend on one parameter $v$, to be chosen in $\mathbb{F}_q$. We bound successively the number of $v \in \mathbb{F}_q$ for which the various conditions are *not* satisfied, assuming that the "bad" values of $v$ are discarded each time we pass from one condition to the next. The hope is that in the end the number of all such $v$ lies (comfortably) below the size of $q$ of at least $[r!\,(r-2)]^2$, i.e. a "good" $v$ is left—even if for practical reasons we count the "bad" $v$ in $\overline{\mathbb{F}}_q$ and replace the indecomposability condition over $\mathbb{F}_q$ with the stricter one over $\overline{\mathbb{F}}_q$.

This works similarly for several parameters $v_i$.

The next chapter will show this strategy in action. For the rest of the current chapter we concern ourselves with some preparations.

## 3.3 Co-primality is never a problem

As mentioned already, dealing with the first condition is always relatively easy.

**Lemma 9**

*Let $g(x) = \alpha_m x^m + \ldots + \alpha_1 x + \alpha_0$ be a polynomial in $\mathbb{K}[x]$ that is not identically zero, and let $h(x) = \beta_n x^n + \ldots + v x^k + \ldots + \beta_1 x + \beta_0 \in \mathbb{K}[x]$ with $k \in \{0, \ldots, n\}$. Further, if $k \geq 1$, suppose that $\alpha_0$ and $\beta_0$ are not both zero.*

*Then there are at most $m$ elements $v$ in $\mathbb{K}$ for which $g$ and $h$ are not co-prime.*

*Proof.* If one of $m, n$ is zero, the statement is trivially true. So assume $m, n \geq 1$. The polynomial $g$ has at most $m$ distinct roots in $\overline{\mathbb{K}}$. Let $\gamma$ be one of them, and assume it is also a root of $h$. Then either $k = 0$, or otherwise, by assumption, $\gamma \neq 0$. In any case, $h(\gamma) = 0$ implies

$$v = -\frac{1}{\gamma^k} \sum_{\substack{i=0 \\ i \neq k}}^{n} \beta_i \gamma^i ,$$

i.e. $\gamma$ determines a unique value $v$ in $\overline{\mathbb{K}}$. (For $\gamma = 0$ it is always $v = 0$, corresponding to $x$ as a common factor.) Hence there can only be as many as $m$ such values in either of $\overline{\mathbb{K}}$ and $\mathbb{K}$. □

(In the presence of several parameters $v_1, \ldots, v_k$, the argument can be used in the way that for every $(k-1)$-tuple of $v_i$ there exists only a bounded number of choices for the remaining parameter such that the two polynomials are not co-prime.)

So, provided $m < |\mathbb{K}|$, there exists always a choice of $v$ in $\mathbb{K}$ that makes $g$ and $h$ co-prime. Of course, if $g(x)$ and $h(x)$ have a power of $x$ in common, they are never co-prime for any $v$. In this case, let $x^l$ denote the highest such power with $l \leq k$, and put $g^*(x) := g(x)/x^l$ and $h^*(x) := h(x)/x^l$. Then Lemma 9 states that there are at most $m - l$ elements $v \in \mathbb{K}$ for which $g^*$ and $h^*$ are not co-prime.

It is clear how the lemma comes into play in our strategy when one of $f_0, f_1$ has non-zero constant term. Should both $f_0(x)$ and $f_1(x)$ be divisible by $x$, work with $f_0^*$ and $f_1^*$, and use Lemma 7 or Lemma 8 to establish that $F_u^* := f_0^* + uf_1^*$ splits; then so does $F_u(x) = x^l F_u^*(x)$.

## 3.4 About decomposition of rational functions

Our studies of indecomposability will rely heavily on the fact that, if $f_0/f_1$ is decomposable at all, one may assume a decomposition with certain "normalisation" properties. Though this has been used by various authors before, it is by no means obvious, and for anyone seeking a rigorous justification (which seems to be lacking in the literature) we include full details here. The proof of the following lemma may be skipped without damage for the understanding of the rest of the thesis.

**Lemma 10 (Normalised Decomposition)**

*Let $g, h \in \mathbb{K}[x]$ be monic polynomials with $\deg g > \deg h$. If $g/h$ is decomposable over $\mathbb{L}$, then there exists a non-trivial decomposition $(Q_1, Q_2, R_1, R_2)$ with the following properties:*

(i) $\omega_1 := \deg Q_1 > \deg Q_2 =: \omega_2$,
$\rho_1 := \deg R_1 > \deg R_2 =: \rho_2$;

(ii) $\omega_1 > 1$, $\rho_1 > 1$;

(iii) $Q_1, Q_2, R_1, R_2$ *are all monic.*

*Proof.* Let $(S_1, S_2, T_1, T_2)$ be a non-trivial decomposition of $f$ over $\mathbb{L}$. Then

$$\frac{g(x)}{h(x)} = \frac{S_1\left(\frac{T_1(x)}{T_2(x)}\right)}{S_2\left(\frac{T_1(x)}{T_2(x)}\right)} \,. \tag{3.1}$$

Write $\deg S_i =: \sigma_i$, $\deg T_i =: \tau_i$ $(i = 1, 2)$. Now multiply numerator and denominator on the right-hand side of (3.1) with $T_2^{\sigma_1+\sigma_2}(x)$ to obtain

$$\frac{g(x)}{h(x)} = \frac{T_2^{\sigma_2}(x) \cdot \left\{T_2^{\sigma_1}(x) \cdot S_1\left(\frac{T_1(x)}{T_2(x)}\right)\right\}}{T_2^{\sigma_1}(x) \cdot \left\{T_2^{\sigma_2}(x) \cdot S_2\left(\frac{T_1(x)}{T_2(x)}\right)\right\}} \,, \tag{3.2}$$

where both expressions in braces are in $\mathbb{L}[x]$. Put

$$d_i := \deg\left\{T_2^{\sigma_i} \cdot S_i\left(\frac{T_1}{T_2}\right)\right\} \qquad (i = 1, 2) \,.$$

From (3.2) and $\deg g > \deg h$ we deduce that (even if $g$ and $h$ are not co-prime)

$$\tau_2\sigma_2 + d_1 \;>\; \tau_2\sigma_1 + d_2 \,. \tag{3.3}$$

Write

$$S_i(x) =: \sum_{\nu=0}^{\sigma_i} s_{i,\nu} x^\nu \qquad (i = 1, 2)\,,$$

then

$$T_2^{\sigma_i}(x) \cdot S_i\left(\frac{T_1(x)}{T_2(x)}\right) = \sum_{\nu=0}^{\sigma_i} s_{i,\nu} T_1^\nu(x) T_2^{\sigma_i - \nu}(x)\,, \tag{3.4}$$

and the degree of the $\nu$-th summand of this is

$$\begin{cases} \tau_1\nu + \tau_2(\sigma_i - \nu)\,, & \text{if } s_{i,\nu} \neq 0\,, \\ -\infty\,, & \text{if } s_{i,\nu} = 0\,. \end{cases} \tag{3.5}$$

In a first step we now replace $(S_1, S_2, T_1, T_2)$ by another non-trivial decomposition $(\hat{S}_1, \hat{S}_2, \hat{T}_1, \hat{T}_2)$ which has property (i) of the lemma. Property (ii) follows then automatically from (i). We consider the cases $\tau_1 = \tau_2$, $\tau_1 > \tau_2$ and $\tau_1 < \tau_2$ separately.

$\boxed{\tau_1 = \tau_2 =: \tau.}$

In this case all non-zero summands in (3.4) have the same degree $\tau\sigma_i$, therefore $d_i \in \{\tau\sigma_i, -\infty\}$ for $i = 1, 2$. But neither of the $d_i$ can be $-\infty$, as this would mean $T_2 = 0$. On the other hand, with both $d_i$ equal to $\tau\sigma_i$, the inequality (3.3) becomes $\tau\sigma_2 + \tau\sigma_1 > \tau\sigma_1 + \tau\sigma_2$, an obvious contradiction. So this case cannot occur.

$\boxed{\tau_1 > \tau_2.}$

Here (3.5) takes its maximal value for $\nu = \sigma_i$ (note $s_{i,\sigma_i} \neq 0$), and since all other summands have strictly lower degree, it follows $d_i = \tau_1\sigma_i$. The inequality (3.3) becomes now $\tau_2\sigma_2 + \tau_1\sigma_1 > \tau_2\sigma_1 + \tau_1\sigma_2$, which is equivalent to $(\tau_1 - \tau_2)(\sigma_1 - \sigma_2) > 0$. Hence we must have $\sigma_1 > \sigma_2$, and we can simply take the original $(S_1, S_2, T_1, T_2)$ as $(\hat{S}_1, \hat{S}_2, \hat{T}_1, \hat{T}_2)$.

$\boxed{\tau_1 < \tau_2.}$

Let $\nu_i \in \{0, \ldots, \sigma_i\}$ be minimal with $s_{i,\nu_i} \neq 0$. Then (3.5) takes its maximum for exactly one index, namely $\nu = \nu_i$, and $d_i = (\tau_1 - \tau_2)\nu_i + \tau_2\sigma_i$, where $\tau_1 - \tau_2 < 0$. Inequality (3.3) becomes now $\tau_2\sigma_2 + \tau_2\sigma_1 + (\tau_1 - \tau_2)\nu_1 > \tau_2\sigma_1 + \tau_2\sigma_2 + (\tau_1 - \tau_2)\nu_2$, which shows that $\nu_1 < \nu_2$, i.e. the lowest power of $x$ occurring in $S_1$ is strictly smaller than the lowest power of $x$ occurring in $S_2$.
Now let $s := \max(\sigma_1, \sigma_2)$ and put $\hat{T}_1(x) := T_2(x)$, $\hat{T}_2(x) := T_1(x)$ and $\hat{S}_i(x) := x^s S_i(x^{-1})$. Then, with $\hat{S} := \hat{S}_1/\hat{S}_2$ and $\hat{T} := \hat{T}_1/\hat{T}_2$,

$$\hat{S}(\hat{T}(x)) = S(\hat{T}^{-1}(x)) = S(T(x)) = f(x)\,.$$

Clearly $\deg \hat{T}_1 > \deg \hat{T}_2$, $\gcd(\hat{T}_1, \hat{T}_2) = 1$, and $\deg \hat{S}_1 = s - \nu_1 > s - \nu_2 = \deg \hat{S}_2$. Moreover, $\gcd(\hat{S}_1, \hat{S}_2) = 1$ (otherwise there exists $\gamma$ in some extension of $\mathbb{L}$ such that $\gamma^s S_1(\gamma^{-1}) = \gamma^s S_2(\gamma^{-1}) = 0$, and since one of $\hat{S}_1, \hat{S}_2$ has non-zero constant term it follows that $\gamma \neq 0$ and $S_1(\gamma^{-1}) = S_2(\gamma^{-1}) = 0$, in contradiction to the

co-primality of $S_1$ and $S_2$). Therefore $(\hat{S}_1, \hat{S}_2, \hat{T}_1, \hat{T}_2)$ is a non-trivial decomposition of $f$ satisfying (i).

In a second and final step we make all polynomials monic without changing their degrees. First write $\hat{T}_1(x)/\hat{T}_2(x)$ as $c \cdot R_1(x)/R_2(x)$ with $R_1, R_2$ monic and $c \in \mathbb{L}^*$. For $i = 1, 2$ write

$$\hat{S}_i(x) \;=:\; \sum_{\nu=0}^{\hat{\sigma}_i} \hat{s}_{i,\nu} x^\nu$$

and define

$$\hat{\hat{S}}_i(x) \;:=\; \sum_{\nu=0}^{\hat{\sigma}_i} c^\nu \hat{s}_{i,\nu} x^\nu \,.$$

Then

$$\hat{S}_i\left(\frac{\hat{T}_1(x)}{\hat{T}_2(x)}\right) \;=\; \sum_{\nu=0}^{\hat{\sigma}_i} \hat{s}_{i,\nu} \left(\frac{\hat{T}_1(x)}{\hat{T}_2(x)}\right)^\nu \;=\; \sum_{\nu=0}^{\hat{\sigma}_i} \hat{s}_{i,\nu} c^\nu \left(\frac{R_1(x)}{R_2(x)}\right)^\nu \;=\; \hat{\hat{S}}_i\left(\frac{R_1(x)}{R_2(x)}\right),$$

so $(\hat{\hat{S}}_1, \hat{\hat{S}}_2, R_1, R_2)$, like $(\hat{S}_1, \hat{S}_2, \hat{T}_1, \hat{T}_2)$, is a non-trivial decomposition of $f$. Let $\mathrm{lc}(H)$ denote the leading coefficient of a polynomial $H$. Define

$$Q_i \;=\; \frac{\hat{\hat{S}}_i}{\mathrm{lc}(\hat{\hat{S}}_1)} \qquad \text{and} \qquad \omega_i := \hat{\sigma}_i = \deg Q_i \qquad (i = 1, 2).$$

Then $Q_1$ is monic, and $Q_1/Q_2 = \hat{\hat{S}}_1/\hat{\hat{S}}_2$, i.e. $(Q_1, Q_2, R_1, R_2)$ is a non-trivial decomposition of $f$. We see that $Q_2$ is also monic as follows. In

$$\frac{g(x)}{h(x)} \;=\; \frac{R_2^{\omega_1}(x) \cdot Q_1\left(\frac{R_1(x)}{R_2(x)}\right)}{R_2^{\omega_1-\omega_2}(x) \cdot \left\{R_2^{\omega_2}(x) \cdot Q_2\left(\frac{R_1(x)}{R_2(x)}\right)\right\}} \tag{3.6}$$

the denominator of the right-hand side must be monic because $g$, $h$ and the numerator of the right-hand side are. Hence

$$1 \;=\; \mathrm{lc}\left[R_2^{\omega_1-\omega_2} \cdot \left\{R_2^{\omega_2} \cdot Q_2\left(\tfrac{R_1}{R_2}\right)\right\}\right] \;=\; \mathrm{lc}(R_2)^{\omega_1-\omega_2} \cdot \mathrm{lc}(Q_2) \cdot \mathrm{lc}(R_1)^{\omega_2} \;=\; \mathrm{lc}(Q_2).$$

□

We show next that, at least over $\overline{\mathbb{K}}$, the normalisation procedure can be carried one step further, allowing us to make assumptions about one of the constant terms of the decomposition. This argument will, in several variations, be crucial in Chapter 4.

Assume a normalised decomposition $f(x) = Q(R(x))$ as in Lemma 10. Because $\omega_1 > 0$, $Q_1$ has a root $\gamma$ in $\overline{\mathbb{K}}$, and one can replace $Q(x)$ by $\hat{Q}(x) := Q(x + \gamma)$ and $R(x)$ by $\hat{R} := R(x) - \gamma$. On polynomial level this means (with the obvious notation)

$$\begin{aligned} \hat{Q}_1(x) &= Q_1(x+\gamma)\,, & \hat{R}_1(x) &= R_1(x) - \gamma \cdot R_2(x)\,, \\ \hat{Q}_2(x) &= Q_2(x+\gamma)\,, & \hat{R}_2(x) &= R_2(x)\,, \end{aligned}$$

and it is easy to see that $\deg \hat{Q}_i = \omega_i$, $\deg \hat{R}_i = \rho_i$ $(i = 1, 2)$ and $(\hat{Q}_1, \hat{Q}_2, \hat{R}_1, \hat{R}_2)$ is a normalised non-trivial decomposition of $f$ over $\overline{\mathbb{K}}$. ($\hat{Q}_1$ and $\hat{Q}_2$ must be co-prime, because if they had a common root $\beta \in \overline{\mathbb{K}}$, then $\beta + \gamma$ would be a common root of $Q_1$ and $Q_2$.) In addition, we have now $\hat{Q}_1(0) = Q_1(\gamma) = 0$. In other words, we may always assume that $Q_1(0) = 0$.

If $\omega_2 > 0$, we may apply the same argument with $\gamma$ being a root of $Q_2$ rather than $Q_1$, and therefore assume $Q_2(0) = 0$ instead.

Finally, we list a few general situations where indecomposability can be established easily.

**Lemma 11** *Let $g, h \in \mathbb{K}[x]$ be monic co-prime polynomials with $\deg g > \deg h$. Then $g/h$ is indecomposable over any extension of $\mathbb{K}$ if one of the following holds.*

(i) *$\deg g$ is prime.*

(ii) *$\deg h = 1$.*

(iii) *$h(x)$ is a power of $x$, and the greatest common divisor of all exponents of $x$ occurring in $g(x)$ and $h(x)$ is 1.*

*Proof.* Suppose $g/h$ is decomposable over some extension field of $\mathbb{K}$. Assume a decomposition as in Lemma 10. Then the right-hand side of (3.6) is in lowest terms. (Suppose not, then there exists $\gamma$ in some extension of $\mathbb{K}$ with

$$R_2^{\omega_1}(\gamma) \cdot Q_1(R(\gamma)) \;=\; R_2^{\omega_1-\omega_2}(\gamma) \cdot R_2^{\omega_2}(\gamma) \cdot Q_2(R(\gamma)) \;=\; 0\,.$$

If $R_2(\gamma) \neq 0$, then $R(\gamma)$ is a common root of $Q_1$ and $Q_2$, contradiction. If $R_2(\gamma) = 0$, then inspection of the expansion of $R_2^{\omega_1}(x) \cdot Q_1(R(x))$ reveals that also $R_1(\gamma) = 0$, again a contradiction.) Hence we can equate numerators and denominators in (3.6) to obtain

$$g \;=\; R_2^{\omega_1} \cdot Q_1\big(\tfrac{R_1}{R_2}\big)\,, \tag{3.7}$$

$$h \;=\; R_2^{\omega_1-\omega_2} \cdot \underbrace{\Big\{R_2^{\omega_2} \cdot Q_2\big(\tfrac{R_1}{R_2}\big)\Big\}}_{=:\,P,\ \text{a polynomial}}\,, \tag{3.8}$$

which gives the "degree equations"

$$\deg g \;=\; \omega_1\rho_1\,, \tag{3.9}$$

$$\deg h \;=\; (\omega_1 - \omega_2)\rho_2 + \omega_2\rho_1\,. \tag{3.10}$$

(i) Suppose $g/h$ is decomposable. Then, from (3.9), $\deg g$ is composite.

(ii) Suppose $g/h$ is decomposable. We show that $\deg h = 1$ is impossible. Assume that $\deg h < 2$. Now look at (3.10): as $\rho_1 \geq 2$, we must have $\omega_2 = 0$, but then $\omega_1 - \omega_2 \geq 2$ and therefore $\rho_2 = 0$, so that $\deg h = 0$ follows.

(iii) Suppose $g/h$ is decomposable and $h(x)$ is a power of $x$. Then both $R_2(x)$ and $P(x)$ are powers of $x$.

We show first that we must have $R_2(x) = 1$. Suppose that $R_2(x) = x^{\rho_2}$ with $\rho_2 > 0$. Then, because of co-primality, $R_1$ must have a non-zero constant term, say $c$.

Writing $Q_2(x) =: \sum_{i=0}^{\omega_2} q_{2,i}x^i$, we find

$$P(x) \;=\; R_1^{\omega_2}(x) \;+\; \sum_{i=0}^{\omega_2-1} q_{2,i}R_1^i(x)x^{\rho_2(\omega_2-i)}\,.$$

The constant term of $R_1^{\omega_2}(x)$ is $c^{\omega_2}$, and the rest of the sum is a multiple of $x$, therefore $P(x)$ has constant term $c^{\omega_2} \neq 0$, a contradiction.

So $R_2(x) = 1$, and consequently

$$P(x) \;=\; Q_2(R_1(x)) \;=\; \sum_{i=0}^{\omega_2} q_{2,i}R_1^i(x).$$

As a power of $x$, this must be equal to $x^{\omega_2\rho_1}$ (the highest power of $x$ occurs in the summand for $i = \omega_2$ but in no other summand). This means that $R_1(x) = x^{\rho_1}$, $h(x) = Q_2(x^{\rho_1})$ and $g(x) = Q_1(x^{\rho_1})$. Hence all exponents of $x$ in $g(x)$ and $h(x)$ must be multiples of $\rho_1 > 1$. □

We remark that criterion (iii) does not work if $h(x)$ is not a power of $x$. For a counterexample in $\mathbb{F}_3$ take $g(x) := x^6$ and $h(x) := x^5 + x^4 + x^3 + 1$; this decomposes non-trivially with $Q_1(x) = x^3$, $Q_2(x) = x^2 + 1$, $R_1(x) = x^2$ and $R_2(x) = x + 1$.

## 3.5 About the discriminant $\Delta_x$

In this section we prove a result that can help us deal with the discriminant condition in Lemma 8 without calculating $\Delta_x$ explicitly.

**Lemma 12** *In the situation of Theorem 7, assume that $f_0$ and $f_1$ are co-prime. Let $f := f_0/f_1$, and write the part of $E_1(x) := f_0'(x)f_1(x) - f_0(x)f_1'(x)$ that is co-prime to $f_1(x)$ as*

$$\prod_{\nu=1}^{s}(x-\beta_\nu)^{d_\nu} \qquad \textit{with distinct } \beta_\nu \in \overline{\mathbb{F}}_q\,.$$

*Then the discriminant $\Delta_x$ of $F_u(x)$ is equal to*

$$c\cdot\prod_{\nu=1}^{s}(u+f(\beta_\nu))^{d_\nu} \qquad \textit{with a constant } c \in \overline{\mathbb{F}}_q^*\,.$$

*Proof.*

(All necessary facts from field theory can be found summed up in [Sti], Appendix A.) For convenience, we write $\overline{\mathbb{F}}_q(u) =: \mathbb{G}$. Let the factorisation of $F_u$ over $\mathbb{T}$ (its splitting field over $\mathbb{G}$) be

$$F_u(x) \;=\; \prod_{i=1}^{r}(x-\alpha_i)\,.$$

We show first that $F_u(x)$ is irreducible over $\mathbb{G}$. Together with $F_u(x) \notin \mathbb{G}[x^p]$, this implies that $F_u$ is separable, i.e. the $\alpha_i$ are all distinct, $\mathbb{G}(\alpha_1) \cong \ldots \cong \mathbb{G}(\alpha_r)$ and $\mathbb{G}(\alpha_1)/\mathbb{G}$ is a Galois extension of degree $r$.

It suffices to show that $F_u(x)$ is irreducible in $\overline{\mathbb{F}}_q[u,x]$. Suppose $F_u = KL$ there. Since the degree of $F_u$ as a polynomial in $u$ is 1, we can assume

$$K \;=\; k_1u + k_0 \qquad \text{and} \qquad L \;=\; l_0 \qquad \text{with} \qquad k_1, k_0, l_0 \in \overline{\mathbb{F}}_q[x]\,.$$

Thus $F_u = k_1l_0u + k_0l_0$, so $f_0 = k_0l_0$ and $f_1 = k_1l_0$, and by co-primality of the two it follows that $l_0$ must be a constant in $\overline{\mathbb{F}}_q$.

The idea of this proof is to express $\Delta_x$ in terms of the norm $N_{\mathbb{G}(\alpha_1)/\mathbb{G}}$ of suitable elements, for which we can, by identifying their minimal polynomials, calculate the norm (to some extent). For a start,

$$\begin{aligned}
\Delta_x &\overset{\text{def}}{=} \prod_{1\le i<j\le r}(\alpha_i-\alpha_j)^2 \overset{(1)}{=} (-1)^{\lfloor\frac{r}{2}\rfloor}\cdot\prod_{1\le i\ne j\le r}(\alpha_i-\alpha_j) \;= \\
&= (-1)^{\lfloor\frac{r}{2}\rfloor}\cdot\prod_{i=1}^{r}\Big(\prod_{\substack{j=1\\ j\ne i}}^{r}(\alpha_i-\alpha_j)\Big) \overset{(2)}{=} \\
&= (-1)^{\lfloor\frac{r}{2}\rfloor}\cdot\prod_{i=1}^{r}F_u'(\alpha_i) \overset{(3)}{=} (-1)^{\lfloor\frac{r}{2}\rfloor}\cdot N_{\mathbb{G}(\alpha_1)/\mathbb{G}}(F_u'(\alpha_1))\,,
\end{aligned}$$

where the steps (1), (2) and (3) are justified as follows:

(1) For each pair $(i,j) \in \{1,\ldots,r\}^2$ with $j > i$ we get a change of sign as $(\alpha_i-\alpha_j)$ is replaced by $(\alpha_j-\alpha_i)$. The number of such pairs is $(r-1)+(r-2)+\ldots+1$; by an obvious induction this is odd if $r \equiv 2$ or $3 \pmod 4$ and even if $r \equiv 0$ or $1 \pmod 4$. Hence the factor $(-1)^{\lfloor\frac{r}{2}\rfloor}$.
(Actually, all we need is that the product on the LHS is a constant multiple of the product on the RHS.)

(2) $F_u(x) \;=\; (x-\alpha_i)\cdot\prod_{\substack{j=1\\ j\ne i}}^{r}(x-\alpha_j) \quad \Rightarrow$

$$\Rightarrow \quad F_u'(x) \;=\; 1\cdot\prod_{\substack{j=1\\ j\ne i}}^{r}(x-\alpha_j) \;+\; (x-\alpha_i)\cdot\frac{d}{dx}\prod_{\substack{j=1\\ j\ne i}}^{r}(x-\alpha_j) \quad \Rightarrow$$

$$\Rightarrow \quad F_u'(\alpha_i) \;=\; \prod_{\substack{j=1\\ j\ne i}}^{r}(\alpha_i-\alpha_j)\,.$$

(3) Let $\{\sigma_1, \ldots, \sigma_r\}$ be the Galois group of $\mathbb{G}(\alpha_1)$ over $\mathbb{G}$. By definition,

$$N_{\mathbb{G}(\alpha_1)/\mathbb{G}}(F_u'(\alpha_1)) \;=\; \prod_{i=1}^{r} \sigma_i\big(F_u'(\alpha_1)\big)\;.$$

So it suffices to show that the $F_u'(\alpha_i)$ are the conjugates of $F_u'(\alpha_1)$. But for $k = 1, \ldots, r$,

$$\sigma_k\big(F_u'(\alpha_1)\big) \;=\; \sigma_k\Big(\prod_{j=2}^{r}(\alpha_1 - \alpha_j)\Big) \;=\; \prod_{j=2}^{r}\big(\sigma_k(\alpha_1) - \sigma_k(\alpha_j)\big) \;=\; F_u'\big(\sigma_k(\alpha_1)\big)\;,$$

which equals $F_u'(\alpha_i)$ for some $i$.

Now write the factorisation of $E_1$ over $\overline{\mathbb{F}}_q$ as

$$E_1(x) \;=\; \prod_{\nu=1}^{\tilde{s}}(x - \beta_\nu)^{d_\nu} \qquad (\beta_\nu \in \overline{\mathbb{F}}_q \text{ dictinct})\,,$$

where, as assumed, the $\beta_\nu$ are numbered in such a way that

$$\begin{aligned} f_1(\beta_\nu) &\neq 0 \quad \text{for } \nu = 1, \ldots, s\,; \\ f_1(\beta_\nu) &= 0 \quad \text{for } \nu = s+1, \ldots, \tilde{s}\,. \end{aligned}$$

From $F_u(\alpha_1) = 0$ it follows that $u = -f_0(\alpha_1)/f_1(\alpha_1)$. Thus we can express the field element $F_u'(\alpha_1) \in \mathbb{G}(\alpha_1)$ as

$$F_u'(\alpha_1) \;=\; f_0'(\alpha_1) - \frac{f_0(\alpha_1)}{f_1(\alpha_1)} f_1'(\alpha_1) \;=\; \frac{E_1(\alpha_1)}{f_1(\alpha_1)} \;=\; \frac{1}{f_1(\alpha_1)} \prod_{\nu=1}^{\tilde{s}}(\alpha_1 - \beta_\nu)^{d_\nu}\;.$$

For $\nu = 1, \ldots, s$ write short $f(\beta_\nu) =: u_\nu$ and define $F_{u,\nu}(x) := F_u(x + \beta_\nu)$. The latter is a monic irreducible polynomial in $\mathbb{G}[x]$ which has $\alpha_1 - \beta_\nu$ as a root, hence it is the minimal polynomial of $\alpha_1 - \beta_\nu$ over $\mathbb{G}$. Its constant term is

$$\begin{aligned} F_{u,\nu}(0) \;&=\; F_u(\beta_\nu) \;=\; f_0(\beta_\nu) + u f_1(\beta_\nu) \;= \\ &=\; f_0(\beta_\nu) - u_\nu f_1(\beta_\nu) + u f_1(\beta_\nu) + u_\nu f_1(\beta_\nu) \;=\; (u + u_\nu) f_1(\beta_\nu)\,. \end{aligned}$$

By [Sti], p. 240, (5), it follows that

$$N_{\mathbb{G}(\alpha_1)/\mathbb{G}}(\alpha_1 - \beta_\nu) \;=\; (-1)^r \cdot (u + u_\nu) f_1(\beta_\nu)\,.$$

We proceed similarly with the remaining parts of $E_1/f_1$. Suppose $\gamma$ is a root of $f_1$ (one of the $\beta_\nu$ with $\nu \geq s+1$ or another root of $f_1$). Define $\hat{F}_u(x) := F_u(x+\gamma)$, then $\hat{F}_u$ is a monic irreducible polynomial in $\mathbb{G}[x]$ with $\alpha_1 - \gamma$ as a root and constant term

$$\hat{F}_u(0) \;=\; F_u(\gamma) \;=\; f_0(\gamma) + u f_1(\gamma) \;=\; f_0(\gamma)\,,$$

so that, by the same argument as above,

$$N_{\mathbb{G}(\alpha_1)/\mathbb{G}}(\alpha_1 - \gamma) \;=\; (-1)^r f_0(\gamma) \;\in \overline{\mathbb{F}}_q^*\,.$$

By multiplicativity of norm, the norm of $f_1(\alpha_1)$ is an element of $\overline{\mathbb{F}}_q^*$, and further

$$\begin{aligned}\Delta_x &= (-1)^{\lfloor \frac{r}{2} \rfloor} \cdot N_{\mathbb{G}(\alpha_1)/\mathbb{G}}(F'_u(\alpha_1)) = (-1)^{\lfloor \frac{r}{2} \rfloor} \cdot N_{\mathbb{G}(\alpha_1)/\mathbb{G}} \left( \frac{E_1(\alpha_1)}{f_1(\alpha_1)} \right) = \\ &= (-1)^{\lfloor \frac{r}{2} \rfloor} \cdot \frac{\prod_{\nu=1}^{s} \left( N_{\mathbb{G}(\alpha_1)/\mathbb{G}}(\alpha_1 - \beta_\nu) \right)^{d_\nu} \cdot [\text{constant in } \overline{\mathbb{F}}_q^*]}{[\text{constant in } \overline{\mathbb{F}}_q^*]} = \\ &= [\text{constant in } \overline{\mathbb{F}}_q^*] \cdot \prod_{\nu=1}^{s} (u + u_\nu)^{d_\nu} ,\end{aligned}$$

as claimed. □

## 3.6 Another preliminary lemma

We conclude this chapter with a little fact which is rather evident, but which will be convenient to have ready for reference later. If one would like to give it a name, "Stairs Lemma" might be appropriate.

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

$$A(x) := \sum_{i=0}^{n} \alpha_i x^i \in \mathbb{K}[x] \qquad \text{and} \qquad B(x) := \sum_{i=0}^{m} \alpha_{i+n-m} x^i \in \mathbb{K}[x] ,$$

i.e. $B$ uses the same coefficients as $A$, shifted by $n - m$ towards the smaller powers; the last $n - m$ coefficients are lost.

**Lemma 13**

*Let $w \in \mathbb{K}$. If $A(x)$ is not identically zero, then neither is $A(x) - wB(x)$.*

*Proof (semi-visual).*

$$\begin{array}{rcl} A(x) &=& \underbrace{\alpha_n x^n + \ldots + \alpha_{m+1} x^{m+1}}_{\text{first } n-m \text{ terms}} + \underbrace{\alpha_m x^m + \ldots + \alpha_{2m-n+1} x^{2m-n+1}}_{\text{second } n-m \text{ terms}} + \ldots \\ wB(x) &=& \qquad\qquad\qquad\qquad\quad w\alpha_n x^m + \ldots + w\alpha_{m+1} x^{2m-n+1} + \ldots \end{array}$$

For $w = 0$ the statement is trivial, so assume $w \in \mathbb{K}^*$. Suppose $A - wB = 0$. Then the first $n - m$ coefficients of $A$ must be zero. Therefore the next $n - m$ coefficients of $A$ must also be zero, etc.—repeat this argument, until after a finite number of steps all coefficients of $A$ have been seen to be zero, in contradiction to $A \neq 0$. □

# Chapter 4

# The Primitive Ternary Case with Designed Distance $\equiv 2 \pmod 3$

Our ambition in this chapter is to use the approach developed so far to identify as many primitive ternary BCH codes as possible with designed distance $\delta \equiv 2 \pmod 3$ which have covering radius $\delta - 1$ (the smallest possible).

In other words, we take $p = 3$, $q = 3^m$ ($m \in \mathbb{N}$) and $r = \delta - 1 = 3t + M$, $M = 1$, in Theorem 6, and we want to find conditions on $q$ under which for every choice of $A_k \in \mathbb{F}_q$ ($k = 2, \ldots, \delta - 1$; $3 \nmid k$), assuming $A_1 = 0$, we can find $s_3, s_6, \ldots, s_{3t}$ in $\mathbb{F}_q$ such that the polynomial $F(x)$ splits completely over $\mathbb{F}_q$. The explicit forms of $F(x)$ and the relevant $g$-polynomials have already been given in Section 2.6.

## 4.1 The central result and set-up of its proof

**Theorem 8** *Let $t \in \mathbb{N}$.*

*The covering radius of the primitive ternary BCH code of length $q - 1$ ($q = 3^m$) and designed distance $\delta = 3t + 2$ is exactly $3t + 1$ whenever $t \equiv 0 \pmod 4$ and $q > [(3t + 1)!\,(3t - 1)]^2$.*

*The same holds for $t \equiv 2 \pmod 4$, provided $m$ is even.*

This adds also to the knowledge about the minimal distance of BCH codes.

**Corollary 1**

*Let $t$, $q$ and $m$ be as in Theorem 8. Then the actual minimal distance of the primitive ternary BCH code of length $q - 1$ and designed distance $3t + 1$ is exactly $3t + 1$.*

*Proof.* Supercode Lemma: $3t + 1 = \rho(\mathcal{C}_{3,q-1,3t+2}) \geq d(\mathcal{C}_{3,q-1,3t+1}) \geq 3t + 1$. □

The proof of Theorem 8 will be given in Sections 4.2–4.4, each covering one of three subcases.

Note that all information about the coefficients $A_k$ is contained in the two polynomials

$$\begin{aligned} g_{t-1}^{(0)}(x^3) &= A_4x^{3t-3} + A_7x^{3t-6} + \ldots + A_{3t-2}x^3 + A_{3t+1}\,, \\ g_{t-1}^{(2)}(x^3) &= A_2x^{3t-3} + A_5x^{3t-6} + \ldots + A_{3t-4}x^3 + A_{3t-1}\,. \end{aligned}$$

We can assume that at least one of the coefficients is non-zero (otherwise simply put $s_3 = s_6 = \ldots = s_{3t} = 0$ to obtain $F(x) = x^r$).

If $A_2 = A_4 = \ldots = A_{3t-4} = A_{3t-2} = 0$, we call this the **Degenerate Case**.

Otherwise, if we are not in the Degenerate Case, let $j \in \{1, \ldots, t-1\}$ be the smallest integer such that $A_{3j-1}$ and $A_{3j+1}$ are not both zero. Put $A_{3j-1} =: C_2$ and $A_{3j+1} =: C_0$ (i.e. $C_i$ is the coefficient that occurs in the $g^{(i)}$-polynomials). With this the $g$-polynomials can be written in the shortened forms

$$\begin{aligned} g_{t-1-\nu}^{(0)}(x^3) &= \sum_{i=0}^{t-\nu-j} A_{3(t-\nu-i)+1}x^{3i} = \\ &= C_0x^{3(t-\nu-j)} + A_{3j+4}x^{3(t-\nu-j-1)} + \ldots + A_{3(t-\nu)-2}x^3 + A_{3(t-\nu)+1}\,, \\ g_{t-1-\nu}^{(2)}(x^3) &= \sum_{i=0}^{t-\nu-j} A_{3(t-\nu-i)-1}x^{3i} = \\ &= C_2x^{3(t-\nu-j)} + A_{3j+2}x^{3(t-\nu-j-1)} + \ldots + A_{3(t-\nu)-4}x^3 + A_{3(t-\nu)-1}\,. \end{aligned}$$

The situation where $C_2g_{t-1}^{(0)} = C_0g_{t-1}^{(2)}$, i.e. one of the two polynomials is a constant (possibly zero) multiple of the other, requires separate treatment, and we label this the **Exceptional Case**. Consequently, by the **Standard Case** we mean the situation where $C_2g_{t-1}^{(0)} \neq C_0g_{t-1}^{(2)}$.

## 4.2 The "Standard Case": $C_2g_{t-1}^{(0)} \neq C_0g_{t-1}^{(2)}$

Note that under the condition for the Standard Case neither of $g_{t-1}^{(0)}$, $g_{t-1}^{(2)}$ can be the zero polynomial.

To start with, put

$$w := \begin{cases} 1, & \text{if } C_0 = 0 = A_{3t+1}\,, \\ 0, & \text{otherwise.} \end{cases}$$

(Instead of $w = 1$ we could take any non-zero element in $\mathbb{F}_q$.) The $w$ will serve as a "switch" to ensure that one of $f_0, f_1$ has a non-vanishing constant term. This avoids the need for further case distinctions when it comes to establishing co-primality.

In the case $w = 1$, choose $l \in \{j+1, \ldots, t-1\}$ with $A_{3l+1} \neq 0$ (exists because

$g^{(0)}_{t-1} \neq 0$). It is helpful to have ready that

$$\begin{aligned} g^{(0)}_{l-1}(x^3) &= \sum_{i=0}^{l-j} A_{3(l-i)+1} x^{3i} = \\ &= C_0 x^{3(l-j)} + A_{3j+4} x^{3(l-j-1)} + \ldots + A_{3l-2} x^3 + A_{3l+1}\,, \\ g^{(2)}_{l-1}(x^3) &= \sum_{i=0}^{l-j} A_{3(l-i)-1} x^{3i} = \\ &= C_2 x^{3(l-j)} + A_{3j+2} x^{3(l-j-1)} + \ldots + A_{3l-4} x^3 + A_{3l-1}\,. \end{aligned}$$

If $w = 0$, all terms involving $l$ will be "switched off", so that in this case it is not necessary to define $l$ at all.

Now choose the $s_{3\nu}$ as follows (all others are understood to be zero):

$$s_{3(t-j)} = u\,, \qquad s_{3(t-l)} = w \ \text{(ignore this if } w = 0)\,, \qquad s_{3t} = v\,.$$

Then $F(x)$ becomes $F_u(x) = f_0(x) + u f_1(x)$ with

$$\begin{aligned} f_0(x) &= x^{3t+1} + x^2 g^{(2)}_{t-1}(x^3) + g^{(0)}_{t-1}(x^3) + w x^{3l+1} + w x^2 g^{(2)}_{l-1}(x^3) + w g^{(0)}_{l-1}(x^3) + vx\,, \\ f_1(x) &= x^{3j+1} + C_2 x^2 + C_0\,. \end{aligned}$$

The $u$ is the indeterminate of Theorem 7, the $v$ is a (single) parameter via which we can influence the properties of $f_0$. We proceed to establish the existence of a choice of $v$ in $\mathbb{F}_q$ such that the conditions of Lemma 7 are satisfied.

## Co-primality

Here the $w$ comes into effect. We have $w = 1$ exactly if $C_0$ and $A_{3t+1}$ are both zero; in this case the constant term of $f_0$ is $A_{3l+1} \neq 0$. Otherwise, if $w = 0$, the constant terms of $f_0$ and $f_1$ are $A_{3t+1}$ and $C_0$, at least one of which is then non-zero.

In each case Lemma 9 applies immediately and shows that there are no more than $\deg f_1 = 3j + 1 \leq 3t - 2$ values $v$ in $\mathbb{F}_q$ for which $f_0$ and $f_1$ are not co-prime. We exclude these from further consideration.

## Indecomposability

Suppose $f_0/f_1 =: f$ is decomposable over $\overline{\mathbb{F}}_q$ with decomposition and notation as in Lemma 10.

We consider first the case $\omega_2 > 0$. To this we can apply the argument described after the proof of Lemma 10 on p. 30, which allows to assume that $Q_2(0) = 0$.

As in (3.8), $f_1 = R_2^{\omega_1 - \omega_2} \cdot \{R_2^{\omega_2} \cdot Q_2(R_1/R_2)\}$. This is a polynomial that does not depend on the choice of $v$, and since $\omega_1 > \omega_2$ it must have $R_2$ as a divisor. Taking all

possible subsets of linear factors of $f_1$, this leaves at most $2^{\deg f_1}$ possibilities for $R_2$. Now $Q_2(0) = 0$ implies that $R_1$ also divides $f_1$, so $R_1$ must also be one out of at most $2^{\deg f_1}$ polynomials. Together there are at most $2^{2\cdot \deg f_1}$ possibilities for $R$. (In fact much fewer, as we have for example not taken into account $\rho_1 > \rho_2$. But there is no need to keep numbers down as long as we arrive at the desired existence result in the end, so we will in general make no effort to do so.)

Now fix one $R$ and suppose it occurs in decompositions of $f_0/f_1$ for two distinct values $v$ and $\tilde{v}$. We introduce the notation $\langle\!\langle\, . \,\rangle\!\rangle$ for the fixed part of a polynomial, i.e. all terms that depend only on the $A_k$. With this, cf. (3.7),

$$R_2^{\omega_1}(x) \cdot Q_1\left(\tfrac{R_1(x)}{R_2(x)}\right) \;=\; \langle\!\langle f_0(x) \rangle\!\rangle \;+\; vx$$

and at the same time

$$R_2^{\omega_1}(x) \cdot \tilde{Q}_1\left(\tfrac{R_1(x)}{R_2(x)}\right) \;=\; \langle\!\langle f_0(x) \rangle\!\rangle \;+\; \tilde{v}x \;\;.$$

Subtraction yields

$$R_2^{\omega_1}(x) \cdot (Q_1 - \tilde{Q}_1)\left(\tfrac{R_1(x)}{R_2(x)}\right) \;=\; (v - \tilde{v})x\,. \tag{4.1}$$

We show that this is impossible. With $q_{1,\mu}$ denoting the coefficient of $x^\mu$ in $Q_1(x)$, and similarly $\tilde{q}_{1,\mu}$ for $\tilde{Q}_1(x)$, the left-hand side of (4.1) reads

$$\sum_{\mu=0}^{\omega_1} (q_{1,\mu} - \tilde{q}_{1,\mu}) R_1^\mu(x) R_2^{\omega_1-\mu}(x)\,.$$

In this, the degree of a non-vanishing summand for some index $\mu$ is $\mu\rho_1 + (\omega_1 - \mu)\rho_2$. This decreases strictly with $\mu$. Hence, if $\mu_0$ is the maximal $\mu$ for which the $\mu$-th summand is non-zero, the degree of the LHS of (4.1) is given by $\mu_0\rho_1 + (\omega_1 - \mu_0)\rho_2$. This has to be equal to 1, but for $\mu_0 \in \{1, \ldots, \omega_1\}$ it is greater than 1, and $\mu_0 = 0$ is also impossible because $\rho_2\omega_1 \neq 1$.

Hence every $R$ occurs for at most one $v$, and consequently the number of $v$ in $\mathbb{F}_q$ allowing a decomposition of $f$ with $\omega_2 > 0$ is limited by $2^{2\cdot \deg f_1} = 2^{6j+2} \leq 2^{6t-4}$.

Secondly, we have to consider decompositions with $\omega_2 = 0$, which are not covered by the "$Q_2(0) = 0$ trick". Here the idea is to use formal differentiation and exploit the characteristic.

If $\omega_2 = 0$ then $Q_2 = 1$ and $f_1 = R_2^{\omega_1}$, i.e.

$$x^{3j+1} + C_2x^2 + C_0 \;=\; R_2^{\omega_1}(x)\,. \tag{4.2}$$

Since $\omega_1\rho_1 = \deg f_0 = 3t + 1$, we have $3 \nmid \omega_1$. Hence taking first derivatives yields

$$x^{3j} - C_2x \;=\; \omega_1 R_2^{\omega_1-1}(x) \cdot R_2'(x)\,.$$

Multiplying both sides of this with $x$ and subtracting from (4.2) gives

$$-C_2x^2 + C_0 \;=\; R_2^{\omega_1-1}(x)\,[R_2(x) - \omega_1 x R_2'(x)]\,.$$

Now equating degrees gives $2 = \rho_2(\omega_1 - 1) + d$, where $d := \deg\,[R_2(x) - \omega_1 x R_2'(x)]$. Since $\omega_1 > 1$, $\rho_2 \neq 0$ and $3 \nmid \omega_1$, the only possibilities for this are

$$\begin{array}{ll} \text{(i)} & \rho_2 = 2,\ \ \omega_1 = 2,\ \ d = 0\,, \\ \text{(ii)} & \rho_2 = 1,\ \ \omega_1 = 2,\ \ d = 1\,. \end{array}$$

Possibility (ii) can be ruled out immediately, because here $\rho_2\omega_1 = 2$ while we have $\rho_2\omega_1 = \deg f_1 = 3j + 1 \geq 4$. With (i), the degree equation for $f_0$ becomes $3t + 1 = 2\rho_1$, which is also impossible, since $t$ is always even in Theorem 8.

Hence decompositions with $\omega_2 = 0$ do not occur at all, and the number of $v$ we must exclude for decomposability reasons remains bounded by $2^{6t-4}$.

We digress for a moment and show that indecomposability can still be shown if we consider also odd $t \geq 3$. This will come in useful later.

All that needs to be modified is the argument for (i) above. In this situation we have $C_0 = C_2^2$ and $R_2(x) = x^2 - C_2$. Suppose $Q(x) = x^2 + \alpha x + \beta$. Replace $Q(x)$ by $\hat{Q}(x) := Q(x + \alpha)$ and $R(x)$ by $\hat{R}(x) := R(x) - \alpha$, then, as on p. 30, $(\hat{Q}_1, \hat{Q}_2, \hat{R}_1, \hat{R}_2)$ is again a normalised non-trivial decomposition of $f$ with the same degrees, now with the additional property that $\hat{Q}$ has vanishing linear term. We drop the "hats" and simply assume that $Q(x)$ is of the form $x^2 + A$. Then

$$f_0(x) \;=\; R_1^2(x) + A(x^2 - C_2)^2\,. \tag{4.3}$$

This may be possible for some $v$. But assume it happens also for $\tilde{v} \neq v$. Then going through the above process again (note that all of $f_0$, $Q$, $R_1$ and $A$ depend on $v$, but $R_2$ does not) yields

$$\tilde{f}_0(x) \;=\; \tilde{R}_1^2(x) + \tilde{A}(x^2 - C_2)^2\,. \tag{4.$\tilde{3}$}$$

Subtract (4.$\tilde{3}$) from (4.3) to obtain

$$(v - \tilde{v})x \;=\; R_1^2(x) - \tilde{R}_1^2(x) + (A - \tilde{A})(x^2 - C_2)^2\,,$$

or

$$(v - \tilde{v})x - (A - \tilde{A})(x^2 - C_2)^2 \;=\; (R_1(x) + \tilde{R}_1(x))(R_1(x) - \tilde{R}_1(x))\,. \tag{4.4}$$

Both $R_1$ and $\tilde{R}_1$ are monic of degree $\rho_1 = \frac{3t+1}{2}$, so $\deg\,(R_1 + \tilde{R}_1) = \frac{3t+1}{2} \geq 5$. Clearly, (4.4) cannot hold with $R_1 = \tilde{R}_1$, therefore the degree of the right-hand side is at least 5. But the degree of the left-hand side is at most 4, a contradiction.

Consequently, at most one more $v \in \mathbb{F}_q$ needs to be excluded to ensure indecomposability also for odd $t \geq 3$.

## Simplicity—Step 1

Making sure that there is always some $\beta \in \overline{\mathbb{F}}_q$ for which $F_\beta$ is simple is more difficult. We do this in three steps. The first is to show that for all but a limited number of $v \in \overline{\mathbb{F}}_q$ there exists $\beta \in \overline{\mathbb{F}}_q$ such that $F_\beta$ has *at least one repeated factor*.

Suppose $v$ is such that $F_\beta$ is square-free for all $\beta \in \overline{\mathbb{F}}_q$. Then the system

$$\begin{aligned} F_u(x) &= f_0(x) + uf_1(x) &&= 0 \\ F'_u(x) &= f'_0(x) + uf'_1(x) &&= 0 \end{aligned} \tag{4.5}$$

has no solution $(u, x) = (\beta, \gamma)$ in $\overline{\mathbb{F}}_q^2$.

Now consider $E_1(x) := f'_0(x)f_1(x) - f_0(x)f'_1(x)$. A root $\gamma$ of $E_1$ with $f'_1(\gamma) \neq 0$ would imply a solution to (4.5) by putting $\beta := -f'_0(\gamma)/f'_1(\gamma)$. Therefore every root of $E_1$ must also be a root of the fixed polynomial $f'_1(x) = x^{3j} - C_2x$, and elementary combinatorics show that the number of possibilities for $E_1$ is at most

$$\sum_{i=1}^{\deg E_1} \binom{\deg f'_1}{i} \binom{(\deg E_1) - 1}{i-1} \;<\; \binom{\deg E_1 + \deg f'_1}{\deg f'_1}.$$

One finds that $E_1(x)$ equals

$$\begin{aligned} &x^{3j}\big(x^2 g^{(2)}_{t-1}(x^3) - g^{(0)}_{t-1}(x^3)\big) \;-\; x^{3t}(C_2x^2 - C_0) \;+\; x\big(C_2 g^{(0)}_{t-1}(x^3) - C_0 g^{(2)}_{t-1}(x^3)\big) \;+ \\ &+\; w\Big[x^{3j}\big(x^2 g^{(2)}_{l-1}(x^3) - g^{(0)}_{l-1}(x^3)\big) \;-\; x^{3l}(C_2x^2 - C_0) \;+\; x\big(C_2 g^{(0)}_{l-1}(x^3) - C_0 g^{(2)}_{l-1}(x^3)\big)\Big] \;- \\ &-\; v(C_2x^2 - C_0)\,, \end{aligned}$$

which turns out to be of degree at most $3t - 1$ (and at least 1).

Thus, since every choice of $E_1$ determines $v$, the maximal number of values $v$ that have to be excluded in this step is

$$\binom{3t - 1 + 3j}{3j} \overset{\text{(easy)}}{\leq} \binom{6t-4}{3t-3},$$

which is still by a factor of at least $(3t+1)!\,(3t-1)^4$ smaller than $[(3t+1)!\,(3t-1)]^2$.

## Simplicity—Step 2

Let $v$ be an element in $\overline{\mathbb{F}}_q$ that has passed through the selection process so far, and let $\beta$ be any element in $\overline{\mathbb{F}}_q$ for which $F_\beta$ has a repeated root (at least one such $\beta$ exists by Step 1). Now assume that $F_\beta$ has a factor of multiplicity $\geq 3$. We show that this can happen only for a relatively small number of $v$.

Let $\gamma$ be a triple of higher root of $F_\beta$, i.e. $F_\beta(\gamma) = F'_\beta(\gamma) = F''_\beta(\gamma) = 0$. From $F_\beta(\gamma) = 0$ it follows that $f_1(\gamma) \neq 0$ (by co-primality of $f_0, f_1$) and $\beta = -f_0(\gamma)/f_1(\gamma)$. Hence

$$0 \;=\; F'_\beta(\gamma) \;=\; f'_0(\gamma) - \frac{f_0(\gamma)}{f_1(\gamma)} f'_1(\gamma)\,,$$

i.e. $\gamma$ is a root of $E_1$ (defined as in Step 1). In the same way, $\gamma$ is also a root of

$$\begin{aligned} E_2(x) \;:=\; & f''_0(x)f_1(x) - f_0(x)f''_1(x) \;= \\ = \;& -x^{3j+1} g^{(2)}_{t-1}(x^3) \;-\; C_0 g^{(2)}_{t-1}(x^3) \;+\; C_2 x^{3t+1} \;+\; C_2 g^{(0)}_{t-1}(x^3) \;+ \\ &+\; w\Big[-x^{3j+1} g^{(2)}_{l-1}(x^3) \;-\; C_0 g^{(2)}_{l-1}(x^3) \;+\; C_2 x^{3l+1} \;+\; C_2 g^{(0)}_{l-1}(x^3)\Big] \;+ \\ &+\; vC_2x\,. \end{aligned}$$

Now consider first the case $C_2 = 0$. Here $E_2(x)$ simplifies to

$$-(x^{3j+1} + C_0)\Big( g_{t-1}^{(2)}(x^3) \; - \; w g_{l-1}^{(2)}(x^3)\Big)\,,$$

which is a fixed polynomial of degree $\leq 3t-2$, not identically zero by Lemma 13. Hence there are at most $3t-2$ possible roots $\gamma$. By $E_1(\gamma) = 0$, each $\gamma$ determines a unique $v \in \overline{\mathbb{F}}_q$.

Next assume $C_2 \neq 0$. Observe that $F'_\beta(0) = v$, hence we can ensure that $\gamma \neq 0$ by excluding $v = 0$. Solving $E_2(x) = 0$ for $v$ and substituting in $E_1(x) = 0$ yields

$$0 \;=\; \frac{f_1(x)}{C_2 x} \cdot \Big[ (C_0 g_{t-1}^{(2)}(x^3) - C_2 g_{t-1}^{(0)}(x^3)) \; + \; w\Big(C_0 g_{l-1}^{(2)}(x^3) - C_2 g_{l-1}^{(0)}(x^3)\Big)\Big]\,.$$

Since $C_2\gamma \neq 0$, we conclude that $\gamma$ must be a root of the polynomial in square brackets, which is fixed of degree $\leq 3t-6$ and not identically zero (using the assumption for the Standard Case and Lemma 13). It is again clear, this time from $E_2(\gamma) = 0$, that $\gamma$ determines a unique $v$.

In each case we count no more than $3t-2$ values of $v$ which must be discarded. For all remaining $v \in \overline{\mathbb{F}}_q$, and all $\beta \in \overline{\mathbb{F}}_q$, all repeated factors have exact multiplicity 2.

It remains to show that in this last situation the polynomial $F_\beta$ has *at most one* repeated factor for at least one choice of $v$ and $\beta$. Again, we count the number of $v$ for which it is possible to have several repeated factors.

This step is less straightforward than the previous two. We show first that it boils down to another decomposition problem for rational functions. This part follows the idea of "Proof of (D)" on pp. 341/342 in [Coh98].

## Simplicity—Step 3: reduction

Write $D(x) := C_2x^2 - C_0$ and define $h_1(x)$ to be such that $E_1(x) = h_1(x) - vD(x)$, i.e. $h_1(x)$ equals

$$\begin{aligned} &x^{3j}\big(x^2 g_{t-1}^{(2)}(x^3) - g_{t-1}^{(0)}(x^3)\big) \; - \; x^{3t}(C_2x^2 - C_0) \; + \; x\big(C_2 g_{t-1}^{(0)}(x^3) - C_0 g_{t-1}^{(2)}(x^3)\big) \; + \\ &+ \; w\Big[x^{3j}\big(x^2 g_{l-1}^{(2)}(x^3) - g_{l-1}^{(0)}(x^3)\big) \; - \; x^{3l}(C_2x^2 - C_0) \; + \; x\big(C_2 g_{l-1}^{(0)}(x^3) - C_0 g_{l-1}^{(2)}(x^3)\big)\Big]\,. \end{aligned}$$

Solving $E_1(x) = 0$ for $v$ then yields

$$v \;=\; \frac{h_1(x)}{D(x)} \;=:\; H_1(x)\,,$$

and if we substitute this for $v$ in $f(x)$ we find

$$f(x) \;=\; \frac{f_0(x)}{f_1(x)} \;=\; \frac{h_2(x)}{D(x)} \;=:\; H_2(x)$$

with

$$h_2(x) \;:=\; \big(x^2 g_{t-1}^{(2)}(x^3) - g_{t-1}^{(0)}(x^3)\big) \; + \; w\big(x^2 g_{l-1}^{(2)}(x^3) - g_{l-1}^{(0)}(x^3)\big)\,.$$

(This polynomial is of a remarkably simple form. Readers may wish to carry out the calculation in detail to convince themselves of the cancellation of $f_1(x)$.)

Further, for any rational function $\varphi(x) = \dfrac{\varphi_1(x)}{\varphi_2(x)}$ in $\overline{\mathbb{F}}_q(x)$ we put

$$B_\varphi(X,Y) \;:=\; \frac{\varphi_1(X)\varphi_2(Y) - \varphi_1(Y)\varphi_2(X)}{X-Y}\,.$$

One verifies easily that $B_\varphi(X,Y)$ is a polynomial in $\overline{\mathbb{F}}_q[X,Y]$ of total degree at most $\deg\,\varphi_1 + \deg\,\varphi_2 - 1$. For example, $\deg\,B_{H_1} \le 3t$ and $\deg\,B_{H_2} \le 3(t-j+1)$.

Now let $\gamma_1, \gamma_2$ be distinct multiple roots of $F_\beta$ in $\overline{\mathbb{F}}_q$. Then $E_1(\gamma_1) = E_1(\gamma_2) = 0$. We claim that $(\gamma_1, \gamma_2)$ is a solution to

$$B_{H_1}(X,Y) \;=\; B_{H_2}(X,Y) \;=\; 0. \tag{4.6}$$

Indeed, the numerator of $B_{H_2}(\gamma_1,\gamma_2)$ is equal to

$$\begin{aligned} & f_0(\gamma_1)f_1(\gamma_2) - f_0(\gamma_2)f_1(\gamma_1) \;= \\ =\; & f_0(\gamma_1)f_1(\gamma_2) + \beta f_1(\gamma_1)f_1(\gamma_2) - \beta f_1(\gamma_1)f_1(\gamma_2) - f_0(\gamma_2)f_1(\gamma_1) \;= \\ =\; & f_1(\gamma_2)F_\beta(\gamma_1) - f_1(\gamma_1)F_\beta(\gamma_2) \;=\; 0\,, \end{aligned}$$

and that of $B_{H_1}(\gamma_1,\gamma_2)$ is equal to

$$\begin{aligned} & h_1(\gamma_1)D(\gamma_2) - vD(\gamma_1)D(\gamma_2) + vD(\gamma_1)D(\gamma_2) - h_1(\gamma_2)D(\gamma_1) \;= \\ =\; & D(\gamma_2)E_1(\gamma_1) - D(\gamma_1)E_1(\gamma_2) \;=\; 0\,, \end{aligned}$$

while the denominators are $\gamma_1 - \gamma_2 \neq 0$.

Moreover, if for at least one $i \in \{1,2\}$ we have $D(\gamma_i) \neq 0$, this determines $v = H_1(\gamma_i)$ uniquely, as $H_1$ is a fixed expression. The only case where a pair $(\gamma_1,\gamma_2)$ can occur with an infinite number of $v$ is when $\{\gamma_1,\gamma_2\} = \{+\sqrt{C_0/C_2}, -\sqrt{C_0/C_2}\}$. But, since we are free to choose $\beta$ for each $v$, we can prevent these values from being roots of $F_\beta$ at all by a change of $\beta$ (note that for fixed $v$ and $\tilde{\beta} \neq \beta$ the polynomials $F_{\tilde{\beta}}$ and $F_\beta$ have no root in common).

The remaining crucial problem is to show that $B_{H_1}$ and $B_{H_2}$ are co-prime. Then it follows from Bezout's Theorem ([Ful], p. 112; applied as in Lemma 3 of [Coh72]) that the number of solution pairs $(\gamma_1,\gamma_2)$ to (4.6), and hence the number of $v$ we have to exclude in this step, is bounded by $\deg\,B_{H_1} \cdot \deg\,B_{H_2} \le 3t \cdot 3(t-j+1) \le 9t^2$.

By Lemma 4 of [Coh72] (or by [FrMR]), if $B_{H_1}$ and $B_{H_2}$ have a common factor, then $H_1$ and $H_2$ must decompose as functions of the same non-trivial rational function. To complete Step 3, we show that this is not the case.

## Simplicity—Step 3: solving the decomposition problem

In order to prove that $H_1$ and $H_2$ do not decompose as functions of the same inner function, we may multiply them individually with a constant to make them monic,

for if a rational function $H$ has a non-trivial decomposition $(Q_1, Q_2, R_1, R_2)$, then $cH$ has non-trivial decomposition $(cQ_1, Q_2, R_1, R_2)$.

We write again $\mathrm{lc}(g)$ for the leading coefficient of a polynomial $g$. Three cases (of increasing difficulty) follow.

$\boxed{C_0 = 0.}$

Then $D(x) = C_2x^2$. The linear coefficient of $h_1$ is $C_2(A_{3t+1} + wA_{3l+1}) \neq 0$, so that $x$, but not $x^2$, cancels from $H_1$. Thus, in reduced form, $H_1$ has linear denominator and is therefore indecomposable by Lemma 11(ii).

$\boxed{C_0 \neq 0,\ C_2 = 0.}$

In this case $w = 0$ and $D(x) = -C_0$, and we have the monic polynomials

$$\frac{-C_0}{\mathrm{lc}(h_1)} H_1(x) \;=\; \frac{1}{\mathrm{lc}(h_1)} \left( x^{3j+2} g_{t-1}^{(2)}(x^3) \;-\; x^{3j} g_{t-1}^{(0)}(x^3) \;+\; C_0 x^{3t} \;-\; C_0 x g_{t-1}^{(2)}(x^3) \right)$$

and

$$\frac{-C_0}{\mathrm{lc}(h_2)} H_2(x) \;=\; \frac{1}{\mathrm{lc}(h_2)} \left( x^2 g_{t-1}^{(2)}(x^3) \;-\; g_{t-1}^{(0)}(x^3) \right) .$$

Suppose the first of these decomposes as $Q(R(x))$ and the second as $S(R(x))$ with $Q(x), R(x), S(x) \in \overline{\mathbb{F}}_q[x]$ all monic of degree $\geq 2$. Upon differentiation,

$$Q'(R(x)) \cdot R'(x) \;=\; \frac{1}{\mathrm{lc}(h_1)} \left( -x^{3j+1} g_{t-1}^{(2)}(x^3) \;-\; C_0 g_{t-1}^{(2)}(x^3) \right) , \tag{4.7}$$

$$S'(R(x)) \cdot R'(x) \;=\; \frac{1}{\mathrm{lc}(h_2)} \left( -x g_{t-1}^{(2)}(x^3) \right) . \tag{4.8}$$

Using the fact that none of $g_{t-1}^{(2)}$, $R'$ and $S'$ is the zero polynomial, we can divide (4.7) by (4.8) to obtain

$$\frac{\mathrm{lc}(h_1)}{\mathrm{lc}(h_2)} \cdot \frac{Q'(R(x))}{S'(R(x))} \;=\; \frac{x^{3j+1} + C_0}{x} .$$

The right-hand side of this is clearly in lowest terms, and if one assumes the same for the left-hand side (after removing any common factor from $Q'$ and $S'$), equating denominators yields the contradiction $\deg R = 1$.

$\boxed{C_0 C_2 \neq 0.}$

Here, dividing by $C_2$ makes both $D$ and $h_2$ monic. Moreover, $\deg h_2 = 3(t-j) + 2$ is strictly greater than $\deg D$. Therefore we can assume for $H_2$ a decomposition as in Lemma 10. Notice again the absence of the $w$-terms, which simplifies matters considerably.

Suppose first that $H_2$ is already in lowest terms. Then equating degrees of the denominators as in (3.10) yields

$$2 \;=\; (\omega_1 - \omega_2)\rho_2 \;+\; \omega_2 \rho_1 .$$

For this there are only two possibilities: $(\omega_1, \omega_2, \rho_1, \rho_2)$ is equal to $(2, 0, \frac{1}{2}\deg h_2, 1)$ or $(\frac{1}{2}\deg h_2, 1, 2, 0)$. We show that in fact neither of these can occur.

- $(2, 0, \frac{1}{2}\deg h_2, 1)$:

  Then $\frac{1}{C_2}D(x) = R_2^2(x)$. If, say, $R_2(x) = x + A$, then this means $C_2x^2 - C_0 = C_2x^2 - C_2Ax + C_2A^2$, which implies both $A = 0$ and $A \neq 0$.

- $(\frac{1}{2}\deg h_2, 1, 2, 0)$:

  Now $\frac{1}{C_2}D(x) = Q_2(R_1(x))$, and by comparing coefficients, the quadratic polynomial $R_1$ must have vanishing linear term. Hence $h_1(x)$ and $h_2(x)$, both also functions of $R_1(x)$, must both involve only even powers of $x$. We claim to the contrary that always one of $h_1(x)$ and $h_2(x)$ contains an odd power of $x$.

  To see this, consider first the part of $h_1(x)$ whose exponents are $\equiv 1 \pmod 3$. Call this $x\Gamma(x)$, then

$$\begin{aligned}\Gamma(x) &= C_2 g_{t-1}^{(0)}(x^3) - C_0 g_{t-1}^{(2)}(x^3) = \\ &= \sum_{i=0}^{t-j} \left(C_2 A_{3(t-i)+1} - C_0 A_{3(t-i)-1}\right) x^{3i}. \end{aligned} \tag{4.9}$$

  By assumption, $\Gamma$ is not identically zero, so let $i_0 \in \{0, \ldots, t-j-1\}$ be such that the $i_0$-th summand in (4.9) does not vanish (the summand for $t-j$ is obviously zero).

  If $i_0$ is even, then $h_1(x)$ contains a non-vanishing $x^{3i_0+1}$-term and we are done.

  So suppose $i_0$ is odd. Now look at

$$\begin{aligned}h_2(x) &= x^2 g_{t-1}^{(2)}(x^3) - g_{t-1}^{(0)}(x^3) = \\ &= \sum_{i=0}^{t-j} A_{3(t-i)-1} x^{3i+2} - \sum_{i=0}^{t-j} A_{3(t-i)+1} x^{3i}. \end{aligned} \tag{4.10}$$

  The coefficient of $x^{3i_0}$ in $\Gamma(x)$ is $0 \neq C_2 A_{3(t-i_0)+1} - C_0 A_{3(t-i_0)-1}$, so at least one of $A_{3(t-i_0)+1}$ and $A_{3(t-i_0)-1}$ is non-zero; hence one of the terms $A_{3(t-i_0)-1}x^{3i_0+2}$ and $A_{3(t-i_0)+1}x^{3i_0}$ in (4.10) contributes the desired odd power of $x$ to $h_2(x)$.

Thus we have established that no non-trivial decomposition of $H_2$ exists if $D$ and $h_2$ are co-prime.

The next case to be considered is that $D$ and $h_2$ have exactly one linear factor in common. Then, after cancellation, the denominator of $H_2$ has degree one, and Lemma 11(ii) shows again that $H_2$ is indecomposable.

Finally, there remains the possibility that $D | h_2$. Then $H_2(x)$ is a monic polynomial of degree exactly $3(t-j)$. A look at

$$h_1(x) = x^{3j} \cdot h_2(x) - x^{3t} \cdot D(x) + x \cdot \left(C_2 g_{t-1}^{(0)}(x^3) - C_0 g_{t-1}^{(2)}(x^3)\right)$$

suggests to examine whether $D$ also divides $h_1$. This is indeed the case, as can be seen from

$$\begin{aligned}
& C_2 g_{t-1}^{(0)}(x^3) - C_0 g_{t-1}^{(2)}(x^3) = \\
= \; & C_2 g_{t-1}^{(0)}(x^3) - C_2 x^2 g_{t-1}^{(2)}(x^3) + C_2 x^2 g_{t-1}^{(2)}(x^3) - C_0 g_{t-1}^{(2)}(x^3) = \\
= \; & C_2 \left( -x^2 g_{t-1}^{(2)}(x^3) + g_{t-1}^{(0)}(x^3) \right) + (C_2 x^2 - C_0)\, g_{t-1}^{(2)}(x^3) = \\
= \; & -C_2 h_2(x) + D(x)\, g_{t-1}^{(2)}(x^3) = -\left( C_2 H_2(x) - g_{t-1}^{(2)}(x^3) \right) D(x) .
\end{aligned}$$

Consequently,

$$H_1(x) = x^{3j} \cdot H_2(x) - x^{3t} - x \cdot \left( C_2 H_2(x) - g_{t-1}^{(2)}(x^3) \right) . \tag{4.11}$$

It is not possible to nail down the leading term of $H_1$, as the $x^{3t}$ cancels with the leading term of $x^{3j} \cdot H_2(x)$, and there may or may not be terms of order higher than $\deg x(C_2 H_2(x) - g_{t-1}^{(2)}(x^3))$. At any rate, $H_1$ is also a polynomial, so if there is a decomposition of $H_1$ and $H_2$ with common inner function, we can assume

$$H_1(x) = \mathrm{lc}(H_1) \cdot Q(R(x)) , \qquad H_2(x) = S(R(x)) \tag{4.12}$$

with $Q(x), R(x), S(x) \in \overline{\mathbb{F}}_q[x]$ all monic of degree $\geq 2$.

The aim is to prove this impossible. To this end we employ a striking intermediate result.

**Lemma 14** *In the situation described,*

$$H_1'(x) = \frac{1}{x} \cdot f_1(x) \cdot H_2'(x) ,$$

*and this is not identically zero.*

*Proof.* We begin by writing $h_2(x)$ in the form

$$h_2(x) = (C_2 x^2 - C_0) x^{3(t-j)} + \sum_{i=0}^{t-j-1} (A_{3(t-i)-1} x^2 - A_{3(t-i)+1})\, x^{3i} .$$

This means that

$$H_2(x) = x^{3(t-j)} + \sum_{i=0}^{t-j-1} \frac{A_{3(t-i)-1} x^2 - A_{3(t-i)+1}}{C_2 x^2 - C_0}\, x^{3i} . \tag{4.13}$$

Using the quotient rule, one finds

$$\begin{aligned}
H_2'(x) &= \sum_{i=0}^{t-j-1} \frac{-A_{3(t-i)-1} x\,(C_2 x^2 - C_0) - (-C_2 x)(A_{3(t-i)-1} x^2 - A_{3(t-i)+1})}{(C_2 x^2 - C_0)^2}\, x^{3i} = \\
&= \frac{x}{(C_2 x^2 - C_0)^2} \sum_{i=0}^{t-j-1} (C_0 A_{3(t-i)-1} - C_2 A_{3(t-i)+1})\, x^{3i} .
\end{aligned} \tag{4.14}$$

It is clear from (4.14) that $H_2'$ vanishes identically if and only if

$$C_0 A_{3(t-i)-1} = C_2 A_{3(t-i)+1} \quad \text{for all } i = 0, \ldots, t-j-1,$$

which is equivalent to $C_0 g_{t-1}^{(2)} = C_2 g_{t-1}^{(0)}$. Since the latter is assumed not to be the case, this proves the "not identically zero" part of the lemma.

Further,

$$\begin{aligned} C_2 H_2(x) - g_{t-1}^{(2)}(x^3) &\overset{(4.13)}{=} C_2 \cdot \left( x^{3(t-j)} + \sum_{i=0}^{t-j-1} \frac{A_{3(t-i)-1}x^2 - A_{3(t-i)+1}}{C_2 x^2 - C_0}\, x^{3i} \right) - \\ &\quad - \left( C_2 x^{(3t-j)} + \sum_{i=0}^{t-j-1} A_{3(t-i)-1} x^{3i} \right) = \\ &= \sum_{i=0}^{t-j-1} \frac{-C_2 A_{3(t-i)+1} + C_0 A_{3(t-i)-1}}{C_2 x^2 - C_0}\, x^{3i} = \\ &\overset{(4.14)}{=} \frac{C_2 x^2 - C_0}{x} \cdot H_2'(x), \end{aligned} \tag{4.15}$$

and it follows then from (4.11) that

$$\begin{aligned} H_1'(x) &= x^{3j} \cdot H_2'(x) - \left[ C_2 H_2(x) - g_{t-1}^{(2)}(x^3) + x C_2 H_2'(x) \right] = \\ &\overset{(4.15)}{=} \left( x^{3j} - \frac{C_2 x^2 - C_0}{x} - C_2 x \right) \cdot H_2'(x) = \frac{1}{x} \cdot f_1(x) \cdot H_2'(x). \end{aligned}$$

□

Coming back to the decomposition (4.12), differentiation yields

$$H_1'(x) = \mathrm{lc}(H_1) \cdot Q'(R(x)) \cdot R'(x), \qquad H_2'(x) = S'(R(x)) \cdot R'(x),$$

and the lemma allows to divide $H_1'(x)$ by $H_2'(x)$. This gives

$$\frac{H_1'(x)}{H_2'(x)} = \frac{f_1(x)}{x},$$

which is clearly in lowest terms, and at the same time

$$\frac{H_1'(x)}{H_2'(x)} = \frac{\mathrm{lc}(H_1) \cdot Q'(R(x)) \cdot R'(x)}{S'(R(x)) \cdot R'(x)} = \mathrm{lc}(H_1) \cdot \frac{Q'(R(x))}{S'(R(x))};$$

equating denominators (after removing common factors of $Q'$ and $S'$, if necessary) leads to the contradiction $\deg R = 1$.

This proves that $H_1(x)$ and $H_2(x)$ cannot decompose non-trivially as functions of the same $R(x)$, and with this Step 3 is complete.

Altogether, the number of $v$ that have been excluded in the various steps to secure the simplicity condition of Lemma 7, together with those for co-primality and indecomposability, is far from exhausting $\mathbb{F}_q$. This proves the splitting of $F(x)$ in the Standard Case.

## 4.3 The "Exceptional Case": $C_0 g_{t-1}^{(2)} = C_2 g_{t-1}^{(0)}$

The condition of the Exceptional Case implies $C_0 g_d^{(2)} = C_2 g_d^{(0)}$ for $d = 0, \ldots, t-1$. It includes the possibility that for one $i \in \{0, 2\}$ the polynomial $g_{t-1}^{(i)}$ is identically zero; note that $g_{t-1}^{(i)} = 0 \Leftrightarrow C_i = 0$. In the case $C_0 \neq 0$ we put $C := C_2/C_0$ so that we can replace $g_d^{(2)}$ by $C g_d^{(0)}$ where useful.

Our choice of parameters is

$$s_3 = u \qquad \text{and} \qquad s_{3t} = uv\,,$$

which results in $F(x) = F_u(x) = f_0(x) + u f_1(x)$ with

$$\begin{aligned} f_0(x) &= x^{3t+1} + x^2 g_{t-1}^{(2)}(x^3) + g_{t-1}^{(0)}(x^3)\,, \\ f_1(x) &= x^{3t-2} + x^2 g_{t-2}^{(2)}(x^3) + g_{t-2}^{(0)}(x^3) + vx\,. \end{aligned}$$

This time the polynomial to be adjusted by the choice of $v$ is $f_1$, and we aim to meet the conditions of Lemma 8.

### Co-primality

The constant term of $f_0$ is $A_{3t+1}$, that of $f_1$ is $A_{3t-2}$. If at least one of these is not zero, Lemma 9 applies immediately and says that at most $3t+1$ values of $v$ need to be excluded to ensure $\gcd(f_0, f_1) = 1$.

If $A_{3t+1} = A_{3t-2} = 0$, we must work with

$$f_i^*(x) := \frac{f_i(x)}{x} \qquad (i = 0, 1)\,.$$

Then, by Lemma 9, it suffices again to exclude at most $3t+1$ values of $v$ (one of them being 0) to obtain $\gcd(f_0^*, f_1^*) = 1$. However, in what follows we must now remember to establish all other conditions both for $f_0^*, f_1^*$ as well as for $f_0, f_1$.

### Indecomposability

Suppose $f_0/f_1$ has a decomposition as in Lemma 10. Subtracting the degree equation $\deg f_1 = (\omega_1 - \omega_2)\rho_2 + \omega_2\rho_1$ from $\deg f_0 = \omega_1\rho_1$ (compare (3.7) to (3.10)) gives $3 = (\omega_1 - \omega_2)(\rho_1 - \rho_2)$, i.e. one of $\omega_1 - \omega_2$ and $\rho_1 - \rho_2$ is equal to 1 and the other to 3. Of course, this works the same way for $f_0^*$ and $f_1^*$.

It is now necessary to distinguish seven subcases:

$A_{3t+1}, A_{3t-2}$ are not both zero (work with $f_0, f_1$)

- $\omega_1 - \omega_2 = 3$
  - $g_{t-1}^{(2)} \neq 0$ — Subcase (1)
  - $g_{t-1}^{(2)} = 0$ — Subcase (2)
- $\omega_1 - \omega_2 = 1$
  - $Q_1$ is not a power of $x$ — Subcase (3)
  - $Q_1$ is a power of $x$ — Subcase (4)

$A_{3t+1} = A_{3t-2} = 0$ (work with $f_0^*, f_1^*$)

- $\omega_1 - \omega_2 = 3$ — Subcase (5)
- $\omega_1 - \omega_2 = 1$
  - $Q_1$ is not a power of $x$ — Subcase (6)
  - $Q_1$ is a power of $x$ — Subcase (7)

**Subcase (1): $\{A_{3t+1}, A_{3t-2}\} \not\subseteq \{0\}$, $\omega_1 - \omega_2 = 3$, $g_{t-1}^{(2)} \neq 0$.**

We have $f_0(x) = R_2^{\omega_1}(x) \cdot Q_1(R_1(x)/R_2(x))$. As detailed after Lemma 10, we can assume $Q_1(0) = 0$, which implies that $R_1$ divides $f_0$. Since $f_0$ is fixed of degree $3t+1$, this limits the number of possibilities for $R_1$ by $2^{3t+1}$.

On the other hand, $f_1(x) = x^{3t-2} + x^2 g_{t-2}^{(2)}(x^3) + g_{t-2}^{(0)}(x^3) + vx$ equals $R_2^3(x)P(x)$ with $P(x) := R_2^{\omega_2}(x) \cdot Q_2(R_1(x)/R_2(x)) \in \overline{\mathbb{F}}_q[x]$. As $R_2^3(x)$ is actually a polynomial in $x^3$, the derivatives of $f_1$ are

$$\begin{aligned} f_1'(x) &= x^{3t-3} - x g_{t-2}^{(2)}(x^3) + v = R_2^3(x) \cdot P'(x), \\ f_1''(x) &= -g_{t-2}^{(2)}(x^3) = R_2^3(x) \cdot P''(x). \end{aligned}$$

The last equation says that $R_2^3(x)$ divides a fixed polynomial in $x^3$ of degree at most $t-j-1$ that is not identically zero by assumption. So there are at most $2^{t-j-1}$ possibilities for $R_2$. This makes together at most $2^{4t-j} \leq 2^{4t-1}$ possibilities for $R$.

It remains to show that each $R$ can only occur with a bounded number of values $v$. We show that in fact every $R$ determines a unique $v$, so that the number of values to be excluded is at most $2^{4t-1}$. To see this, fix $R$ and assume it occurs with distinct $v$ and $\tilde{v}$. Then, from $f_1'$,

$$\begin{aligned} x^{3t-3} - x g_{t-2}^{(2)}(x^3) + v &= R_2^3(x) \cdot P'(x), \\ x^{3t-3} - x g_{t-2}^{(2)}(x^3) + \tilde{v} &= R_2^3(x) \cdot \tilde{P}'(x) \end{aligned}$$

(with obvious $\tilde{P}$), and after subtracting,

$$0 \neq v - \tilde{v} = R_2^3(x) \cdot \big(P'(x) - \tilde{P}'(x)\big).$$

This would require $R_2$ to be a constant, but then $\rho_2 = 0$ and $\rho_1 = \rho_2 + 1 = 1$, in contradiction to our decomposition assumptions.

**Subcase (2): $\{A_{3t+1}, A_{3t-2}\} \not\subseteq \{0\}$, $\omega_1 - \omega_2 = 3$, $g_{t-1}^{(2)} = 0$.**

As in the previous subcase, the number of possible $R_1$ can be limited to at most $2^{3t+1}$ by assuming $Q_1(0) = 0$.

Now we have

$$\begin{aligned} f_1(x) &= x^{3t-2} + g_{t-2}^{(0)}(x^3) + vx = R_2^3(x) \cdot P(x)\,, \\ f_1'(x) &= x^{3t-3} + v = R_2^3(x) \cdot P'(x)\,, \end{aligned}$$

and therefore $R_2^3(x)$ divides $f_1(x) - xf_1'(x) = g_{t-2}^{(0)}(x^3)$, a fixed polynomial in $x^3$ of degree $\leq t-j-1$, which is not identically zero (because its counterpart already is). Again the choices for $R_2$ are bounded by $2^{t-j-1}$, and hence those for $R$ by $2^{4t-1}$.

If a fixed $R$ comes with $\tilde{v} \neq v$, then $x^{3t-3} + v = R_2^3(x) \cdot P'(x)$ and $x^{3t-3} + \tilde{v} = R_2^3(x) \cdot \tilde{P}'(x)$, hence again $v - \tilde{v} = R_2^3(x) \cdot \big(P'(x) - \tilde{P}'(x)\big)$. The rest of the argument which shows that every $R$ leads to the exclusion of at most one $v$ is as above.

**Subcase (3): $\{A_{3t+1}, A_{3t-2}\} \not\subseteq \{0\}$, $\omega_1 - \omega_2 = 1$, $Q_1$ not a power of $x$.**

Assume again $Q_1(0) = 0$ to bound the choices for $R_1$ by $2^{3t+1}$.

Now let $R_1$ be fixed. Because $Q_1$ is not a power of $x$, one obtains by subtracting $R_1^{\omega_1}(x)$ from $f_0(x) = R_2^{\omega_1}(x) \cdot Q_1(R_1(x)/R_2(x))$ a fixed non-zero polynomial of degree at most

$$(\omega_1 - 1)\rho_1 + \rho_2 = \omega_1\rho_1 - (\rho_1 - \rho_2) = (\deg f_0) - 3 = 3t - 2$$

that is divisible by $R_2$. Thus there are at most $2^{3t-2}$ choices of $R_2$ for any given $R_1$, allowing at most $2^{6t-1}$ choices for $R$.

Next fix $R$ and assume it occurs with distinct $v$ and $\tilde{v}$. Write

$$Q_2(x) =: \sum_{i=0}^{\omega_2} q_{2,i} x^i \qquad \text{and} \qquad \tilde{Q}_2(x) =: \sum_{i=0}^{\omega_2} \tilde{q}_{2,i} x^i\,.$$

Then, with $\langle\!\langle\, . \,\rangle\!\rangle$ denoting the fixed part,

$$\begin{aligned} \langle\!\langle f_1(x) \rangle\!\rangle + vx &= R_2(x) \cdot \sum_{i=0}^{\omega_2} q_{2,i} R_1^i(x) R_2^{\omega_2 - i}(x)\,, \\ \langle\!\langle f_1(x) \rangle\!\rangle + \tilde{v}x &= R_2(x) \cdot \sum_{i=0}^{\omega_2} \tilde{q}_{2,i} R_1^i(x) R_2^{\omega_2 - i}(x)\,, \end{aligned}$$

and after subtracting (note $q_{2,\omega_2} = \tilde{q}_{2,\omega_2} = 1$)

$$(v - \tilde{v})x = R_2(x) \cdot \sum_{i=0}^{\omega_2 - 1} (q_{2,i} - \tilde{q}_{2,i}) R_1^i(x) R_2^{\omega_2 - i}(x)\,.$$

Now equating degrees shows (because the degrees of the summands on the RHS strictly increase with $i$)

$$1 = \rho_2 + \rho_1 i_0 + \rho_2(\omega_2 - i_0) \quad \text{for some } i_0 \in \{0, \ldots, \omega_2 - 1\}\,.$$

Here clearly $\rho_2 = 0$ or $\rho_2 = 1$. But in the first case it follows that $\rho_1 = 1$, and in the second case $i_0 = \omega_2 = 0$ and therefore $\omega_1 = 1$; both contradict the properties of the assumed decomposition.

Therefore there can be no more than one $v$ for every $R$, and we have to exclude at most $2^{6t-1}$ values to avoid that $f_0/f_1$ has a decomposition of this type.

Subcase (4): $\{A_{3t+1}, A_{3t-2}\} \not\subseteq \{0\}$, $\omega_1 - \omega_2 = 1$, $Q_1$ a power of $x$.

We show that this subcase cannot occur.

Since $A_{3t+1}$ and $A_{3t-2}$ are not both zero, we must have $g_{t-1}^{(0)} \neq 0$ and $C_0 \neq 0$. Therefore (with $C = C_2/C_0$)

$$f_0(x) \;=\; x^{3t+1} + (Cx^2+1)g_{t-1}^{(0)}(x^3)\,.$$

This must now be equal to $R_1^{\omega_1}(x)$. Consequently,

$$f_0'(x) \;=\; x^{3t} - Cxg_{t-1}^{(0)}(x^3) \;=\; \omega_1 R_1^{\omega_1-1}(x)\cdot R_1'(x)\,,$$

where $3 \nmid \omega_1$, $\omega_1 \geq 2$ and $\rho_1 \geq 3$.

From $R_1^{\omega_1}(x) = x^{3t+1} + (Cx^2+1)g_{t-1}^{(0)}(x^3)$ it follows moreover that the greatest common divisor of $R_1(x)$ and $g_{t-1}^{(0)}(x^3)$ is a power of $x$. Write $R_1(x) = x^l S(x)$ with $S(x) \in \overline{\mathbb{F}}_q[x]$ and $S(0) \neq 0$. Then $\gcd\big(S(x), g_{t-1}^{(0)}(x^3)\big)$ divides $\gcd\big(R_1(x), g_{t-1}^{(0)}(x^3)\big)$ but is itself not divisible by $x$, hence

$$\gcd\big(S(x), g_{t-1}^{(0)}(x^3)\big) \;=\; 1\,.$$

Now $f_0(x) - xf_0'(x) = (1-Cx^2)g_{t-1}^{(0)}(x^3)$ is divisible by $R^{\omega_1-1}(x) = [x^l S(x)]^{\omega_1-1}$. Since $S(x)$ is co-prime to $g_{t-1}^{(0)}(x^3)$, we must have $S^{\omega_1-1}(x) \mid 1 - Cx^2$, so that

$$(\omega_1 - 1)\cdot \deg S \;\leq\; 2\,.$$

This leaves three possibilities:

(i) $\deg S = 0$;
(ii) $\deg S = 1$, $\omega_1 = 2$;
(iii) $\deg S = 2$, $\omega_1 = 2$.

If $S(x)$ were equal to 1, then $R_1(x)$, and therefore $f_0(x)$, would be a power of $x$, but this is not so. This excludes (i). But (ii) and (iii) are also impossible, because here $\omega_1 \mid \deg f_0$ implies that $t$ is odd, whereas we are considering only even $t$.

However, we will need later that this part of the proof holds also for odd $t \geq 3$. So here is an alternative argument that still rules out (ii) and (iii) in this situation: then $\rho_1 = \frac{\deg f_0}{\omega_1} \geq 5$, so $l = \rho_1 - \deg S \geq 3$; this means that $f_0(x) = R_1^2(x)$ is divisible by $x^6$, contrary to the assumption that $A_{3t+1}$ and $A_{3t-2}$ are not both zero.

In Subcases (5) to (7) we assume for $f_0^*/f_1^*$ a decomposition as in Lemma 10.

Subcase (5): $A_{3t+1} = A_{3t-2} = 0,\ \omega_1 - \omega_2 = 3.$

The polynomial $f_0^*$ is fixed of degree $3t$. Applying the usual $Q_1(0) = 0$ argument bounds the number of choices for $R_1$ by $2^{3t}$.

Under the given conditions,

$$f_1^*(x) \;=\; x^{3t-3} + x g_{t-2}^{(2)}(x^3) + x^2 g_{t-3}^{(0)}(x^3) + v \;=\; R_2^3(x) \cdot P(x)\,,$$

where $P$ is defined as in Subcase (1). From

$$f_1^{*\prime}(x) \;=\; g_{t-2}^{(2)}(x^3) - x g_{t-3}^{(0)}(x^3) \;=\; R_2^3(x) \cdot P'(x)$$

one sees that $R_2^3(x)$ divides the fixed polynomial $f_1^{*\prime}(x)$. This cannot be zero, because all coefficients $A_2, A_4, \ldots, A_{3t-4}$ appear in it, one of which must be non-zero (since $A_{3t-2}$ *is* zero). Further deg $f_1^{*\prime} \le 3(t-j-1)$, so that deg $R_2 \le t-j-1$. This gives at most $2^{t-j-1}$ possibilities for $R_2$ and at most $2^{4t-j-1} \le 2^{4t-2}$ for $R$.

Limiting the number of $v$ per $R$ to one is analogous to Subcase (1), using $v - \tilde{v} = R_2^3(x) \cdot \big(P(x) - \tilde{P}(x)\big)$; replacing $P', \tilde{P}'$ with $P, \tilde{P}$ does not affect this argument.

Subcase (6): $A_{3t+1} = A_{3t-2} = 0,\ \omega_1 - \omega_2 = 1,\ Q_1$ not a power of $x$.

Once again we use the assumption $Q_1(0) = 0$ to bound the number of possible $R_1$ by $2^{3t}$. Arguing analogously to the beginning of Subcase (3), we further limit the number of possible $R$ to $2^{6t-3}$. Once again we will also show that only one $v$ is possible for each $R$, but this time this turns out to be more difficult than in previous situations.

Suppose $R$ occurs with distinct $v$ and $\tilde{v}$. Subtracting $\langle\!\langle f_1^*(x) \rangle\!\rangle + \tilde{v} = R_2(x)\tilde{P}(x)$ from $\langle\!\langle f_1^*(x) \rangle\!\rangle + v = R_2(x)P(x)$, with the appropriate $P$ and $\tilde{P}$, yields

$$v - \tilde{v} \;=\; R_2(x) \cdot \big(P(x) - \tilde{P}(x)\big)\,,$$

which again forces $R_2$ to be a constant. Unlike in previous cases with $\omega_1 - \omega_2 = 3$, this does now not immediately lead to a contradiction, but merely implies $\rho_1 = 3$.

Assume first that $g_{t-1}^{(0)} = 0$. Then

$$\begin{aligned} f_0^*(x) &= x^{3t} + x g_{t-1}^{(2)}(x^3) \;=\; Q_1(R_1(x))\,, \\ f_1^*(x) &= x^{3t-3} + x g_{t-2}^{(2)}(x^3) + v \;=\; Q_2(R_1(x))\,. \end{aligned}$$

Put $\hat{f}_1^*(x) := f_1^*(x) - v$. With $\hat{Q}_2(x) := Q_2(x) - v$, this decomposes as

$$\hat{f}_1^*(x) \;=\; x^{3t-3} + x g_{t-2}^{(2)}(x^3) \;=\; \hat{Q}_2(R_1(x))\,.$$

(Of course, we could have used $\tilde{v}$ here as well.) Differentiation yields

$$f_0^{*\prime}(x) \;=\; g_{t-1}^{(2)}(x^3) \;=\; x^3 g_{t-2}^{(2)}(x^3) + A_{3t-1} \;=\; Q_1'(R_1(x)) \cdot R_1'(x)$$

and

$$\hat{f}_1^{*\prime}(x) \;=\; g_{t-2}^{(2)}(x^3) \;=\; \hat{Q}_2'(R_1(x)) \cdot R_1'(x)\,. \tag{4.16}$$

It follows that $R_1'(x)$ divides $f_0^{*\prime}(x) - x^3\hat{f}_1^{*\prime}(x) = A_{3t-1}$, i.e. if $A_{3t-1} \neq 0$ then $R_1'$ is a constant $K \in \overline{\mathbb{F}}_q^*$. Define

$$\hat{Q}_1(x) := \begin{cases} Q_1'(x), & \text{if } A_{3t-1} = 0, \\ Q_1'(x) - \dfrac{A_{3t-1}}{K}, & \text{if } A_{3t-1} \neq 0. \end{cases}$$

Then it is easily verified that in both cases

$$\hat{Q}_1(R_1(x)) \cdot R_1'(x) = x^3 g_{t-2}^{(2)}(x^3). \tag{4.17}$$

One checks further that none of $g_{t-2}^{(2)}$, $R_1'$ and $\hat{Q}_2'$ is the zero polynomial (in the case of the latter two this would mean that $R_1(x)$ or $\hat{Q}_2(x)$ had to be in $\overline{\mathbb{F}}_q[x^3]$). Hence, as a result of dividing (4.17) by (4.16),

$$\frac{\hat{Q}_1}{\hat{Q}_2'}(R_1(x)) = x^3.$$

This implies that $R_1(x)$ is a function of $x^3$, and we have a contradiction.

Now consider the situation for $g_{t-1}^{(0)} \neq 0$. Here one finds

$$\begin{aligned} f_0^*(x) &= x^{3t} + (Cx^7 + x^5) g_{t-3}^{(0)}(x^3) = Q_1(R_1(x)), \\ f_1^*(x) &= x^{3t-3} + (Cx^4 + x^2) g_{t-3}^{(0)}(x^3) + v = Q_2(R_1(x)). \end{aligned}$$

The rest of the procedure is similar to that in the case $g_{t-1}^{(0)} = 0$: define $\hat{f}_1^*$ and $\hat{Q}_2$ as above, look at $f_0^{*\prime}/\hat{f}_1^{*\prime}$ (this time $g_{t-3}^{(0)}$ is non-zero and can be cancelled) and find $(Q_1'/\hat{Q}_2')(R_1(x)) = x^3$, which leads to the same contradiction.

**Subcase (7): $A_{3t+1} = A_{3t-2} = 0$, $\omega_1 - \omega_2 = 1$, $Q_1$ a power of $x$.**

As it will turn out, similarly to Subcase (4), this type of decomposition is again impossible.

Since $\omega_2 = \omega_1 - 1 \geq 1$, we assume $Q_2(0) = 0$ (the reason will become clear later). Write $R_1(x) = x^l S(x)$ with $S(x) \in \overline{\mathbb{F}}_q[x]$, $S(0) \neq 0$.

Assume first that $g_{t-1}^{(0)} = 0$. Then

$$f_0^*(x) = x^{3t} + x g_{t-1}^{(2)}(x^3) = R_1^{\omega_1}(x).$$

Therefore $\gcd\big(R_1(x), g_{t-1}^{(2)}(x^3)\big)$ is a power of $x$ and $\gcd\big(S(x), g_{t-1}^{(2)}(x^3)\big) = 1$. The derivative of $f_0^*$ is

$$f_0^{*\prime}(x) = g_{t-1}^{(2)}(x^3) = \omega_1 R_1^{\omega_1 - 1}(x) \cdot R_1'(x),$$

not identically zero. Hence $S(x)$ divides $g_{t-1}^{(2)}(x^3)$ and we must have $S(x) = 1$. But then $R_1(x)$, and consequently $f_0^*(x)$, is a power of $x$, contradiction.

Now we come to the case $g_{t-1}^{(0)} \neq 0$. Here

$$f_0^*(x) \;=\; x^{3t} + (Cx^7 + x^5)g_{t-3}^{(0)}(x^3) \;=\; R_1^{\omega_1}(x)\,.$$

Therefore $\gcd\big(R_1(x), g_{t-3}^{(0)}(x^3)\big)$ is a power of $x$ and $\gcd\big(S(x), g_{t-3}^{(0)}(x^3)\big) = 1$. This time it is necessary to differentiate twice:

$$\begin{aligned} f_0^{*\prime}(x) \;&=\; (Cx^6 - x^4)g_{t-3}^{(0)}(x^3) \;=\; \omega_1 R_1^{\omega_1 - 1}(x)\cdot R_1'(x)\,, \\ f_0^{*\prime\prime}(x) \;&=\; -x^3 g_{t-3}^{(0)}(x^3) \;=\; \omega_1 R_1^{\omega_1 - 2}(x)\cdot\big[(\omega_1 - 1)\big(R_1'(x)\big)^2 \;+\; R_1(x)R_1''(x)\big]. \end{aligned}$$

Provided $\omega_1 > 2$, this shows that $S(x)$ divides the non-zero polynomial $g_{t-3}^{(0)}(x^3)$, and a contradiction is obtained as above.

What if $\omega_2 = 2$? Here our assumption $Q_2(0) = 0$ comes in and results in $R_1(x)$ dividing

$$R_2(x)\cdot\left\{R_2^{\omega_2}(x)\cdot Q_2\left(\frac{R_1(x)}{R_2(x)}\right)\right\} \;=\; f_1^*(x) \;=\; x^{3t-3} + (Cx^4 + x^2)g_{t-3}^{(0)}(x^3) + v\,.$$

As a consequence, $R_1(x)$ divides also $f_0^*(x) - x^3 f_1^*(x) = -vx^3$, and this implies $R_1(x) = x^3$ and again the contradiction that $f_0^*(x)$ is a power of $x$. This completes the final subcase.

In summary, at most $2^{4t} + 2^{6t-1}$ values $v$ must be discarded in the case that one of $A_{3t+1}, A_{3t-2}$ is non-zero, and at most $2^{4t-2} + 2^{6t-3}$ in the case $A_{3t+1} = A_{3t-2} = 0$, in order to ensure indecomposability of $f_0/f_1$ over $\overline{\mathbb{F}}_q$. Both numbers are smaller than $2^{6t}$, which is, already for small $t$, only a fraction of $[(3t+1)!\,(3t-1)]^2$.

## Condition (iii)

The degree difference is clear (also for $f_0^*$ and $f_1^*$). As for the second part, we show that $f_1$ is square-free for all but a small number of $v$. Then the same is true for $f_1^*$ (because $f_1^* | f_1$, except for $v = 0$, which we have already excluded anyway in this case).

Let $\gamma$ be a multiple root of $f_1$. Then $\gamma$ satisfies

$$0 \;=\; f_1(\gamma) \;=\; \gamma^{3t-2} + \gamma^2 g_{t-2}^{(2)}(\gamma^3) + g_{t-2}^{(0)}(\gamma^3) + v\gamma \tag{4.18}$$

and

$$0 \;=\; f_1'(\gamma) \;=\; \gamma^{3t-3} - \gamma g_{t-2}^{(2)}(\gamma^3) + v\,. \tag{4.19}$$

Solving (4.19) for $v$ and substituting into (4.18) shows that

$$-\gamma^2 g_{t-2}^{(2)}(\gamma^3) + g_{t-2}^{(0)}(\gamma^3) \;=\; 0\,,$$

i.e. $\gamma$ is a root of a fixed non-zero polynomial of degree at most $3t - 3j - 1 \leq 3t - 4$. So this is the maximal number of elements in $\overline{\mathbb{F}}_q$ which can be a multiple root of $f_1$ at all, and clearly each can be so only for a single choice of $v$. Hence, by excluding at most $3t - 4$ values of $v$, we can guarantee that $f_1$ is square-free.

## Discriminants

Here our Lemma 12 comes finally into action. First observe that if $\alpha$ is a common root of $E_1$ and $f_1$, then $f_0(\alpha)f_1'(\alpha) = 0$, and, since $f_0(\alpha) \neq 0$, this means that $\alpha$ must be a repeated root of $f_1$. But we have just arranged that such roots do not exist, therefore we can assume that $E_1$ and $f_1$ are co-prime. Consequently, in Lemma 12, all factors of $E_1$ contribute to the degree of $\Delta_x$ in $u$, and to prove that $\Delta_x$ is a non-square in $\overline{\mathbb{F}}_q(u)$ it suffices to show that $E_1$ has odd degree. We begin by identifying all situations where this is the case.

We divide this analysis into three subcases. Here our conclusions use the fact that $t$ is even in Theorem 8.

$\boxed{A_{3t+1}, A_{3t-2} \text{ are not both zero.}}$

This implies $C_0 \neq 0$, and with $C = C_2/C_0$ (possibly zero) one finds

$$E_1(x) \;=\; (Cx^2 - 1)\underbrace{\left(A_{3t+1}x^{3t-3} + vg_{t-1}^{(0)}(x^3)\right)}_{=:\,\Omega(x)},$$

which has odd degree if and only if $\Omega$ has odd degree.

Suppose $A_{3t+1} \neq 0$. Then the leading term of $\Omega(x)$ is $A_{3t+1}x^{3t-3}$ for $j > 1$, and for $j = 1$ it is $(A_{3t+1} + vC_0)x^{3t-3}$ provided we avoid $v = -A_{3t+1}/C_0$. In both cases $\deg \Omega = 3t - 3$ is odd.

Suppose otherwise that $A_{3t+1} = 0$, in which case $\Omega(x) = vg_{t-1}^{(0)}(x^3)$. This has degree $3(t - j)$ (assuming that we exclude $v = 0$), so $\deg \Omega$ is odd if and only if $j$ is odd.

$\boxed{A_{3t+1} = A_{3t-2} = 0,\ C_0 \neq 0.}$

Here $E_1^*(x) := f_0^{*\prime}(x)f_1^*(x) - f_0^*(x)f_1^{*\prime}(x) = vx^4(Cx^2 - 1)g_{t-3}^{(0)}(x^3)$, with $C$ zero or not. The degree of $E_1^*$ is odd if and only if $\deg g_{t-3}^{(0)}(x^3) = 3(t - j - 2)$ is odd, i.e. iff $j$ is odd.

$\boxed{A_{3t+1} = A_{3t-2} = C_0 = 0.}$

In this final subcase $E_1^*(x) = A_{3t-1}x^{3t-3} + vg_{t-1}^{(2)}(x^3)$, and the rest is analogous to the first subcase: $\deg E_1^*$ is always odd for $A_{3t-1} \neq 0$, while for $A_{3t-1} = 0$ it is odd exactly if $j$ is odd.

In all of the above cases where $\deg E_1$ (or $\deg E_1^*$, respectively) is odd we are done, having established Lemma 8 with "Situation S".

The remaining open problems are:

$$\begin{array}{ll} \text{(I)} & A_{3t+1} = 0,\ A_{3t-2} \neq 0,\ j \text{ even;} \\ \text{(II)} & A_{3t+1} = A_{3t-2} = 0,\ C_0 \neq 0,\ j \text{ even;} \\ \text{(III)} & A_{3t+1} = A_{3t-2} = C_0 = A_{3t-1} = 0,\ j \text{ even.} \end{array}$$

In these situations, a closer examination of $E_1$ or $E_1^*$ reveals that $\Delta_x$ is still a non-square in *most* cases, but there are always circumstances in which $\Delta_x$ is definitely a square in $\overline{\mathbb{F}}_q(u)$. The details for (I) can serve as an example.

In (I) we have $E_1(x) = (Cx^2 - 1)vx^3 g_{t-2}^{(0)}(x^3)$ with $C = C_2/C_0$ and $g_{t-2}^{(0)}(0) \neq 0$.

Assume first $C_2 = 0$. Then the roots of $E_1$ are the triple root zero and certain $\alpha_\mu \neq 0$ (all with multiplicity divisible by 3). Now $f(0) = 0$, but $f_0(\alpha_\mu) = \alpha_\mu^{3t+1} \neq 0$ and so $f(\alpha_\mu) \neq 0$ for all other roots. Therefore, by Lemma 12, the factor $u$ appears precisely with multiplicity 3 in $\Delta_x$, making $\Delta_x$ a non-square in $\overline{\mathbb{F}}_q(u)$, and we are in "Situation S".

So far everything is fine. But now consider $C_2 \neq 0$. Then we have the additional non-zero roots $\pm\eta$ of $Cx^2 - 1$, and a problem occurs: if $f(-\eta) = 0$, $\alpha_1 = \eta$, and all other $\alpha_\mu$ have even multiplicity, then $\Delta_x$ *is* a square. Take

$$g_{t-1}^{(2)}(x^3),\ g_{t-1}^{(0)}(x^3) \ := \ x^{18} - x^{15} - x^{12} + x^9 + x^6 - x^3$$

(with $\eta = 1$) for an example where this actually happens.

Similar phenomena occur with (II) and (III). In these sporadic situations where $\Delta_x$ is a square, the problem would be solved by showing that $\Delta_x \in \mathbb{F}_q(u)$, thus establishing "Situation A". But we do not know of a general way how to do this. Therefore we must think of alternatives: changing the choice of $s_3, \ldots, s_{3t}$ and/or replacing the discriminant criterion by another argument that shows $\overline{G}$ is a full symmetric group. This is the topic of the next and final subsection.

## Discriminants—the remaining open cases

We treat (I) to (III) in reverse order. In all cases, because $j$ is even and $j \leq t-1$, we can assume $t \geq 4$ (even).

**(III): $A_{3t+1} = A_{3t-2} = 0,\ C_0 = 0,\ A_{3t-1} = 0,\ j$ even.**

In this case we have $g_d^{(0)} = 0$ for all $d$ and $g_0^{(2)} = 0$ (because $j > 1$). Therefore

$$\begin{aligned} F(x) \ = \ & [x^{3t+1} + x^2 g_{t-1}^{(2)}(x^3)] + \\ & + \ s_3 \cdot [x^{3t-2} + x^2 g_{t-2}^{(2)}(x^3)] + \\ & + \ \ldots \ + \\ & + \ s_{3t-3} \cdot [x^4] + \\ & + \ s_{3t} \cdot [x]\,. \end{aligned}$$

Write $g^{(2)}_{t-1}(x^3)$ as $x^{3z}g^*(x^3)$ with $g^*(0) \neq 0$. Under the given circumstances we have $z \in \{1, \ldots, t-2\}$. Now put

$$l := \begin{cases} 1, & \text{if } z \text{ is odd,} \\ 4, & \text{if } z \text{ is even,} \end{cases}$$

and choose $s_3 = u$, $s_{3t-l+1} = uv$. Then $F_u(x) = f_0(x) + uf_1(x)$ with

$$\begin{aligned} f_0(x) &= x^{3t+1} + x^{3z+2}\, g^*(x^3)\,, \\ f_1(x) &= x^{3t-2} + x^{3z-1}\, g^*(x^3) + vx^l\,, \end{aligned}$$

where (after excluding $v = 0$) the highest common power of $x$ is $x^l$ for all $z$. Hence we work with

$$\begin{aligned} f^*_0(x) &= x^{3t+1-l} + x^{3z+2-l}\, g^*(x^3)\,, \\ f^*_1(x) &= x^{3t-2-l} + x^{3z-1-l}\, g^*(x^3) + v\,. \end{aligned}$$

For $l = 1$ this parameter choice is the same as earlier in this section. For $l = 4$ we have, with $T := t - 1$,

$$\begin{aligned} f^*_0(x) &= x^{3T} + x\, g^{(2)}_{T-1}(x^3)\,, \\ f^*_1(x) &= x^{3T-3} + x\, g^{(2)}_{T-2}(x^3) + v\,, \end{aligned}$$

i.e. again the same polynomials as before, only with $T$ instead of $t$. Since $T \geq 3$, and since the first three subsections of this section do not depend on $t$ being even, the arguments for conditions (i)–(iii) of Lemma 8 remain valid in all cases. This establishes that $\overline{G}$ contains the alternating group of appropriate degree.

To conclude that $\overline{G}$ is the full symmetric group, it remains to find an odd permutation in it.

For $u = 0$ we get

$$F^*_0(x) = f^*_0(x) = x^{3z+2-l} \underbrace{\left[x^{3(t-z)-1} + g^*(x^3)\right]}_{=:\,\Lambda(x)}\,.$$

Thus we are done if we can show that $\Lambda(x)$ is square-free. Then $F^*_0$ factorises over any intermediate field of $\mathbb{T}/\overline{\mathbb{F}}_q$, in particular the inertia field of a place $\mathfrak{P}$ over $P_0$ (cf. page 25), as

$$F^*_0(x) = \prod_{i=1}^{r(0)} H_i(x)^{e_i}\,,$$

where $e_1 = 3z + 2 - l$ (for $H_1(x) := x$) and all other $e_i$ are 1. It follows, because

$$3z + 2 - l = \begin{cases} 3z + 1\,, & \text{if } z \text{ is odd,} \\ 3z - 2\,, & \text{if } z \text{ is even,} \end{cases}$$

is even and not divisible by the characteristic 3, that $\overline{G}$ contains a $(3z+2-l)$-cycle (the generator of $G^{(0)}$ with cycle pattern $e_1 \cdots e_{r(0)}$), which is an odd permutation.

We have $\Lambda'(x) = -x^{3(t-z)-2}$. This is clearly co-prime to $\Lambda(x)$. So $\Lambda(x)$ cannot have a repeated root, and the above conclusions apply.

(II): $A_{3t+1} = A_{3t-2} = 0,\ C_0 \neq 0,\ j$ even.

Basically, everything here is as in the previous case, except that the place of $x^2 g_d^{(2)}$ is taken by $(Cx^2+1)g_d^{(0)}$, and it is now $g_{t-1}^{(0)}(x^3)$ which we write as $x^{3z}g^*(x^3)$ with $g^*(0) \neq 0$. This time $z \in \{2, \ldots, t-2\}$.

For $z \geq 3$ proceed exactly as before. This leads to

$$F_0^*(x) \;=\; x^{3z-l} \underbrace{\left[x^{3(t-z)+1} + (Cx^2+1)g^*(x^3)\right]}_{=:\,\Lambda(x)},$$

where

$$3z-l \;=\; \begin{cases} 3z-1, & \text{if } z \text{ is odd,} \\ 3z-4, & \text{if } z \text{ is even,} \end{cases}$$

is again even and $\not\equiv 0 \pmod 3$, so that square-freeness of $\Lambda$, when established, yields a $(3z-l)$-cycle in $\overline{G}$ and proves it to be a symmetric group.

If $C = 0$, square-freeness of $\Lambda$ is shown as above. Difficulties arise when $C \neq 0$. A multiple root $\gamma$ of $\Lambda$ is also a root of

$$\Lambda(x) - x\Lambda'(x) \;=\; -(Cx^2-1)g^*(x^3)\,.$$

But $\gamma$ cannot be a root of $g^*$ (then $\gamma \neq 0$ and at the same time $0 = \Lambda(\gamma) = \gamma^{3(t-z)+1}$), so we must have $\gamma = \pm\sqrt{C_0/C_2}$. Since $\Lambda''(x) = -Cg^*(x^3)$ is co-prime to $\Lambda'(x)$, the only possibilities for multiple roots of $\Lambda$ are $\pm\sqrt{C_0/C_2}$ as double roots.

Coming back to the cycle pattern $e_1 \cdots e_{r(0)}$, there is a problem if and only if *exactly one* of $\pm\sqrt{C_0/C_2}$ is a double root of $\Lambda$: then the aforementioned generator of the inertia group is a composition of an even cycle and a transposition, thus even, and we do not get an odd permutation in $\overline{G}$.

Here is the way out of this: replace $l$ by

$$L \;:=\; \begin{cases} 4, & \text{if } z \text{ is odd,} \\ 1, & \text{if } z \text{ is even,} \end{cases}$$

(the opposite choice of before). Then $x^L$ cancels from $f_0(x)$ and $f_1(x)$, and $f_0^*, f_1^*, F_0^*$ are as before, but with $L$ instead of $l$; $\Lambda$ remains unaffected. As a result, $\overline{G}$ contains now the composition of an *odd* $(3z-L)$-cycle with the transposition corresponding to the single square factor of $\Lambda$. This is the desired odd permutation.

The case $z = 2$ needs to be treated separately, because here, with the same $l$ and parameter choice as before, the highest common power of $x$ in $f_0(x)$ and $f_1(x)$ is only $x^3 = x^{l-1}$. After cancelling this, we have

$$\begin{aligned} f_0^*(x) &= x^{3T+1} + x^2\, g_{T-1}^{(2)}(x^3) + g_{T-1}^{(0)}(x^3)\,, \\ f_1^*(x) &= x^{3T-2} + x^2\, g_{T-2}^{(2)}(x^3) + g_{T-2}^{(0)}(x^3) + vx\,, \end{aligned}$$

again with $T := t-1 \geq 3$. Up to replacing $t$ with $T$, these polynomials are identical with $f_0(x), f_1(x)$ in the original case where one of $A_{3t+1}, A_{3t-2}$ is non-zero (p. 48), so again the proofs of conditions (i)–(iii) from the earlier subsections carry over.

To show that $\overline{G}$ is the full symmetric group, we can this time use the familiar argument: $E_1^*(x) = vx^3(Cx^2-1)g_{t-3}^{(0)}(x^3)$, and $x^3 g_{t-3}^{(0)}(x^3)$ has odd degree $3(t-j-1)$, implying "Situation S".

**(I): $A_{3t+1} = 0$, $A_{3t-2} \neq 0$, $j$ even.**

The reader be warned that we have left the most tedious problem to the end! The strategy of the previous two cases fails here, because with it we get $f_1(0) \neq 0$ and

$$F_0(x) \;=\; x^3\left[x^{3t-2} + (Cx^2+1)g^*(x^3)\right]\,,$$

where the multiplicity of $x$ is divisible by 3. Instead, we will have to make completely new parameter choices and go through all conditions again. To complicate matters further, different arguments are needed for $j = 2$ and $j \geq 4$.

We begin with $j = 2$. Choose $s_{3t-3} = v$ and $s_{3t} = u$, then

$$\begin{aligned} f_0(x) &= x^{3t+1} + (Cx^2+1)g_{t-1}^{(0)}(x^3) + vx^4\,,\\ f_1(x) &= x\,, \end{aligned}$$

and after cancelling $x$,

$$\begin{aligned} f_0^*(x) &= x^{3t} + (Cx^2+1)x^2 g^*(x^3) + vx^3\,,\\ f_1^*(x) &= 1\,, \end{aligned}$$

where we write $g^*$ for $g_{t-2}^{(0)}$ to indicate its non-zero constant term. We aim at the conditions of Lemma 7.

**Co-primality.** Trivial.

**Indecomposability.** Assume $f_0^*(x) = Q(R(x))$ with $Q(x), R(x) \in \overline{\mathbb{F}}_q[x]$ monic of degrees $\omega > 1$ and $\rho > 1$, respectively.

Suppose first that $\deg Q' \geq 1$. Then we can play yet another variation of the "transformation trick" on page 30. Let $\gamma$ be a root of $Q'$ and put $\hat{Q}(x) := Q(x+\gamma)$, $\hat{R}(x) := R(x) - \gamma$. Then $\hat{Q}(\hat{R}(x))$ is another decomposition of $f_0^*(x)$ with the same degrees, and by the chain rule we have $\hat{Q}'(x) = Q'(x+\gamma)$, so $\hat{Q}'(0) = Q'(\gamma) = 0$. As always, we drop the "hats" and assume $Q'(0) = 0$. As a consequence, $R(x)$ divides $f_0^{*\prime}(x) = Q'(R(x)) \cdot R'(x)$, a fixed non-zero polynomial of degree at most $3t-6$, so there are at most $2^{3t-6}$ possibilities for $R$. Assume that one such $R$ appears with distinct $v$ and $\tilde{v}$. Then, in the usual way, we find $(v-\tilde{v})x^3 = Q(R(x)) - \tilde{Q}(R(x))$, which implies $\rho \cdot \deg(Q - \tilde{Q}) = 3$. Since $\rho > 1$, we must have $\deg(Q-\tilde{Q}) = 1$. But if $Q - \tilde{Q}$ is linear, it follows that $R(x) \in \overline{\mathbb{F}}_q[x^3]$, in contradiction to $R' \neq 0$. So there are at most $2^{3t-6}$ values of $v$ to exclude to avoid a decomposition with $\deg Q' \geq 1$.

Now we turn to the case $Q'(x) = q_1 \in \overline{\mathbb{F}}_q^*$, i.e. $Q(x)$ is of the form $Q_1(x^3) + q_1 x$. Then

$$x^{3t} + (Cx^2+1)x^2 g_{t-2}^{(0)}(x^3) + vx^3 \;=\; f_0^*(x) \;=\; Q_1(R^3(x)) + q_1 R(x)\,. \qquad (4.20)$$

Let us temporarily consider general $j \geq 2$ (even). From the expression on the left-hand side of (4.20), $f_0^*(x)$ contains the non-vanishing term $C_0 x^{3(t-j)-1}$, and on the

right-hand side the highest term that is not a term in $x^3$ can have degree at most $\rho$. Therefore, and because $Q'(x) = q_1$ forces $3|\omega$, we have

$$3(t-j)-1 \;\le\; \rho \;=\; \frac{3t}{\omega} \;\le\; \frac{3t}{3} \;=\; t\,,$$

from which it follows that $j \ge \frac{1}{3}(2t-1)$. Thus a decomposition with $Q'(x) = q_1$ is impossible for $j < \frac{1}{3}(2t-1)$. In particular, since $t \ge 4$, this holds always for $j = 2$.

**Simplicity.** Taking $\beta = 0$, we have

$$F_0^*(x) \;=\; x^2 \underbrace{\left[x^{3t-2} + (Cx^2+1)g^*(x^3) + vx\right]}_{=:\,\Lambda(x)}$$

and must make $\Lambda$ square-free by choice of $v$. A multiple root of $\Lambda$ is also a root of $\Lambda(x) - x\Lambda'(x) = -(Cx^2-1)g^*(x^3)$, a fixed non-zero polynomial of degree at most $2+3(t-j-1)$, and each such (non-zero) root $\gamma$ determines $v$ uniquely by $\Lambda'(\gamma) = 0$. Hence, for $j = 2$, exclusion of at most $3t-7$ values $v$ suffices to guarantee simplicity of $F_0^*$, so that all conditions of Lemma 7 are satisfied.

To cope with the case $j \ge 4$ we need two parameters $v_i$. Note that $j > 2$ means $g_1^{(0)} = 0$, so the choice $s_{3t-6} = v_1$, $s_{3t-3} = u$ and $s_{3t} = uv_2$ yields

$$\begin{aligned} f_0(x) &= x^{3t+1} + (Cx^2+1)g_{t-1}^{(0)}(x^3) + v_1x^7\,,\\ f_1(x) &= x^4 + v_2x\,. \end{aligned}$$

Therefore the polynomials to work with, after excluding $v_2 = 0$ and cancelling $x$, are (again with $g_{t-2}^{(0)} =: g^*$ because $g^*(0) \neq 0$)

$$\begin{aligned} f_0^*(x) &= x^{3t} + (Cx^2+1)x^2g^*(x^3) + v_1x^6\,,\\ f_1^*(x) &= x^3 + v_2\,. \end{aligned}$$

We aim again at Lemma 7.

**Co-primality.** For each $v_1 \in \overline{\mathbb{F}}_q$ there are at most $\deg f_0^* = 3t$ values of $v_2$ in $\overline{\mathbb{F}}_q$ for which we do not have co-primality. Exclude these $(v_1, v_2)$.

**Indecomposability.** Assume that $f_0^*/f_1^*$ has a decomposition as in Lemma 10 with $Q_1(0) = 0$. Analysis of the degrees shows that there are only two possibilities: $(\omega_1, \omega_2, \rho_1, \rho_2) = (t, 1, 3, 0)$ or $(3, 0, t, 1)$. A decomposition of the first type leads to the contradiction $f_0^*(x) = Q_1(R_1(x))$ with $R_1(x) \in \overline{\mathbb{F}}_q[x^3]$. So the challenge is to restrict for each $v_1 \in \overline{\mathbb{F}}_q$ the number of $v_2 \in \overline{\mathbb{F}}_q$ which allow a decomposition with $(\omega_1, \omega_2, \rho_1, \rho_2) = (3, 0, t, 1)$. Here we have $R_2(x) = x + V_2$, where $V_2 := \sqrt[3]{v_2}$. From $Q_1(0) = 0$ we conclude that $R_1(x)$ divides

$$f_0^*(x) \;=\; R_2^3(x) \cdot Q_1\left(\frac{R_1(x)}{R_2(x)}\right)\,,$$

which is a fixed polynomial (for fixed $v_1$). Hence the number of choices for $R_1$ is at most $2^{3t}$. We show that no two $v_2, \tilde{v}_2 \in \overline{\mathbb{F}}_q$, $v_2 \neq \tilde{v}_2$, can occur with the same $R_1$. Suppose

$$\frac{f_0^*(x)}{x^3+v_2} \;=\; Q_1\left(\frac{R_1(x)}{x+V_2}\right) \quad \text{and} \quad \frac{f_0^*(x)}{x^3+\tilde{v}_2} \;=\; \tilde{Q}_1\left(\frac{R_1(x)}{x+\tilde{V}_2}\right)\,,$$

where $\tilde{V}_2 := \sqrt[3]{\tilde{v}_2}$ and both $Q_1$ and $\tilde{Q}_1$ are cubic polynomials with zero constant term, say $Q_1(x) := x^3 + Ax^2 + Bx$ and $\tilde{Q}_1(x) := x^3 + \tilde{A}x^2 + \tilde{B}x$. Then

$$\begin{aligned} f_0^*(x) &= R_1^3(x) + A \cdot R_1^2(x) \cdot (x + V_2) + B \cdot R_1(x) \cdot (x + V_2)^2 = \\ &= R_1^3(x) + \tilde{A} \cdot R_1^2(x) \cdot (x + \tilde{V}_2) + \tilde{B} \cdot R_1(x) \cdot (x + \tilde{V}_2)^2 , \end{aligned}$$

and upon dividing through $R_1 \neq 0$ and re-arranging,

$$\begin{aligned} R_1(x) \left[ (A - \tilde{A})x + (AV_2 - \tilde{A}\tilde{V}_2) \right] &= \\ &= -\left[ (B - \tilde{B})x^2 - (BV_2 - \tilde{B}\tilde{V}_2)x + (BV_2^2 - \tilde{B}\tilde{V}_2^2) \right] . \end{aligned}$$

This is impossible:

- If $A \neq \tilde{A}$, then the degree of the LHS is exactly $t + 1 \geq 5$, while that of the RHS is 2 or smaller.
- If $A = \tilde{A} \neq 0$, then $AV_2 - \tilde{A}\tilde{V}_2 \neq 0$, so the degree of the LHS is exactly $t \geq 4$, still greater than that of the RHS.
- If $A = \tilde{A} = 0$, then both sides are identically zero. This implies $B = \tilde{B} = 0$, so that $Q_1(x), \tilde{Q}_1(x)$ and consequently $f_0^*(x)$ are in $\overline{\mathbb{F}}_q[x^3]$, which is not the case.

Thus indecomposability can be achieved by excluding at most $2^{3t}$ values of $v_2$ for each $v_1$.

**Simplicity.** Taking $\beta = 0$, we have

$$F_0^*(x) = x^2 \underbrace{\left[ x^{3t-2} + (Cx^2 + 1)g^*(x^3) + v_1 x^4 \right]}_{=: \Lambda(x)}$$

and must make $\Lambda$ square-free by choice of $v_1$. This can be done in exactly the same way as for $j = 2$. Thus, exclusion of at most $3t - 3j - 1$ values $v_1$ suffices to ensure the simplicity of $F_0^*$ (for a suitable $v_2$). So all conditions of Lemma 7 are satisfied, and the last part of the proof for the Exceptional Case is complete.

## 4.4 The "Degenerate Case": all coefficients $A_k$ with $k \leq 3t - 2$ are zero

This is the easiest of our three cases—and also the hardest! While the vanishing of all coefficients $A_2, \ldots, A_{3t-2}$ makes the calculations very simple, it also greatly reduces the scope for manipulating $f_0$ and $f_1$ to endow them with the desired properties. Indeed, as will become clear, the cases where only one of $A_{3t-1}, A_{3t+1}$ is non-zero each alone account for all restrictions on $t$ and $m$ in Theorem 8.

We distinguish again three obvious subcases.

$\boxed{A_{3t-1}A_{3t+1} \neq 0.}$

Choosing $s_3 = u$, one obtains $F(x) = F_u(x) = f_0(x) + uf_1(x)$ with

$$\begin{aligned} f_0(x) &= x^{3t+1} + A_{3t-1}x^2 + A_{3t+1}\,, \\ f_1(x) &= x^{3t-2}\,. \end{aligned}$$

For these, conditions (i)–(iii) of Lemma 8 are trivially satisfied (use Lemma 11(iii) for indecomposability). For the discriminant there is Lemma 12:

$$E_1(x) = (A_{3t-1}x^2 - A_{3t+1})x^{3t-3}\,,$$

and the part of this which is co-prime to $f_1(x)$ is $A_{3t-1}x^2 - A_{3t+1}$. Hence

$$\Delta_x = c \cdot (u + f(\alpha))(u + f(-\alpha)) \qquad \text{where } \alpha \in \overline{\mathbb{F}}_q^* \text{ with } \alpha^2 = \frac{A_{3t+1}}{A_{3t-1}}\,.$$

To show that $\Delta_x$ is a non-square in $\overline{\mathbb{F}}_q(u)$ one has to rule out $f(\alpha) = f(-\alpha)$. With $t$ even, we have $f_1(\alpha) = f_1(-\alpha)$ and $f_0(\alpha) \neq f_0(-\alpha)$, because $f_0(\alpha) = f_0(-\alpha)$ would be equivalent to $\alpha = 0$. Therefore Lemma 8 holds with "Situation S".

$\boxed{\text{The only non-zero coefficient is } A_{3t+1}.}$

With $s_3 = u$ as above, we have now

$$\begin{aligned} f_0(x) &= x^{3t+1} + A_{3t+1}\,, \\ f_1(x) &= x^{3t-2}\,. \end{aligned}$$

Again we aim at Lemma 8, where conditions (i)–(iii) hold trivially. But this time $E_1(x) = -A_{3t+1}x^{3t-3}$, so that the part co-prime to $f_1(x)$ is a constant (reflecting the fact that $\Delta_x \in \mathbb{F}_q$) and Lemma 12 is of no use. Instead, $F_u(x) = x^{3t+1} + ux^{3t-2} + A_{3t+1}$ is a trinomial, and we can use Theorem 3.87 of [LiNi] to determine its discriminant explicitly:

$$\Delta_x = (-1)^{3t(3t+1)/2}A_{3t+1}^{3t} = \begin{cases} A_{3t+1}^{3t}\,, & \text{if } t \equiv 0 \text{ or } 1 \pmod 4\,, \\ -A_{3t+1}^{3t}\,, & \text{if } t \equiv 2 \text{ or } 3 \pmod 4\,. \end{cases}$$

Since $\Delta_x$ is now an element of $\mathbb{F}_q$, the only situations where we get a result are those where $\Delta_x$ is a square in $\mathbb{F}_q$. We check the above cases individually for this:

- If $t \equiv 0 \pmod 4$, then $\Delta_x$ is an even power of $A_{3t+1}$ and thus always a square in $\mathbb{F}_q$.
- If $t \equiv 2 \pmod 4$, then $\Delta_x$ is the negative of an even power of $A_{3t+1}$. This is a square in $\mathbb{F}_q$ if and only if $-1$ is a square in $\mathbb{F}_q$, i.e. if and only if the degree $m$ of $\mathbb{F}_q$ over $\mathbb{F}_3$ is even.
- If $t \equiv 1 \pmod 4$, then $\Delta_x$ is an odd power of $A_{3t+1}$. This is a square in $\mathbb{F}_q$ if and only if $A_{3t+1}$ is already one.

- If $t \equiv 3 \pmod 4$, then $\Delta_x$ is the negative of an odd power of $A_{3t+1}$. This is a square in $\mathbb{F}_q$ if and only if $-1$ and $A_{3t+1}$ are both squares or both non-squares in $\mathbb{F}_q$.

In the first two cases we have successfully established Lemma 8 (with "Situation A") under the conditions of Theorem 8. In the latter two cases it depends on $A_{3t+1}$ whether or not $\Delta_x$ is a square in $\mathbb{F}_q$, and for half of the elements in $\mathbb{F}_q^*$ it will actually be a non-square. This forces us to disregard odd values of $t$.

As far as the last point is concerned, the question springs to mind whether Lemma 7 would offer a better alternative to Lemma 8. Unfortunately, it turns out that Lemma 7 cannot be used at all: for any possible choice of $s_3, \ldots, s_{3t}$ one gets

$$\begin{aligned} f_0(x) &= x^{3t+1} + A_{3t+1} + \Sigma_0(x)\,, \\ f_1(x) &= x^{3t-2} + \Sigma_1(x)\,, \end{aligned}$$

where $\Sigma_0(x)$ and $\Sigma_1(x)$ are polynomials in $\mathbb{F}_q[x]$ with all exponents $\equiv 1 \pmod 3$. Hence $F''_\beta(x) = f''_0(x) + \beta f''_1(x)$ is identically zero, so that every double root of $F_\beta$ is automatically a triple root and $F_\beta$ cannot be simple.

Therefore, in the Degenerate Case, we rely on the use of Lemma 8 with "Situation A"!

The only non-zero coefficient is $A_{3t-1}$.

This is very similar to the previous situation. After choosing $s_3 = u$ we must first cancel $x^2$ from $f_0(x)$ and $f_1(x)$ to arrange co-primality. Then we have

$$\begin{aligned} f_0^*(x) &= x^{3t-1} + A_{3t-1}\,, \\ f_1^*(x) &= x^{3t-4}\,. \end{aligned}$$

As before, conditions (i)–(iii) of Lemma 8 are obviously met, and by Theorem 3.87 of [LiNi] the discriminant $\Delta_x^*$ of $F_u^*(x) = x^{3t-1} + ux^{3t-4} + A_{3t-1}$ is

$$\Delta_x^* = (-1)^{3t(3t-1)/2} A_{3t-1}^{3t-2} = \begin{cases} A_{3t-1}^{3t-2}\,, & \text{if } t \equiv 0 \text{ or } 3 \pmod 4\,, \\ -A_{3t-1}^{3t-2}\,, & \text{if } t \equiv 1 \text{ or } 2 \pmod 4\,. \end{cases}$$

The cases individually:

- If $t \equiv 0 \pmod 4$, then $\Delta_x^*$ is an even power of $A_{3t-1}$, always a square in $\mathbb{F}_q$.
- If $t \equiv 2 \pmod 4$, then $\Delta_x^*$ is the negative of an even power of $A_{3t-1}$, a square in $\mathbb{F}_q$ if and only if $m$ is even.

With this the proof of Theorem 8 is complete. The remaining cases pose the same problems as before (with their roles interchanged):

- If $t \equiv 1 \pmod 4$, then $\Delta_x^*$ is the negative of an odd power of $A_{3t-1}$, a square in $\mathbb{F}_q$ iff $-1$ and $A_{3t-1}$ are both squares or both non-squares in $\mathbb{F}_q$.

- If $t \equiv 3 \pmod 4$, then $\Delta_x^*$ is an odd power of $A_{3t-1}$, a square in $\mathbb{F}_q$ iff $A_{3t-1}$ is one.

The use of Lemma 7 is again prohibited by the fact that for every choice of $s_3, \ldots, s_{3t}$ either $F_\beta'' = 0$ or $F_\beta'(x) = x \cdot F_\beta''(x)$, so that a non-zero root of $F_\beta'$ is automatically a root of $F_\beta''$, and $F_\beta$ cannot be simple. Thus Theorem 8 is the best possible result we can get on the basis of Theorem 6.

## 4.5 Failure of the approach for $t = 1$

Attempting to prove an analogous statement to Theorem 8 for $t = 1$ can serve as a simple example for a situation where the approach of Theorem 6 fails, i.e. where there exists always a choice of coefficients $A_k$ for which $F(x)$ does not split.

For $t = 1$ we find

$$F(x) \;=\; (x^4 + A_1x^3 + A_2x^2 + A_4) \;+\; s_3\,(x + A_1)$$

(where $A_1$ may be zero or not). The only influence we can exert here is by varying $s_3$.

Suppose $A_4 = -A_2A_1^2$, in which case

$$x^4 + A_1x^3 + A_2x^2 + A_4 \;=\; (x^3 + A_2x - A_2A_1)(x + A_1)\,,$$

so that

$$F(x) \;=\; (x^3 + A_2x - A_2A_1 + s_3)(x + A_1)\,.$$

Put $s_3 - A_2A_1 =: \hat{u}$. Then $F(x)$ splits if and only if $x^3 + A_2x + \hat{u}$ splits for some $\hat{u} \in \mathbb{F}_q$.

Now let in this situation $A_2$ be the negative of a non-square in $\mathbb{F}_q$. Then $F(x)$ cannot split, for if it did, say with

$$x^3 + A_2x + \hat{u} \;=\; (x - \alpha)(x - \beta)(x - \gamma)\,, \qquad \alpha, \beta, \gamma \in \mathbb{F}_q\,,$$

then comparing coefficients shows that $\alpha + \beta + \gamma = 0$ and $\alpha\beta + \alpha\gamma + \beta\gamma = A_2$, and substituting $\gamma = -(\alpha + \beta)$ in the second equation results in $-A_2 = (\alpha - \beta)^2$, contradicting the assumption made about $A_2$.

# Chapter 5

# Refinement of the Approach

Since failure of our method, like in the example of the last section of Chapter 4, does not allow any conclusions about the covering radius of BCH codes, there is a case for trying to strengthen the method. An obvious way for this which we will not pursue here would be to try to expand further the toolbar for showing $\overline{G} = G$. We begin with a heuristic argument why even without this it is reasonable to hope for more results.

## 5.1 Why there should be room for improvement

Recall from the end of Section 2.1 that we need to find for all $(a_k) \in \mathbb{F}_q^S$ an $\varepsilon \in \mathbb{F}_p$ such that the Helleseth system $\varepsilon\sigma_k = a_k$ has a solution $x_1, \ldots, x_r \in \mathbb{F}_q$.

The space $\mathbb{F}_p^* \times \mathbb{F}_q^S =: \mathcal{H}$ with $(p-1)q^s$ elements describes all possible Helleseth systems with $\varepsilon_1 = \ldots = \varepsilon_r =: \varepsilon$ for fixed $q$, $\delta$ and $r$ (all notation as in Section 2.1). Call $(\varepsilon, (a_k)) \in \mathcal{H}$ **solvable** if there exists a solution to the system $\varepsilon\sigma_k = a_k$ in $\mathbb{F}_q$. The aim is to identify a solvable subset of $\mathcal{H}$ which covers all $(a_k) \in \mathbb{F}_q^S$.

Clearly, if $(\varepsilon, (a_k))$ is solvable, then so is $(\vartheta\varepsilon, (\vartheta a_k))$ for any $\vartheta \in \mathbb{F}_p^*$.

For any $c \in \mathbb{F}_q$ and $\varepsilon \in \mathbb{F}_p^*$, the recursion

$$b_k := a_k - c^k r\varepsilon - \sum_{\mu=1}^{k-1} \binom{k}{\mu} c^{k-\mu} b_\mu \qquad (k = 1, \ldots, \delta - 1) \tag{5.1}$$

induces on $\mathbb{F}_q^S$ a well-defined bijection

$$\varphi_{c,\varepsilon} : \quad (a_k) \longmapsto (b_k) .$$

The inverse of $\varphi_{c,\varepsilon}$ is $\varphi_{-c,\varepsilon}$; more generally, for fixed $\varepsilon$ the group $(\{\varphi_{c,\varepsilon} : c \in \mathbb{F}_q\}, \circ)$ is isomorphic to $(\mathbb{F}_q, +)$. A non-recursive alternative expression for the $b_k$ is

$$b_k = (-c)^k r\varepsilon + \sum_{\nu=0}^{k-1} \binom{k}{\nu} (-c)^\nu a_{k-\nu} . \tag{5.2}$$

(We omit the proofs of these statements, which, where not obvious, are mainly exercises in the arithmetic of binomial coefficients.)

The map $\varphi_{c,\varepsilon}$ stems from Lemma 4 about linear change of variables: (5.1) is identical with (2.6). The lemma states that for a solvable $(\varepsilon,(a_k)) \in \mathcal{H}$ and $c \in \mathbb{F}_q$ the element $(\varepsilon, \varphi_{c,\varepsilon}(a_k)) \in \mathcal{H}$ is also solvable.

There are two obvious equivalence relations on $\mathcal{H}$:

$$(\varepsilon_z,(z_k)) \approx (\varepsilon_a,(a_k)) \quad :\Leftrightarrow \quad (\varepsilon_z,(z_k)) = (\vartheta\varepsilon_a,(\vartheta a_k)) \text{ for some } \vartheta \in \mathbb{F}_p^*,$$
$$(\varepsilon_z,(z_k)) \lessapprox (\varepsilon_a,(a_k)) \quad :\Leftrightarrow \quad \varepsilon_z = \varepsilon_a \ \wedge\ (z_k) = \varphi_{c,\varepsilon}(a_k) \text{ for some } c \in \mathbb{F}_q\,.$$

As seen above, both preserve solvability. Each $\approx$-class has size $p-1$. The $\lessapprox$-classes in the case $r \not\equiv 0 \pmod p$ all have size $q$, so there are $(p-1)q^{s-1}$ such classes, and as a system of representatives one can take $\mathbb{F}_p^* \times \{a_1\} \times \mathbb{F}_q^{S\setminus\{1\}}$ with any fixed element $a_1$ of $\mathbb{F}_q$. (For $r \equiv 0$ the situation is different, e.g. the equivalence classes have different sizes, and the map $\varphi_{c,\varepsilon}$ no longer depends on $\varepsilon$. We neglect this case.)

We want to combine $\approx$ and $\lessapprox$ into one equivalence relation to get large classes of solvable elements. For this some preparations.

**Lemma 15**

(i) *If* $(\varepsilon_a,(a_k)) \lessapprox (\varepsilon_b,(b_k)) \approx (\varepsilon_z,(z_k))$ *in* $\mathcal{H}$, *then there exists* $(\varepsilon_w,(w_k))$ *with* $(\varepsilon_a,(a_k)) \approx (\varepsilon_w,(w_k)) \lessapprox (\varepsilon_z,(z_k))$.

(ii) *For all* $(\varepsilon_a,(a_k)),(\varepsilon_z,(z_k)) \in \mathcal{H}$:
$$(\varepsilon_a,(a_k)) \approx (\varepsilon_z,(z_k)) \ \wedge\ (\varepsilon_a,(a_k)) \lessapprox (\varepsilon_z,(z_k)) \quad \Rightarrow \quad (\varepsilon_a,(a_k)) = (\varepsilon_z,(z_k))\,.$$

*Proof.*

(i) The hypothesis $(\varepsilon_a,(a_k)) \lessapprox (\varepsilon_b,(b_k)) \approx (\varepsilon_z,(z_k))$ means, using (5.2), that there are $c \in \mathbb{F}_q$ and $\vartheta \in \mathbb{F}_p^*$ with

$$\varepsilon_z = \vartheta\varepsilon_a \quad \text{and} \quad z_k = \vartheta b_k = (-c)^k r\vartheta\varepsilon_a + \sum_{\nu=0}^{k-1} \binom{k}{\nu} (-c)^\nu \vartheta a_{k-\nu} \quad (k \in S)\,.$$

Now it is easily checked that $(\varepsilon_w,(w_k)) := (\vartheta\varepsilon_a,(\vartheta a_k))$ has the desired property.

(ii) From $\lessapprox$ it follows that $\varepsilon_z = \varepsilon_a$, and from $\approx$ that $\varepsilon_z = \vartheta\varepsilon_a$. Therefore $\vartheta = 1$ and $(\varepsilon_z,(z_k)) = (\varepsilon_a,(a_k))$. □

From the first part of the lemma it is clear that

$$(\varepsilon_z,(z_k)) \sim (\varepsilon_a,(a_k)) \quad :\Leftrightarrow \quad \exists\,(\varepsilon_b,(b_k)) \in \mathcal{H}:\ (\varepsilon_a,(a_k)) \approx (\varepsilon_b,(b_k)) \lessapprox (\varepsilon_z,(z_k))$$

is again an equivalence relation on $\mathcal{H}$. Denote the $\sim$-class of $(\varepsilon,(a_k))$ by $[\varepsilon,(a_k)]_\sim$, similarly those of $\approx$ and $\lessapprox$. If $(\varepsilon,(a_k))$ is solvable then all elements of $[\varepsilon,(a_k)]_\sim$ are solvable. Part (i) of Lemma 15 shows further that

$$[\varepsilon,(a_k)]_\sim \;=\; \bigcup_{(\varepsilon_b,(b_k))\in[\varepsilon,(a_k)]_\approx} [\varepsilon_b,(b_k)]_{\lessapprox}\,,$$

and by part (ii) the union is disjoint, so that for $r \not\equiv 0 \pmod p$ all $\sim$-classes have size $(p-1)q$ and $|\mathcal{H}/\!\sim| = q^{s-1}$. A system of representatives for $\mathcal{H}/\!\sim$ is given by $\{\varepsilon\}\times\{a_1\}\times\mathbb{F}_q^{S\setminus\{1\}}$, where $\varepsilon$ is any fixed element of $\mathbb{F}_p^*$ and $a_1$ of $\mathbb{F}_q$.

In Section 2.4 we associated with each system $(\varepsilon,(a_k))$ an $(A_k)\in\mathbb{F}_q^S$ such that solving the splitting problem for $F(x)$ with the $A_k$ as coefficients implies solvability of the system. Let us write this now as

$$\begin{array}{rccc} \psi: & \mathcal{H} & \longrightarrow & \mathbb{F}_q^S \\ & (\varepsilon,(a_k)) & \longmapsto & \psi_\varepsilon(a_k) \end{array}$$

with

$$\begin{array}{rccc} \psi_\varepsilon: & \mathbb{F}_q^S & \longrightarrow & \mathbb{F}_q^S \\ & (a_k) & \longmapsto & (A_k)\,, \end{array}$$

where the $A_k$ are defined as in (2.9). For every $\varepsilon\in\mathbb{F}_p^*$ the map $\psi_\varepsilon$ is bijective; the inverse was given in (2.10).

For $r \not\equiv 0 \pmod p$, represent $\mathcal{H}/\!\sim$ as described with fixed $\varepsilon\in\mathbb{F}_p^*$ and $a_1\in\mathbb{F}_q$. With these, $A_1 = -\varepsilon^{-1}a_1$ is also fix, and $\psi$ induces a map

$$\begin{array}{rccc} \tilde\psi: & \mathcal{H}/\!\sim & \longrightarrow & \{A_1\}\times\mathbb{F}_q^{S\setminus\{1\}} \\ & [\varepsilon,(a_k)]_\sim & \longmapsto & (A_k)\,. \end{array} \tag{5.3}$$

This is the approach of Theorem 6. If in this situation the polynomial problem can be solved for all $(A_k)\in\{A_1\}\times\mathbb{F}_q^{S\setminus\{1\}}$, as for example under the hypotheses of Theorem 8, this shows that one representative of each $\sim$-class—and therefore *all* $(\varepsilon,(a_k))\in\mathcal{H}$—are solvable. But this is far more than necessary, since it would suffice to find *one* $\varepsilon$ for each $(a_k)$ such that $(\varepsilon,(a_k))$ is solvable!

Thus, in a sense, with Theorem 6 we prove a result that is $p-1$ times as strong as necessary. There is no reason to assume that this stronger result holds whenever the desired bound for the covering radius does; some situations may fall in between the two and require a more subtle approach.

Note that the considerations of this section make a difference only for $p>2$, in other words: here is a real novelty of the non-binary case.

## 5.2 A refined approach

The map $\tilde\psi$ in (5.3), which depends on $\varepsilon$ and $a_1$, is a bijection. On the other hand, the example in Section 4.5 shows that there can be $A_k\in\{A_1\}\times\mathbb{F}_q^{S\setminus\{1\}}$ for which

$F(x)$ does definitely *not* split over $\mathbb{F}_q$. Therefore, if we want to extend our results, we must abandon the unified choice of $\varepsilon$ and $a_1$.

Instead, in order to exploit as fully as possible the potential of our approach, the idea is to go through the process of Chapter 2 for each $(\varepsilon, (a_k)) \in \mathcal{H}$ *individually*. In practice this can be immediately realised as follows.

1. Start from $(a_k) \in \mathbb{F}_q^S$.
2. Take unspecified $\varepsilon \in \mathbb{F}_p^*$, $c \in \mathbb{F}_q$ and map $(a_k)$ to $(A_k)$ via $\psi_\varepsilon \circ \varphi_{c,\varepsilon}$:

$$\begin{array}{ccccc} \mathbb{F}_q^S & \xrightarrow{\varphi_{c,\varepsilon}} & \mathbb{F}_q^S & \xrightarrow{\psi_\varepsilon} & \mathbb{F}_q^S \\ (a_k) & \longmapsto & (b_k) & \longmapsto & (A_k)\,. \end{array}$$

3. Try to find a choice of $\varepsilon$ and $c$ for which splitting of $F(x)$ with coefficients $A_k$ can be proved.

If this procedure can be completed successfully for all $(a_k)$, the bound for the covering radius follows.

In the subsequent sections this refined approach will be put to the test in the case of ternary codes with designed distance $\delta \equiv 2 \pmod 3$. In this case $\varepsilon$ is one of $+1$ or $-1$, in particular $\varepsilon^{-1} = \varepsilon$, and the $(b_k) = \varphi_{c,\varepsilon}(a_k)$ and $(A_k) = \psi_\varepsilon(b_k)$ are given by (5.2) and

$$A_k \;=\; -k\varepsilon \sum_{\mu=0}^{k-1} b_{k-\mu} A_\mu$$

(with auxiliary element $A_0 = 1$). Moreover, we write $\square$ for the set of squares in $\mathbb{F}_q$ and $\boxslash$ for the set of non-squares, i.e. $\mathbb{F}_q$ is the disjoint union of $\square$, $\boxslash$ and $\{0\}$.

## 5.3 The case $t = 1$ ($\delta = 3t + 2$) re-visited

Let us first come back to the case $t = 1$ of Section 4.5. It turns out that with the improved method it is possible to obtain an analogue of Theorem 8 for odd $m$. Motivated by the fact that the method still fails for even $m$, we show then by more direct means that in this case the hypothesis $\rho = \delta - 1$ is actually false.

**Theorem 9** *Consider the primitive ternary BCH code $\mathcal{C}_{3,q-1,5}$ of length $q-1$ ($q = 3^m$) and designed distance 5.*

(i) *If $m \geq 9$ is odd, then the covering radius of $\mathcal{C}_{3,q-1,5}$ is exactly 4.*

(ii) *If $m \geq 4$ is even, then the covering radius of $\mathcal{C}_{3,q-1,5}$ is exactly 5.*

*Proof.*

We begin with part (i). Let $(a_1, a_2, a_4) \in \mathbb{F}_q^S$. The transformation $(A_k) = \psi_\varepsilon \circ \varphi_{c,\varepsilon}(a_k)$ yields

$$\begin{aligned} A_1 &= c - \varepsilon a_1\,, \\ A_2 &= \varepsilon a_2 - a_1^2\,, \\ A_4 &= -\varepsilon a_4 - \varepsilon c^2 a_2 + \varepsilon c a_1^3 + \varepsilon a_2 a_1^2 + c^2 a_1^2 - c a_2 a_1 - a_2^2 + a_1^4\,. \end{aligned}$$

The polynomial to split is $F(x) = x^4 + A_1 x^3 + A_2 x^2 + A_4 + s_3(x + A_1)$. With $s_3 = u$ this becomes $F_u(x) = f_0(x) + u f_1(x)$ with

$$\begin{aligned} f_0(x) &= x^4 + A_1 x^3 + A_2 x^2 + A_4\,, \\ f_1(x) &= x + A_1\,. \end{aligned}$$

One finds

$$f_0, f_1 \text{ are co-prime} \quad \Leftrightarrow \quad A_4 \neq -A_2 A_1^2 \quad \Leftrightarrow \quad \varepsilon a_4 + \varepsilon a_2 a_1^2 + a_2^2 \neq 0\,.$$

The last condition cannot be "transformed away" by choice of $c$. Thus, if we want to work with a particular $\varepsilon$, it is inevitable to distinguish two cases according to whether the last expression is zero or not. Two other expressions will govern our division into subcases in a similar way (their significance will become clear immediately in the course of the proof). The complete list is:

$$\begin{aligned} I_1 &:= \varepsilon a_4 + \varepsilon a_2 a_1^2 + a_2^2 &&= -A_4 - A_2 A_1^2\,, \\ I_2 &:= \varepsilon a_2 - a_1^2 &&= A_2\,, \\ I_3 &:= -\varepsilon a_4 + \varepsilon a_2 a_1^2 + a_2^2 - a_1^4 &&= A_4 - A_2^2 \text{ when } c = \varepsilon a_1. \end{aligned}$$

Seven subcases follow. These are not all disjoint, but the reader will quickly verify that they cover all possible $(a_1, a_2, a_4) \in \mathbb{F}_q^S$. By "$I_1$ for $-\varepsilon$" we mean the expression $-\varepsilon a_4 - \varepsilon a_2 a_1^2 + a_2^2$, etc.

Case 1: $I_1 I_2 I_3 \neq 0$ for some $\varepsilon \in \{1, -1\}$.

Take $\varepsilon$ with $I_1 I_2 I_3 \neq 0$, then use Lemma 8. Co-primality of $f_0$ and $f_1$ is settled by $I_1 \neq 0$, indecomposability of $f_0/f_1$ follows from the fact that $f_1$ is linear, and condition (iii) is obvious. The only serious problem is the discriminant $\Delta_x$.

For this employ Lemma 12. One finds $E_1(x) = A_2 x^2 - A_2 A_1 x - A_4$, and further $E_1(-A_1) = -A_4 - A_2 A_1^2 = I_1$, not zero by assumption, so that $E_1$ is co-prime to $f_1$ and all factors contribute to the discriminant.

By putting $c := \varepsilon a_1$ we can arrange $A_1 = 0$. Then $A_4$ becomes equal to $I_1 \neq 0$, and by assumption $A_2 \neq 0$, so that $E_1$ has roots $\pm\alpha \in \overline{\mathbb{F}}_q^*$, where $\alpha^2 = A_4/A_2$. To show that $\Delta_x$ is a non-square in $\overline{\mathbb{F}}_q(u)$, and thus establish "Situation S" of Lemma 8, it suffices now to show $f(\alpha) \neq f(-\alpha)$. But $f(\alpha) = f(-\alpha)$ is equivalent (for $c = \varepsilon a_1$) to $A_4 - A_2^2 = 0$, which was excluded by $I_3 \neq 0$.

Case 2: $I_1 \neq 0 \wedge I_2 = 0$ for one $\varepsilon \in \{1,-1\}$, $I_1 = 0$ for $-\varepsilon$.

Together the conditions imply $a_4 = 0$ and $a_2 = \varepsilon a_1^2 \neq 0$. Work with $\varepsilon$ as described (and arbitrary $c$). As before, co-primality, indecomposability and condition (iii) of Lemma 8 are clear. The general expression for the discriminant (computed with Maple) is

$$\Delta_x \;=\; -A_2^3u^2 + (A_2^4A_1 - A_2^3A_1^3)u + (A_4^3 + A_4^2A_2^2 + A_4A_2^4 - A_4A_2^3A_1^2)\,.$$

Here, thanks to $A_2 = 0$, this reduces to $\Delta_x = A_4^3 = a_1^{12}$, a square in $\mathbb{F}_q$, so that we have "Situation A".

Case 3: $I_1 = 0$ for both $\varepsilon \in \{1,-1\}$.

(We can take any $\varepsilon$.) The condition implies $a_2 = a_4 = 0$; in particular, $A_2$ simplifies to $-a_1^2$. After cancelling $x + A_1$ from $f_0(x)$ and $f_1(x)$ we must show that

$$F_u^*(x) \;=\; f_0^*(x) + uf_1^*(x) \;=\; x^3 + A_2x - A_2A_1 + u$$

splits. Choose $u = A_2A_1$, then

$$F_u^*(x) \;=\; x^3 - a_1^2x \;=\; x(x + a_1)(x - a_1)\,.$$

Case 4: $I_2 = 0$ for both $\varepsilon \in \{1,-1\}$.

The condition implies $a_1 = a_2 = 0$, so assume $a_4 \neq 0$. Then $f_0$ and $f_1$ are co-prime. Conditions (i)–(iii) of Lemma 8 hold, and the discriminant is

$$\Delta_x \;=\; A_4^3 \;=\; \begin{cases} -a_4^3 & \text{for } \varepsilon = 1\,, \\ a_4^3 & \text{for } \varepsilon = -1\,. \end{cases}$$

In the case of odd $m$ one of these is a square in $\mathbb{F}_q$ and we can arrange "Situation A" by choice of $\varepsilon$. (For even $m$ there is nothing we can do if $a_4 \in \square$.)

Case 5: $I_2 = 0$ for one $\varepsilon \in \{1,-1\}$, $I_3 = 0$ for $-\varepsilon$.

From the conditions it follows that $\varepsilon a_4 = a_1^4$ and further $A_1 = c - \varepsilon a_1$, $A_2 = A_4 = 0$. Thus with $c := \varepsilon a_1$ we get $A_k = 0$ for all $k$, and $F_0(x) = f_0(x) = x^4$ splits trivially.

Case 6: $I_2 \neq 0 \wedge I_3 = 0$ for both $\varepsilon \in \{1,-1\}$.

$I_3 = 0$ for both $\varepsilon$ implies $a_1^4 - a_2^2 = 0$ and therefore $a_2 = \pm a_1^2$. This contradicts $I_2 \neq 0$ for one of the $\varepsilon$. So this case cannot occur.

Case 7: $I_1 = 0$ for one $\varepsilon \in \{1,-1\}$, $I_1 \neq 0 \wedge I_2 \neq 0 \wedge I_3 = 0$ for $-\varepsilon$.

Here we find $0 = -\varepsilon a_2 a_1^2 - a_1^4 = a_1^2(-\varepsilon a_2 - a_1^2)$ and so $a_1 = 0$, $-\varepsilon a_4 = a_2^2 \neq 0$. With $c := 0$ the $A_k$ become $A_1 = A_4 = 0$ and $A_2 = \varepsilon a_2$. For odd $m$, exactly one of $\pm\varepsilon a_2$ is a square.

Assume first that $-\varepsilon a_2 = \beta^2$, $\beta \in \mathbb{F}_q$. With $u := 0$ this means $F_u(x) = x^4 - \beta^2 x^2$, which splits as $x^2(x+\beta)(x-\beta)$.

Now suppose $\varepsilon a_2 = \beta^2$ with $\beta \in \mathbb{F}_q$. In this case work with $-\varepsilon$ instead. Then $I_1 \neq 0$ implies co-primality of $f_0, f_1$; conditions (ii) and (iii) of Lemma 8 are satisfied as usual. In addition we have now $A_1 = 0$, $A_2 = -\varepsilon a_2$ and $A_4 = a_2^2 = A_2^2$, so the discriminant is (cf. Case 2)

$$\Delta_x \;=\; -A_2^3 u^2 \;=\; \varepsilon a_2^3 u^2 \;=\; (\beta^3 u)^2$$

and we are in "Situation A". This completes the proof of part (i) of Theorem 9.

To prove part (ii), we establish first that the covering radius of $\mathcal{C}_{3,q-1,5}$ must be greater than 4 when $m$ is even. By Theorem 4, this follows if there exists one $(a_1, a_2, a_4) \in \mathbb{F}_q^S$ for which the system

$$\begin{aligned} \varepsilon_1 x_1 + \varepsilon_2 x_2 + \varepsilon_3 x_3 + \varepsilon_4 x_4 &= a_1 \\ \varepsilon_1 x_1^2 + \varepsilon_2 x_2^2 + \varepsilon_3 x_3^2 + \varepsilon_4 x_4^2 &= a_2 \\ \varepsilon_1 x_1^4 + \varepsilon_2 x_2^4 + \varepsilon_3 x_3^4 + \varepsilon_4 x_4^4 &= a_4 \end{aligned} \tag{5.4}$$

has no solution with $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4 \in \{1,-1\}$ and $x_1, x_2, x_3, x_4 \in \mathbb{F}_q$.

A look back at the proof of part (i) shows that in most cases the system *does* have a solution (with all $\varepsilon$ equal), for both odd and even $m$. Only in two situations we could not show splitting of $F(x)$ for even $m$:

- $a_1 = a_2 = 0$, $a_4 \in \square$ (Case 4), (5.5)
- $a_1 = 0$, $a_2 \in \square$, $a_4 = \pm a_2^2$ (Case 7). (5.6)

We prove now that for (5.5) the system (5.4) has indeed no $(3,q)$-solution when $m$ is even. Since even $m$ implies that the negative of $a_4$ is also a non-square, we need only consider the possibilities that exactly four, three or two of the $\varepsilon_i$ are $+1$.

Let us first deal with the system

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 0 \\ x_1^2 + x_2^2 + x_3^2 + x_4^2 &= 0 \\ x_1^4 + x_2^4 + x_3^4 + x_4^4 &= a_4 \in \square . \end{aligned}$$

The first equation says $x_1 + x_2 = -x_3 - x_4$. Squaring both sides and subtracting the second equation yields $x_1 x_2 = (x_3 - x_4)^2$. Therefore

$$x_1^4 + x_2^4 \;=\; (x_1^2 + x_2^2)^2 + (x_1 x_2)^2 \;=\; (-x_3^2 - x_4^2)^2 + (x_3 - x_4)^4 ,$$

and substituting this in the last equation gives $a_4 = -x_3x_4(x_3-x_4)^2$. By symmetry,

$$a_4 = -x_ix_j(x_i-x_j)^2 \qquad \text{for } i,j \in \{1,2,3,4\},\ i \neq j\,.$$

Now let $x_1 = \alpha$, $x_2 = \beta$, $x_3 = \gamma$, $x_4 = \delta$ be a solution in $\overline{\mathbb{F}}_q$. For fixed $\alpha$, the other values are roots of

$$a_4 = -\alpha x(\alpha - x)^2 = -\alpha x^3 - \alpha^2 x^2 - \alpha^3 x\,,$$

or (since we can assume without loss $\alpha \neq 0$) of

$$x^3 + \alpha x^2 + \alpha^2 x + \frac{a_4}{\alpha} = 0\,.$$

Dividing out the factor $x-\beta$ (say), and applying the formula for quadratic equations to the remainder, we find that

$$\gamma, \delta = \alpha + \beta \pm \sqrt{\alpha\beta}\,.$$

Suppose the solution lies completely in $\mathbb{F}_q$. This is the case if and only if $\alpha\beta \in \square$. But then $a_4 = -\alpha\beta(\alpha-\beta)^2 \in \square$, a contradiction.

Next we turn to

$$\begin{aligned} x_1 + x_2 + x_3 - x_4 &= 0 \\ x_1^2 + x_2^2 + x_3^2 - x_4^2 &= 0 \\ x_1^4 + x_2^4 + x_3^4 - x_4^4 &= a_4 \in \boxslash\,. \end{aligned}$$

Proceeding as above, we find $x_1x_2 = x_3^2 - x_3x_4$, $a_4 = x_3^2x_4(x_3-x_4)$, and by symmetry

$$a_4 = x_i^2x_4(x_i - x_4) \qquad \text{for } i = 1,2,3\,.$$

Let $x_1 = \alpha$, $x_2 = \beta$, $x_3 = \gamma$, $x_4 = \delta$ be a solution in $\overline{\mathbb{F}}_q$. Fix $\delta$. We can assume $\delta \neq 0$, since we cannot have a solution in $\mathbb{F}_q$ with $\delta = 0$ (this would also be one for all $\varepsilon$ equal to 1). Continue as above by dividing out $x - \alpha$ from $x^3 - \delta x^2 - a_4/\delta$ to obtain

$$\beta, \gamma = \alpha - \delta \pm \sqrt{\delta(\delta - \alpha)}\,.$$

Suppose the solution is in $\mathbb{F}_q$. That is the case exactly if $\delta(\delta-\alpha) \in \square$ or $\delta - \alpha = 0$. But then $a_4 = -\alpha^2 \cdot \delta(\delta-\alpha) \in \square$ or $a_4 = 0$, either of which is a contradiction.

The last possibility to examine is that

$$\begin{aligned} x_1 + x_2 - x_3 - x_4 &= 0 \\ x_1^2 + x_2^2 - x_3^2 - x_4^2 &= 0 \\ x_1^4 + x_2^4 - x_3^4 - x_4^4 &= a_4 \in \boxslash \end{aligned}$$

has a solution. But here the above strategy leads (via $x_1x_2 = x_3x_4$) immediately to the contradiction $a_4 = 0$.

It remains to show that the covering radius of $\mathcal{C}_{3,q-1,5}$ is at most 5. To do this, it suffices to show that the system

$$\begin{aligned} \varepsilon x_1 + \varepsilon x_2 + \varepsilon x_3 + \varepsilon x_4 + \varepsilon x_5 &= a_1 \\ \varepsilon x_1^2 + \varepsilon x_2^2 + \varepsilon x_3^2 + \varepsilon x_4^2 + \varepsilon x_5^2 &= a_2 \\ \varepsilon x_1^4 + \varepsilon x_2^4 + \varepsilon x_3^4 + \varepsilon x_4^4 + \varepsilon x_5^4 &= a_4 \end{aligned} \tag{5.7}$$

has a $(3,q)$-solution for $(a_1,a_2,a_4)$ as in (5.5) and (5.6); for all other $(a_k) \in \mathbb{F}_q^S$ such a solution exists with $x_5 = 0$, as the proof of part (i) shows.

We use Theorem 6. The relevant polynomial is $F_u(x) = f_0(x) + uf_1(x)$ with

$$\begin{aligned} f_0(x) &= x^5 + A_1x^4 + A_2x^3 + A_4x + A_5\,, \\ f_1(x) &= x^2 + A_1x + A_2\,, \end{aligned}$$

where $(A_k) = \psi_\varepsilon \circ \varphi_{c,\varepsilon}(a_k)$ for some $\varepsilon \in \mathbb{F}_p^*$, $c \in \mathbb{F}_q$. The map $\varphi_{c,\varepsilon}$ depends on $r$, which is now $\equiv 2 \pmod 3$, but this does not come into effect if we choose $c := 0$. Then the $A_k$ are, as for $r = 4$,

$$\begin{aligned} A_1 &= -\varepsilon a_1\,, \\ A_2 &= \varepsilon a_2 - a_1^2\,, \\ A_4 &= -\varepsilon a_4 + \varepsilon a_2 a_1^2 - a_2^2 + a_1^4\,. \end{aligned}$$

The "fill-up element" $a_5$, and thus $A_5$, can be chosen arbitrarily in $\mathbb{F}_q$.

$\boxed{a_1 = a_2 = 0,\ a_4 \in \square.}$

This is the case $A_1 = A_2 = 0$, $A_4 = \pm a_4 \neq 0$.

$$\begin{aligned} f_0(x) &= x^5 + A_4x + A_5\,, \\ f_1(x) &= x^2\,. \end{aligned}$$

Use Lemma 8. Avoid $A_5 = 0$ to make $f_0$ and $f_1$ co-prime. Indecomposability follows from $\deg f_0 = 5$ prime, and condition (iii) is clear. $E_1(x) = -A_4x^2 + A_5x$; the part of this which is co-prime to $f_1(x)$ is $-A_4x + A_5$, of odd degree. Therefore we are in "Situation S".

$\boxed{a_1 = 0,\ a_2 \in \square,\ a_4 = \pm a_2^2.}$

Take $\varepsilon := -1$ for $a_4 = a_2^2$ and $\varepsilon := 1$ for $a_4 = -a_2^2$, then this is the case $A_1 = A_4 = 0$, $A_2 \neq 0$.

$$\begin{aligned} f_0(x) &= x^5 + A_2x^3 + A_5\,, \\ f_1(x) &= x^2 + A_2\,. \end{aligned}$$

As before, "Situation S" of Lemma 8 can be established: $f_0(x) = x^2 f_1(x) + A_5$, so to ensure co-primality one must avoid $A_5 = 0$, and then $E_1 = A_5x$ has automatically degree 1.

It follows that for even $m$ the covering radius of $\mathcal{C}_{3,q-1,5}$ is exactly 5. □

When $A_1 = 0$, (5.5) corresponds for $t > 1$ to the Degenerate Case with $A_{3t+1}$ as the only non-zero coefficient. The proof of part (ii) shows that it is definitely a "spoiler". Coding-theoretically, this means that certain codewords of the supercode $\mathcal{C}_{3,q-1,3}$, all of which satisfy the first two equations of (5.4), fail to have distance $\leq 4$ from a codeword of $\mathcal{C}_{3,q-1,5}$.

Despite trying similar techniques, it has not been possible to decide the same problem for (5.6), which corresponds in the same way to the remaining two instances of the Degenerate Case. Neither is it clear what the coding-theoretical interpretation of this situation is.

## 5.4 No new results for even $t$

The case $t = 1$ shows that our refined approach can indeed lead to better results. We ask next: is it possible with the flexible choice of $\varepsilon$ and $c$ to get rid of the restriction in Theorem 8 that $m$ must be even for $t \equiv 2 \pmod 4$?

The answer is no. To see this, take again, as in (5.5), $(a_k) \in \mathbb{F}_q^S$ with $a_k = 0$ for $k = 1, \ldots, 3t$[†] and $a_{3t+1} \neq 0$. Then $(A_k) = \psi_\varepsilon \circ \varphi_{c,\varepsilon}(a_k)$ is given for general $\varepsilon$ and $c$ by

$$\begin{aligned} A_1 &= c\,, \\ A_k &= 0 \quad \text{for } k = 2, \ldots, 3t\,, \\ A_{3t+1} &= -\varepsilon a_{3t+1} \neq 0\,. \end{aligned}$$

For $c = 0$ we are exactly in the situation of the Degenerate Case (Section 4.4), second subcase, which led to the restriction on $m$ in the first place.

For $c \neq 0$ we have $A_1 \neq 0$ and

$$\begin{aligned} F(x) \;=\; & [x^{3t+1} + A_1 x^{3t} + A_{3t+1}] \;+ \\ & + \quad s_3 \cdot [x^{3t-2} + A_1 x^{3t-3}] \;+ \\ & + \quad \ldots \quad + \\ & + \; s_{3t-3} \cdot [x^4 + A_1 x^3] \;+ \\ & + \quad s_{3t} \cdot [x + A_1]\,. \end{aligned}$$

With this neither Lemma 7 nor Lemma 8 can cope:

- There is no power of $x$ of the form $x^{3i+2}$, and for no choice of $s_3, \ldots, s_{3t}$ a common factor of $f_0, f_1$ that could cancel. Therefore, in Lemma 7, we have always $F_\beta'' = 0$ identically, so $F_\beta$ cannot be simple.
- The polynomial $f_1(x)$ is inevitably of the form

$$\sum_\nu s_{3\nu}(x^{3t+1-3\nu} + A_1 x^{3t-3\nu}) \;=\; (x + A_1) \cdot \underbrace{\sum_\nu s_{3\nu} x^{3(t-\nu)}}_{\text{a polynomial in } x^3}\,,$$

  so condition (iii) of Lemma 8 is always violated.

This puts already an end to our hopes of extending Theorem 8 when $t$ is even. Nevertheless, to get a more complete picture, it is worth having a look also at the case $a_{3t-1} \neq 0$ and $a_k = 0$ for all other $k$. (This is the analogue of (5.6) for $t > 1$, as Section 5.7 will show.) Here the $A_k$ have values

$$\begin{aligned} A_1 &= c\,, \\ A_k &= 0 \quad \text{for } k = 2, \ldots, 3t-2 \text{ and } k = 3t\,, \\ A_{3t-1} &= \varepsilon a_{3t-1} \neq 0\,, \\ A_{3t+1} &= -c^2 \varepsilon a_{3t-1}\,. \end{aligned}$$

[†]Here and in the sequel we tacitly identify an $S$-tuple $(a_k) \in \mathbb{F}_q^S$ with the appropriate $r$-tuple $(a_k) \in \mathbb{F}_q^r$.

Again $c = 0$ reproduces the situation of Section 4.4 (third subcase), and for $c \neq 0$ we find

$$\begin{aligned} F(x) \;=\; & [x^{3t+1} + A_1 x^{3t} + A_{3t-1} x^2 + A_{3t+1}] \;+ \\ & + \quad s_3 \cdot [x^{3t-2} + A_1 x^{3t-3}] \;+ \\ & + \quad \ldots \quad + \\ & + \; s_{3t-3} \cdot [x^4 + A_1 x^3] \;+ \\ & + \quad s_{3t} \cdot [x + A_1]\,. \end{aligned} \tag{5.8}$$

Each expression in square brackets is divisible by $x+c$, so co-prime polynomials can only be obtained from

$$F^*(x) \;:=\; \frac{F(x)}{x+c} \;=\; [x^{3t} + \varepsilon a_{3t-1} x - c\varepsilon a_{3t-1}] + \sum_{\nu=1}^{t} s_{3\nu} x^{3(t-\nu)}\,.$$

But for this the same considerations as above apply.

Not willing to give up yet, we will return to these particular problem situations in the next section to try out an idea that suggests itself exclusively to $(a_k) \in \mathbb{F}_q^S$ with very few non-zero entries.

## 5.5 Working with one $\varepsilon_i$ different from the rest

A disadvantage of our method is that it can only handle the special case of (2.3) where all $\varepsilon_i$ are equal. Thus we gain no information about the solvability of systems with general $\varepsilon_1, \ldots, \varepsilon_r \in \mathbb{F}_p^*$, and much of the power of Theorem 4 is lost.

In general, nothing can be done about this. However, for individual $(a_k) \in \mathbb{F}_q^S$ it is at least possible to choose *one* $\varepsilon_i$ different from the rest, provided almost all of the $a_k$ are zero. In this section we describe first a general way how to do this, and then examine whether this additional degree of freedom finally allows to find a solution for the $(a_k)$ of the last section.

In the system (2.3), take $\varepsilon_i := \varepsilon$ for $i = 1, \ldots, r-1$ and $\varepsilon_r := \vartheta\varepsilon$ with $\vartheta \in \mathbb{F}_p^*$:

$$\begin{array}{ccccccccc} \varepsilon x_1 & + & \ldots & + & \varepsilon x_{r-1} & + & \vartheta\varepsilon x_r & = & a_1 \\ \vdots & & & & \vdots & & \vdots & & \vdots \\ \varepsilon x_1^k & + & \ldots & + & \varepsilon x_{r-1}^k & + & \vartheta\varepsilon x_r^k & = & a_k \\ \vdots & & & & \vdots & & \vdots & & \vdots \\ \varepsilon x_1^{\delta-1} & + & \ldots & + & \varepsilon x_{r-1}^{\delta-1} & + & \vartheta\varepsilon x_r^{\delta-1} & = & a_{\delta-1}\,. \end{array}$$

Put $y_i := x_i / x_r$ $(i = 1, \ldots, r-1)$ and divide the $k$-th equation by $x_r^k$ to obtain the equivalent system

$$\begin{array}{ccccccc} \varepsilon y_1 & + & \ldots & + & \varepsilon y_{r-1} & = & \hat{a}_1 \\ \vdots & & & & \vdots & & \vdots \\ \varepsilon y_1^k & + & \ldots & + & \varepsilon y_{r-1}^k & = & \hat{a}_k \\ \vdots & & & & \vdots & & \vdots \\ \varepsilon y_1^{\delta-1} & + & \ldots & + & \varepsilon y_{r-1}^{\delta-1} & = & \hat{a}_{\delta-1} \end{array} \tag{5.9}$$

with

$$\hat{a}_k := \frac{a_k}{x_r^k} - \vartheta\varepsilon .$$

Then go through the procedure of Section 2.3 and use the usual method.

There is a slight difficulty with the last step. In Section 2.3 we took into consideration the possibility that a system may have more variables than equations ($r > \delta - 1$); this was compensated by adding "dummy equations". With (5.9), in contrast, it is possible to have *one more equation* than variables, namely if $r = \delta - 1$. In this case the Newton identities impose an additional condition on $s_p, s_{2p}, \ldots, s_{tp}$. The practical examples below will illuminate this point.

The "eliminated" variable $x_r$ is still implicitly present in the $\hat{a}_k$ for $a_k \neq 0$. This is the reason why this approach becomes quickly unusable when several $a_k$ are non-zero.

Now let $p = 3$, $\delta = 3t + 2$ and $r = \delta - 1$. Assume $t \geq 2$ (even or odd). We apply the above to $(a_k)$ with either $a_{3t-1}$ or $a_{3t+1}$ as the only non-zero entry, beginning with the former case.

In $\mathbb{F}_3$ we can take without loss $\varepsilon = 1$ and $\vartheta = -1$. Then the modified system is

$$y_1^k + \ldots + y_{3t}^k = \hat{a}_k \qquad (k \in S) \tag{5.10}$$

where $\hat{a}_{3t-1} = X + 1$ with $X := a_{3t-1}/x_{3t+1}^{3t-1}$, $x_{3t+1}$ arbitrary in $\mathbb{F}_q$, and all other $\hat{a}_k$ are equal to 1.

Going through the procedure in the frame on p. 13, but taking $k$ only up to $3t$ in the second step, yields inductively (with details left to the reader)

$$s_{3j+1} = s_{3j}\,, \quad s_{3j+2} = 0 \qquad \text{for } j = 0, \ldots, t-2$$

and

$$s_{3t-2} = s_{3t-3}\,, \quad s_{3t-1} = (-1)^{t+1}X\,.$$

By Lemma 5, the roots of

$$F(y) = \sum_{l=0}^{3t}(-1)^l s_l y^{3t-l} = (-1)^t s_{3t} + (y-1)\sum_{l=0}^{t-1}(-1)^l s_{3l} y^{3t-3l-1} + Xy$$

solve the equations of (5.10) for $k \leq 3t$. It remains to satisfy also the equation for $k = 3t + 1$. Theorem 5 with $3t$ variables states for $k = 3t + 1$ that

$$\sum_{l=0}^{3t}(-1)^l \sigma_{3t+1-l} S_l = 0\,;$$

by replacing $S_l$ with $s_l$ and $\sigma_k$ with $\varepsilon^{-1}\hat{a}_k$ we get the condition

$$s_{3t} = (-1)^{t+1}\frac{1}{\hat{a}_1}\sum_{l=0}^{3t-1}(-1)^l \hat{a}_{3t+1-l} s_l \tag{5.11}$$

that extends $\varepsilon\sigma_k(y_1,\ldots,y_{3t}) = \hat{a}_k$ also to $k = 3t+1$ (cf. the proof of Lemma 5). This determines $s_{3t}$ completely as $(-1)^{t+1}X$. So the polynomial for which we have to study the splitting problem finally takes the form

$$F(y) \;=\; (y-1)\cdot\Big[X + \sum_{l=0}^{t-1}(-1)^l s_{3l} y^{3t-3l-1}\Big].$$

From this an interesting conclusion can be drawn. Obviously, 1 is always a root of $F(y)$, so any solution to (5.10) found in this way will have $y_i = 1$ for at least one $i \in \{1,\ldots,3t\}$. This means that $x_i = x_{3t+1}$, and the columns for $x_i$ and $x_{3t+1}$ in the original system cancel out with each other. In other words, any solution from $F(y)$ implies one to the original system with all $\varepsilon_i$ equal and at most $3t-1$ variables different from zero. Such a solution would already have been found in the previous section, where the original system was studied with arbitrary $\varepsilon$ and $c$ (in fact, already in Section 4.4, since (5.8) can have a double root 0 only if $c = 0$). In a word: the change of only one $\varepsilon_i$ produces nothing new here.

If $a_{3t+1}$ is the only non-zero entry of $(a_k)$, the result is less conclusive. In the analogous way one calculates

$$F(y) \;=\; -X \;+\; (y-1)\cdot\sum_{l=0}^{t-1}(-1)^l s_{3l} y^{3t-3l-1}$$

with $X := a_{3t+1}/x_{3t+1}^{3t+1}$ and $x_{3t+1}$ arbitrary in $\mathbb{F}_q$. We show that the use of Lemma 8 (assuming that conditions (i)–(iii) are satisfied) can never lead to "Situation S" because the discriminant $\Delta_y$ is always an element of $\mathbb{F}_q$. Let $I, J \subseteq \{1,\ldots,t-1\}$ with $I \cap J = \emptyset$ and $1 \in J$ be sets such that for $\nu \in I$ we put $(-1)^\nu s_{3\nu} := w_\nu$ and for $\mu \in J$ we put $(-1)^\mu s_{3\mu} := uv_\mu$. Then $F(y) = F_u(y) = f_0(y) + uf_1(y)$ with

$$f_0(y) \;=\; (y-1)\cdot\underbrace{\Big[y^{3t-1} + \sum_{\nu\in I} w_\nu y^{3(t-\nu)-1}\Big]}_{=:\,y^2G_0(y^3)} - X\,,$$

$$f_1(y) \;=\; (y-1)\cdot\underbrace{\sum_{\mu\in J} v_\mu y^{3(t-\mu)-1}}_{=:\,y^2G_1(y^3)}\,,$$

and in Lemma 12 one finds $E_1(y) = XyG_1(y^3)$. Since $yG_1(y^3)$ divides $f_1(y)$, the part of $E_1$ co-prime to $f_1$ is a constant, and consequently $\Delta_y \in \mathbb{F}_q$.

Calculations with Maple for small values of $t$ yielded $\Delta_y = (-1)^{3t(3t-1)/2}X^{3t-2}$ in all cases (regardless of the choice of the $s_{3l}$). If true in general, this formula would again only confirm the results already found in Section 4.4: $X^{3t-2}$ is a square iff $a_{3t+1}^{3t-2}$ is a square, and so $\Delta_y$ would be always a square for $t \equiv 0 \pmod 4$, for $t \equiv 2$ exactly if $m$ is even, and for odd $t$ for exactly half of the values of $a_{3t+1}$ (for either parity of $m$).

In particular, another attempt to extend Theorem 8 to $t \equiv 2$ and odd $m$ has been unsuccessful. The persistent failure of these attempts seems to be an indication

that in this situation the hypothesis $\rho = \delta - 1$ may, as for $t = 1$, actually be false. We will not endeavour in this work to prove a generalisation of Theorem 9(ii).

## 5.6 Economical change of $\varepsilon$

The final challenge of this chapter will be to see whether with the refined method Theorem 8 can be extended to some odd $t > 1$.

However, the immediate way in which the method was realized in Section 5.2 is not very clever in our circumstances. Starting from $(a_k) \in \mathbb{F}_q^S$ makes it difficult to make optimal use of the considerable knowledge from Chapter 4 about coefficient tuples $(A_k)$ for which the polynomial splitting problem has already been solved. Another important observation is that $A_1 = 0$ seems to be crucial in our proof of Theorem 8 (all attempts of a modification for $A_1 \neq 0$ have so far met with serious obstacles).

The subject of this section is a systematic study how the concept of working with individual $\varepsilon$ and $c$ for each $(a_k) \in \mathbb{F}_q^S$ can be put into practice as economically as possible in view of these two points. We allow arbitrary odd characteristic and use the terminology of Section 5.1.

Let the space $\mathcal{H}/\!\sim$ be represented as $\{\varepsilon_0\} \times \{b_1\} \times \mathbb{F}_q^{S\setminus\{1\}}$. Refer to this as the **normalised** representation of $\mathcal{H}/\!\sim$. In the light of the remark about $A_1 = 0$ above, we are only interested in the case $b_1 = 0$.

Recall from (5.3) the bijection

$$\begin{array}{rccc} \tilde{\psi}: & \mathcal{H}/\!\sim & \longrightarrow & \{0\} \times \mathbb{F}_q^{S\setminus\{1\}} \\ & [\varepsilon_0, (b_k)]_\sim & \longmapsto & (A_k)\,. \end{array}$$

Now assume that for some $(A_k) \in \mathbb{F}_q^S$ with $A_1 = 0$ the splitting of the polynomial $F(x)$ cannot be established. The aim is to identify exactly all $(a_k) \in \mathbb{F}_q^S$ which would have been covered by this, and try each of them again with a different $\varepsilon$.

The equivalence class of Helleseth systems for which the problem remains open is $\tilde{\psi}^{-1}(A_k) = [\varepsilon_0, (b_k)]_\sim$ with $(b_k) = \psi_{\varepsilon_0}^{-1}(A_k)$, already in normalised representation. The elements of this class can be listed as

$$\begin{array}{ccccc} (\varepsilon_0, (b_k)) & \cdots & (\varepsilon_0, \varphi_{c,\varepsilon_0}(b_k)) & \cdots \\ \vdots & & \vdots & \\ (\vartheta\varepsilon_0, (\vartheta b_k)) & \cdots & (\vartheta\varepsilon_0, \varphi_{c,\vartheta\varepsilon_0}(\vartheta b_k)) & \cdots \\ \vdots & & \vdots & \end{array}$$

($p-1$ rows corresponding to the elements $\vartheta \in \mathbb{F}_p^*$, and $q$ columns corresponding to the elements $c \in \mathbb{F}_q$). A general element in this class is $(\vartheta\varepsilon_0, (a_k))$ with

$$(a_k) \;=\; \varphi_{c,\vartheta\varepsilon_0}(\vartheta b_k) \;=\; \vartheta \cdot \varphi_{c,\varepsilon_0}(b_k) \;=\; \vartheta \cdot \varphi_{c,\varepsilon_0} \circ \psi_{\varepsilon_0}^{-1}(A_k)\,.$$

For this, the change of $\varepsilon$ is now executed by replacing $\vartheta\varepsilon_0$ with another $\varepsilon \in \mathbb{F}_p^*$. Since this is arbitrary, we can write it in the form $\vartheta\varepsilon_1$ with arbitrary $\varepsilon_1 \in \mathbb{F}_p^*$. So the new system to be considered is $(\vartheta\varepsilon_1, (a_k))$.

To find the image under $\tilde{\psi}$ of the $\sim$-equivalence class of $(\vartheta\varepsilon_1, (a_k))$, it is necessary to find first the normalised representative of this class. The $\varepsilon$-component is normalised by

$$(\vartheta\varepsilon_1, (a_k)) \approx \left(\varepsilon_0, \left(\tfrac{\varepsilon_0}{\vartheta\varepsilon_1} \cdot a_k\right)\right) ,$$

and then $d \in \mathbb{F}_q$ has to be determined such that $\varphi_{d,\varepsilon_0}\left(\tfrac{\varepsilon_0}{\vartheta\varepsilon_1} \cdot a_k\right) =: (\bar{b}_k)$ has $\bar{b}_1 = 0$. One calculates

$$\bar{b}_1 = -dr\varepsilon_0 + \frac{\varepsilon_0 a_1}{\vartheta\varepsilon_1} \qquad \text{and} \qquad a_1 = -\vartheta cr\varepsilon_0 ,$$

so $d = -\dfrac{\varepsilon_0}{\varepsilon_1}c$, and in normalised representation our class is $[\varepsilon_0, (\bar{b}_k)]_\sim$ with

$$(\bar{b}_k) = \varphi_{-\frac{\varepsilon_0}{\varepsilon_1}c,\varepsilon_0}\left(\tfrac{\varepsilon_0}{\varepsilon_1} \cdot \varphi_{c,\varepsilon_0} \circ \psi_{\varepsilon_0}^{-1}(A_k)\right) .$$

Finally, by applying the map $\tilde{\psi}$, we find that the new coefficients for the polynomial splitting problem are

$$(\overline{A}_k) := \psi_{\varepsilon_0} \circ \varphi_{-\frac{\varepsilon_0}{\varepsilon_1}c,\varepsilon_0}\left(\tfrac{\varepsilon_0}{\varepsilon_1} \cdot \varphi_{c,\varepsilon_0} \circ \psi_{\varepsilon_0}^{-1}(A_k)\right) .$$

The problem must now be solved for all $c \in \mathbb{F}_q$, but $\varepsilon_1 \in \mathbb{F}_p^*$ can be chosen individually in each case.

In all this we can take w.l.o.g. $\varepsilon_0 := 1$. If we further replace $\varepsilon_1$ with $\varepsilon_1^{-1}$, the expression for the $\overline{A}_k$ simplifies to

$$(\overline{A}_k) = \psi_1 \circ \varphi_{-\varepsilon_1 c,1}\left(\varepsilon_1 \cdot \varphi_{c,1} \circ \psi_1^{-1}(A_k)\right) . \tag{5.12}$$

One checks that the choice $\varepsilon_1 := 1$ yields the identity. In the special case $p = 3$, the only real change of $\varepsilon$ is obtained from $\varepsilon_1 := -1$.

## 5.7 Notes on the case of odd $t$ greater than 1

**Conjecture** *Let $t \in \mathbb{N}$ be odd.*

(i) *For odd $m$ the analogue of Theorem 8 holds, i.e. the covering radius of the primitive ternary BCH code of length $q-1$ ($q = 3^m$) with designed distance $3t+2$ is exactly $3t+1$ whenever $q > [(3t+1)!(3t-1)]^2$.*

(ii) *If $m$ is even and sufficiently large, then the covering radius of $C_{3,q-1,3t+2}$ is exactly $3t+2$.*

What evidence is there to support this?

Part (ii), like the corresponding statement for $t \equiv 2 \pmod 4$ and odd $m$, is based on the repeated failure of our attempts to lower the bound for the covering radius to $3t+1$. As we will see below, the refined version of our method still fails in the Degenerate Case for even $m$. From Section 5.5 we know that varying one of the $\varepsilon_i$ does nothing to help either.

And of course there is the example of $t = 1$, where the statement is positively true by Theorem 9(ii). (But one has to be careful with $t = 1$, as this case may be pathological—see below.)

Beyond that, the second part of the Conjecture is mere speculation.

When it comes to the more interesting part (i) of the Conjecture, we have more concrete evidence to go by.

First of all, it is again true for $t = 1$ by Theorem 9(i).

Secondly, for $t \geq 3$ the proof of the Standard Case from Section 4.2 carries over completely. The only place where the parity of $t$ has any relevance is at the end of the subsection "Indecomposability", and there an extra argument was included to cover odd $t$.

So let us turn to the Degenerate Case. We have the results from Section 4.4, obtained with some fixed $\varepsilon \in \mathbb{F}_3^*$. Where this leads to nothing, we use now (5.12) with $\varepsilon_1 := -1$ to see what happens if we change to "the other $\varepsilon$".

The step-by-step calculation of $(\overline{A}_k)$ is shown on page 83. The mapping of $(a_k)$ to $(\overline{b}_k)$ via $\varphi_{c,1}$ uses that, modulo 3,

$$\sum_{\nu=1}^{k-1} \binom{k}{\nu} \equiv \left\{ \begin{array}{ll} 0\,, & \text{if } k \text{ is odd} \\ -1\,, & \text{if } k \text{ is even} \end{array} \right\} \quad \text{for all } k \in \mathbb{N}\,,$$

a fact easily proved using $\binom{k}{\nu} = \binom{k-1}{\nu-1} + \binom{k-1}{\nu}$. All other steps are straightforward with induction. The list gives the results for arbitrary $t \in \mathbb{N}$. Those for $t = 1$ differ to some extent from those for larger $t$. This is the reason why we said that the situation of Theorem 9 may be misleading and must be taken with caution.

For odd $t \geq 3$ we re-visit now the cases of Section 4.4. All notation from there is retained, and where a change of $\varepsilon$ is carried out the new objects are denoted by overlined symbols ($\overline{f}_0$, $\overline{E}_1$, $\overline{\Delta}_x$, ...) in the obvious way.

$\boxed{A_{3t-1}A_{3t+1} \neq 0.}$

Unlike for even $t$ (see p. 62), it is now possible to have $f(\alpha) = f(-\alpha)$. This happens exactly when $A_{3t-1}^s = A_{3t+1}^{s-1}$, where $s := (3t+1)/2$. Only in this case it is necessary to employ the change of $\varepsilon$.

Then the only $\overline{A}_k$ that can be different from zero are

$$\overline{A}_2 = c^2\,, \quad \overline{A}_4 = c^4\,, \quad \overline{A}_{3t-1} = -A_{3t-1}\,, \quad \overline{A}_{3t+1} = -A_{3t+1} - c^2 A_{3t-1}\,.$$

For $c \neq 0$ this is the Standard Case (because $\overline{A}_2\overline{A}_{3t+1} \neq \overline{A}_4\overline{A}_{3t-1}$). For $c = 0$ the change of $\varepsilon$ amounts to multiplying the $A_k$ through with $-1$. Since it is not possible

to have $A_{3t-1}^s = A_{3t+1}^{s-1}$ and $(-A_{3t-1})^s = (-A_{3t+1})^{s-1}$ at the same time, this must lead to Situation S. This settles the case $A_{3t-1}A_{3t+1} \neq 0$ (for all $m$).

The only non-zero coefficient is $A_{3t+1}$.

Change of $\varepsilon$ yields either the Standard Case (for $c \neq 0$, then $\overline{A}_2\overline{A}_{3t+1} - \overline{A}_4\overline{A}_{3t-1} = -c^2 A_{3t+1} \neq 0$) or, for $c = 0$, the Degenerate Case with $\overline{A}_{3t+1} = -A_{3t+1}$ as the only non-zero coefficient. In the latter situation,

$$\begin{aligned} \Delta_x &= A_{3t+1}^{3t}\,, \quad \overline{\Delta}_x = -A_{3t+1}^{3t} && \text{when } t \equiv 1 \pmod 4\,, \\ \Delta_x &= -A_{3t+1}^{3t}\,, \quad \overline{\Delta}_x = A_{3t+1}^{3t} && \text{when } t \equiv 3 \pmod 4\,. \end{aligned}$$

Hence, for odd $m$, one of $\Delta_x, \overline{\Delta}_x$ is always a square. (It is obvious that we lose the even $m$ here.)

Assuming that the remaining case will be analogous to this, we seem to be close to reducing the proof of part (i) of the Conjecture to the Exceptional Case. But at this point we are in for a nasty surprise.

The only non-zero coefficient is $A_{3t-1}$.

Use the change of $\varepsilon$ again. For $c = 0$ this yields, similarly as above, $\overline{\Delta}_x^* = -\Delta_x^*$, so that for odd $m$ either $\overline{\Delta}_x^* \in \square$ or $\Delta_x^* \in \square$. But for $c \neq 0$ we find

$$\overline{A}_2\overline{A}_{3t+1} \;=\; -c^4 A_{3t-1} \;=\; \overline{A}_4\overline{A}_{3t-1}\,,$$

so this time we are in the Exceptional Case! With the parameter choice of p. 48 we get

$$\begin{aligned} \overline{f}_0(x) &= x^{3t+1} + (\overline{C}x^2+1)\,\overline{g}_{t-1}^{(0)}(x^3) \;=\; x^{3t+1} + (x^2+c^2)(c^2x^{3t-3} - A_{3t-1})\,, \\ \overline{f}_1(x) &= x^{3t-2} + (\overline{C}x^2+1)\,\overline{g}_{t-2}^{(0)}(x^3) + vx \;=\; x^{3t-2} + (x^2+c^2)\,c^2x^{3t-6} + vx\,, \end{aligned}$$

and so

$$\overline{E}_1 \;=\; (x^2-c^2)\left[-A_{3t-1}x^{3t-3} + v(c^2x^{3t-3} - A_{3t-1})\right];$$

this is a particularly obstinate instance of the Exceptional Case which is not covered by any of the arguments in Section 4.3 and has so far also withstood all other attempts of a general solution.

Yet we can uphold our conjecture on the following basis. Calculations with Maple for $t = 3, 5, 7$ suggest that the general form of the discriminant $\overline{\Delta}_x$ (with $\overline{f}_0, \overline{f}_1$ as above) is

$$\begin{aligned} \overline{\Delta}_x \;=\; (-1)^{(t-1)/2}\Big[&(c^{6t}v^{3t-1} - c^{6t-6}v^{3t-4}A^3)\,u^{3t-1} + \\ &+ (c^{6t-4}v^{3t-6}A^5 - c^{6t+2}v^{3t-3}A^2)\,u^{3t-3} - A^{3t}v^2u^2 + c^2A^{3t+2}\Big]. \end{aligned}$$

Can this be a square in $\overline{\mathbb{F}}_q(u)$? For $t \equiv 1 \pmod 4$ this can definitely be ruled out by elementary means. For $t \equiv 3 \pmod 4$ the same appears to lead to $R(v) = 0$ with

a rational expression $R(x) \in \overline{\mathbb{F}}_q(x)$ of reasonably small degree, but not constant. Thus, excluding a number of $v$ would settle the problem.

A rigorous (and possibly more elegant) proof remains to be found.

Suppose these last difficulties can be overcome. Then all that separates part (i) of the Conjecture from a theorem is the Exceptional Case.

Again, with the little addition at the end of Subcase (4) of "Indecomposability", the arguments for the first three conditions of Lemma 8 from Section 4.3 cover also odd $t \geq 3$. Only for the discriminant the ice becomes thin. Here Section 4.3 leaves numerous cases open for odd $t$. The hope that in all these the change of $\varepsilon$ would make it possible to avoid the Exceptional Case altogether comes only partly true—this works for most situations, but some serious problem cases remain. We look at some examples.

For $\mu, \nu, \in \{1, \ldots, t\}$ with $\mu < \nu$ define the **Exceptional Case indicator (ECI) for $(\mu, \nu)$** by

$$\mathrm{ECI}(\mu, \nu) := A_{3\mu-1}A_{3\nu+1} - A_{3\mu+1}A_{3\nu-1} .$$

We are in the Exceptional Case if and only if all possible ECIs are zero.

Now take the case $t = 3$ with $A_5 = A_7 = 0$ and $A_2A_4A_8 \neq 0$, $A_{10} = A_4A_8/A_2$. This is an example of the Exceptional Case for which Section 4.3 offers no way to decide the discriminant problem. After changing $\varepsilon$ (the full general expressions for the $\overline{A}_k$ up to index 13, calculated by computer, are listed on p. 84), we find for $c = 0$

$$\begin{aligned} \overline{\mathrm{ECI}}(1,2) &= \overline{\mathrm{ECI}}(2,3) = 0 , \\ \overline{\mathrm{ECI}}(1,3) &= -A_8A_2^2 + A_4^3 - A_4^2A_2^2 . \end{aligned}$$

Thus if $-A_8A_2^2 + A_4^3 - A_4^2A_2^2 = 0$ we are again in the Exceptional Case. Further,

$$\begin{aligned} &\overline{A}_2 = -A_2 , && \overline{A}_5 = 0 , && \overline{A}_8 = -A_8 + A_4^2 - A_4A_2^2 , \\ &\overline{A}_4 = -A_4 + A_2^2 , && \overline{A}_7 = 0 , && \overline{A}_{10} = \overline{A}_4\overline{A}_8/\overline{A}_2 ; \end{aligned}$$

therefore, if $\overline{A}_4$ and $\overline{A}_8$ are different from zero, we get the same difficulties as before. This can happen: take $A_2, A_4 \in \mathbb{F}_q^*$ with $A_4 \neq \pm A_2^2$, and put $A_8 := A_4^2(A_4 - A_2^2)/A_2^2$.

Examples for $t \geq 5$ can be obtained by taking as non-zero $A_k$ exactly those with $k \in \{3t-4, 3t-2, 3t-1, 3t+1\}$. Then $E_1(x)$ has even degree, and for $c = 0$ one gets $(\overline{A}_k) = (-A_k)$ and thus again the Exceptional Case.

Unfortunately, the expressions for the discriminants are, already for $t = 3$, too voluminous to be studied directly.

Tiresome as such examples may be, the do not seem to indicate a fundamental obstacle. In judging the problems it must be taken into account that, apart from Lemma 12, we have not given much emphasis to developing powerful tools to deal with general discriminants. Future research in this direction may be the key to settling part (i) of the Conjecture.

$$A_k = 0 \quad \text{for } k = 1, \ldots, 3t-2 \text{ and } k = 3t$$

$$\Big\downarrow \psi_1^{-1}$$

*values for $t = 1$ where different:*

$$b_k = 0 \quad \text{for } k = 1, \ldots, 3t-2 \text{ and } k = 3t$$
$$b_{3t-1} = A_{3t-1}$$
$$b_{3t+1} = -A_{3t+1} \qquad b_4 = -A_2^2 - A_4$$

$$\Big\downarrow -\varphi_{c,1}$$

$$a_k = -(-c)^k \quad \text{for } k = 1, \ldots, 3t-2 \text{ and } k = 3t$$
$$a_{3t-1} = -(-c)^{3t-1} - A_{3t-1}$$
$$a_{3t+1} = -(-c)^{3t+1} + A_{3t+1} \qquad a_4 = -c^4 + A_2^2 + A_4$$

$$\Big\downarrow \varphi_{c,1}$$

$$\bar{b}_k = 0 \quad \text{for odd } k \in \{1, \ldots, 3t-2\} \cup \{3t\}$$
$$\bar{b}_k = c^k \quad \text{for even } k \in \{1, \ldots, 3t-2\} \cup \{3t\}$$
$$\bar{b}_{3t-1} = \begin{cases} c^{3t-1} - A_{3t-1} & \text{if } t \text{ is odd} \\ -A_{3t-1} & \text{if } t \text{ is even} \end{cases}$$
$$\bar{b}_{3t+1} = \begin{cases} c^{3t+1} + A_{3t+1} & \text{if } t \text{ is odd} \\ A_{3t+1} & \text{if } t \text{ is even} \end{cases} \qquad \bar{b}_4 = c^4 + A_2^2 + A_4$$

$$\Big\downarrow \psi_1$$

$$\overline{A}_1 = 0$$
$$\overline{A}_2 = c^2 \qquad \overline{A}_2 = c^2 - A_2$$
$$\overline{A}_3 = 0$$
$$\overline{A}_4 = c^4 \qquad \overline{A}_4 = c^4 - c^2 A_2 + A_2^2 - A_4$$
$$\overline{A}_k = 0 \quad \text{for } k = 5, \ldots, 3t-2 \text{ and } k = 3t$$
$$\overline{A}_{3t-1} = -A_{3t-1}$$
$$\overline{A}_{3t+1} = -A_{3t+1} - c^2 A_{3t-1}$$

**Table 1:** Calculation of $(\overline{A}_k)$ in the Degenerate Case ($p = 3$, $r \equiv 1$)

$$
\begin{aligned}
\overline{A}_1 &= 0\,,\\
\overline{A}_2 &= -A_2 + c^2\,,\\
\overline{A}_3 &= 0\,,\\
\overline{A}_4 &= -A_4 + A_2^2 - c^2 A_2 + c^4\,,\\
\overline{A}_5 &= -A_5 - cA_4 - cA_2^2 - c^3 A_2\,,\\
\overline{A}_6 &= 0\,,\\
\overline{A}_7 &= -A_7 - A_5 A_2 + cA_4 A_2 - cA_2^3 - c^2 A_5 + c^3 A_2^2 - c^5 A_2\,,\\
\overline{A}_8 &= -A_8 + A_4^2 - A_4 A_2^2 - cA_7 + cA_5 A_2 - c^2 A_4 A_2 + c^3 A_5 + c^4 A_4\,,\\
\overline{A}_9 &= 0\,,\\
\overline{A}_{10} &= -A_{10} - A_8 A_2 + A_5^2 - A_4^2 A_2 + A_4 A_2^3 + cA_7 A_2 + cA_5 A_4 - c^2 A_8 - c^2 A_4^2\\
&\quad - c^2 A_4 A_2^2 - c^2 A_2^4 - c^3 A_7 + c^3 A_5 A_2 + c^4 A_4 A_2 + c^4 A_2^3 + c^5 A_5 - c^6 A_4\\
&\quad - c^6 A_2^2\,,\\
\overline{A}_{11} &= -A_{11} - A_7 A_4 - A_7 A_2^2 + A_5 A_4 A_2 - cA_{10} + cA_8 A_2 - cA_5^2 - c^2 A_7 A_2\\
&\quad - c^2 A_5 A_4 + c^3 A_4^2 + c^3 A_4 A_2^2 - c^3 A_2^4 + c^4 A_5 A_2 + c^5 A_4 A_2 + c^5 A_2^3 + c^6 A_5\\
&\quad + c^7 A_4 - c^7 A_2^2\,,\\
\overline{A}_{12} &= 0\,,\\
\overline{A}_{13} &= -A_{13} - A_{11} A_2 - A_8 A_5 + A_7 A_4 A_2 + A_7 A_2^3 - A_5 A_4^2 + cA_{10} A_2 + cA_8 A_4\\
&\quad + cA_7 A_5 - cA_4^3 + cA_4 A_2^4 - c^2 A_{11} + c^2 A_7 A_4 - c^2 A_7 A_2^2 + c^2 A_5 A_4 A_2\\
&\quad - c^2 A_5 A_2^3 + c^3 A_{10} - c^4 A_7 A_2 - c^4 A_5 A_4 + c^5 A_4^2 - c^5 A_4 A_2^2 + c^5 A_2^4\\
&\quad + c^6 A_7 - c^6 A_5 A_2 - c^7 A_2^3 + c^8 A_5 - c^9 A_4 + c^9 A_2^2\,,\\
&\vdots
\end{aligned}
$$

**Table 2:** General values of the $\overline{A}_k$ in (5.12) up to index $k = 13$ ($p = 3$, $r \equiv 1$)

# Chapter 6

# And next?

In Chapters 2 and 3 we have achieved our aim of presenting a workable method for estimating the covering radius of long primitive BCH codes over arbitrary prime alphabets. In each case, however, the main burden remains the technical manipulation and analysis of the individual polynomials (recall that we get no result at all if we fail to establish splitting of $F(x)$ for a single coefficient tuple $(A_k) \in \mathbb{F}_q^S$). Due to the limited scope of the present project, it was not possible to apply our method to many different situations; on the contrary, we focussed instead on ternary codes with designed distance $\delta \equiv 2 \pmod 3$ and treated this quasi as an in-depth example and a playing-field to explore various ideas.

We devote this last chapter to an outlook towards what else could (and should?) be done with the machinery developed in this work. It is also an appropriate place for a brief recapitulation of the problems that remained open in Chapters 4 and 5.

## 6.1 Open questions for $p = 3$ and $\delta \equiv 2 \pmod 3$

For ternary codes with designed distance $\delta = 3t + 2$, getting a result means showing $\rho \leq \delta - 1$ for large enough $q$.

The cases where our method repeatedly failed to do this, and where we believe this is because $\rho \leq \delta - 1$ may not be generally true (or may even be generally false), are

- $t \equiv 2 \pmod 4$, $m$ odd;
- $t$ odd, $m$ even.

Here it would be desirable to have ways of *raising the lower bound* for the covering radius, something for which our method is unsuitable (and for which also the techniques of the proof of Theorem 9(ii) are not feasible when $t > 1$).

A key to this may lie in the question what is with (5.6), or more generally with the Degenerate Case with $A_{3t-1} \neq 0$. Do these force $\rho > \delta - 1$? What is their coding-theoretical meaning?

In the remaining open case, when $t$ and $m$ are both odd, we conjecture that the result holds and a proof will eventually be found, as discussed in detail in Section 5.7. The role of the discriminant in this context has already been stressed. Independently,

one may also try to open up new ways by improving on the function-field-theoretic results in Section 3.2.

## 6.2 The ternary case with $\delta \equiv 0 \pmod 3$

The obvious next step in the use of our method is to apply it to "the other" ternary case, namely codes with designed distance $\delta \equiv 0 \pmod 3$.

In view of Theorem 2(i), the first challenge must be to see if there are conditions under which one can show $\rho = \delta - 1$. For this write $\delta = 3t + 3$ with $t \in \mathbb{N}_0$. Then, by Theorem 6 with $r = 3t + 2$, for every choice of $A_k \in \mathbb{F}_q$ $(k = 1, \ldots, 3t + 2;\ 3 \nmid k)$ there have to be found elements $s_3, s_6, \ldots, s_{3t}$ in $\mathbb{F}_q$ such that

$$\begin{aligned} F(x) \;=\; & [x^{3t+2} + x\, g_{t-1}^{(1)}(x^3) + g_{t-1}^{(0)}(x^3)] \;+ \\ & + \quad s_3 \cdot [x^{3t-1} + x\, g_{t-2}^{(1)}(x^3) + g_{t-2}^{(0)}(x^3)] \;+ \\ & + \quad \ldots \quad + \\ & + s_{3t-6} \cdot [x^8 + x\, g_1^{(1)}(x^3) + g_1^{(0)}(x^3)] \;+ \\ & + s_{3t-3} \cdot [x^5 + x\, g_0^{(1)}(x^3) + g_0^{(0)}(x^3)] \;+ \\ & + \quad s_{3t} \cdot [x^2 + A_1 x + A_2] \end{aligned}$$

with

$$\begin{aligned} g_{t-1}^{(1)}(x^3) &= A_1 x^{3t} + A_4 x^{3t-3} + A_7 x^{3t-6} + \ldots + A_{3t-2} x^3 + A_{3t+1}\,, \quad \text{etc.} \\ g_{t-1}^{(0)}(x^3) &= A_2 x^{3t} + A_5 x^{3t-3} + A_8 x^{3t-6} + \ldots + A_{3t-1} x^3 + A_{3t+2}\,, \quad \text{etc.} \end{aligned}$$

splits completely over $\mathbb{F}_q$. (One may assume $A_1 = 0$.)

Where this fails, one will go for $\rho \leq \delta$. Then, with $r = \delta = 3t$ $(t \in \mathbb{N})$, one must find for every choice of $A_k \in \mathbb{F}_q$ $(k = 1, \ldots, 3t - 1;\ 3 \nmid k)$ elements $s_3, s_6, \ldots, s_{3t} \in \mathbb{F}_q$ such that

$$\begin{aligned} F(x) \;=\; & [x^{3t} + x^2\, g_{t-1}^{(2)}(x^3) + x\, g_{t-1}^{(1)}(x^3)] \;+ \\ & + \quad s_3 \cdot [x^{3(t-1)} + x^2\, g_{t-2}^{(2)}(x^3) + x\, g_{t-2}^{(1)}(x^3)] \;+ \\ & + \quad \ldots \quad + \\ & + s_{3t-6} \cdot [x^6 + x^2\, g_1^{(2)}(x^3) + x\, g_1^{(1)}(x^3)] \;+ \\ & + s_{3t-3} \cdot [x^3 + A_1 x^2 + A_2 x] \;+ \\ & + \quad s_{3t} \end{aligned}$$

with

$$\begin{aligned} g_{t-1}^{(2)}(x^3) &= A_1 x^{3t-3} + A_4 x^{3t-6} + \ldots + A_{3t-5} x^3 + A_{3t-2}\,, \quad \text{etc.} \\ g_{t-1}^{(1)}(x^3) &= A_2 x^{3t-3} + A_5 x^{3t-6} + \ldots + A_{3t-4} x^3 + A_{3t-1}\,, \quad \text{etc.} \end{aligned}$$

splits completely over $\mathbb{F}_q$. (Here $A_1 = 0$ can *not* be assumed.)

At the time of writing this account, these problems are under investigation. They had initially been put back because of apparent fundamental obstacles. For instance, if in the "$\rho = \delta - 1$" case $-A_2 \in \square$ and all other $A_k$ are zero, then $F(x)$ contains an irreducible factor $x^2 + A_2$. It takes a combination of linear transformation and choice of $\varepsilon$—a concept which was only fully developed at a late stage—to resolve this (for odd $m$). Although now all such serious problems seem to have been eliminated, much technical work still has to be done to catch up with the case $\delta \equiv 2$. So, for the time being, the question whether the upper bound for $\rho(\mathcal{C}_{3,q-1,3t})$ can be improved must remain open.

However, building upon the work of Chapter 4, we can show with little extra effort:

**Corollary 2** *Let $t \in \mathbb{N}$.*

*The covering radius of the primitive ternary BCH code of length $q-1$ ($q = 3^m$) and designed distance $\delta = 3t$ is at most $3t+1$ whenever $q > [(3t+1)!\,(3t-1)]^2$.*

*Proof.*

We have $r = 3t+1$ in Theorem 6, and the $g$-polynomials and $F(x)$ are the same as in Section 2.6 and Chapter 4 (we normalise again $A_1$ to 0). The important difference is that the element $A_{3t+1}$ is now at our disposal. We go again through the proof of Theorem 8 and see what improvements and simplifications this brings.

The proof for the Standard Case carries over completely (i.e. for all $t > 1$). In fact, it simplifies greatly, as by choosing $A_{3t+1} \neq 0$ we can get rid of the $w$-terms.

The Exceptional Case can be avoided altogether by choosing $A_{3t+1}$ different from $A_{3t-1}C_0/C_2$. (No more awkward discriminants!)

Finally, consider the Degenerate Case for $t > 1$. If $A_{3t-1} = 0$ take also $A_{3t+1} = 0$, then $F(x) = x^r$ splits trivially. Otherwise we may assume $A_{3t-1}A_{3t+1} \neq 0$; for even $t$ argue as on page 62, for odd $t$ we have $f(\alpha) = f(-\alpha)$ iff $A_{3t-1}^s = A_{3t+1}^{s-1}$ (cf. page 80), which can be avoided by choice of $A_{3t+1}$.

It remains $t = 1$. Here $F(x) = (x^4 + A_2x^2 + A_4) + ux$. The case $A_2 = 0$ is settled by choosing $A_4 = u = 0$, so assume $A_2 \neq 0$. Avoid $A_4 = 0$ to arrange co-primality, then the first three conditions of Lemma 8 are satisfied. Further, $E_1(x) = A_2x^2 - A_4$ has roots $\pm\alpha$ with $\alpha^2 = A_4/A_2$; by avoiding $A_4 = A_2^2$ one gets $0 \neq f(\alpha) = -f(-\alpha)$ and thus "Situation S". □

This confirms the result of Kaipainen in Theorem 2(i), with the bonus of yielding the explicit value $[(3t+1)!\,(3t-1)]^2$ for $q_0$.

## 6.3 Primes $p$ greater than 3

Since Theorem 6 holds for arbitrary primes $p$ (this being an essential point of it), it would be a waste not to try at some stage to use it for $p > 3$.

The main difference with $p$-ary[†] codes is that there are now $p-1$ cases to distinguish, namely with designed distances $\delta \equiv 0$ (or 1) and $\delta \equiv 2, \ldots, p-1 \pmod{p}$. It is certainly reasonable to pick the case $\delta = pt+2$ with hypothesis "$\rho \leq \delta - 1$" before any other, because here the shape of $F(x)$ is closest possible in similarity to that of the ternary case in Chapter 4, and one may hope to retain some of the structure of the proof of Theorem 8. The polynomial $F(x)$ for $\delta = pt+2$ and "$\rho \leq \delta - 1$" is the same as for $\delta = pt+2-m$ and "$\rho \leq \delta - 1 + m$" ($m = 2, 3, \ldots$); each has $r = pt+1$. Therefore, whenever the first of these cases is completed successfully, this brings with it a whole flag of corollaries, similarly to Theorem 8 and Corollary 2.

The next obvious difference is that the coefficients $A_k$ are now distributed over $p-1$ polynomials $g_{t-1}^{(i)}$ (for $r = pt+1$ these are $g_{t-1}^{(0)}$ and $g_{t-1}^{(2)}, \ldots, g_{t-1}^{(p-1)}$) and their descendants.

From now on, fix $r = pt+1$ and assume $A_1 = 0$, so that all $g_{t-1}^{(i)}$ have degree at most $p(t-1)$. In generalisation of the ternary situation in Chapter 4, we speak of the **Degenerate Case** when only $A_k$ with $k \geq p(t-1)+2$ (the constant terms of the $g_{t-1}^{(i)}$) are allowed to be non-zero. This leads potentially to $2^{p-1}-1$ different subcases, though these can possibly be grouped in some efficient way. The details of this case have not yet been pursued.

More significantly, the **Standard Case** can be generalised in a sensible way. For $p = 3$ it was characterised by $C_2 g_{t-1}^{(0)} - C_0 g_{t-1}^{(2)} \neq 0$, and it turns out that the appropriate generalisation (with the obvious definition of $C_i$ for $i = 0, 2, 3, \ldots, p-1$) is

$$\sum_{\substack{i,k \in \{0,2,3,\ldots,p-1\} \\ i>k}} a_{ik}\, x^{i+k-2} \left( C_k g_{t-1}^{(i)}(x^p) - C_i g_{t-1}^{(k)}(x^p) \right) \neq 0,$$

where the $a_{ik}$ are certain elements in $\mathbb{F}_p^*$. With this definition, and the same choice of $s_p, \ldots, s_{pt}$ as for $p = 3$ in Section 4.2, everything from co-primality to Step 1 of showing simplicity extends to arbitrary $p > 3$ (with some extra work in the case of indecomposability). A new difficulty occurs in Step 2 of the simplicity part, when, in a certain situation, $f_1''(x)$ and $f_1(x) - x f_1'(x)$ have a common non-zero root. Without elaborating further on this, we mention it as an example that even beyond $p = 3$ structural phenomena exist which do not emerge for smaller alphabet sizes.

Another serious problem comes with Step 3. It is an encouraging sign that, at least when $w = 0$, the reduction from Section 4.2 works for $p > 3$ exactly as in the ternary case, in particular $f_1$ cancels from $H_2$ (cf. page 42). Yet, the polynomial $D$ in the denominator of $H_2$ is now of degree up to $p-1$, and any part of this may cancel with the numerator—so we have not one, but $p$ indecomposability problems to solve (with denominator degrees from $p-1$ to 0)!

The **Exceptional Case** (defined as everything not covered by the other two cases) has not been looked into so far. However, the fact that by a theorem of Jordan ([Wie], Thm. 13.9; see also the extensions in [Zie]) a $p$-ary version of Lemma 8 is possible gives rise to some optimism.

Of course, the ultimate goal would be to prove a theorem for all (or at least an infinite number of) primes $p$. Whether it will ever be possible to find the necessary

[†]For simplicity, we assume $p > 3$ in this section. The reader will easily check that all statements for which this is meaningful extend the case $p = 3$.

generic arguments—e.g. in cases like the Step-3-problem above, where an infinite number of decomposition problems would have be solved—the author is sceptic, but the mental challenge is undoubtedly intriguing.

But even attempts on individual small primes should be worthwhile and provide both results on covering radii and new insights into the fascinating structure of the problem. Clearly, the effort involved grows quickly with the size of $p$, so $p = 5$ may give a first indication. An early step in this direction was made in [KaiSu], where it was shown soon after the publication of Cohen's method that the covering radius of the 5-ary BCH code of length $5^m - 1$ with designed distance 7 is at most 7 when $m \geq 13$; for odd $m \geq 9$, this was subsequently improved by Cohen to $\rho(\mathcal{C}_{5,5^m-1,7}) = 6$. Another possibility to be considered is whether parts of the method could be formulated algorithmically and implemented on a computer, so that larger primes become accessible.

## 6.4 Non-primitive codes

For completeness, we remark that the method should also allow a generalisation to non-primitive codes. For this, by Section 2.1, one has to substitute the variable $x$ by $y := x^N$, where $N$ is the degree of primitivity. Making $F(y)$ split in $\mathbb{F}_q[x]$ will presumably require a deeper analysis of the Galois-theoretic aspects ([CohSt] may come in useful) and is left for future treatment elsewhere.

# Bibliography


[Surv] G. D. Cohen, M. G. Karpovsky, H. F. Mattson Jr. and J. R. Schatz, "Covering Radius—Survey and Recent Results", *IEEE Transactions on Information Theory*, Vol. IT-31 No. 3 (1985), 328–343.

[Coh72] S. D. Cohen, "Uniform Distribution of Polynomials over Finite Fields", *J. London Math. Soc. (2)* **6** (1972), 93–102.

[Coh91] S. D. Cohen, "Permutation polynomials and primitive permutation groups", *Arch. Math.* **57** (1991), 417–423.

[Coh97] S. D. Cohen, "The Length of Primitive BCH Codes with Minimal Covering Radius", *Designs, Codes and Cryptography* **10** (1997), 5–16.

[Coh98] S. D. Cohen, "Polynomial Factorisation and an Application to Regular Directed Graphs", *Finite Fields and Their Applications* **4** (1998), 316–346.

[Coh99] S. D. Cohen, "Some Function Field Estimates with Applications", in: "Number theory and applications", Lecture Notes in Pure and Applied Mathematics 204, Dekker, (1999), 23–45.

[CohSt] S. D. Cohen and W. W. Stothers, "The Galois group of $f(x^r)$", *Glasgow Math. J.* **25** (1984), 75–91.

[FrCoh] R. Franken and S. D. Cohen, "The Covering Radius of Some Primitive Ternary BCH Codes", *Finite Fields and Applications. 7th International Conference Fq7, Toulouse, France, May 2003. Revised Papers*, Springer LNCS 2948 (2004), 166–180.

[Fri] M. Fried, "On a conjecture of Schur", *Michigan Math. J.* **17** (1970), 41–55.

[FrMR] M. D. Fried and R. E. MacRae, "On Curves with Separated Variables", *Math. Ann.* **180** (1969), 220–226.

[Ful] W. Fulton, "Algebraic Curves", Benjamin, New York, 1969.

[GPZ] D. C. Gorenstein, W. W. Peterson and N. Zierler, "Two-error correcting Bose-Chaudhury codes are quasi-perfect", *Inform. and Control* **3** (1960), 291–294.

[Hel] T. Helleseth, "On the Covering Radius of Cyclic Linear Codes and Arithmetic Codes", *Discr. Appl. Math.* **11** (1985), 157–173.

[Kai] Y. Kaipainen, "On the Covering Radius of Long Non-binary BCH Codes", dissertation, University of Turku, 1995.

[KaiSu] Y. Kaipainen and K. Suominen, "On the Covering Radius of Long 5-ary BCH Codes with Minimum Distance 7", *Appl. Algebra Egnrg. Comm. Comput.* **8** (1997), 403–410.

[LiNi] R. Lidl and H. Niederreiter, "Introduction to Finite Fields and Their Applications", Cambridge University Press, 1994.

[vLi] J. H. van Lint, "Introduction to Coding Theory" (second edition), Graduate Texts in Mathematics 86, Springer, Berlin/Heidelberg, 1992.

[MoMo] C.J. Moreno and O. Moreno, "Constructive Elementary Approach to the Covering Radius of Long BCH Codes", conference abstract, 1993.

[Shp] I. E. Shparlinski, addition to the Russian translation of [Hel], *Kiberneticheskii Sbornik* **25** (1985), 82–84.

[Sti] H. Stichtenoth, "Algebraic Function Fields and Codes", Springer, Berlin/Heidelberg, 1993.

[Tie] A. Tietäväinen, "On the Covering Radius of Long Binary BCH Codes", *Discr. Appl. Math.* **16** (1987), 75–77.

[SkVl] A. N. Skorobogatov and S. G. Vladuts, "The Covering Radius of Long Binary BCH Codes", *Problemy Peredachi Inform.* **25**(1) (1989), 38–45. In Russian.

[Wie] H. Wielandt, "Finite Permutation Groups", Academic Press, New York/London, 1968.

[Zie] T. E. Zieschang, "Primitive permutation groups containing a $p$-cycle", *Arch. Math.* **64** (1995), 471–474.