QUATERNARY CODES

SERIES ON APPLIED MATHEMATICS

Editor-in-Chief: Frank Hwang
Associate Editors-in-Chief: Zhong-ci Shi and U Rothblum

Series on
Applied Mathematics
Volume 8

QUATERNARY CODES

Zhe-Xian Wan
Chinese Academy of Sciences, China
and
Lund University, Sweden

World Scientific
Singapore • New Jersey • London • Hong Kong

Published by

World Scientific Publishing Co. Pte. Ltd.

P O Box 128, Farrer Road, Singapore 912805

USA office: Suite 1B, 1060 Main Street, River Edge, NJ 07661

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

Library of Congress Cataloging-in-Publication Data
Wan, Zhe-Xian
Quaternary codes / by Zhe-Xian Wan.
p. cm. -- (Series on applied mathematics ; v. 8)
Includes bibliographical references and index.
ISBN 9810232748 (alk. paper)
1. Coding theory. I. Title. II. Series.
QA268.W35 1997
003'.54--dc21 97-29178
CIP

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

This book is printed on acid-free paper.

Printed in Singapore by Uto-Print

To Shi-Xian

This page is intentionally left blank

PREFACE

A binary error-correcting code of length n is just a subset of the vector space $\mathbb{F}_2^n$ and linear codes are subspaces of $\mathbb{F}_2^n$. The vectors in a code are called codewords and the Hamming distance between two codewords is the number of positions in which they differ. The rate of a code of length n is defined to be the logarithm to the base 2 of the number of codewords in the code divided by n. One of the fundamental problems in coding theory is to construct and study codes of length n with large rate subject to the condition that the minimum of the distances between any two different codewords is some given integer d, the minimal distance of the code.

Historically, linear codes have been the most important codes since they are easier to construct, encode, and decode. Around 1970 several binary nonlinear codes having at least twice as many codewords as any linear code with the same length and minimal distance have been constructed. Among them are the Nordstrom–Robinson code, the Preparata codes, the Kerdock codes, the Goethals codes, the Delsarte–Goethals codes, etc. However, these binary nonlinear codes are not so easy to describe, to encode and decode as the linear codes. It is also discovered that the weight enumerator of the Preparata code is the MacWilliams transform of that of the Kerdock code of the same length, though they are not dual to each other, which seems to be a mystery in coding theory.

A surprising breakthrough in coding theory is that the Kerdock codes can be viewed as cyclic codes over $\mathbb{Z}_4$ (Nechaev (1989) and Hammons *et al.* (1994)) and the binary image of the $\mathbb{Z}_4$-dual of the Kerdock code over $\mathbb{Z}_4$ can be regarded as a variant of the Preparata code (Hammons *et al.* (1994)). This leads to a new direction in coding theory, the study of cyclic codes over $\mathbb{Z}_4$.

This book aims to be an introduction to this new direction. The first draft was prepared for several lectures at the Department of Mathematics, Shaanxi Normal University, Xi'an, China in May 1996 and the second draft for a series of lectures at the Department of Information Technology, Lund University, Lund, Sweden. Then these drafts were revised completely to the present form. The Hensel lemma and Galois rings which are important tools for the study of $\mathbb{Z}_4$-codes are included. The Gray map being a connection between $\mathbb{Z}_4$-codes and their binary images is introduced. The quaternary Kerdock codes and Preparata codes and their binary images are studied in detail. The construction of lattices from $\mathbb{Z}_4$-codes and the weight enumerators of self-dual $\mathbb{Z}_4$-codes are mentioned. To read the book only a rudiment of binary codes is necessary.

The author is indebted to Rolf Johannesson who supported the author's work in many aspects and created an active and productive atmosphere in the Information Theory Group in Lund where the present book was written. The author is also indebted to Anupama Pawar K. and Babitha Yadav for their beautiful typesetting and to E. H. Chionh for her helpful and careful editorial work. Without their support and help the book could not have appeared so soon.

Zhe-Xian Wan

CONTENTS

CHAPTER 1

QUATERNARY LINEAR CODES AND THEIR GENERATOR MATRICES

1.1. Definition

Let $\mathbb{Z}_4$ be the ring of integers mod 4, n be a positive integer, and $\mathbb{Z}_4^n$ be the set of n-tuples over $\mathbb{Z}_4$, i.e.

$$\mathbb{Z}_4^n = \{(x_1, \dots, x_n) \mid x_i \in \mathbb{Z}_4 \quad \text{for} \quad i = 1, \dots, n\}.$$

The all "0" n-tuple $(0, \dots, 0)$ and the all "1" n-tuple $(1, \dots, 1)$ will be denoted by 0^n and 1^n, respectively.

Any non-empty subset C of $\mathbb{Z}_4^n$ is called a *quaternary code*[1] or, simply and more precisely, a $\mathbb{Z}_4$-code or a code over $\mathbb{Z}_4$, and n is called the *length* of the code. n-tuples in $\mathbb{Z}_4^n$ are called *words* and n-tuples in a quaternary code C are called *codewords* of C.

Let both C and C' be quaternary codes of length n. If $C' \subseteq C$, C' is called a *subcode* of C.

For all $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n) \in \mathbb{Z}_4^n$ define a componentwise addition

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n),$$

then $\mathbb{Z}_4^n$ becomes an additive abelian group of order 4^n.

Any subgroup of $\mathbb{Z}_4^n$ is called a *quaternary linear code*, or simply, $\mathbb{Z}_4$-*linear code*.

[1] There is some ambiguity in the terminology "quaternary code", because codes over $\mathbb{F}_{2^2}$ are also called quaternary codes. But in this book quaternary codes always mean codes over $\mathbb{Z}_4$.

For all $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{Z}_4^n$ define

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_n y_n ,$$

which is called the *inner product* of $\mathbf{x}$ and $\mathbf{y}$. If $\mathbf{x} \cdot \mathbf{y} = 0$, then $\mathbf{x}$ and $\mathbf{y}$ are said to be *orthogonal.*

Let C be a quaternary linear code of length n. Define

$$C^{\perp} = \{\mathbf{x} \in \mathbb{Z}_4^n \,|\, \mathbf{x} \cdot \mathbf{y} = 0 \ \text{for all } \mathbf{y} \in C\}.$$

It is easy to verify that $C^{\perp}$ is a subgroup of $\mathbb{Z}_4^n$. Hence $C^{\perp}$ is also a quaternary linear code, called the *dual code* of C. If $C \subset C^{\perp}$, C is called a *self-orthogonal code.* If $C = C^{\perp}$, C is called a *self-dual code.*

Two quaternary codes C_1 and C_2 both of length n are said to be *equivalent,* if one can be obtained from the other by permuting the coordinates and (if necessary) changing the signs of certain coordinates. Quaternary codes differ only by a permutation of coordinates are said to be *permutation-equivalent.* The *automorphism group* $\mathrm{Aut}(C)$ of a quaternary code C is the group generated by all permutations and sign-changes of the coordinates that preserve the set of codewords of C.

Let us recall that an additive abelian group of prime power order p^m, where p is a prime and $m > 0$, can be written uniquely as a direct sum of m_1 cyclic subgroups of order p^{e_1}, ... , and m_r cyclic subgroups of order p^{e_r}, where m_1, e_1, ... , m_r, e_r are positive integers and $e_1 > \cdots > e_r$. Then we say that the group is of *type* $(p^{e_1})^{m_1} \cdots (p^{e_r})^{m_r}$. Clearly, $m = m_1 e_1 + \cdots + m_r e_r$. We also agree that an abelian group consisting of the identity element alone is of type p^0

For example, the additive group $\mathbb{Z}_4^n$ is of type $(2^2)^m$, since it is a direct sum of n cyclic subgroups of order 2^2. We have

$$\mathbb{Z}_4^n = \dot{+}_{i=1}^{n} \left\{ \left(0, \dots, 0, \underset{i}{x}, 0, \dots, 0\right) | \, x \in \mathbb{Z}_4 \right\},$$

where each

$$\{(0, \dots, 0, x, 0, \dots, 0) \,|\, x \in \mathbb{Z}_4\}$$

is a cyclic subgroup of order 2^2

A quaternary linear code is a subgroup of some $\mathbb{Z}_4^n$, where n is the length of the code, and its order is a power of 2. So we can say the *type* of a quaternary linear code. Clearly, equivalent quaternary linear codes are of the same type. The type of a quaternary linear code is of the form $(2^2)^m$, $(2^2)^{m_1} 2^{m_2}$, 2^m, or 2^0. In the following we simply write the type $(2^2)^m$ as 4^m and the type $(2^2)^{m_1}2^{m_2}$ as $4^{m_1}2^{m_2}$.

$\mathbb{Z}_4$ has only three subgroups, which are of type $4^1, 2^1$, or 2^0, respectively. Thus there are three quaternary linear codes of length 1, and they are

$$\{(0), (1), (2), (3)\}, \ \{(0), (2)\} \text{ and } \{(0)\}.$$

Now let us enumerate the quaternary linear codes of length 2. Clearly, subgroups of $\mathbb{Z}_4^2$ are of type 4^2, 4^12^1, 2^2, 4^1, 2^1 or 2^0. There is only one quaternary linear code of length 2 and type 4^2, which is $\mathbb{Z}_4^2$ and is generated by the rows of the 2×2 matrix over $\mathbb{Z}_4$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

This matrix is called a *generator matrix* of the quaternary linear code $\mathbb{Z}_4^2$.

There are four subgroups of $\mathbb{Z}_4^2$, which are of type 4^12^1 and each of them is generated by the rows of one of the following 2×2 matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}.$$

Clearly

$$\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}$$

generate the same subgroup, so the second matrix is not listed. The quaternary linear codes generated by the first matrix and the second matrix, respectively, are permutation-equivalent; so are the quaternary linear codes generated by the third matrix and the fourth matrix, respectively. Therefore there are only two inequivalent quaternary linear codes of length 2 and type 4^12^1.

There is only one quaternary linear code of length 2 and type 2^2, which has a generator matrix of the form

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

This is a self-dual code.

Quaternary linear codes of length 2 and type 4^1 are generated by any one of the following 1×2 matrices:

$$(1 \ 0), \ (0 \ 1), \ (1 \ 1), \ (1 \ 2), \ (2 \ 1), \ (1 \ 3), \ (3 \ 1).$$

Clearly, the first two matrices generate permutation-equivalent quaternary linear codes, so are the fourth and fifth matrices, and the sixth and seventh matrices. Moreover, the quaternary linear codes generated by the third and sixth matrices are equivalent. Therefore there are three inequivalent quaternary linear codes of length 2 and type 4^1

Quaternary linear codes of length 2 and type 2^1 are generated by any one of the following 1×2 matrices:

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

Clearly, the first two matrices generate equivalent quaternary linear codes. Therefore there are two inequivalent quaternary linear codes of length 2 and type 2^1 Both of them are self-orthogonal.

Finally there is only one quaternary linear code of length 2 and type 2^0, which is $\{(0,\ 0)\}$.

Therefore altogether there are $1 + 2 + 1 + 3 + 2 + 1 = 10$ inequivalent quaternary linear codes of length 2.

1.2. Generator Matrices

Throughout the book if it is clear from the context we make the convention that the elements 0 and 1 of $\mathbb{Z}_2$ are regarded also as elements 0 and 1 of $\mathbb{Z}_4$, respectively, a word $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}_2^n$ is also regarded as a word $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}_4^n$, and a $\mathbb{Z}_2$-matrix M (i.e. a matrix over $\mathbb{Z}_2$) is also regarded as a $\mathbb{Z}_4$-matrix (i.e. a matrix over $\mathbb{Z}_4$). Thus if M is a $\mathbb{Z}_2$-matrix then $2M$ is a well-defined $\mathbb{Z}_4$-matrix.

Let C be a $\mathbb{Z}_4$-linear code of length n. A $k \times n$ matrix G over $\mathbb{Z}_4$ is called a *generator matrix* of C if the rows of G generate C and no proper subset of the rows of G generates C.

Proposition 1.1. *Any $\mathbb{Z}_4$-linear code C containing some nonzero codewords is permutation-equivalent to a $\mathbb{Z}_4$-linear code with a generator matrix of the form*

$$\begin{pmatrix} I_{k_1} & A & B \\ 0 & 2I_{k_2} & 2C \end{pmatrix}, \tag{1.1}$$

where I_{k_1} and I_{k_2} denote the $k_1 \times k_1$ and $k_2 \times k_2$ identity matrices, respectively, A and C are $\mathbb{Z}_2$-matrices, and B is a $\mathbb{Z}_4$-matrix. Then C is an abelian group of type $4^{k_1}2^{k_2}$, C contains $2^{2k_1+k_2}$ codewords, and C is a free $\mathbb{Z}_4$-module if and only if $k_2 = 0$.

Proof. We apply induction on the code length n. We distinguish the following two cases:

(a) There is a codeword of order 4 in C. After permuting the coordinates of the codeword and (if necessary) multiplying the codeword by -1, we can assume that the codeword of order 4 is of the form

$$(1, c_2, \ldots, c_n).$$

Let

$$C' = \{(0, x_2, \ldots, x_n) \in C\}.$$

Clearly C' is also a $\mathbb{Z}_4$-linear code and can be regarded as a code of length $n-1$ by deleting the first coordinate. By induction hypothesis, C' has a generator matrix of the form

$$\begin{pmatrix} 0 & I_{k_1-1} & A_1 & B_1 \\ 0 & 0 & 2I_{k_2} & 2C \end{pmatrix},$$

where A_1 and C are $\mathbb{Z}_2$ matrices and B_1 is a $\mathbb{Z}_4$ matrix. Then C has a generator matrix of the form

$$\begin{pmatrix} 1 & c_2 \cdots c_{k_1} & c_{k_1+1} \cdots c_{k_1+k_2} & c_{k_1+k_2+1} \cdots c_n \\ 0 & I_{k_1-1} & A_1 & B_1 \\ 0 & 0 & 2I_{k_2} & 2C \end{pmatrix}.$$

After adding a certain linear combination of the last $k_1 + k_2 - 1$ rows of the above matrix to the first row, we can assume that it is carried into a matrix of the form (1.1).

(b) There is no codeword of order 4 in C. Then all nonzero codewords in C are of order 2. Since $C \neq \{0^n\}$, there is a codeword of order 2 in C. As in (a) we can assume that this codeword is of the form

$$(2,\ 2c_2, \ldots, 2c_n).$$

Define C' as in (a). Then C' is also a $\mathbb{Z}_4$-linear code without codewords of order 4. C' can be regarded as a code of length $n-1$. By induction hypothesis, C' has a generator matrix of the form

$$(0\ \ 2I_{k_2-1}\ \ 2C_1),$$

where C_1 is a $\mathbb{Z}_2$ matrix. Then C has a generator matrix of the form

$$\begin{pmatrix} 2 & 2c_2 \cdots 2c_{k_2} & 2c_{k_2+1} \cdots 2c_n \\ 0 & 2I_{k_2-1} & 2C_1 \end{pmatrix}$$

After adding a certain linear combination of the last $k_2 - 1$ rows of the above matrix to the first row, we can assume that it is carried into a matrix of the form

$$(2I_{k_2}\ \ 2C)\,,$$

which is a matrix of the form (1.1) with $k_1 = 0$. $\square$

Let $u_1, \ldots, u_{k_1} \in \mathbb{Z}_4$ and $u_{k_1+1}, \ldots, u_{k_1+k_2} \in \mathbb{Z}_2$. We may regard $u_1, \ldots, u_{k_1}, u_{k_1+1}, \ldots, u_{k_1+k_2}$ as *information symbols.* Then *encoding* is carried out by matrix multiplication

$$(u_1, \ldots, u_{k_1}, u_{k_1+1}, \ldots, u_{k_1+k_2})\,G.$$

Proposition 1.2. *The dual code $C^\perp$ of the $\mathbb{Z}_4$-linear code C with generator matrix* (1.1) *has generator matrix*

$$\begin{pmatrix} -{}^tB - {}^tC\,{}^tA & {}^tC & I_{n-k_1-k_2} \\ 2\,{}^tA & 2I_{k_2} & 0 \end{pmatrix}, \tag{1.2}$$

where n is the code length of C. $C^\perp$ is an abelian group of type $4^{n-k_1-k_2}2^{k_2}$ and $C^\perp$ contains $2^{2n-2k_1-k_2}$ codewords.

Proof. Denote the $\mathbb{Z}_4$-linear code with generator matrix (1.2) by C' Clearly $C' \subset C^\perp$. Let $\mathbf{c} = (c_1, c_2, \ldots, c_n) \in C^\perp$. After adding a certain linear combination of the first $n - k_1 - k_2$ rows of (1.2) to $\mathbf{c}$, we can obtain a codeword of $C^\perp$, which is of the form

$$\mathbf{c}' = (c_1, \ldots, c_{k_1}, c_{k_1+1}, \ldots, c_{k_1+k_2}, 0, \ldots, 0).$$

Since $\mathbf{c}'$ is orthogonal to the last k_2 rows of (1.1), each of $c_{k_1+1}, \ldots, c_{k_1+k_2}$ is 0 or 2. After adding a certain linear combination of the last k_2 rows of (1.2) to $\mathbf{c}'$ we can obtain a codeword of $C^\perp$, which is of the form

$$\mathbf{c}'' = (c_1, \ldots, c_k, 0, \ldots, 0)\,.$$

Since $\mathbf{c}''$ is orthogonal to the first k_1 rows of (1.1), $c_1 = \cdots = c_k = 0$. Therefore $\mathbf{c} \in C'$. $\square$

The matrix (1.2) is called a *parity check matrix* of the $\mathbb{Z}_4$-linear code C generated by the rows of the matrix (1.1). A word $\mathbf{c} = (c_1, \ldots, c_n)$ belongs to C if and only if $\mathbf{c}$ is orthogonal to every row of (1.2).

Corollary 1.3. *Any self-dual $\mathbb{Z}_4$-code of length n contains 2^n codewords.*

Proof. Let C be a self-dual $\mathbb{Z}_4$-code of length n with generator matrix (1.1). By Proposition 1.1, $|C| = 2^{2k_1+k_2}$ and by Proposition 1.2, $|C^\perp| = 2^{2n-2k_1-k_2}$. Since $C^\perp = C$, we have $2^{2n-2k_1-k_2} = 2^{2k_1+k_2}$. Therefore $n = 2k_1 + k_2$ and $|C| = 2^n$ □

1.3. Examples

Example 1.1. Let $\mathcal{K}_4$ denote the $\mathbb{Z}_4$-linear code with generator matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 2 & 2 \end{pmatrix}. \tag{1.3}$$

By Proposition 1.1, $\mathcal{K}_4$ is of type $4^1 2^2$ Therefore $|\mathcal{K}_4| = 16$. It follows from Proposition 1.2 that $\mathcal{K}_4^\perp$ is also of type $4^1 2^2$. Therefore $|\mathcal{K}_4^\perp| = 16$. It is obvious that any two rows of (1.3), distinct or not, are orthogonal. Therefore $\mathcal{K}_4 \subseteq \mathcal{K}_4^\perp$. Hence $\mathcal{K}_4 = \mathcal{K}_4^\perp$ and $\mathcal{K}_4$ is a self-dual code. □

Example 1.2. Let C_1 be the $\mathbb{Z}_4$-linear code with generator matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 2 \end{pmatrix} \tag{1.4}$$

It is clear that C_1 is self-orthogonal. By Proposition 1.1, C_1 is of type $4^1 2^1$ and by Proposition 1.2 $C_1^\perp$ is of type $4^2 2^1$. $C_1^\perp$ has generator matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & 0 & 1 \\ 2 & 2 & 0 & 0 \end{pmatrix}. \tag{1.5}$$

□

Example 1.3. Let $\mathcal{O}_8$ be the $\mathbb{Z}_4$-linear code with generator matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 3 & 1 & 2 & 1 \\ 0 & 1 & 0 & 0 & 1 & 2 & 3 & 1 \\ 0 & 0 & 1 & 0 & 3 & 3 & 3 & 2 \\ 0 & 0 & 0 & 1 & 2 & 3 & 1 & 1 \end{pmatrix} \tag{1.6}$$

By Proposition 1.1 $\mathcal{O}_8$ is of type 4^4 and then by Proposition 1.2 $\mathcal{O}_8^\perp$ is also of type 4^4 It is easy to check that any two rows of the generator matrix, distinct or not, are orthogonal. Therefore $\mathcal{O}_8 = \mathcal{O}_8^\perp$, i.e. $\mathcal{O}_8$ is self-dual. $\mathcal{O}_8$ is called the *octacode*. $\square$

Example 1.4. Let $\mathcal{K}_8$ be the $\mathbb{Z}_4$-linear code with generator matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 \end{pmatrix} . \tag{1.7}$$

By Proposition 1.1 $\mathcal{K}_8$ is of type $4^1 2^6$ and then by Proposition 1.2 $\mathcal{K}_8^\perp$ is also of type $4^1 2^6$. Clearly, any two rows of the generator matrix, distinct or not, are orthogonal. Therefore $\mathcal{K}_8 = \mathcal{K}_8^\perp$ and $\mathcal{K}_8$ is self-dual. $\square$

CHAPTER 2

WEIGHT ENUMERATORS

2.1. Weight Enumerators of Quaternary Codes

Let C be a $\mathbb{Z}_4$-code and n be its length. Let a be an element of $\mathbb{Z}_4$, i.e. $a = 0$, 1, 2 or 3. For all $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}_4^n$, define the *weight* of $\mathbf{x}$ at a to be

$$w_a(\mathbf{x}) = |\{i \mid x_i = a\}|.$$

Then the *complete weight enumerator* of C is defined to be the homogeneous polynomial of degree n in four indeterminates X_0, X_1, X_2 and X_3

$$W_C(X_0, X_1, X_2, X_3) = \sum_{\mathbf{c} \in C} X_0^{w_0(\mathbf{c})} X_1^{w_1(\mathbf{c})} X_2^{w_2(\mathbf{c})} X_3^{w_3(\mathbf{c})}, \tag{2.1}$$

(see Klemm (1987)).

Example 2.1. Let C_2 be the $\mathbb{Z}_4$-linear codes with generator matrix

$$\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$$

Then $|C_2| = 8$ and C_2 consists of the following eight codewords

$$(0,\ 0),\ (1,\ 1),\ (2,\ 2),\ (3,\ 3),\ (0,\ 2),\ (1,\ 3),\ (2,\ 0),\ (3,\ 1).$$

The numbers $w_a(\mathbf{c})$, where $a \in \mathbb{Z}_4$ and $\mathbf{c} \in C_2$ are shown in the following table.

Table 2.1.

	w_0	w_1	w_2	w_3
$(0,0)$	2	0	0	0
$(1,1)$	0	2	0	0
$(2,2)$	0	0	2	0
$(3,3)$	0	0	0	2
$(0,2)$	1	0	1	0
$(1,3)$	0	1	0	1
$(2,0)$	1	0	1	0
$(3,1)$	0	1	0	1

Therefore by (2.1) we have

$$W_{C_2}(X_0, X_1, X_2, X_3) = X_0^2 + X_1^2 + X_2^2 + X_3^2 + 2X_0X_2 + 2X_1X_3. \qquad (2.2)$$

□

Example 2.2. Let $\mathcal{K}_4$ be the $\mathbb{Z}_4$-linear code introduced in Example 1.1. $\mathcal{K}_4$ has 16 codewords and the numbers $w_a(\mathbf{c})$, where $a \in \mathbb{Z}_4$ and $\mathbf{c} \in \mathcal{K}_4$, are shown in the following table.

Table 2.2.

	w_0	w_1	w_2	w_3
0000	4	0	0	0
1111	0	4	0	0
2222	0	0	4	0
3333	0	0	0	4
0202	2	0	2	0
1313	0	2	0	2
2020	2	0	2	0
3131	0	2	0	2
0022	2	0	2	0
1133	0	2	0	2
2200	2	0	2	0
3311	0	2	0	2
0220	2	0	2	0
1331	0	2	0	2
2002	2	0	2	0
3113	0	2	0	2

Therefore

$$W_{\mathcal{K}_4}(X_0, X_1, X_2, X_3) = X_0^4 + X_1^4 + X_2^4 + X_3^4 + 6X_0^2X_2^2 + 6X_1^2X_3^2 . \qquad (2.3)$$

□

Let f be a function defined on $\mathbb{Z}_4^n$ with values in $\mathbb{C}[X_0, X_1, X_2, X_3]$. The *Hadamard transform* of f, denoted by $\hat{f}$, is defined by

$$\hat{f}(\mathbf{x}) = \sum_{\mathbf{y}\in\mathbb{Z}_4^n} i^{\,\mathbf{x}\cdot\mathbf{y}} f(\mathbf{y}) \quad \text{for all} \quad \mathbf{x} \in \mathbb{Z}_4^n , \qquad (2.4)$$

where $i = \sqrt{-1}$.

Lemma 2.1. *Let C be a $\mathbb{Z}_4$-linear code of length n. Then*

$$\sum_{\mathbf{x}\in C^\perp} f(\mathbf{x}) = \frac{1}{|C|} \sum_{\mathbf{x}\in C} \hat{f}(\mathbf{x}).$$

Proof. We have

$$\begin{aligned} \sum_{\mathbf{x}\in C} \hat{f}(\mathbf{x}) &= \sum_{\mathbf{x}\in C} \sum_{\mathbf{y}\in\mathbb{Z}_4^n} i^{\mathbf{x}\cdot\mathbf{y}} f(\mathbf{y}) \\ &= \sum_{\mathbf{y}\in\mathbb{Z}_4^n} f(\mathbf{y}) \sum_{\mathbf{x}\in C} i^{\mathbf{x}\cdot\mathbf{y}}. \end{aligned}$$

For $\mathbf{y} \in C^\perp$, $\mathbf{x} \cdot \mathbf{y} = 0$ and $i^{\mathbf{x}\cdot\mathbf{y}} = i^0 = 1$ for all $\mathbf{x} \in C$, then the inner sum is equal to $|C|$. For $\mathbf{y} \notin C^\perp$, as $\mathbf{x}$ runs through C, either $\mathbf{x} \cdot \mathbf{y}$ takes values 0, 1, 2, 3 equally often or only values 0, 2 equally often. But $i^0 + i^1 + i^2 + i^3 = 0$ and $i^0 + i^2 = 0$, so the inner sum is zero. Therefore

$$\sum_{\mathbf{x}\in C} \hat{f}(\mathbf{x}) = |C| \sum_{\mathbf{y}\in C^\perp} f(\mathbf{y}) .$$

□

We have the following generalization of MacWilliams identity to $\mathbb{Z}_4$-linear codes.

Theorem 2.2. *Let C be a $\mathbb{Z}_4$-linear code, then*

$$\begin{aligned} W_{C^\perp}(X_0, X_1, X_2, X_3) = \frac{1}{|C|} W_C\,(&X_0 + X_1 + X_2 + X_3, X_0 + iX_1 - X_2 - iX_3, \\ &X_0 - X_1 + X_2 - X_3, X_0 - iX_1 - X_2 + iX_3). \end{aligned}$$

Proof. Let $f(\mathbf{x}) = X_0^{w_0(\mathbf{x})} X_1^{w_1(\mathbf{x})} X_2^{w_2(\mathbf{x})} X_3^{w_3(\mathbf{x})}$ for all $\mathbf{x} \in \mathbb{Z}_4^n$. Let us compute the Hadamard transform $\hat{f}(\mathbf{x})$ of $f(\mathbf{x})$. By (2.4),

$$\hat{f}(\mathbf{x}) = \sum_{\mathbf{y}\in\mathbb{Z}_4^n} i^{\mathbf{x}\cdot\mathbf{y}} X_0^{w_0(\mathbf{y})} X_1^{w_1(\mathbf{y})} X_2^{w_2(\mathbf{y})} X_3^{w_3(\mathbf{y})}.$$

Clearly

$$i^{\mathbf{x}\cdot\mathbf{y}} = i^{x_1 y_1} i^{x_2 y_2} \cdots i^{x_n y_n}$$

and for $a \in \mathbb{Z}_4$,

$$w_a(\mathbf{y}) = \delta_{a,y_1} + \delta_{a,y_2} + \cdots + \delta_{a,y_n},$$

where δ is the Kronecker delta. Then $\hat{f}(\mathbf{x})$ can be written as

$$\begin{aligned}
\hat{f}(\mathbf{x}) &= \sum_{\mathbf{y}\in\mathbb{Z}_4^n} (i^{x_1 y_1} X_0^{\delta_{0,y_1}} X_1^{\delta_{1,y_1}} X_2^{\delta_{2,y_1}} X_3^{\delta_{3,y_1}}) \cdots \\
&\quad \times (i^{x_n y_n} X_0^{\delta_{0,y_n}} X_1^{\delta_{1,y_n}} X_2^{\delta_{2,y_n}} X_3^{\delta_{3,y_n}}) \\
&= \left(\sum_{y_1\in\mathbb{Z}_4} i^{x_1 y_1} X_0^{\delta_{0,y_1}} X_1^{\delta_{1,y_1}} X_2^{\delta_{2,y_1}} X_3^{\delta_{3,y_1}} \right) \cdots \\
&\quad \times \left(\sum_{y_n\in\mathbb{Z}_4} i^{x_n y_n} X_0^{\delta_{0,y_n}} X_1^{\delta_{1,y_n}} X_2^{\delta_{2,y_n}} X_3^{\delta_{3,y_n}} \right) \\
&= \left(\sum_{k=0}^{3} i^{x_1 k} X_k \right) \cdots \left(\sum_{k=0}^{3} i^{x_n k} X_k \right) \\
&= \prod_{j=0}^{3} \left(\sum_{k=0}^{3} i^{jk} X_k \right)^{w_j(\mathbf{x})} \qquad (2.5)
\end{aligned}$$

The last equality follows from the observation that when $x_l = j$, $\sum_{k=0}^{3} i^{x_l k} X_k = \sum_{k=0}^{3} i^{jk} X_k$, and there are $w_j(\mathbf{x})$'s x_l equal to j, which contribute together $(\sum_{k=0}^{3} i^{jk} X_k)^{w_j(\mathbf{x})}$

For $\mathbf{c} \in \mathcal{C}$, $f(\mathbf{c}) = X_0^{w_0(\mathbf{c})} X_1^{w_1(\mathbf{c})} X_2^{w_2(\mathbf{c})} X_3^{w_3(\mathbf{c})}$, then by Lemma 2.1 and (2.5), we have

$$\begin{aligned}
W_{\mathcal{C}^\perp}(X_0, X_1, X_2, X_3) &= \sum_{\mathbf{c}\in\mathcal{C}^\perp} X_0^{w_0(\mathbf{c})} X_1^{w_1(\mathbf{c})} X_2^{w_2(\mathbf{c})} X_3^{w_3(\mathbf{c})} \\
&= \sum_{\mathbf{c}\in\mathcal{C}^\perp} f(\mathbf{c})
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{|C|} \sum_{\mathbf{c} \in C} \hat{f}(\mathbf{c}) \\
&= \frac{1}{|C|} \sum_{\mathbf{c} \in C} \left(\sum_{k=0}^{3} i^{0k} X_k \right)^{w_0(\mathbf{c})} \left(\sum_{k=0}^{3} i^{1k} X_k \right)^{w_1(\mathbf{c})} \\
&\quad \times \left(\sum_{k=0}^{3} i^{2k} X_k \right)^{w_2(\mathbf{c})} \left(\sum_{k=0}^{3} i^{3k} X_k \right)^{w_3(\mathbf{c})} \\
&= \frac{1}{|C|} W_C \left(\sum_{k=0}^{3} i^{0k} X_k, \sum_{k=0}^{3} i^{1k} X_k, \sum_{i=0}^{3} i^{2k} X_k, \sum_{i=0}^{3} i^{3k} X_k \right) \\
&= \frac{1}{|C|} W_C(X_0 + X_1 + X_2 + X_3, X_0 + iX_1 - X_2 - iX_3, \\
&\quad X_0 - X_1 + X_2 - X_3, X_0 - iX_1 - X_2 + iX_3).
\end{aligned}
$$

□

Theorem 2.2 is from Klemm (1987).

Example 2.3. Let

$$C_3 = \{(0, 0), (2, 2)\}.$$

Clearly, C_3 is a $\mathbb{Z}_4$-linear code with weight enumerator

$$W_{C_3}(X_0, X_1, X_2, X_3) = X_0^2 + X_2^2.$$

It is easy to verify that $C_3^{\perp} = C_2$ where C_2 is the $\mathbb{Z}_4$-linear code appeared in Example 2.1. By Theorem 2.2,

$$
\begin{aligned}
W_{C_2}(X_0, X_1, X_2, X_3) &= W_{C_3^{\perp}}(X_0, X_1, X_2, X_3) \\
&= \frac{1}{2} W_{C_3}(X_0 + X_1 + X_2 + X_3, X_0 + iX_1 - X_2 - iX_3, \\
&\quad X_0 - X_1 + X_2 - X_3, X_0 - iX_1 - X_2 + iX_3) \\
&= \frac{1}{2}[(X_0 + X_1 + X_2 + X_3)^2 + (X_0 - X_1 + X_2 - X_3)^2] \\
&= X_0^2 + X_1^2 + X_2^2 + X_3^2 + 2X_0X_2 + 2X_1X_3,
\end{aligned}
$$

which coincides with (2.2). □

Example 2.4. The $\mathbb{Z}_4$-linear code $\mathcal{C}_1$ in Example 1.2 has eight codewords. It is easy to compute the weight enumerator of $\mathcal{C}_1$.

$$W_{\mathcal{C}_1}(X_0, X_1, X_2, X_3) = X_0^4 + X_1^4 + X_2^4 + X_3^4 + 2X_0^2 X_2^2 + 2X_1^2 X_3^2$$

$\mathcal{C}_1^\perp$ has 32 codewords, but the weight enumerator can be computed by Theorem 2.2.

$$\begin{aligned}
&W_{\mathcal{C}_1^\perp}(X_0, X_1, X_2, X_3) \\
&= \frac{1}{8} W_{\mathcal{C}_1}(X_0 + X_1 + X_2 + X_3, X_0 + iX_1 - X_2 - iX_3, \\
&\qquad X_0 - X_1 + X_2 - X_3, X_0 - iX_1 - X_2 + iX_3) \\
&= \frac{1}{8}\big[(X_0 + X_1 + X_2 + X_3)^4 + (X_0 + iX_1 - X_2 - iX_3)^4 \\
&\quad + (X_0 - X_1 + X_2 - X_3)^4 + (X_0 - iX_1 - X_2 + iX_3)^4 \\
&\quad + 2(X_0 + X_1 + X_2 + X_3)^2 (X_0 - X_1 + X_2 - X_3)^2 \\
&\quad + 2(X_0 + iX_1 - X_2 - iX_3)^2 (X_0 - iX_1 - X_2 + iX_3)^2\big].
\end{aligned}$$

□

The complete weight enumerator of a $\mathbb{Z}_4$-code $\mathcal{C}$ is usually denoted by

$$\begin{aligned}
\mathrm{cwe}_{\mathcal{C}}(X_0, X_1, X_2, X_3) &= W_{\mathcal{C}}(X_0, X_1, X_2, X_3) \\
&= \sum_{c \in \mathcal{C}} X_0^{w_0(c)} X_1^{w_1(c)} X_2^{w_2(c)} X_3^{w_3(c)}
\end{aligned} \tag{2.6}$$

Permutation equivalent codes have the same complete weight enumerator but equivalent codes may have distinct complete weight enumerators. The appropriate weight enumerator for an equivalence class of codes is the *symmetrized weight enumerator*, obtained by identifying X_1 and X_3 in (2.6)

$$\begin{aligned}
\mathrm{swe}_{\mathcal{C}}(X_0, X_1, X_2) &= \mathrm{cwe}_{\mathcal{C}}(X_0, X_1, X_2, X_1) \\
&= \sum_{c \in \mathcal{C}} X_0^{w_0(c)} X_1^{w_1(c)+w_3(c)} X_2^{w_2(c)},
\end{aligned} \tag{2.7}$$

which is a homogeneous polynomial of degree n in X_0, X_1 and X_2 (see Conway and Sloane (1993a)).

Example 2.5. The symmetrized weight enumerators of $\mathcal{K}_4$, $\mathcal{C}_2$ and $\mathcal{C}_3$ are:

$$\mathrm{swe}_{\mathcal{K}_4}(X_0, X_1, X_2) = X_0^4 + 8X_1^4 + X_2^4 + 6X_0^2 X_2^2, \tag{2.8}$$

$$\mathrm{swe}_{C_2}(X_0, X_1, X_2) = X_0^2 + 4X_1^2 + X_2^2 + 2X_0X_2 \tag{2.9}$$

and

$$\mathrm{swe}_{C_3}(X_0, X_1, X_2) = X_0^2 + X_2^2, \tag{2.10}$$

respectively. □

Example 2.6. The complete weight enumerator of the octacode $\mathcal{O}_8$ is

$$\begin{aligned}\mathrm{cwe}_{\mathcal{O}_8}(X_0, X_1, X_2, X_3) = {} & X_0^8 + X_1^8 + X_2^8 + X_3^8 + 14\,(X_0^4X_2^4 + X_1^4X_3^4) \\ & + 56\,(X_0^3X_1^3X_2X_3 + X_0^3X_1X_2X_3^3 \\ & + X_0X_1^3X_2^3X_3 + X_0X_1X_2^3X_3^3)\end{aligned} \tag{2.11}$$

and the symmetrized weight enumerator of $\mathcal{O}_8$ is

$$\begin{aligned}\mathrm{swe}_{\mathcal{O}_8}(X_0, X_1, X_2) = {} & X_0^8 + 16X_1^8 + X_2^8 + 14X_0^4X_2^4 \\ & + 112X_0X_1^4X_2\,(X_0^2 + X_2^2).\end{aligned} \tag{2.12}$$

□

From Theorem 2.2 follows the following generalization of MacWilliams identity for swe_C.

Theorem 2.3. *Let C be a $\mathbb{Z}_4$-linear code, then*

$$\mathrm{swe}_{C^\perp}(X_0, X_1, X_2) = \frac{1}{|C|}\,\mathrm{swe}_C(X_0 + 2X_1 + X_2,\ X_0 - X_2,\ X_0 - 2X_1 + X_2).$$

Proof. By (2.7) and Theorem 2.2,

$$\begin{aligned}\mathrm{swe}_{C^\perp}(X_0, X_1, X_2) &= \mathrm{cwe}_{C^\perp}(X_0, X_1, X_2, X_1) \\ &= \frac{1}{|C|}\,\mathrm{cwe}_C(X_0 + X_1 + X_2 + X_1,\ X_0 + iX_1 - X_2 - iX_1, \\ &\qquad X_0 - X_1 + X_2 - X_1,\ X_0 - iX_1 - X_2 + iX_1) \\ &= \frac{1}{|C|}\,\mathrm{cwe}_C(X_0 + 2X_1 + X_2, X_0 - X_2, \\ &\qquad X_0 - 2X_1 + X_2,\ X_0 - X_2) \\ &= \frac{1}{|C|}\,\mathrm{swe}_C(X_0 + 2X_1 + X_2,\ X_0 - X_2,\ X_0 - 2X_1 + X_2).\end{aligned}$$

□

Example 2.7. The symmetrized weight enumerator of $\mathcal{C}_3$ is given by (2.10). We know that $\mathcal{C}_2 = \mathcal{C}_3^{\perp}$. Therefore by Theorem 2.3, we have

$$\begin{aligned}\mathrm{swe}_{\mathcal{C}_2}(X_0, X_1, X_2) &= \frac{1}{|\mathcal{C}_3|}\,\mathrm{swe}_{\mathcal{C}_3}(X_0 + 2X_1 + X_2,\, X_0 - X_2,\, X_0 - 2X_1 + X_2)\\ &= \frac{1}{2}\,[(X_0 + 2X_1 + X_2)^2 + (X_0 - 2X_1 + X_2)^2]\\ &= X_0^2 + 4X_1^2 + X_2^2 + 2X_0X_2,\end{aligned}$$

which coincides with (2.9). □

The *Lee weights* of 0, 1, 2, 3 $\in \mathbb{Z}_4$, denoted by $w_{\mathrm{L}}(0)$, $w_{\mathrm{L}}(1)$, $w_{\mathrm{L}}(2)$, $w_{\mathrm{L}}(3)$, respectively, are defined by

$$w_{\mathrm{L}}(0) = 0,\ w_{\mathrm{L}}(1) = w_{\mathrm{L}}(3) = 1,\ w_{\mathrm{L}}(2) = 2.$$

The *Lee weight* $w_{\mathrm{L}}(\mathbf{x})$ of $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}_4^n$ is defined to be the integral sum of the Lee weights of its components

$$w_{\mathrm{L}}(\mathbf{x}) = \sum_{i=1}^{n} w_{\mathrm{L}}(x_i).$$

This weight function defines a distance function

$$d_{\mathrm{L}}(\mathbf{x}, \mathbf{y}) = w_{\mathrm{L}}(\mathbf{x} - \mathbf{y})$$

on $\mathbb{Z}_4^n$, which is called the *Lee distance*.

Actually when we use $\mathbb{Z}_4$-codes in communication, the four alphabets 0, 1, 2, 3 are usually used to represent the signal points $i^0 = 1$, $i^1 = i$, $i^2 = -1$, $i^3 = -i$, respectively, in the complex plane. Denote by $d_{\mathrm{E}}^2\,(i^a, i^b)$ the square of the *Euclidean distance* between i^a and i^b. Then

$$d_{\mathrm{L}}(a, b) = \frac{1}{2}\, d_{\mathrm{E}}^2\,(i^a, i^b).$$

More generally, to any $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}_4^n$ there corresponds a complex vector

$$i^{\mathbf{x}} = (i^{x_1}, \ldots, i^{x_n}).$$

For any $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_4^n$, the square of the *Euclidean distance* between $i^{\mathbf{x}}$ and $i^{\mathbf{y}}$ is given by

$$d_{\mathrm{E}}^2(i^{\mathbf{x}}, i^{\mathbf{y}}) = \sum_{i=1}^{n} d_{\mathrm{E}}^2(i^{x_i}, i^{y_i}).$$

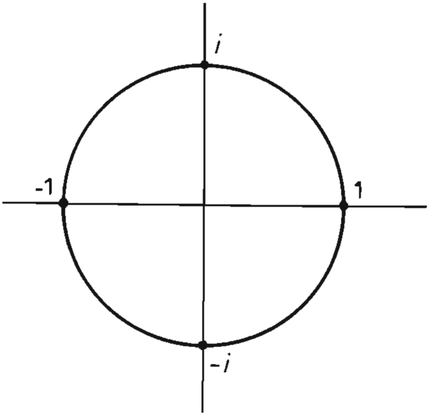

Fig. 2.1.

Then

$$d_{\mathrm{L}}(\mathbf{x}, \mathbf{y}) = \frac{1}{2} d_{\mathrm{E}}^2(i^{\mathbf{x}}, i^{\mathbf{y}}). \tag{2.13}$$

This explains why we introduce the Lee weight and Lee distance in $\mathbb{Z}_4^n$.

The *Lee weight enumerator* of a $\mathbb{Z}_4$-code C of length n is defined to be

$$\mathrm{Lee}_C(X, Y) = \sum_{\mathbf{c} \in C} X^{2n - w_{\mathrm{L}}(\mathbf{c})} Y^{w_{\mathrm{L}}(\mathbf{c})}, \tag{2.14}$$

(see Hammons *et al.* (1994)). It is obvious that

$$w_{\mathrm{L}}(\mathbf{x}) = w_1(\mathbf{x}) + 2w_2(\mathbf{x}) + w_3(\mathbf{x}) \quad \text{for all} \quad \mathbf{x} \in \mathbb{Z}_4^n.$$

From (2.7) and (2.14) we deduce that

$$\mathrm{Lee}_C(X, Y) = \mathrm{swe}_C(X^2, XY, Y^2), \tag{2.15}$$

which is a homogeneous polynomial of degree $2n$. From Theorem 2.3 follows also the following generalization of MacWilliams identity for Lee_C.

Theorem 2.4. *Let C be a $\mathbb{Z}_4$-linear code, then*

$$\mathrm{Lee}_{C^\perp}(X, Y) = \frac{1}{|C|} \mathrm{Lee}_C(X + Y, X - Y).$$

Proof. By (2.15) and Theorem 2.3,

$$\begin{aligned}
\mathrm{Lee}_{C^\perp}(X, Y) &= \mathrm{swe}_{C^\perp}(X^2, XY, Y^2) \\
&= \frac{1}{|C|}\, \mathrm{swe}_C(X^2 + 2XY + Y^2, X^2 - Y^2, X^2 - 2XY + Y^2) \\
&= \frac{1}{|C|}\, \mathrm{swe}_C[(X+Y)^2, (X+Y)(X-Y), (X-Y)^2] \\
&= \frac{1}{|C|}\, \mathrm{Lee}_C(X+Y, X-Y).
\end{aligned}$$

□

The *Hamming weight* $w_{\mathrm{H}}(\mathbf{x})$ of $\mathbf{x} \in \mathbb{Z}^n$ is defined to be

$$w_{\mathrm{H}}(\mathbf{x}) = w_1(\mathbf{x}) + w_2(\mathbf{x}) + w_3(\mathbf{x}).$$

This weight function defines also a distance function

$$d_{\mathrm{H}}(\mathbf{x}, \mathbf{y}) = w_{\mathrm{H}}(\mathbf{x} - \mathbf{y})$$

on $\mathbb{Z}_4^n$, which is called the *Hamming distance* between $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_4^n$. The *Hamming weight enumerator* of a $\mathbb{Z}_4$-code C of length n is defined to be

$$\mathrm{Ham}_C(X, Y) = \sum_{\mathbf{c} \in C} X^{n - w_{\mathrm{H}}(\mathbf{x})} Y^{w_{\mathrm{H}}(\mathbf{x})}, \tag{2.16}$$

(see Conway and Sloane (1993a)). It is obvious that

$$\begin{aligned}
\mathrm{Ham}_C(X, Y) &= \mathrm{cwe}_e(X, Y, Y, Y) \\
&= \mathrm{swe}_c(X, Y, Y).
\end{aligned} \tag{2.17}$$

We also have

Theorem 2.5. *Let C be a $\mathbb{Z}_4$-linear code, then*

$$\mathrm{Ham}_{C^\perp}(X, Y) = \frac{1}{|C|}\, \mathrm{Ham}_C(X + 3Y, X - Y).$$

□

2.2. Krawtchouk Polynomials

Let n be a fixed positive integer, q a prime power, and x an indeterminate. The polynomials

$$K_k(x) = K_k(x, n) = \sum_{j=0}^{k} (-1)^j (q-1)^{k-j} \binom{x}{j} \binom{n-x}{k-j}, \quad k = 0, 1, 2 \ldots, \tag{2.18}$$

are called the *Krawtchouk polynomials*, where

$$\binom{x}{j} = \begin{cases} \dfrac{x(x-1)\cdots(x-j+1)}{j!}, & \text{if } j \text{ is a positive integer,} \\ 1, & \text{if } j = 0, \\ 0, & \text{otherwise}, \end{cases}$$

see Krawtchouk (1929), (1933).

Let C be a code of length n over $\mathbb{F}_q$, not necessarily linear. For any $\mathbf{c} = (c_1, c_2, \ldots, c_n) \in C$, define the *Hamming weight* of $\mathbf{c}$ to be

$$w(\mathbf{c}) = |\{j \mid c_j \neq 0\}|.$$

Let A_i be the number of codewords of Hamming weight i in C, then $\{A_0, A_1, \ldots, A_n\}$ is called the *weight distribution* of code C. Define

$$W_C(X, Y) = \sum_{i=0}^{n} A_i X^{n-i} Y^i$$

and call it the *weight enumerator* of code C.

Proposition 2.6. *Let C and C' be codes of length n over $\mathbb{F}_q$, and A_i and A'_i be the number of codewords of weight i in C and C', respectively. If*

$$W_{C'}(X,Y) = \frac{1}{|C|} W_C(X + (q-1)Y, X - Y), \tag{2.19}$$

then

$$A'_k = \frac{1}{|C|} \sum_{i=0}^{n} A_i K_k(i), \quad k = 0, 1, 2, \ldots, n, \tag{2.20}$$

and conversely.

Proof. By definition,

$$W_C(X + (q-1)Y, X - Y) = \sum_{i=0}^{n} A_i (X + (q-1)Y)^{n-i} (X - Y)^i$$

Expanding, we obtain

$$\begin{aligned} W_C(X + (q-1)Y, X - Y) = \sum_{i=0}^{n} A_i \sum_{j=0}^{n-i} \binom{n-i}{j} X^{n-i-j} ((q-1)Y)^j \\ \times \sum_{l=0}^{i} (-1)^l \binom{i}{l} X^{i-l} Y^l \end{aligned}$$

$$= \sum_{i=0}^{n} A_i \sum_{k=0}^{n} \sum_{j=0}^{k} (-1)^j (q-1)^{k-j} \binom{i}{j} \binom{n-i}{k-j} X^{n-k} Y^k$$

$$= \sum_{i=0}^{n} A_i \sum_{k=0}^{n} K_k(i) X^{n-k} Y^k$$

$$= \sum_{k=0}^{n} \sum_{i=0}^{n} A_i K_k(i) X^{n-k} Y^k$$

Therefore (2.19) holds if and only if (2.20) holds. □

In particular, let C be a linear code of length n over $\mathbb{F}_q$ and $C^\perp$ be its dual code, then their weight enumerators $W_C(X, Y)$ and $W_{C^\perp}(X, Y)$ are connected by the MacWilliams identity

$$W_{C^\perp}(X, Y) = \frac{1}{|C|} W_C(X + (q-1)Y, X - Y).$$

Thus by Proposition 2.6 their weight distributions $\{A_0, A_1, \ldots, A_n\}$ and $\{A'_0, A'_1, \ldots, A'_n\}$ are connected by (2.20), in which the values of the Krawtchouk polynomials appear.

It is worthwhile to exploit some properties of the Krawtchouk polynomials. First, the generating function of the Krawtchouk polynomials is given by

Proposition 2.7. (Generating Function) *Let z be an indeterminate, then*

$$\sum_{k=0}^{\infty} K_k(x) z^k = (1 + (q-1) z)^{n-x} (1 - z)^x. \tag{2.21}$$

When $x = i$ is a non-negative integer $\leq n$,

$$\sum_{k=0}^{n} K_k(i) z^k = (1 + (q-1) z)^{n-i} (1 - z)^i \tag{2.22}$$

Proof. We have the binomial series

$$(1 + (q-1) z)^{n-x} = \sum_{j=0}^{\infty} \binom{n-x}{j} ((q-1) z)^j,$$

$$(1 - z)^x = \sum_{l=0}^{\infty} (-1)^l \binom{x}{l} z^l.$$

Multiplying the above two expressions together, we obtain

$$(1+(q-1)\,z)^{n-x}\,(1-z)^{x} = \sum_{k=0}^{\infty}\sum_{j=0}^{k}(-1)^j\,(q-1)^{k-j}\binom{x}{j}\binom{n-x}{k-j}\,z^k$$
$$= \sum_{k=0}^{\infty} K_k(x)\,z^k.$$

When $x = i$ is a non-negative integer $\le n$, all the summations in the above deduction are finite and finally we obtain (2.22). □

We also have alternative expressions of $K_k(x)$.

Proposition 2.8. (Alternative Expressions)

(i) $$K_k(x) = \sum_{j=0}^{k} (-q)^j\,(q-1)^{k-j}\binom{n-j}{k-j}\binom{x}{j}, \tag{2.23}$$

(ii) $$K_k(x) = \sum_{j=0}^{k} (-1)^j\,q^{k-j}\binom{n-k+j}{j}\binom{n-x}{k-j}. \tag{2.24}$$

Proof. (i) By Proposition 2.7,

$$\sum_{k=0}^{\infty} K_k(x)\,z^k = [1+(q-1)\,z]^{n-x}\,(1-z)^x$$
$$= [1+(q-1)\,z]^n\left(\frac{1-z}{1+(q-1)\,z}\right)^x$$
$$= [1+(q-1)\,z]^n\left(1-\frac{qz}{1+(q-1)\,z}\right)^x$$
$$= [1+(q-1)\,z]^n\sum_{j=0}^{\infty}(-1)^j\binom{x}{j}\left(\frac{qz}{1+(q-1)\,z}\right)^j$$
$$= \sum_{j=0}^{\infty}(-q)^j\binom{x}{j}\,z^j\,(1+(q-1)\,z)^{n-j}$$

$$= \sum_{j=0}^{\infty} (-q)^j \binom{x}{j} z^j \sum_{l=0}^{\infty} \binom{n-j}{l} ((q-1)\, z)^l$$

$$= \sum_{k=0}^{\infty} \left(\sum_{j=0}^{k} (-q)^j (q-1)^{k-j} \binom{n-j}{k-j} \binom{x}{j} \right) z^k .$$

Equating the coefficients of z^k, we obtain (2.23).

(ii) can be proved in a similar way, by starting from

$$\sum_{k=0}^{\infty} K_k(x)\, z^k = (1-z)^n \left(\frac{1-(q-1)\,z}{1-z} \right)^{n-x} \qquad \square$$

Corollary 2.9. *$K_k(x)$ is a polynomial of degree k, whose leading coefficient is $(-1)^k \frac{q^k}{k!}$ and constant term is $(q-1)^k \binom{n}{k}$.*

Proof. The terms with $j < k$ on the R.H.S. of (2.23) are polynomials of degree $\leq j < k$, and the term with $j = k$ is $(-q)^k \binom{x}{k}$, which is a polynomial of degree k with leading coefficient $(-1)^k \frac{q^k}{k!}$. Putting $x = 0$ in (2.23), we obtain the constant term $K_k(0) = (q-1)^k \binom{n}{k}$. $\square$

Proposition 2.10. (Three Terms Recurrence)

$$(k+1)\, K_{k+1}(x) + (qx - n(q-1) + k(q-2))\, K_k(x) + (n-k+1)\, (q-1)\, K_{k-1}(x) = 0. \tag{2.25}$$

Proof. Differentiating both sides of (2.21) with respect to z, we obtain

$$\sum_{k=1}^{\infty} K_k(x)\, k\, z^{k-1} = (n-x)\, [1+(q-1)\, z]^{n-x-1}\, (q-1)\, (1-z)^x$$
$$+\ [1+(q-1)\, z]^{n-x}\, x(1-z)^{x-1}(-1).$$

Multiplying both sides of the above equation by $[1+(q-1)\, z]\, (1-z)$ and then substituting (2.21) into it, we get

$$\left(\sum_{k=1}^{\infty} K_k(x)\, k\, z^{k-1} \right) [1+(q-1)\, z]\ (1-z)$$

$$= \left(\sum_{k=0}^{\infty} K_k(x)\, z^k \right) [(n-x)\, (q-1)\, (1-z) - (1+(q-1)\, z)\, x] .$$

Equating the coefficients of z^k, we have

$$\begin{aligned}&(k+1)\,K_{k+1}(x) + k(q-2)\,K_k(x) - (k-1)\,(q-1)\,K_{k-1}(x)\\ &\quad = ((n-x)\,(q-1) - x)\,K_k(x) + (-(n-x)\,(q-1) - (q-1)\,x)\,K_{k-1}(x).\end{aligned}$$

Transposing and simplifying, we obtain (2.25). □

Proposition 2.11. (Orthogonality Relation) *For non-negative integers r and s*

$$\sum_{i=0}^{n} \binom{n}{i} (q-1)^i\, K_r(i)\, K_s(i) = q^n (q-1)^r \binom{n}{r} \delta_{rs}\,, \tag{2.26}$$

where δ_{rs} is the Kronecker delta.

Proof. By (2.22), the L.H.S. of (2.26) is the coefficient of $y^r z^s$ in

$$\sum_{i=0}^{n} \binom{n}{i} (q-1)^i\, [1+(q-1)\,y]^{n-i}\,(1-y)^i\,[1+(q-1)\,z]^{n-i}\,(1-z)^i\,. \tag{2.27}$$

Clearly, the expression (2.27) is equal to

$$\begin{aligned}&\sum_{i=0}^{n} \binom{n}{i} [(1+(q-1)\,y)\,(1+(q-1)\,z)]^{n-i}\,[(q-1)\,(1-y)\,(1-z)]^i\\ &\quad = [(1+(q-1)\,y)\,(1+(q-1)\,z) + (q-1)\,(1-y)\,(1-z)]^n\\ &\quad = q^n\,[1+(q-1)\,yz]^n\\ &\quad = q^n \sum_{r=0}^{n} \binom{n}{r} (q-1)^r y^r z^r\,.\end{aligned}$$

The coefficient of $y^r z^s$ in the above expression is equal to the R.H.S. of (2.26). □

Proposition 2.12. *For non-negative integers r and s*

$$\binom{n}{r} (q-1)^r\, K_s(r) = \binom{n}{s} (q-1)^s\, K_r(s).$$

Proof. This follows from (2.18) by rearranging the binomial coefficients. □

Corollary 2.13. *For non-negative integers r and s,*

$$\sum_{i=0}^{n} K_r(i)\, K_i(s) = q^n\, \delta_{rs}\,.$$

Proof. This follows from Propositions 2.11 and 2.12. □

Proposition 2.14. *Let $\alpha(x)$ be a polynomial of degree m, then $\alpha(x)$ can be expressed as*

$$\alpha(x) = \sum_{k=0}^{m} \alpha_k\, K_k(x), \tag{2.28}$$

where

$$\alpha_k = q^{-n} \sum_{i=0}^{n} \alpha(i)\, K_i(k), \quad k = 0, 1, \ldots, m. \tag{2.29}$$

Proof. By Corollary 2.9, $K_k(x)$ is a polynomial of degree k, thus $\alpha(x)$ can be expressed as (2.28). Substituting $x = i$ into (2.28), then multiplying it by $K_i(l)$, and then summing on i, by Corollary 2.13 we obtain (2.29). □

(2.28) is called the *Krawtchouk expansion* of the polynomial $\alpha(x)$ of degree m and the coefficients α_k ($k = 0, 1, \ldots, m$) in (2.28) are called the *Krawtchouk coefficients* of the expansion.

In the following we are mainly interested in Krawtchouk polynomials $K_k(x)$ for the case $q = 2$. From the proceeding results we have

Proposition 2.15. *Let $q = 2$. Then*

(i) (Definition)

$$K_k(x) = K_k(x,\, n) = \sum_{j=0}^{k} (-1)^j \binom{x}{j} \binom{n-x}{k-j}$$

for any integer $k \geq 0$.

(ii) (Generating Function)

$$\sum_{k=0}^{\infty} K_k(x) z^k = (1+z)^{n-x} (1-z)^x\,,$$

$$\sum_{k=0}^{n} K_k(i)z^k = (1+z)^{n-i}(1-z)^i$$

for any integer i with $0 \le i \le n$.

(iii) (Alternative Expressions)

$$K_k(x) = \sum_{j=0}^{k} (-2)^j \binom{n-j}{k-j}\binom{x}{j},$$

$$K_k(x) = \sum_{j=0}^{k} (-1)^j 2^{k-j} \binom{n-k+j}{j}\binom{n-x}{k-j}$$

(iv) *Leading coefficient of $K_k(x) = (-1)^k \frac{2^k}{k!}$ Constant term of $K_k(x)$ $= \binom{n}{k}$.*

(v) (Three Terms Recurrence)

$$(k+1)K_{k+1}(x) + (2x-n)K_k(x) + (n-k+1)K_{k-1}(x) = 0\,.$$

(vi) (Orthogonality Relation)

$$\sum_{i=0}^{n} \binom{n}{i} K_r(i)K_s(i) = 2^n \binom{n}{r}\,\delta_{rs}$$

for integers r, $s \ge 0$.

(vii) $\binom{n}{r}K_s(r) = \binom{n}{s}K_r(s)$ *for integers r, $s \ge 0$.*

(viii) (Orthogonality Relation)

$$\sum_{i=0}^{n} K_r(i)K_i(s) = 2^n\,\delta_{rs}$$

for integers r, $s \ge 0$.

(ix) (Krawtchouk Expansion) *For any polynomial $\alpha(x)$ of degree m, if*

$$\alpha(x) = \sum_{k=0}^{m} \alpha_k K_k(x),$$

then

$$\alpha_k = 2^{-n} \sum_{i=0}^{n} \alpha(i)\,K_i(k), \quad k = 0,\,1,\ldots,m.$$

□

For latter purpose, we write down the first seven Krawtchouk polynomials for $q = 2$.

Proposition 2.16. *For $q = 2$, we have*

$$K_0(x) = 1,$$

$$K_1(x) = -2x + n,$$

$$K_2(x) = 2x^2 - 2nx + \binom{n}{2},$$

$$K_3(x) = -\frac{4}{3}x^3 + 2nx^2 - \left(n^2 - n + \frac{2}{3}\right)x + \binom{n}{3},$$

$$K_4(x) = \frac{2}{3}x^4 - \frac{4}{3}nx^3 + \left(n^2 - n + \frac{4}{3}\right)x^2 - \left(\frac{1}{3}n^3 - n^2 + \frac{4}{3}n\right)x + \binom{n}{4},$$

$$\begin{aligned} K_5(x) = &-\frac{4}{15}x^5 + \frac{2}{3}nx^4 - \left(\frac{2}{3}n^2 - \frac{2}{3}n + \frac{4}{3}\right)x^3 \\ &+ \left(\frac{1}{3}n^3 - n^2 + 2n\right)x^2 \\ &- \left(\frac{1}{12}n^4 - \frac{1}{2}n^3 + \frac{5}{4}n^2 - \frac{5}{6}n + \frac{2}{5}\right)x + \binom{n}{5}, \end{aligned}$$

$$\begin{aligned} K_6(x) = &\frac{4}{45}x^6 - \frac{4}{15}nx^5 + \left(\frac{1}{3}n^2 - \frac{1}{3}n + \frac{8}{9}\right)x^4 \\ &- \left(\frac{2}{9}n^3 - \frac{2}{3}n^2 + \frac{16}{9}n\right)x^3 \\ &+ \left(\frac{1}{12}n^4 - \frac{1}{2}n^3 + \frac{19}{12}n^2 - \frac{7}{6}n + \frac{46}{45}\right)x^2 \\ &- \left(\frac{1}{60}n^5 - \frac{1}{6}n^4 + \frac{25}{36}n^3 - \frac{7}{6}n^2 + \frac{46}{45}n\right)x + \binom{n}{6} \end{aligned}$$

Proof. The expressions of $K_0(x)$ and $K_1(x)$ can be obtained directly from the definition of Krawtchouk polynomials (Proposition 2.15(i)). The expressions of the latter five can be derived from the three terms recurrence (Proposition 2.15(v)). □

2.3. Distance Enumerators of Binary Codes

In later chapters we shall study several binary nonlinear codes, for which the distance distributions and distance enumerators play an important role.

Let C be a binary code of length n, which is not necessarily linear. Define

$$B_i = |C|^{-1} |\{(\mathbf{c}, \mathbf{c}') \,|\, \mathbf{c}, \mathbf{c}' \in C,\ d(\mathbf{c}, \mathbf{c}') = i\}|,\ i = 0, 1, \ldots, n,$$

where d is the Hamming distance on $\mathbb{F}_2^n$. Clearly $B_0 = 1$ and $\sum_{i=0}^{n} B_i = |C|$. $\{B_0, B_1, \ldots, B_n\}$ is called the *distance distribution* of code C and the polynomial

$$D_C(X, Y) = \sum_{i=0}^{n} B_i X^{n-i} Y^i$$

is called the *distance enumerator* of C. Define

$$d = \min \{i \,|\, i > 0,\ B_i > 0\}$$

and

$$s = |\{i \,|\, i > 0,\ B_i > 0\}|,$$

d is called the *minimum distance* of C and s is the number of distinct nonzero distances between codewords of C.

Recall that

$$A_i = |\{\mathbf{c} \in C \,|\, w(\mathbf{c}) = i\}|,\ i = 0, 1, \ldots, n,$$

where w is the Hamming weight on $\mathbb{F}_2^n$. Then $A_0 = 1$ when $\mathbf{0} = 0^n \in C$, and $\sum_{i=0}^{n} A_i = |C|$, $\{A_0, A_1, \ldots, A_n\}$ is the weight distribution of code C and the polynomial

$$W_C(X, Y) = \sum_{i=0}^{n} A_i X^{n-i} Y^i$$

is the weight enumerator of C.

More generally, for all $\mathbf{c} \in C$, define

$$A_i(\mathbf{c}) = \{\mathbf{c}' \in C \,|\, w(\mathbf{c}' - \mathbf{c}) = i\},\ i = 0, 1, \ldots, n.$$

Then we also have $A_0(\mathbf{c}) = 1$ and $\sum_{i=0}^{n} A_i(\mathbf{c}) = n$. $\{A_0(\mathbf{c}), A_1(\mathbf{c}), \ldots, A_n(\mathbf{c})\}$ is called the *weight distribution* of $C - \mathbf{c}$ and the polynomial

$$W_{C-\mathbf{c}}(X, Y) = \sum_{i=0}^{n} A_i(\mathbf{c}) X^{n-i} Y^i$$

is called the *weight enumerator* of $C - \mathbf{c}$.

If C is a linear code, clearly we have $B_i = A_i = A_i(\mathbf{c})$, $i = 0, 1, \ldots, n$ and $D_C(X, Y) = W_C(X, Y) = W_{C-\mathbf{c}}(X, Y)$ for all $\mathbf{c} \in C$. In general, if for a binary code C of length n we have $A_i = A_i(\mathbf{c})$, $i = 0, 1, \ldots, n$, for all $\mathbf{c} \in C$ or we have, equivalently, $W_C(X, Y) = W_{C-\mathbf{c}}(X, Y)$ for all $\mathbf{c} \in C$, then C is called *distance invariant.* For such a code, we also have $D_i = A_i$, $i = 0, 1, \ldots, n$ and $D_C(X, Y) = W_C(X, Y)$. Linear codes are distance invariant.

Let $\{B_0, B_1, \ldots, B_n\}$ be the distance distribution of a binary code C of length n. Define

$$B'_k = |C|^{-1} \sum_{i=0}^{n} B_i K_k(i), \quad k = 0, 1, \ldots, n, \tag{2.30}$$

where $K_k(i)$ is the value the Krawtchouk polynomial $K_k(x)$ when $q = 2$ at the point $x = i$, and define

$$D'_C(X, Y) = |C|^{-1} D_C(X + Y, X - Y).$$

$\{B'_0, B'_1, \ldots, B'_n\}$ is called the *MacWilliams transform* of $\{B_0, B_1, \ldots, B_n\}$ and $D'_C(X, Y)$ is called the *MacWilliams transform* of $D_C(X, Y)$. By the proof of Proposition 2.6,

$$D'_C(X, Y) = \sum_{i=0}^{n} B'_i X^{n-i} Y^i.$$

Lemma 2.17. *For any vector* $\mathbf{x} \in \mathbb{F}_2^n$ *with* $w(\mathbf{x}) = i$,

$$\sum_{\substack{\mathbf{y} \in \mathbb{F}_2^n \\ w(\mathbf{y}) = k}} (-1)^{\mathbf{x}\cdot\mathbf{y}} = K_k(i).$$

Proof. For each j, $0 \le j \le k$, we count the number of vectors $\mathbf{y} \in \mathbb{F}_2^n$ with $w(\mathbf{y}) = k$ such that $\mathbf{x} \cdot \mathbf{y} = j$. The number is $\binom{i}{j}\binom{n-i}{k-j}$. Then by Proposition 2.15 (i),

$$\sum_{\substack{\mathbf{y} \in \mathbb{F}_2^n \\ w(\mathbf{y}) = k}} (-1)^{\mathbf{x}\cdot\mathbf{y}} = \sum_{j=0}^{k} (-1)^j \binom{i}{j} \binom{n-i}{k-j} = K_k(i). \qquad \square$$

Proposition 2.18. *Let C be a binary code with distance distribution $\{B_0, B_1, \ldots, B_n\}$, and $\{B'_0, B'_1, \ldots, B'_n\}$ be its MacWilliams transform. Then $B'_0 = 1$ and $B'_k \ge 0$ for $k = 1, 2, \ldots, n$.*

Proof. By Lemma 2.17,

$$
\begin{aligned}
|C|^2 B'_k &= |C| \sum_{i=0}^{n} B_i K_k(i) \\
&= \sum_{i=0}^{n} \sum_{\substack{\mathrm{x},\,\mathrm{y}\in C \\ d(\mathrm{x},\,\mathrm{y})=i}} \sum_{\substack{\mathrm{z}\in\mathbb{F}_2^n \\ w(\mathrm{z})=k}} (-1)^{(\mathrm{x}-\mathrm{y})\cdot\mathrm{z}} \\
&= \sum_{\substack{\mathrm{z}\in\mathbb{F}_2^n \\ w(\mathrm{z})=k}} \left(\sum_{\mathrm{x}\in C} (-1)^{\mathrm{x}\cdot\mathrm{z}} \right)^2 \\
&\geq 0.
\end{aligned}
$$

Moreover, by Proposition 2.16 $\mathcal{K}_0(x) = 1$, so

$$
B'_0 = |C|^{-1} \sum_{i=0}^{n} B_i K_0(i) = |C|^{-1} \sum_{i=0}^{n} B_i = 1. \qquad \square
$$

For the weight distribution $\{A_0, A_1, \ldots, A_n\}$ and weight enumerator $W_C(X, Y)$ of a binary code C of length n we can define their *MacWilliams transforms* $\{A'_0, A'_1, \ldots, A'_n\}$ and $W'_C(X, Y)$, respectively, in a similar way, i.e.

$$
A'_k = |C|^{-1} \sum_{i=0}^{n} A_i K_k(i), \quad k = 0, 1, \ldots, n
$$

and

$$
W'_C(X, Y) = |C|^{-1}\, W_C(X+Y, X-Y)\,.
$$

The proof of Proposition 2.18 has the following corollary.

Corollary 2.19. *Let C be a binary code of length n with weight distribution $\{A_0, A_1, \ldots, A_n\}$ and distance distribution $\{B_0, B_1, \ldots, B_n\}$, and let $\{A'_0, A'_1, \ldots, A'_n\}$ and $\{B'_0, B'_1, \ldots, B'_n\}$ be their MacWilliams transforms, respectively. Assume that $B'_k = 0$ for some k where $0 \leq k \leq n$, then*

$$
\sum_{\mathrm{x}\in C} (-1)^{\mathrm{x}\cdot\mathrm{z}} = 0 \text{ for every } \mathrm{z} \in \mathbb{F}_2^n \text{ with } w(\mathrm{z}) = k
$$

and $A'_k = 0$. $\square$

Define

$$
d' = \min\,\{i \mid i > 0,\ B'_i > 0\}
$$

and

$$s' = |\{i \mid i > 0,\ B'_i > 0\}|.$$

d' is called the *dual distance* and s' the *external distance* of code C. If C is linear, d' is the minimum distance of the dual code $C^{\perp}$ of C. If C is nonlinear, the following proposition gives a combinational interpretation of d'.

Proposition 2.20. *Let C be a binary code of length n and dual distance d'. Let $[C]$ be the $|C| \times n$ array with the codewords of C as rows. Then if $r < d'$ any set of r columns of $[C]$ contains each r-tuple exactly $2^{-r}|C|$ times.*

Proof. Since $B'_k = 0$ for $1 \le k < d'$, by Corollary 2.19 we have $\sum_{\mathbf{x} \in C} (-1)^{\mathbf{x} \cdot \mathbf{z}} = 0$ for every $\mathbf{z} \in \mathbb{F}_2^n$ with $w(\mathbf{z}) = k$. Taking any $\mathbf{z} \in \mathbb{F}_2^n$ with $w(\mathbf{z}) = 1$ we see that every column of $[C]$ must have $2^{-1}|C|$ ones and $2^{-1}|C|$ zeros. Then taking any $\mathbf{z} \in \mathbb{F}_2^n$ with $w(\mathbf{z}) = 2$ we conclude that every pair of columns of $[C]$ must contain each of the four possible pairs $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$ exactly $2^{-2}|C|$ times. Proceeding in this way, the proposition will be proved. □

Let C be a binary code of length n. For any $\mathbf{v} \in \mathbb{F}_2^n$, $C + \mathbf{v} = \{\mathbf{c} + \mathbf{v} \mid \mathbf{c} \in C\}$ is called the *translate* of C by $\mathbf{v}$. Define

$$A_i(\mathbf{v}) = |\{\mathbf{c} + \mathbf{v} \mid \mathbf{c} \in C,\ w(\mathbf{c} + \mathbf{v}) = i\}|. \tag{2.31}$$

Then $\{A_0(\mathbf{v}), A_1(\mathbf{v}), \ldots, A_n(\mathbf{v})\}$ is called the *weight distribution* of the translate $C + \mathbf{v}$. We also have $\sum_{i=0}^{n} A_i(\mathbf{v}) = |C|$. Define

$$A'_k(\mathbf{v}) = |C|^{-1} \sum_{i=0}^{n} A_i(\mathbf{v}) K_k(i), \quad k = 0, 1, \ldots, n, \tag{2.32}$$

then $\{A'_0(\mathbf{v}), A'_1(\mathbf{v}), \ldots, A'_n(\mathbf{v})\}$ is called the *MacWilliams transform* of $\{A_0(\mathbf{v}), A_1(\mathbf{v}), \ldots, A_n(\mathbf{v})\}$.

Proposition 2.21. *Let C be a binary code of length n and $\mathbf{v}$ be an arbitrary vector of $\mathbb{F}_2^n$. Then*

(i) $A'_0(\mathbf{v}) = 1$.

(ii) $\sum_{k=0}^{n} A'_k(\mathbf{v}) = 2^n\, |C|^{-1} A_0(\mathbf{v})$.

(iii) $\sum_{\mathbf{v} \in \mathbb{F}_2^n} A'_k(\mathbf{v}) A'_l(\mathbf{v}) = 2^n\, B'_k\, \delta_{kl}$.

(iv) $B'_k = 0$ *if and only if* $A'_k(\mathbf{v}) = 0$ *for all* $\mathbf{v} \in \mathbb{F}_2^n$.

Proof. (i) By Proposition 2.16, $K_0(x) = 1$. Then by (2.32) and $\sum_{i=0}^{n} A_i(\mathbf{v}) = |C|$,

$$A'_0(\mathbf{v}) = |C|^{-1} \sum_{i=0}^{n} A_i(\mathbf{v})\, K_0(i) = |C|^{-1} \sum_{i=0}^{n} A_i(\mathbf{v}) = 1.$$

(ii) By (2.32),

$$\sum_{k=0}^{n} A'_k(\mathbf{v}) = \sum_{k=0}^{n} |C|^{-1} \sum_{i=0}^{n} A_i(\mathbf{v})\, K_k(i)$$

$$= |C|^{-1} \sum_{i=0}^{n} A_i(\mathbf{v}) \sum_{k=0}^{n} K_k(i).$$

In the second formula in Proposition 2.15 (ii) let $z = 1$, we obtain

$$\sum_{k=0}^{n} K_k(i) = \begin{cases} 0 & \text{for any integer } i > 0, \\ 2^n & \text{for } i = 0. \end{cases}$$

Substituting into the above equation, we obtain

$$\sum_{k=0}^{n} A'_k(\mathbf{v}) = 2^n\, |C|^{-1}\, A_0(\mathbf{v}).$$

(iii) By (2.31), (2.32) and Lemma 2.19,

$$\sum_{\mathbf{v}\in\mathbb{F}_2^n} A'_k(\mathbf{v})\, A'_l(\mathbf{v}) = |C|^{-2} \sum_{\mathbf{v}\in\mathbb{F}_2^n} \sum_{i=0}^{n} A_i(\mathbf{v})\, K_k(i) \sum_{j=0}^{n} A_j(\mathbf{v})\, K_l(j)$$

$$= |C|^{-2} \sum_{\mathbf{v}\in\mathbb{F}_2^n} \sum_{i=0}^{n} \sum_{\substack{\mathbf{c}\in C \\ w(\mathbf{c}+\mathbf{v})=i}} \sum_{\substack{\mathbf{y}\in\mathbb{F}_2^n \\ w(\mathbf{y})=k}} (-1)^{(\mathbf{c}+\mathbf{v})\cdot\mathbf{y}}$$

$$\times \sum_{j=0}^{n} \sum_{\substack{\mathbf{c}'\in C \\ w(\mathbf{c}'+\mathbf{v})=j}} \sum_{\substack{\mathbf{z}\in\mathbb{F}_2^n \\ w(\mathbf{z})=l}} (-1)^{(\mathbf{c}'+\mathbf{v})\cdot\mathbf{z}}$$

$$= |C|^{-2} \sum_{\mathbf{v}\in\mathbb{F}_2^n} \sum_{\mathbf{c}\in C} \sum_{\substack{\mathbf{y}\in\mathbb{F}_2^n \\ w(\mathbf{y})=k}} (-1)^{(\mathbf{c}+\mathbf{v})\cdot\mathbf{y}}$$

$$\times \sum_{\mathbf{c}'\in C} \sum_{\substack{\mathbf{z}\in\mathbb{F}_2^n \\ w(\mathbf{z})=l}} (-1)^{(\mathbf{c}'+\mathbf{v})\cdot\mathbf{z}}$$

$$= |C|^{-2} \sum_{\mathbf{c}\in C} \sum_{\mathbf{c}'\in C} \sum_{\substack{\mathbf{y}\in\mathbb{F}_2^n \\ w(\mathbf{y})=k}} \sum_{\substack{\mathbf{z}\in\mathbb{F}_2^n \\ w(\mathbf{z})=l}} (-1)^{\mathbf{c}\cdot\mathbf{y}+\mathbf{c}'\cdot\mathbf{z}}$$

$$\times \sum_{\mathbf{v}\in\mathbb{F}_2^n} (-1)^{\mathbf{v}\cdot(\mathbf{y}+\mathbf{z})}.$$

But for any $\mathbf{x} \in \mathbb{F}_2^n$, we have

$$\sum_{\mathbf{v}\in\mathbb{F}_2^n} (-1)^{\mathbf{v}\cdot\mathbf{x}} = \begin{cases} 0 & \text{if} \quad \mathbf{x} \neq 0, \\ 2^n & \text{if} \quad \mathbf{x} = 0. \end{cases}$$

Therefore when $k \neq l$, we have

$$\sum_{\mathbf{v}\in\mathbb{F}_2^n} A'_k(\mathbf{v})\, A'_l(\mathbf{v}) = 0,$$

and when $k = l$, we have

$$\sum_{\mathbf{v}\in\mathbb{F}_2^n} A'_k(\mathbf{v})^2 = 2^n\, |C|^{-2} \sum_{\mathbf{c}\in C} \sum_{\mathbf{c}'\in C} \sum_{\substack{\mathbf{y}\in\mathbb{F}_2^n \\ w(\mathbf{y})=k}} (-1)^{(\mathbf{c}+\mathbf{c}')\mathbf{y}}$$

$$= 2^n\, |C|^{-2} \sum_{i=0}^{n} \sum_{\substack{(\mathbf{c},\,\mathbf{c}')\in C^2 \\ d(\mathbf{c},\,\mathbf{c}')=i}} \sum_{\substack{\mathbf{y}\in\mathbb{F}_2^n \\ w(\mathbf{y})=k}} (-1)^{(\mathbf{c}+\mathbf{c}')\cdot\mathbf{y}}$$

$$= 2^n\, |C|^{-2} \sum_{i=0}^{n} \sum_{\substack{(\mathbf{c},\,\mathbf{c}')\in C^2 \\ d(\mathbf{c},\,\mathbf{c}')=i}} K_k(i)$$

$$= 2^n\, |C|^{-1} \sum_{i=0}^{n} B_i\, K_k(i)$$

$$= 2^n\, B'_k\,.$$

Hence (iii) is proved.

(iv) By (iii),

$$\sum_{\mathbf{v}\in\mathbb{F}_2^n} A'_k(\mathbf{v})^2 = 2^n\, B'_k$$

from which it follows that $B'_k = 0$ if and only if $A'_k(\mathbf{v}) = 0$ for all $\mathbf{v} \in \mathbb{F}_2^n$. $\square$

Let C be a binary code of length n. As before denote the external distance of C by s'. Let $0, \sigma_1, \sigma_2, \ldots, \sigma_{s'}$, be the subscripts i for which $B'_i \neq 0$. The *annihilator polynomial* $\alpha(x)$ of C is defined to be

$$\alpha(x) = 2^n \,|\, C \,|^{-1} \prod_{j=1}^{s'} (1 - \sigma_j^{-1}\, x).$$

Clearly, $\deg \alpha(x) = s'$ and for $0 < i \leq n$ either $\alpha(i) = 0$ or $B'_i = 0$. Let

$$\alpha(x) = \sum_{k=0}^{s'} \alpha_k \, K_k(x)$$

be the Krawtchouk expansion of $\alpha(x)$, where the Krawtchouk coefficients are given by

$$\alpha_k = 2^{-n} \sum_{l=0}^{n} \alpha(l)\, K_l(k), \quad k = 0, 1, \ldots, s'.$$

We have $\alpha_{s'} \neq 0$ since $\deg \alpha(s) = s'$.

Proposition 2.22. $\sum_{k=0}^{s'} \alpha_k \, A_k(\mathrm{v}) = 1$ *for all* $\mathrm{v} \in \mathbb{F}_2^n$.

Proof. We may write $\alpha(x) = \sum_{k=0}^{n} \alpha_k \, K_k(x)$, where $\alpha_{s'+1} = \alpha_{s'+2} = \cdots = \alpha_n = 0$. Then as in the proof of Proposition 2.14, we also have $\alpha_k = 2^{-n} \sum_{l=0}^{n} \alpha(l)\, K_l(k)$ for all $k = 0, 1, \ldots, n$. We compute

$$\begin{aligned}
\sum_{k=0}^{s'} \alpha_k \, A_k(\mathrm{v}) &= \sum_{k=0}^{n} \alpha_k \, A_k(\mathrm{v}) \\
&= 2^{-n} \sum_{k=0}^{n} \sum_{l=0}^{n} \alpha(l) \; K_l(k) \; A_k(\mathrm{v}) \\
&= 2^{-n} \sum_{l=0}^{n} \alpha(l) \sum_{k=0}^{n} K_l(k) \; A_k(\mathrm{v}) \\
&= 2^{-n} \sum_{l=0}^{n} \alpha(l) \,|\, C \,|\; A'_l(\mathrm{v}).
\end{aligned}$$

For $l = \sigma_1, \sigma_2, \ldots,$ or $\sigma_{s'}$, $\alpha(l) = 0$; for $l \neq 0$, σ_1, $\sigma_2, \ldots,$ and $\sigma_{s'}$, $B'_l = 0$ which, by Proposition 2.21(iv), implies that $A'_l(\mathrm{v}) = 0$ for all $\mathrm{v} \in \mathbb{F}_2^n$; for $l = 0$, $\alpha(0) = 2^n \,|\, C \,|^{-1}$ and by Proposition 2.21(i), $A'_0(\mathrm{v}) = 1$. Therefore

$$\sum_{k=0}^{s'} \alpha_k \, A_k(\mathrm{v}) = 1 \quad \text{for all } \mathrm{v} \in \mathbb{F}_2^n. \qquad \square$$

Corollary 2.23. *For any* $\mathbf{v} \in \mathbb{F}_2^n$ *there exists at least one* $\mathbf{c} \in C$ *such that* $d(\mathbf{c}, \mathbf{v}) \leq s'$

Proof. Given any $\mathbf{v} \in \mathbb{F}_2^n$, by Proposition 2.22 there is at least one $A_{k_0}(\mathbf{v}) \neq 0$ where $0 \leq k_0 \leq s'$. Then there is at least one $\mathbf{c} \in C$ such that $w(\mathbf{c} + \mathbf{v}) = k_0 \leq s'$ That is $d(\mathbf{c}, \mathbf{v}) \leq s'$ □

This corollary explains why s' is called the external distance of C.

Most of this section are due to Delsarte (1973), but some proofs are different. The minimum distance, the number of distinct nonzero distances, the dual distance, and the external distance of a binary code are called the *four fundamental parameters* of the code by Delsarte.

CHAPTER 3

THE GRAY MAP

3.1. The Gray Map

In communication systems employing quadrature phase-shift keying (QPSK), the preferred assignment of two information bits to the four possible phases is the one shown in Fig. 3.1, in which adjacent phases differ by only one binary digit. This map is called the *Gray map* and has the advantage that, when a codeword over $\mathbb{Z}_4$ is transmitted across an additive white Gaussian noise channel, the errors most likely to occur are those causing a single erroneously decoded information bit. The Gray map is usually denoted by ϕ, i.e.

$$\begin{aligned} \phi \,:\, \mathbb{Z}_4 &\to \mathbb{Z}_2^2 \\ 0 &\mapsto 00 \\ 1 &\mapsto 01 \\ 2 &\mapsto 11 \\ 3 &\mapsto 10 \end{aligned}$$

Clearly, ϕ is a bijection from $\mathbb{Z}_4$ to $\mathbb{Z}_2^2$. Denote the Hamming weight of a binary vector $\mathbf{v}$ by $w(\mathbf{v})$ and the Hamming distance between two binary vectors $\mathbf{u}$ and $\mathbf{v}$ of the same length by $d(\mathbf{u}, \mathbf{v})$. Clearly,

$$w_{\mathrm{L}}(x) = w(\phi(x)) \quad \text{for all } x \in \mathbb{Z}_4, \tag{3.1}$$

and we can easily verify that

$$d_{\mathrm{L}}(x, y) = d(\phi(x), \phi(y)) \quad \text{for all } x, y \in \mathbb{Z}_4. \tag{3.2}$$

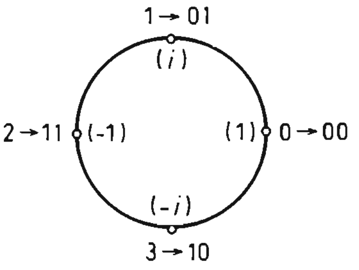

Fig. 3.1.

But ϕ is not an additive group homomorphism from $\mathbb{Z}_4$ to $\mathbb{Z}_2^2$.

It will be helpful to introduce the following three maps α, β, γ from $\mathbb{Z}_4$ to $\mathbb{Z}_2$ by the following table.

Table 3.1.

$\mathbb{Z}_4$	α	β	γ
0	0	0	0
1	1	0	1
2	0	1	1
3	1	1	0

Clearly, α is an additive group homomorphism from $\mathbb{Z}_4$ to $\mathbb{Z}_2$, but β and γ are not. Each element $x \in \mathbb{Z}_4$ has a 2-adic expansion

$$x = \alpha(x) + 2\beta(x).$$

We also have

$$\alpha(x) + \beta(x) + \gamma(x) = 0 \quad \text{for all } x \in \mathbb{Z}_4.$$

The Gray map ϕ can be expressed in terms of β and γ as follows:

$$\phi(x) = (\beta(x), \gamma(x)) \quad \text{for all } x \in \mathbb{Z}_4.$$

The maps α, β, γ can be extended to $\mathbb{Z}_4^n$ in an obvious way. For $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}_4^n$, define

$$\alpha(\mathbf{x}) = (\alpha(x_1), \ldots, \alpha(x_n)),$$

$$\beta(\mathbf{x}) = (\beta(x_1), \ldots, \beta(x_n)),$$
$$\gamma(\mathbf{x}) = (\gamma(x_1), \ldots, \gamma(x_n)).$$

Then ϕ is extended to $\mathbb{Z}_n^4$ as follows:

$$\phi(\mathbf{x}) = (\beta(\mathbf{x}), \gamma(\mathbf{x})) \quad \text{for all } \mathbf{x} \in \mathbb{Z}_4^n. \tag{3.3}$$

Clearly, the extended ϕ is a bijection from $\mathbb{Z}_4^n$ to $\mathbb{Z}_2^{2n}$. For any $\mathbf{x} \in \mathbb{Z}_4^n$ $\phi(\mathbf{x})$ is called the *binary image* of $\mathbf{x}$ under ϕ.

Theorem 3.1. *ϕ is a weight-preserving map from*

$$(\mathbb{Z}_4^n, \textit{ Lee weight}) \quad \textit{to} \quad (\mathbb{Z}_2^{2n}, \textit{ Hamming weight}),$$

i.e.

$$w_{\mathrm{L}}(\mathbf{x}) = w(\phi(\mathbf{x})) \quad \textit{for all } \mathbf{x} \in \mathbb{Z}_4^n, \tag{3.4}$$

and ϕ is also a distance-preserving map from

$$(\mathbb{Z}_4^n, \textit{ Lee distance}) \quad \textit{to} \quad (\mathbb{Z}_2^{2n}, \textit{ Hamming distance}),$$

i.e.

$$d_{\mathrm{L}}(\mathbf{x}, \mathbf{y}) = d(\phi(\mathbf{x}), \phi(\mathbf{y})) \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{Z}_4^n. \tag{3.5}$$

Proof. For any $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}_4^n$,

$$w_{\mathrm{L}}(\mathbf{x}) = \sum_{i=1}^{n} w_{\mathrm{L}}(x_i)$$

and

$$\begin{aligned} w(\phi(\mathbf{x})) &= w((\beta(\mathbf{x}), \gamma(\mathbf{x}))) = w(\beta(\mathbf{x})) + w(\gamma(\mathbf{x})) \\ &= \sum_{i=1}^{n} w(\beta(x_i)) + \sum_{i=1}^{n} w(\gamma(x_i)) \\ &= \sum_{i=1}^{n} w((\beta(x_i), \gamma(x_i))) \\ &= \sum_{i=1}^{n} w(\phi(x_i)). \end{aligned}$$

By (3.1), $w_{\mathrm{L}}(x_i) = w(\phi(x_i))$, $i = 1, 2, \ldots, n$. Therefore we have (3.4). Similarly, from (3.2) we deduce (3.5). □

From (2.13) and (3.5) it follows that for any $\mathbf{x}$, $\mathbf{y} \in \mathbb{Z}_4^n$,

$$d(\phi(\mathbf{x}), \phi(\mathbf{y})) = \frac{1}{2} d_{\mathrm{E}}^2 \left(i^{\mathbf{x}}, i^{\mathbf{y}}\right).$$

The following proposition is obvious.

Proposition 3.2. *For any* $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}_4^n$, *we have*

$$w_{\mathrm{L}}(\mathbf{x}) \equiv \sum_{i=1}^{n} x_i \,(\mathrm{mod}\ 2). \tag{3.6}$$

Let $\phi(\mathbf{x}) = (y_1, y_2, \ldots, y_{2n}) \in \mathbb{Z}_2^{2n}$, *then*

$$\sum_{i=1}^{n} x_i \equiv \sum_{i=1}^{2n} y_i \ (\mathrm{mod}\ 2). \tag{3.7}$$

In particular, if $\sum_{i=1}^{n} x_i = 0$ *or* 2 *in* $\mathbb{Z}_4$, *then* $\mathbf{x}$ *is an even Lee weight word in* $\mathbb{Z}_4^n$ *and* $\phi(\mathbf{x})$ *is an even (Hamming) weight word in* $\mathbb{Z}_2^{2n}$

Proof. If we regard $\sum_{i=1}^{n} x_i$ as a sum in $\mathbb{Z}_4$, then

$$\sum_{i=1}^{n} x_i = w_1(\mathbf{x}) + 2w_2(\mathbf{x}) + 3w_3(\mathbf{x}).$$

But

$$w_{\mathrm{L}}(\mathbf{x}) = w_1(\mathbf{x}) + 2w_2(\mathbf{x}) + w_3(\mathbf{x}),$$

where the R.H.S. is regarded as a sum in $\mathbb{Z}$. Therefore we have (3.6). Moreover, if we regard $\sum_{i=1}^{2n} y_i$ as a sum in $\mathbb{Z}$, by Theorem 3.1 we have

$$\sum_{i=1}^{2n} y_i = w\left(\phi(\mathbf{x})\right) = w_{\mathrm{L}}(\mathbf{x}). \tag{3.8}$$

From (3.6) and (3.8) we deduce (3.7). □

3.2. Binary Images of $\mathbb{Z}_4$-Codes

Let $\mathcal{C}$ be a $\mathbb{Z}_4$-code. Define

$$C = \phi(\mathcal{C}) = \{\phi(\mathbf{c}) \,|\, \mathbf{c} \in \mathcal{C}\},$$

which is called the *binary image* of $\mathcal{C}$ under the Gray map or, simply, the binary image of $\mathcal{C}$. If $\mathcal{C}$ is of length n, then $C \subseteq \mathbb{Z}_2^{2n}$, i.e. C is a binary code of length $2n$. We recall that

$$\min\{w(\phi(\mathbf{c})) \mid \mathbf{c} \in \mathcal{C},\ \mathbf{c} \neq 0^n\}$$

and

$$\min\{d(\phi(\mathbf{c}),\ \phi(\mathbf{c}')) \mid \mathbf{c},\ \mathbf{c}' \in \mathcal{C},\ \mathbf{c} \neq \mathbf{c}'\}$$

are the minimum (Hamming) weight and distance of C, respectively. Similarly we define

$$\min\{w_{\mathrm{L}}(\mathbf{c}) \mid \mathbf{c} \in \mathcal{C},\ \mathbf{c} \neq 0^n\}$$

and

$$\min\{d_{\mathrm{L}}(\mathbf{c},\ \mathbf{c}') \mid \mathbf{c},\ \mathbf{c}' \in \mathcal{C},\ \mathbf{c} \neq \mathbf{c}'\}$$

to be the *minimum Lee weight* and *distance* of $\mathcal{C}$, respectively. Theorem 3.1 implies, in particular,

Proposition 3.3. *Let $\mathcal{C}$ be a $\mathbb{Z}_4$-code and $C = \phi(\mathcal{C})$. Then the minimum Lee weight and distance of $\mathcal{C}$ are equal to the minimum (Hamming) weight and distance of $C = \phi(\mathcal{C})$, respectively.* □

From Proposition 3.2 we deduce immediately.

Proposition 3.4. *Let $\mathcal{C}$ be a $\mathbb{Z}_4$-code of length n and assume that for all codewords $\mathbf{c} = (c_1, \ldots, c_n)$ of $\mathcal{C}$, $\sum_{i=1}^{n} c_i \equiv 0 \pmod 2$, then all codewords of $\mathcal{C}$ are of even Lee weight and all codewords of its binary image $\phi(\mathcal{C})$ are of even (Hamming) weight.* □

Example 3.1. Consider the binary image of the $\mathbb{Z}_4$-linear code $\mathcal{C}_3 = \{(0,\ 0),\ (2,\ 2)\}$ appeared in Example 2.3. We have

$$\phi(0,\ 0) = (\beta(0,\ 0),\ \gamma(0,\ 0)) = (0,\ 0,\ 0,\ 0),$$

$$\phi(2,\ 2) = (\beta(2,\ 2),\ \gamma(2,\ 2)) = (1,\ 1,\ 1,\ 1).$$

Therefore

$$\phi(\mathcal{C}_3) = \{(0,\ 0,\ 0,\ 0),\ (1,\ 1,\ 1,\ 1)\},$$

which is a binary linear code. □

Example 3.2. The binary image $\varphi(\mathcal{K}_4)$ of the $\mathbb{Z}_4$-linear code $\mathcal{K}_4$ with generator matrix (1.3)

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 2 & 2 \end{pmatrix}$$

consists of the following 16 codewords

$$\begin{array}{cccccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \end{array} \qquad \begin{array}{cccccccc} 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{array}$$

It is easy to see that $\phi(\mathcal{K}_4)$ is a binary linear code with minimum distance 4. Hence $\phi(\mathcal{K}_4)$ is the extended binary Hamming code of length 8. It has the following generator matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}. \tag{3.9}$$

□

Example 3.3. The binary image $\phi(\mathcal{C}_1)$ of the linear code $\mathcal{C}_1$ appeared in Example 1.2 with generator matrix (1.4)

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 2 \end{pmatrix}$$

consists of the following eight codewords

$$\begin{array}{cccccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{array} \qquad \begin{array}{cccccccc} 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \end{array}$$

$\phi(\mathcal{C}_1)$ is also a binary linear code with generator matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}. \tag{3.10}$$

□

Denote the images of all the rows of a matrix M over $\mathbb{Z}_4$ under the maps α, β, γ, φ by $\alpha(M)$, $\beta(M)$, $\gamma(M)$, $\varphi(M)$, respectively. Then $\phi(M) = (\beta(M), \gamma(M))$.

Proposition 3.5. *Let $C = \phi(\mathcal{C})$ be the binary image of a $\mathbb{Z}_4$-linear code $\mathcal{C}$ with generator matrix* (1.1). *If C is linear, then C has generator matrix*

$$\begin{pmatrix} I_{k_1} & A & \alpha(B) & I_{k_1} & A & \alpha(B) \\ 0 & I_{k_2} & C & 0 & I_{k_2} & C \\ 0 & 0 & \beta(B) & I_{k_1} & A & \gamma(B) \end{pmatrix} \tag{3.11}$$

Proof. Assume that C is linear, then C is generated by the binary image of the rows of the matrix

$$\begin{pmatrix} I_{k_1} & A & B \\ 2I_{k_1} & 2A & 2B \\ 3I_{k_1} & 3A & 3B \\ 0 & 2I_{k_2} & 2C \end{pmatrix},$$

where A and C are matrices over $\mathbb{Z}_2$ and B is a matrix over $\mathbb{Z}_4$. Clearly we have

$$\begin{aligned}\phi(I_{k_1}\ A\ B) &= (\beta(I_{k_1}\ A\ B)\ \gamma(I_{k_1}\ A\ B)) \\ &= (0\ 0\ \beta(B)\ I_{k_1}\ A\ \gamma(B)).\end{aligned}$$

$$\begin{aligned}\phi(2I_{k_1}\ 2A\ 2B) &= (\beta(2I_{k_1}\ 2A\ 2B)\ \gamma(2I_{k_1}\ 2A\ 2B)) \\ &= (I_{k_1}\ A\ \alpha(B)\ I_{k_1}\ A\ \alpha(B)),\end{aligned}$$

$$\begin{aligned}\phi(3I_{k_1}\ 3A\ 3B) &= (\beta(3I_{k_1}\ 3A\ 3B)\ \gamma(3I_{k_1}\ 3A\ 3B)) \\ &= (I_{k_1}\ A\ \gamma(B)\ 0\ 0\ \beta(B)),\end{aligned}$$

$$\begin{aligned}\phi(0\ 2I_{k_2}\ 2C) &= (\beta(0\ 2I_{k_2}\ 2C)\ \gamma(0\ 2I_{k_2}\ 2C)) \\ &= (0\ I_{k_2}\ C\ 0\ I_k\ C).\end{aligned}$$

From $\alpha(B) + \beta(B) + \gamma(B) = 0$, we deduce

$(I_{k_1}\ A\ \gamma(B)\ 0\ 0\ \beta(B)) = (0\ 0\ \beta(B)\ I_{k_1}\ A\ \gamma(B)) + (I_{k_1}\ A\ \alpha(B)\ I_{k_1}\ A\ \alpha(B))$. Hence C is generated by the rows of (3.11). Clearly the rows of (3.11) are linearly independent. Therefore (3.11) is a generator matrix of C. □

Notice that the generator matrix of $\varphi(\mathcal{K}_4)$ given in Example 3.2 is precisely the one given by Proposition 3.5.

We recall that a binary code C is said to be *distance invariant* if the Hamming weight enumerator of its translators $\mathbf{u} + C$ are the same for all $\mathbf{u} \in C$. Clearly, binary linear codes are distance invariant. Moreover, we have

Theorem 3.6. *For any $\mathbb{Z}_4$-linear code $\mathcal{C}$, its binary image $C = \phi(\mathcal{C})$ is distance invariant.*

Proof. Since $\mathcal{C}$ is linear, $\mathbf{u} + \mathcal{C} = \mathcal{C}$ for all $\mathbf{u} \in \mathcal{C}$. Hence $\mathcal{C}$ is distance invariant with respect to the Lee weight, i.e.

$$\{w_{\mathrm{L}}(\mathbf{u}+\mathbf{c}) \mid \mathbf{c} \in \mathcal{C}\} = \{w_{\mathrm{L}}(\mathbf{c}) \mid \mathbf{c} \in \mathcal{C}\} \qquad \text{for all } \mathbf{u} \in \mathcal{C}. \tag{3.12}$$

Then

$$\begin{aligned}
\{w(\phi(\mathbf{u}) + \phi(\mathbf{c})) \mid \mathbf{c} \in \mathcal{C}\} &= \{w(\phi(\mathbf{u}) - \phi(\mathbf{c})) \mid \mathbf{c} \in \mathcal{C}\} \qquad (\text{in } \mathbb{Z}_2,\ -1 = 1) \\
&= \{d(\phi(\mathbf{u}),\ \phi(\mathbf{c})) \mid \mathbf{c} \in \mathcal{C}\} \\
&= \{d_{\mathrm{L}}(\mathbf{u},\ \mathbf{c}) \mid \mathbf{c} \in \mathcal{C}\} \qquad (\text{by Theorem 3.1}) \\
&= \{w_{\mathrm{L}}(\mathbf{u} - \mathbf{c}) \mid \mathbf{c} \in \mathcal{C}\} \\
&= \{w_{\mathrm{L}}(\mathbf{u} + \mathbf{c}) \mid \mathbf{c} \in \mathcal{C}\} \qquad (\mathcal{C} \text{ is linear}) \\
&= \{w_{\mathrm{L}}(\mathbf{c}) \mid \mathbf{c} \in \mathcal{C}\} \qquad (\text{by (3.12)}) \\
&= \{w(\phi(\mathbf{c})) \mid \mathbf{c} \in \mathcal{C}\} \qquad (\text{by (3.4)})
\end{aligned}$$

for all $\phi(\mathbf{u}) \in \phi(\mathcal{C})$. Therefore C is distance invariant. □

Let $\mathcal{C}$ be a $\mathbb{Z}_4$-linear code. Its binary image $C = \phi(\mathcal{C})$ is, in general, not linear and it need not have a dual code. We define the $\mathbb{Z}_4$-*dual* of $C = \phi(\mathcal{C})$ to be $C_\perp = \phi(\mathcal{C}^\perp)$. In the diagram

$$\begin{array}{ccc}
\mathcal{C} & \longrightarrow & C = \phi(\mathcal{C}) \\
\downarrow & & \\
\mathcal{C}^\perp & \longrightarrow & C_\perp = \phi(\mathcal{C}^\perp)
\end{array}$$

Fig. 3.2.

we cannot always add an arrow on the right to produce a commuting diagram. But we have

Theorem 3.7. *Let $\mathcal{C}$ and $\mathcal{C}^{\perp}$ be dual $\mathbb{Z}_4$-linear codes, and $C = \phi(\mathcal{C})$ and $C_{\perp} = \phi(\mathcal{C}^{\perp})$ be their binary images. Then the weight enumerators $W_C(X, Y)$ and $W_{C_\perp}(X, Y)$ of C and $C_\perp$, respectively, are related by the binary MacWilliam identity*

$$W_{C_\perp}(X, Y) = \frac{1}{|C|} W_C(X+Y, X-Y). \tag{3.13}$$

Proof. By Theorems 3.1 and 2.4, and $|\mathcal{C}| = |C|$, we have

$$\begin{aligned} W_{C_\perp}(X, Y) &= \mathrm{Lee}_{\mathcal{C}^\perp}(X, Y) \\ &= \frac{1}{|\mathcal{C}|} \mathrm{Lee}_{\mathcal{C}}(X+Y, X-Y) \\ &= \frac{1}{|C|} W_C(X+Y, X-Y). \end{aligned}$$ □

So, we call the binary codes $C = \phi(\mathcal{C})$ and $C_\perp = \phi(\mathcal{C}^\perp)$ *formally dual*. If $\mathcal{C}$ is self-dual, i.e. $\mathcal{C}^\perp = \mathcal{C}$, then $C = C_\perp$ and we call C *formally self-dual*.

Proposition 3.8. *Let $\mathcal{C}$ be a $\mathbb{Z}_4$-linear code of length n, $\mathcal{C}^\perp$ be its dual code, and $C = \phi(\mathcal{C})$ and $C_\perp = \phi(\mathcal{C}^\perp)$ be their binary images, respectively. Let $\{A_0, A_1, \ldots, A_{2n}\}$ be the weight distribution of C. Then the MacWilliams transform of $\{A_0, A_1, \ldots, A_{2n}\}$ is the weight distribution $\{A'_0, A'_1, \ldots, A'_{2n}\}$ of $C_\perp$ and the MacWilliams transform of $\{A'_0, A'_1, \ldots, A'_{2n}\}$ is $\{A_0, A_1, \ldots, A_{2n}\}$.*

Proof. By Theorem 3.7, we have (3.13)

$$W_{C_\perp}(X, Y) = |C|^{-1} W_C(X+Y, X-Y).$$

Then the first assertion follows from Proposition 2.6. For the proof of the second assertion, we compute

$$|C_\perp|^{-1} \sum_{l=0}^{n} A'_l K_k(l) = |C_\perp|^{-1}\, |C|^{-1} \sum_{l=0}^{n} \left(\sum_{i=0}^{n} A_i K_l(i) \right) K_k(l)$$

$$= 2^{-2n} \sum_{i=0}^{n} A_i \sum_{l=0}^{n} K_l(i)\, K_k(l) \qquad (|\, C_\perp \,|\, |\, C \,| = 2^{2n})$$

$$= 2^{-2n} \sum_{i=0}^{n} A_i\, 2^{2n}\, \delta_{ik} \qquad \text{(Proposition 2.15(viii))}$$

$$= A_k. \qquad \square$$

3.3. Linearity Conditions

A binary code C is called *$\mathbb{Z}_4$-linear* if after a permutation of its coordinates, it is the binary image of a $\mathbb{Z}_4$-linear code $\mathcal{C}$. Now we want to study the following problems.

(i) When is a given binary code $\mathbb{Z}_4$-linear?
(ii) When is the binary image of a $\mathbb{Z}_4$-linear code linear?
A trivial necessary condition for a binary code to be $\mathbb{Z}_4$-linear is

Proposition 3.9. *If a binary code is $\mathbb{Z}_4$-linear, then its length is even.* $\square$

Define a permutation σ on the $2n$-dimensional vector $(x_1, \ldots, x_{2n})$ as follows:

$$\sigma : (x_1, \ldots, x_n, x_{n+1}, \ldots, x_{2n}) \rightarrow (x_{n+1}, \ldots, x_{2n}, x_1, \ldots, x_n). \tag{3.14}$$

We call σ the "swap" map. Clearly,

$$\sigma = (1\ n+1)\,(2\ n+2) \cdots (n\ 2n).$$

Then for any $\mathbf{x} \in \mathbb{Z}_4^n$,

$$\sigma(\phi(\mathbf{x})) = \sigma(\beta(\mathbf{x})\ \gamma(\mathbf{x})) = (\gamma(\mathbf{x}),\ \beta(\mathbf{x})) = \phi(-\mathbf{x}). \tag{3.15}$$

Therefore we have

Proposition 3.10. *If a binary code C is $\mathbb{Z}_4$-linear, then after a permutation of its coordinates, $\sigma(C) = C$.* $\square$

Denote by $*$ the componentwise multiplication of two vectors, i.e.

$$(x_1, \ldots, x_n) * (y_1, \ldots, y_n) = (x_1 y_1, \ldots, x_n y_n).$$

Lemma 3.11. *For all* $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_4^n$, *we have*

$$(\phi(\mathbf{x}) + \sigma(\phi(\mathbf{x}))) * (\phi(\mathbf{y}) + \sigma(\phi(\mathbf{y}))) = \phi(2\,\alpha(\mathbf{x}) * \alpha(\mathbf{y}))\,,$$

where the multiplication of $\alpha(\mathbf{x}) * \alpha(\mathbf{y})$ *by 2 is performed in* $\mathbb{Z}_4$.

Proof. By (3.14),

$$\begin{aligned}
&(\phi(\mathbf{x}) + \sigma(\phi(\mathbf{x}))) * (\phi(\mathbf{y}) + \sigma(\phi(\mathbf{y}))) \\
&\quad = ((\beta(\mathbf{x}), \gamma(\mathbf{x})) + (\gamma(\mathbf{x}), \beta(\mathbf{x}))) * ((\beta(\mathbf{y}), \gamma(\mathbf{y})) + (\gamma(\mathbf{y}), \beta(\mathbf{y}))) \\
&\quad = (\beta(\mathbf{x}) + \gamma(\mathbf{x}), \gamma(\mathbf{x}) + \beta(\mathbf{x})) * (\beta(\mathbf{y}) + \gamma(\mathbf{y}), \gamma(\mathbf{y}) + \beta(\mathbf{y})) \\
&\quad = (\alpha(\mathbf{x}), \alpha(\mathbf{x})) * (\alpha(\mathbf{y}), \alpha(\mathbf{y})) \\
&\quad = (\alpha(\mathbf{x}) * \alpha(\mathbf{y}),\ \alpha(\mathbf{x}) * \alpha(\mathbf{y})) \\
&\quad = \phi(2\,\alpha(\mathbf{x}) * \alpha(\mathbf{y}))\,.
\end{aligned}$$

□

Lemma 3.12. *For all* $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_4^n$, *we have*

$$\phi(\mathbf{x} + \mathbf{y}) = \phi(\mathbf{x}) + \phi(\mathbf{y}) + (\phi(\mathbf{x}) + \sigma(\phi(\mathbf{x}))) * (\phi(\mathbf{y}) + \sigma(\phi(\mathbf{y})))\,. \tag{3.16}$$

Proof. By Lemma 3.11, (3.16) is equivalent to

$$\phi(\mathbf{x}) + \phi(\mathbf{y}) + \phi(\mathbf{x} + \mathbf{y}) = \phi\,(2\alpha(\mathbf{x}) * \alpha(\mathbf{y}))\,. \tag{3.17}$$

Therefore it is sufficient to verify (3.17). We have

$$\text{L.H.S. of } (3.17) = (\beta(\mathbf{x}) + \beta(\mathbf{y}) + \beta(\mathbf{x} + \mathbf{y}),\ \gamma(\mathbf{x}) + \gamma(\mathbf{y}) + \gamma(\mathbf{x} + \mathbf{y}))\,.$$

$$\text{R.H.S. of } (3.17) = (\alpha(\mathbf{x}) * \alpha(\mathbf{y}),\ \alpha(\mathbf{x}) * \alpha(\mathbf{y}))\,.$$

Thus we need to show that

$$\begin{aligned}
\beta(\mathbf{x}) + \beta(\mathbf{y}) + \beta(\mathbf{x} + \mathbf{y}) &= \gamma(\mathbf{x}) + \gamma(\mathbf{y}) + \gamma(\mathbf{x} + \mathbf{y}) \\
&= \alpha(\mathbf{x}) * \alpha(\mathbf{y}) \text{ for all } \mathbf{x}, \mathbf{y} \in \mathbb{Z}_4^n.
\end{aligned}$$

It is enough to check the above identity for the case $n = 1$. Using Table 3.1 we can check it easily. □

Corollary 3.13. *For all* $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_4^n$, *we have*

$$\phi(\mathbf{x} + \mathbf{y}) = \phi(\mathbf{x}) + \phi(\mathbf{y}) + \phi(2\alpha(\mathbf{x}) * \alpha(\mathbf{y}))\,. \tag{3.18}$$

□

Now we can answer the problems proposed at the beginning of this section.

Proposition 3.14. *A binary, not necessarily linear, code C of even length is $\mathbb{Z}_4$-linear if and only if after a permutation of its coordinates,*

$$\mathbf{u},\, \mathbf{v} \in C \Rightarrow \mathbf{u} + \mathbf{v} + (\mathbf{u} + \sigma(\mathbf{u})) * (\mathbf{v} + \sigma(\mathbf{v})) \in C. \tag{3.19}$$

Proof. Assume that $C = \phi(\mathcal{C})$, where $\mathcal{C}$ is a $\mathbb{Z}_4$-linear code. Let $\mathbf{u},\, \mathbf{v} \in C$, then there are $\mathbf{x},\, \mathbf{y} \in \mathcal{C}$ such that $\mathbf{u} = \phi(\mathbf{x}), \mathbf{v} = \phi(\mathbf{y})$. Since $\mathcal{C}$ is linear, $\mathbf{x} + \mathbf{y} \in \mathcal{C}$. By Lemma 3.12,

$$\begin{aligned} \mathbf{u} + \mathbf{v} + (\mathbf{u} + \sigma(\mathbf{u})) * (\mathbf{v} + \sigma(\mathbf{v})) &= \phi(\mathbf{x}) + \phi(\mathbf{y}) \\ &\quad + (\phi(\mathbf{x}) + \sigma(\phi(\mathbf{x}))) * (\phi(\mathbf{y}) + \sigma(\phi(\mathbf{y}))) \\ &= \phi(\mathbf{x} + \mathbf{y}) \in \phi(\mathcal{C}) = C. \end{aligned}$$

Conversely, assume that condition (3.19) holds. Let $\dim C = 2n$. Define

$$\mathcal{C} = \{\mathbf{c} \in \mathbb{Z}_4^n \,|\, \phi(\mathbf{c}) \in C\}$$

Let us prove that $\mathcal{C}$ is a $\mathbb{Z}_4$-linear code. Let $\mathbf{x},\, \mathbf{y} \in \mathcal{C}$. Then $\phi(\mathbf{x}),\, \phi(\mathbf{y}) \in C$. By (3.19),

$$\phi(\mathbf{x}) + \phi(\mathbf{y}) + (\phi(\mathbf{x}) + \sigma(\phi(\mathbf{x}))) * (\phi(\mathbf{y}) + \sigma(\phi(\mathbf{y}))) \in C.$$

By (3.16), $\phi(\mathbf{x} + \mathbf{y}) \in C$. Therefore $\mathbf{x} + \mathbf{y} \in \mathcal{C}$. □

Corollary 3.15. *A binary linear code C of even length is $\mathbb{Z}_4$-linear if and only if after a permutation of its coordinates,*

$$\mathbf{u},\, \mathbf{v} \in C \Rightarrow (\mathbf{u} + \sigma(\mathbf{u})) * (\mathbf{v} + \sigma(\mathbf{v})) \in C.$$ □

Proposition 3.16. *The binary image $C = \phi(\mathcal{C})$ of a $\mathbb{Z}_4$-linear code $\mathcal{C}$ is linear if and only if*

$$\mathbf{x},\, \mathbf{y} \in \mathcal{C} \Rightarrow 2\,\alpha(\mathbf{x}) * \alpha(\mathbf{y}) \in \mathcal{C}. \tag{3.20}$$

Proof. Assume that C is linear. Since $\mathcal{C}$ is linear, for any $\mathbf{x},\, \mathbf{y} \in \mathcal{C}$, $\mathbf{x}+\mathbf{y} \in \mathcal{C}$. Then $\phi(\mathbf{x})$, $\phi(\mathbf{y})$, $\phi(\mathbf{x}+\mathbf{y}) \in C$. Since C is linear, $\phi(\mathbf{x})+\phi(\mathbf{y})+\phi(\mathbf{x}+\mathbf{y}) \in C$. By Corollary 3.13, $\phi(2\,\alpha(\mathbf{x}) * \alpha(\mathbf{y})) \in C$. Since ϕ is a bijection, $2\alpha(\mathbf{x}) * \alpha(\mathbf{y}) \in \mathcal{C}$.

Conversely, assume that condition (3.20) holds. Let $\mathbf{u},\, \mathbf{v} \in C$. There are $\mathbf{x},\, \mathbf{y} \in \mathcal{C}$ such that $\mathbf{u} = \phi(\mathbf{x})$, $\mathbf{v} = \phi(\mathbf{y})$. By (3.20), $2\alpha(\mathbf{x}) * \alpha(\mathbf{y}) \in \mathcal{C}$. Since $\mathcal{C}$ is linear, $\mathbf{x} + \mathbf{y} + 2\alpha(\mathbf{x}) * \alpha(\mathbf{y}) \in \mathcal{C}$ and $\phi(\mathbf{x} + \mathbf{y} + 2\alpha(\mathbf{x}) * \alpha(\mathbf{y})) \in C$. By Corollary 3.13,

$$\phi(\mathrm{x}+\mathrm{y})+\phi(2\alpha(\mathrm{x}) * \alpha(\mathrm{y}))+\phi(2\alpha(\mathrm{x}+\mathrm{y}) * \alpha(2\alpha(\mathrm{x}) * \alpha(\mathrm{y})))$$
$$=\phi(\mathrm{x}+\mathrm{y}+2\alpha(\mathrm{x}) * \alpha(\mathrm{y})) \in C.$$

Clearly, $\alpha(2\alpha(\mathrm{x}) * \alpha(\mathrm{y}))=0$. Therefore

$$\phi(\mathrm{x}+\mathrm{y})+\phi(2\alpha(\mathrm{x}) * \alpha(\mathrm{y})) \in C.$$

Again by Corollary 3.13,

$$\phi(2\alpha(\mathrm{x}) * \alpha(\mathrm{y}))=\phi(\mathrm{x}+\mathrm{y})+\phi(\mathrm{x})+\phi(\mathrm{y}).$$

Hence

$$\begin{aligned}\mathrm{u}+\mathrm{v} &= \phi(\mathrm{x})+\phi(\mathrm{y}) \\ &= \phi(\mathrm{x})+\phi(\mathrm{y})+\phi(\mathrm{x}+\mathrm{y})+\phi(\mathrm{x}+\mathrm{y}) \\ &= \phi(2\alpha(\mathrm{x}) * \alpha(\mathrm{y}))+\phi(\mathrm{x}+\mathrm{y}) \in C.\end{aligned}$$

This proves that C is linear. □

Corollary 3.17. *Let $\mathcal{C}$ be a $\mathbb{Z}_4$-linear code, $\mathrm{x}_1, \ldots, \mathrm{x}_m$ be a set of generators of $\mathcal{C}$, and $C=\phi(\mathcal{C})$. Then C is linear if and only if $2\alpha(\mathrm{x}_i) * \alpha(\mathrm{x}_j) \in \mathcal{C}$ for all i, j satisfying $1 \leq i \leq j \leq m$.*

Proof. Because α is a group homomorphism. □

Example 3.4. Consider the octacode $\mathcal{O}_8$ introduced in Example 1.3. It has generator matrix (1.6). Denote the first and second rows of (1.6) by x_1 and x_2, respectively, i.e.

$$\mathrm{x}_1=(1\ 0\ 0\ 0\ 3\ 1\ 2\ 1),$$
$$\mathrm{x}_2=(0\ 1\ 0\ 0\ 1\ 2\ 3\ 1).$$

Clearly,

$$2\alpha(\mathrm{x}_1) * \alpha(\mathrm{x}_2)=(0\ 0\ 0\ 0\ 2\ 0\ 0\ 2) \notin \mathcal{O}_8.$$

By Proposition 3.16, $\phi(\mathcal{O}_8)$ is nonlinear. Since $\mathcal{O}_8$ is self-dual, $\phi(\mathcal{O}_8)$ is formally self-dual. $\phi(\mathcal{O}_8)$ is called the *Nordstrom–Robinson code.* It is a nonlinear binary code of length 16 and has 256 codewords. It is easy to check that the sum of elements of each row of generator matrix (1.6) is equal to 0 in $\mathbb{Z}_4$, from which we deduce that the sum of the components of every codeword of $\mathcal{O}_8$ is equal to 0 in $\mathbb{Z}_4$. By Proposition 3.4 all codewords of $\phi(\mathcal{O}_8)$ are of even weight. By checking the weights of all the codewords of $\phi(\mathcal{O}_8)$ we know that $\phi(\mathcal{O}_8)$ has minimum weight 6. Since the zero word $0^{16} \in \phi(\mathcal{O}_8)$ and by Proposition 3.6 $\phi(\mathcal{O}_8)$ is distance invariant, $\phi(\mathcal{O}_8)$ has minimum distance 6. Puncturing the coordinates of the codewords of $\phi(\mathcal{O}_8)$ at a fixed position, we obtain a binary nonlinear code of length 15, with 256 codewords and minimum distance 5. But, the 2-error-correcting BCH code of length 15 and minimum distance 5 contains only 128 codewords. $\square$

Example 3.5. Consider the $\mathbb{Z}_4$-code $\mathcal{K}_8$ introduced in Example 1.4. It has generator matrix (1.7). It can be readily checked that for any two rows $\mathbf{x}$ and $\mathbf{y}$, $2\alpha(\mathbf{x}) * \alpha(\mathbf{y}) \in \mathcal{K}_8$. By Corollary 3.17, $\phi(\mathcal{K}_8)$ is a binary linear code. Since $\mathcal{K}_8$ is a self-dual $\mathbb{Z}_4$-code, $\phi(\mathcal{K}_8)$ is formally self-dual. But it can be verified directly that $\phi(\mathcal{K}_8)$ is a self-dual binary linear code. $\square$

Most propositions of Secs. 3.1–3.3 are due to Hammons *et al.* (1994) but now the proofs of them are complete.

3.4. Binary Codes Associated with a $\mathbb{Z}_4$-Linear Code

Let C be a $\mathbb{Z}_4$-linear code. Besides the binary image $\phi(C)$, there are two binary codes $C^{(1)}$ and $C^{(2)}$ which are canonically associated with C. They are defined by

$$C^{(1)} = \{\alpha(\mathbf{c}) \mid \mathbf{c} \in \mathcal{C}\} \tag{3.21}$$

and

$$C^{(2)} = \{\beta(\mathbf{c}) \mid \mathbf{c} \in C,\, \alpha(\mathbf{c}) = 0\}, \tag{3.22}$$

respectively.

Proposition 3.18. *Let C be a $\mathbb{Z}_4$-linear code of length n, and $C^{(1)}$ and $C^{(2)}$ be the binary codes defined by* (3.21) *and* (3.22), *respectively. Then*

(i) *Both $C^{(1)}$ and $C^{(2)}$ and binary linear code, and $C^{(1)} \subseteq C^{(2)}$.*
(ii) *If C is of type $4^{k_1}2^{k_2}$ and has generator matrix* (1.1), *$C^{(1)}$ is a binary linear $[n,\, k_1]$-code with generator matrix*

$$(I_{k_1} \;\; A \;\; \alpha(B)) \tag{3.23}$$

and $C^{(2)}$ is a binary linear $[n,\, k_1 + k_2]$-code with generator matrix

$$\begin{pmatrix} I_{k_1} & A & \alpha(B) \\ & I_{k_2} & C \end{pmatrix}. \tag{3.24}$$

Proof. (i) Since $\alpha : \mathbb{Z}_4 \to \mathbb{Z}_2$ is a group homomorphism, the extended map $\alpha : \mathbb{Z}_4^n \to \mathbb{Z}_2^n$ is also a group homomorphism. $C^{(1)}$ is the image of the map $\alpha : \mathcal{C} \to \mathbb{Z}_2^n$, therefore $C^{(1)}$ is a subgroup of $\mathbb{Z}_2^n$, i.e. $C^{(1)}$ is linear.

Let $\beta(\mathbf{c}), \beta(\mathbf{c}') \in C^{(2)}$, where $\mathbf{c}, \mathbf{c}' \in \mathcal{C}$ and $\alpha(\mathbf{c}) = \alpha(\mathbf{c}') = 0$. Then $\alpha(\mathbf{c} + \mathbf{c}') = 0$ and all components of $\mathbf{c}$ and $\mathbf{c}'$ are either 0 or 2. If we restrict β to the subgroup $\{0, 2\}$ of the additive group of $\mathbb{Z}_4$, then $\beta : \{0, 2\} \to \mathbb{Z}_2$ is an isomorphism of groups and the extension $\beta : \{0, 2\}^n \to \mathbb{Z}_2^n$ is also an isomorphism of groups. Therefore $\beta(\mathbf{c}) + \beta(\mathbf{c}') = \beta(\mathbf{c} + \mathbf{c}') \in C^{(2)}$. Hence $C^{(2)}$ is also linear.

Let $\alpha(\mathbf{c}) \in C^{(1)}$, where $\mathbf{c} \in \mathcal{C}$, then $2\mathbf{c} \in \mathcal{C}$, $\alpha(2\mathbf{c}) = 0$ and $\alpha(\mathbf{c}) = \beta(2\mathbf{c}) \in C^{(2)}$. Therefore $C^{(1)} \subseteq C^{(2)}$.

(ii) is obvious. □

We have the following converse of Proposition 3.18.

Proposition 3.19. *Given two binary linear codes C' and C'', both of length n, with $C' \subseteq C''$, there is a $\mathbb{Z}_4$-linear code $\mathcal{C}$ with $C^{(1)} = C'$ and $C^{(2)} = C''$. If, in addition, C' is doubly even, and $C'' \subseteq C'^{\perp}$, then there is a self-orthogonal $\mathbb{Z}_4$-linear code $\mathcal{C}$ with $C^{(1)} = C'$ and $C^{(2)} = C''$. Furthermore, if $C'' = C'^{\perp}$, then $\mathcal{C}$ is self-dual.*

Proof. Let $\dim C' = k_1$, $\dim C'' = k_1 + k_2$. Without loss of generality we may assume that C' and C'' have generator matrices

$$(I_{k_1} \;\; A \;\; B)$$

and

$$\begin{pmatrix} I_{k_1} & A & B \\ & I_{k_2} & C \end{pmatrix},$$

respectively. Let $\mathcal{C}$ be the $\mathbb{Z}_4$-linear code with generator matrix

$$\begin{pmatrix} I_{k_1} & A & B \\ & 2I_{k_2} & 2C \end{pmatrix}, \tag{3.25}$$

then by Proposition 3.18 (ii), $C^{(1)} = C'$ and $C^{(2)} = C''$. The first assertion is proved.

Now we assume that C' is doubly even and that $C'' \subset C'^{\perp}$ From $C'' \subset C'^{\perp}$ we deduce that any one of the first k_1 rows of (3.25) and any one of its last k_2 rows as words in $\mathbb{Z}_4^n$ are orthogonal. Clearly, any two of its last k_2 rows as words in $\mathbb{Z}_4^n$ are orthogonal. Since C' is doubly even, any one of its first k_1

rows as a word in $\mathbb{Z}_4^n$ is orthogonal to itself. But two distinct rows of the first k_1 rows as words in $\mathbb{Z}_n^4$ are not necessarily orthogonal. So, the $\mathbb{Z}_4$-linear code having generator matrix (3.25) is not necessarily self-orthogonal. We have to modify (3.25) so that the $\mathbb{Z}_4$-linear code it generates is self-orthogonal. For any pair (i, j) with $1 \le j < i \le k_1$ we replace the (i, j)th entry of (3.25) by the inner product mod 4 of the ith row and jth row. From $C' \subseteq C''$ and $C'' \subseteq C'^{\perp}$ we deduce $C' \subseteq C'^{\perp}$, so such an inner product mod 4 is either 0 or 2. Denote the matrix so obtained by G, then it is easy to see that any two rows of G are orthogonal. Let C_G be the $\mathbb{Z}_4$-linear code generated by G, then C_G is self-orthogonal and clearly $C_G^{(1)} = C'$ and $C_G^{(2)} = C''$ The second assertion is also proved.

Assume further that $C'' = C'^{\perp}$ Then

$$k_1 + k_2 = \dim C'' = \dim C'^{\perp} = n - \dim C' = n - k_1.$$

By Propositions 1.1 and 1.2, $|C_G| = 2^{2k_1+k_2}$ and $|C_G^{\perp}| = 2^{2n-2k_1-k_2}$ Hence $|C_G| = |C_G^{\perp}|$. But C_G is self-orthogonal, so C_G is self-dual. □

Proposition 3.20. *Let C' and C'' be two binary linear codes of length n and $C' \subseteq C''$ Define*

$$C = C' + 2C'' = \{\mathbf{a} + 2\mathbf{b} \mid \mathbf{a} \in C',\ \mathbf{b} \in C''\}. \tag{3.26}$$

Then C is a $\mathbb{Z}_4$-linear code if and only if

$$\mathbf{a},\ \mathbf{a}' \in C' \Rightarrow \mathbf{a} * \mathbf{a}' \in C''. \tag{3.27}$$

In this case,

(i) $C^{(1)} = C'$ *and* $C^{(2)} = C''$

(ii) $\phi(C) = \{(\mathbf{u},\ \mathbf{u} + \mathbf{v}) \mid \mathbf{u} \in C'',\ \mathbf{v} \in C'\}$.

(iii) *Assume that C' is doubly even, and that $C'' \subseteq C'^{\perp}$ Then C is self-orthogonal if and only if*

$$\mathbf{a},\ \mathbf{a}' \in C' \Rightarrow w(\mathbf{a} * \mathbf{a}') \equiv 0 \pmod 4. \tag{3.28}$$

In this case, if $C'' = C'^{\perp}$, then C is self-dual.

Proof. Denote the addition in $\mathbb{Z}_2$ by $\oplus$ and addition in $\mathbb{Z}_4$ by $+$. For all $\mathbf{a},\ \mathbf{a}_1 \in C'$ and $\mathbf{b},\ \mathbf{b}_1 \in C''$, we have the identity

$$(\mathbf{a} + 2\mathbf{b}) + (\mathbf{a}_1 + 2\mathbf{b}_1) = (\mathbf{a} \oplus \mathbf{a}_1) + 2(\mathbf{b} \oplus \mathbf{b}_1 \oplus (\mathbf{a} * \mathbf{a}_1)),$$

from which it follows that C is a $\mathbb{Z}_4$-linear code if and only if (3.27) holds.

If (3.27) holds, (i) and (ii) are obvious. Computed in $\mathbb{Z}_4$, $(\mathbf{a}+2\mathbf{b}) \quad (\mathbf{a}'+2\mathbf{b}') = \mathbf{a}\cdot\mathbf{a}'$. Computed in $\mathbb{Z}$, $\mathbf{a}\cdot\mathbf{a}' = w(\mathbf{a} * \mathbf{a}')$. Therefore (iii) holds. □

Example 3.6. Let $\mathcal{C} = \mathcal{K}_4$ be the $\mathbb{Z}_4$-linear code studied in Example 1.1. It has generator matrix (1.3). By Proposition 3.18, $C^{(1)}$ has generator matrix

$$(1\ 1\ 1\ 1)$$

and $C^{(2)}$ has generator matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Therefore $C^{(1)}$ is the repetition code of length 4, $C^{(2)}$ is the parity check code (or the even weight code) of length 4, and $C^{(1)} \subseteq C^{(2)}$. It is clear that condition (3.27) is trivially fulfilled for $C' = C^{(1)}$ and $C'' = C^{(2)}$, so $C^{(1)} + 2C^{(2)}$ is a $\mathbb{Z}_4$-linear code. Clearly $C^{(1)} + 2C^{(2)} = \mathcal{K}_4$. Moreover, $C^{(1)}$ is clearly doubly even, (3.28) is also trivially fulfilled, and $C^{(2)} = C^{(1)\perp}$, we deduce again that $\mathcal{K}_4$ is self-dual. □

Example 3.7. Let $\mathcal{C} = \mathcal{K}_8$ be the $\mathbb{Z}_4$-linear code studied in Example 1.4. It has generator matrix (1.7). By Proposition 3.18, $C^{(1)}$ is a binary linear code with generator matrix

$$(1\ 1\ 1\ 1\ 1\ 1\ 1\ 1)$$

and $C^{(2)}$ is a binary linear code with generator matrix

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

$C^{(1)}$ is the repetition code of length 8, $C^{(2)}$ is the parity check code of length 8 and $C^{(1)} \subseteq C^{(2)}$. Clearly, we have $\mathcal{K}_8 = C^{(1)} + 2C^{(2)}$ By Corollary 3.20, we deduce again that $\mathcal{K}_8$ is self-dual. □

Example 3.8. Let $\mathcal{C} = \mathcal{C}_1$ be the $\mathbb{Z}_4$-linear code studied in Example 1.2. It has generator matrix (1.4). By Proposition 3.18, $C^{(1)}$ is a binary linear code with generator matrix

$$(1\ 1\ 1\ 1)$$

and $C^{(2)}$ is a binary linear code with generator matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

Clearly, we have $\mathcal{C}_1 = C^{(1)} + 2C^{(2)}$. By Corollary 3.20 we deduce again that $\mathcal{C}_1$ is self-orthogonal.

Moreover, from Example 1.2 we know that $\mathcal{C}_1^\perp$ has generator matrix (1.5). Denote the binary codes associated with $\mathcal{C}_1^\perp$ by C' and C'', then by Proposition 3.18 C' and C'' has generator matrices

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix},$$

respectively. Clearly, (3.27) is fulfilled and $\mathcal{C}_2^\perp = C' + 2C''$. $\square$

Propositions 3.18 and 3.19 are due to Conway and Sloane (1993). Most of Proposition 3.20 can be found in Bonnecaze *et al.* (1995).

CHAPTER 4

$\mathbb{Z}_4$-LINEARITY AND $\mathbb{Z}_4$-NONLINEARITY OF SOME BINARY LINEAR CODES

4.1. A Review of Reed–Muller Codes

Let m be a positive integer,

$$G_2 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

and

$$G_{2^m} = \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \otimes \cdots \otimes \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}}_{m \text{ in number}},$$

where $\otimes$ denotes the Kronecker product of matrices. It can be easily verified that G_{2^m} is a $2^m \times 2^m$ nonsingular matrix whose entries are either 0 or 1, that the Hamming weight of each row vector of G_{2^m} is a power of 2, and that the number of row vectors of Hamming weight $2^r (0 \leq r \leq m)$ is $\frac{m!}{r!(m-r)!}$. The row vectors of G_{2^m} of Hamming weight $\geq 2^{m-r}$ generate a binary linear code of length 2^m, dimension

$$\sum_{i=0}^{r} \frac{m!}{i!\,(m-i)!},$$

and minimum distance 2^{m-r}, which is called the *rth order Reed–Muller code* of length 2^m and is denoted by $\mathrm{RM}(r, m)$. The generator matrix of $\mathrm{RM}(r, m)$ formed by the row vectors of Hamming weight $\geq 2^{m-r}$ of G_{2^m} will be denoted by $G(r, m)$.

It is known that RM(r, m) and RM($m-r-1$, m) are dual to each other. It is also clear that RM(m, m) $= \mathbb{F}_2^{2^m}$, RM($m-1$, m) consists of all even weight words of length 2^m, RM($m-2$, m) is the extended binary Hamming code H_{2^m} of length 2^m when $m \geq 3$, RM(1, m) is the first-order Reed–Muller code of length 2^m and RM(0, m) $= \{0^{2^m}, 1^{2^m}\}$.

We agree that RM(-1, m) $=$ RM($m+1$, m) $= \{0^{2^m}\}$ and that $G(-1, m) = G(m+1, m) = 0^{2^m}$ for any $m > 0$.

As we remarked above, the number of row vectors of Hamming weight 2^{m-1} of G_{2^m} is m. It is easy to see that they are

$$(0\ 1)^{2^{m-1}},\ (0\ 0\ 1\ 1)^{2^{m-2}}, \ldots, 0^{2^{m-1}}1^{2^{m-1}}.$$

Denote them by $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$, respectively. Then

$$G(1, m) = \begin{pmatrix} 1^{2^m} \\ \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_m \end{pmatrix} \tag{4.1}$$

is a generator matrix of RM(1, m). The row vectors of Hamming weights $2^{m-r}\,(0 \leq r \leq m)$ are

$$\mathbf{v}_{i_1} * \mathbf{v}_{i_2} * \cdots * \mathbf{v}_{i_r}, \quad 1 \leq i_1 < i_2 < \cdots < i_r \leq m.$$

We understand that when $r = 0$, $\mathbf{v}_{i_1} * \mathbf{v}_{i_2} * \cdots * \mathbf{v}_{i_r} = 1^{2^m}$ Then the row vectors

$$\mathbf{v}_{i_1} * \mathbf{v}_{i_2} * \cdots * \mathbf{v}_{i_s}, \quad 1 \leq i_1 < i_2 < \cdots < i_s \leq m, \quad 0 \leq s \leq r$$

form the generator matrix $G(r, m)$ of the rth-order Reed–Muller code RM(r, m).

We agree that any binary linear code equivalent to RM(r, m) will also be called the rth-order Reed–Muller code and denoted by RM(r, m). In particular, let ξ be a primitive (2^m-1)th root of unity in the finite field $\mathbb{F}_{2^m}$, and form the $m \times 2^m$ matrix

$$M_m = (0\ 1\ \xi\ \xi^2\ \cdots\ \xi^{2^m-2}),$$

where each ξ^j is replaced by ${}^t(a_{1j}, \ldots, a_{mj})$ if $\xi^j = a_{1j} + a_{2j}\,\xi + \cdots + a_{mj}\,\xi^{m-1}$. Denote the rows of M_m by $\mathbf{u}_1, \ldots, \mathbf{u}_m$ in succession. It is known that the row vectors

$$\mathbf{u}_{i_1} * \mathbf{u}_{i_2} * \cdots * \mathbf{u}_{i_s}, \quad 1 \le i_1 < i_2 < \cdots < i_s \le m, \quad 0 \le s \le r$$

generate a binary linear code of length 2^m which is equivalent to the above defined rth-order Reed–Muller code RM(r, m). Then it will also be called the rth-order Reed-Muller code of length 2^m and denoted by RM(r, m) also. This definition of RM(r, m) has the advantage that when the components at the leftmost position of its codewords are deleted, we get a cyclic code, which will be called the *shortened rth-order Reed–Muller code* and denoted by RM$(r, m)^-$.

For more details on Reed–Muller codes, see MacWilliams and Sloane (1977), Chap. 14.

4.2. The $\mathbb{Z}_4$-Linearity of Some RM(r, m)

Let m be a non-negative integer and $0 \le r \le m$. The $\mathbb{Z}_4$-linear code of length 2^{m-1} generated by the matrix

$$\begin{pmatrix} G(r-1, m-1) \\ 2G(r, m-1) \end{pmatrix}$$

over $\mathbb{Z}_4$ will be denoted by ZRM(r, $m-1$). □

Example 4.1. The matrix

$$\begin{pmatrix} G(0, 2) \\ 2G(1, 2) \end{pmatrix}$$

generates the $\mathbb{Z}_4$-linear code ZRM(1, 2). Clearly, ZRM(1, 2) has generator matrix (1.3)

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 2 & 2 \end{pmatrix}.$$

Therefore ZRM(1, 2) is the $\mathbb{Z}_4$-linear code $\mathcal{K}_4$ introduced in Example 1.1. By Example 3.2 we know that φ(ZRM(1, 2)) is the extended binary linear Hamming code H_8 = RM(2, 3) of length $8 = 2^3$. Hence H_8 is $\mathbb{Z}_4$-linear. □

Example 4.2. The matrix

$$\begin{pmatrix} G(0, 3) \\ 2G(1, 3) \end{pmatrix}$$

generates the $\mathbb{Z}_4$-linear code ZRM(1, 3). ZRM(1, 3) has generator matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 2 & 0 & 2 & 0 & 2 \\ 0 & 0 & 2 & 2 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 2 & 2 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 1^8 \\ 2\mathbf{v}_1 \\ 2\mathbf{v}_2 \\ 2\mathbf{v}_3 \end{pmatrix} \qquad \square$$

Example 4.3. The matrix

$$\begin{pmatrix} G(1, 3) \\ 2G(2, 3) \end{pmatrix}$$

generates the $\mathbb{Z}_4$-linear ZRM(2, 3). ZRM(2, 3) has generator matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 1^8 \\ \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \\ 2\mathbf{v}_1 * \mathbf{v}_2 \\ 2\mathbf{v}_1 * \mathbf{v}_3 \\ 2\mathbf{v}_2 * \mathbf{v}_3 \end{pmatrix} \qquad \square$$

Proposition 4.1. *The binary rth-order Reed-Muller code* RM(r, m) *of length* $n = 2^m$ *is* $\mathbb{Z}_4$*-linear for* $r = 0, 1, 2, m-1$ *and* m. *More precisely, it is the binary image of the* $\mathbb{Z}_4$*-linear code* ZRM$(r, m-1)$ *of length* 2^{m-1} *for* $r = 0, 1, 2, m-1$ *and* m.

Proof. For $r = 0$, ZRM$(0, m-1)$ is generated by

$$\begin{pmatrix} G(-1, m-1) \\ 2G(0, m-1) \end{pmatrix}$$

and, hence, has generator matrix

$$(2^{2^{m-1}}).$$

Thus

$$\mathrm{ZRM}(0, m-1) = \{0^{2^{m-1}}, 2^{2^{m-1}}\}.$$

Under φ

$$0^{2^{m-1}} \mapsto 0^{2^m},$$

$$2^{2^{m-1}} \mapsto 1^{2^m}.$$

Hence

$$\varphi(\mathrm{ZRM}(0,\, m-1)) = \mathrm{RM}(0,\, m).$$

For $r = 1$, ZRM$(1, m-1)$ is generated by

$$\begin{pmatrix} G(0,\, m-1) \\ 2G(1,\, m-1) \end{pmatrix}.$$

Hence it has generator matrix

$$\begin{pmatrix} 1^{2^{m-1}} \\ 2\mathbf{v}_1 \\ \vdots \\ 2\mathbf{v}_{m-1} \end{pmatrix},$$

where $\mathbf{v}_1, \ldots, \mathbf{v}_{m-1}$ are 2^{m-1}-dimensional vectors. It follows that $|\,\mathrm{ZRM}(1,\, m-1)\,| = 4\cdot 2^{m-1} = 2^{m+1}$. Obviously, the condition of Corollary 3.16 is fulfilled for the generators $1^{2^{m-1}}, 2\mathbf{v}_1, \ldots, 2\mathbf{v}_{m-1}$ of ZRM$(1,\, m-1)$. By Corollary 3.16, $\phi(\mathrm{ZRM}(1,\, m-1))$ is linear. Under ϕ,

$$\begin{aligned} 1^{2^{m-1}} &\mapsto (0\ 1)^{2^{m-1}} = \mathbf{v}_1 \text{ in } \mathbb{Z}_2^{2^m}, \\ 2\cdot 1^{2^{m-1}} &\mapsto (1\ 1)^{2^{m-1}} = 1^{2^m}, \\ 2\mathbf{v}_i &\mapsto \mathbf{v}_{i+1} \text{ in } \mathbb{Z}_2^{2^m} \quad (i = 1, \ldots, m-1). \end{aligned}$$

Therefore $\phi(\mathrm{ZRM}(1,\, m-1)) \supseteq \mathrm{RM}(1,\, m)$. But $|\,\mathrm{RM}(1,\, m)\,| = 2^{m+1}$ Hence $\phi(\mathrm{ZRM}(1,\, m)) = \mathrm{RM}(1,\, m)$.

The cases $r = 2$ and $r = m-1$ can be proved in the same way, and the case $r = m$ is trivial. □

4.3. The $\mathbb{Z}_4$-Nonlinearity of Extended Binary Hamming Codes H_{2^m} when $m \geq 5$

We mentioned in Sec. 4.1 that the $(m-2)$th-order Reed–Muller code of length 2^m, RM$(m-2,\, m)$, is the extended binary Hamming code H_{2^m} when $m \geq 3$. In Example 3.2 we showed that H_8 is $\mathbb{Z}_4$-linear. By Proposition 4.1, $H_{2^4} = \mathrm{RM}(4-2,\, 4) = \mathrm{RM}(2,\, 4)$ is also $\mathbb{Z}_4$-linear. In the following we will show that when $m \geq 5$, $H_{2^m} = \mathrm{RM}(m-2,\, m)$ is not $\mathbb{Z}_4$-linear. We begin with some lemmas.

Lemma 4.2. *Let H_{2^m} be a $[2^m,\, 2^m - m - 1,\, 4]$ extended binary Hamming code and $m \geq 4$. Then H_{2^m} contains at least two codewords of weight 4 that meet in just one coordinate.*

Proof. It is well known that the matrix (4.1)

$$G(1,\,m) = \begin{pmatrix} 1^{2^m} \\ \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_m \end{pmatrix}$$

is a parity-check matrix of H_{2^m}. For illustration we write this matrix down explicitly for the case $m = 4$.

$$\begin{pmatrix} 1^{16} \\ \mathbf{v}_1 \\ \mathbf{v}_2 \\ \mathbf{v}_3 \\ \mathbf{v}_4 \end{pmatrix} = \begin{pmatrix} 1&1&1&1&1&1&1&1&1&1&1&1&1&1&1&1 \\ 0&1&0&1&0&1&0&1&0&1&0&1&0&1&0&1 \\ 0&0&1&1&0&0&1&1&0&0&1&1&0&0&1&1 \\ 0&0&0&0&1&1&1&1&0&0&0&0&1&1&1&1 \\ 0&0&0&0&0&0&0&0&1&1&1&1&1&1&1&1 \end{pmatrix}.$$

The columns of $G(1,\,m)$ will be numbered by $0, 1, 2, \ldots, 2^m - 1$. Clearly, the sum of the zeroth, first and second columns of $G(1,\,m)$ is equal to the third column and the sum of the zeroth, fourth and eighth columns of $G(1,\,m)$ is equal to the twelfth column. Therefore

$$(1\,1\,1\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0^{2^m-4})$$

and

$$(1\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,0^{2^m-4})$$

are two codewords of H_{2^m}, which have weight 4 and meet in just one coordinate.

Lemma 4.3. *Let m be an integer ≥ 4, A_2 and A_3 be non-negative integers satisfying the condition $2A_2 + 3A_3 < 2^{m-1}$. Then there does not exist binary linear code of length $2^{m-1} - 2A_2 - 3A_3$, dimension $\geq 2^{m-1} - m - A_2 - A_3$, and minimum distance 4 unless $A_2 = A_3 = 0$ and the code is the extended binary Hamming code of length 2^{m-1}*

Proof. Assume that there is a binary linear code of length $2^{m-1} - 2A_2 - 3A_3$, dimension $\geq 2^{m-1} - m - A_2 - A_3$, and minimum distance 4 and denote it by C. Clearly, we must have $2^{m-1} - 2A_2 - 3A_3 > 2^{m-1} - m - A_2 - A_3$, and hence, $A_2 + 2A_3 < m$. We can delete the coordinates of all codewords of C at a fixed position in such a way that we obtain a binary linear code C^- of length $2^{m-1} - 2A_2 - 2A_3 - 1$, dimension $\geq 2^{m-1} - m - A_2 - A_3$, and minimum distance ≥ 3. By sphere-packing bound for C^-

$$2^{2^{m-1}-m-A_2-A_3}\left(1+\binom{2^{m-1}-2A_2-3A_3-1}{1}\right) \leq 2^{2^{m-1}-2A_2-3A_3-1}. \tag{4.2}$$

Then

$$\begin{aligned}\text{L.H.S. of (4.2)} &= 2^{2^{m-1}-m-A_2-A_3}(2^{m-1}-2A_2-3A_3)\\ &= 2^{2^{m-1}-2A_2-3A_3-1}\,2^{A_2+2A_3-m+1}(2^{m-1}-2A_2-3A_3).\end{aligned}$$

We have $2A_2+3A_3 \leq 2(A_2+2A_3) < 2m$, so $2^{m-1}-2A_2-3A_3 > 2^{m-1}-2m$.

For $m \geq 6$, we have $2^{m-1}-2m > 2^{m-2}$ and then $2^{m-1}-2A_2-3A_3 > 2^{m-2}$ Consequently

$$\begin{aligned}\text{L.H.S. of (4.2)} &> 2^{2^{m-1}-2A_2-3A_3-1}\,2^{A_2+2A_3-1}\\ &> \text{R.H.S. of (4.2)},\end{aligned}$$

unless $A_2 = A_3 = 0$.

When $m = 5$, then $A_2 + 2A_3 < 5$. If $(A_2, A_3) \neq (0, 0)$, then there are seven possibilities:

$$(A_2, A_3) = (0, 1), (0, 2), (1, 0), (1, 1), (2, 0), (3, 0), (4, 0).$$

For any one of these possibilities, we always have

$$2^{A_2+2A_3-m+1}(2^{m-1}-2A_2-3A_3) > 1. \tag{4.3}$$

Therefore we also have

$$\text{L.H.S. of (4.2)} > \text{R.H.S. of (4.2)}. \tag{4.4}$$

When $m = 4$, then $A_2 + 2A_3 < 4$. If $(A_2, A_3) \neq (0, 0)$, there are five possibilities

$$(A_2, A_3) = (0, 1), (1, 0), (1, 1), (2, 0), (3, 0).$$

For any one of these possibilities we also have (4.3) and hence (4.4).

Therefore we conclude that when $m \geq 4$, we must have $A_2 = A_3 = 0$. When $A_2 = A_3 = 0$, the code C is of length 2^{m-1}, dimension $\geq 2^{m-1} - m$, and minimum distance 4. From the sphere packing bound for C^- we deduce that dim $C = 2^{m-1} - m$. It is well known that there is a unique binary linear

$[2^{m-1}, 2^{m-1}-m, 4]$-code, the extended binary Hamming code of length 2^{m-1}, within equivalence. So C is the extended binary Hamming code of length 2^{m-1}. $\square$

Proposition 4.4. *The extended binary Hamming code H_{2^m} of length 2^m is not $\mathbb{Z}_4$-linear for $m \geq 5$.*

Proof. We will prove by contradiction. Assume that H_{2^m} is a $[2^m, 2^m - m - 1, 4]$ extended binary Hamming code with its coordinates so arranged that $H_{2^m} = \phi(\mathcal{H})$ for some $\mathbb{Z}_4$-linear code $\mathcal{H}$. Let

$$F = \{\mathbf{c} \in H_{2^m} \,|\, \sigma(\mathbf{c}) = \mathbf{c}\},$$

where σ is the map defined by (3.14). Clearly, F is a linear subcode of H_{2^m}. Since $(1, 0^{2^{m-1}-1}, 1, 0^{2^{m-1}-1}) \notin H_{2^m}$, it does not belong to F either. It follows that dim $F \leq 2^{m-1} - 1$.

Define a map

$$\begin{aligned} \psi \,:\, H_{2^m} &\to F \\ \mathbf{c} &\mapsto \mathbf{c} + \sigma(\mathbf{c}). \end{aligned}$$

Clearly, ψ is a group homomorphism and Im $\psi \subset$ Ker $\psi = F$. Since dim Ker $\psi =$ dim $F \leq 2^{m-1} - 1$ and dim Ker ψ + dim Im ψ = dim $H_{2^m} = 2^m - m - 1$, we have dim Im $\psi \geq 2^{m-1} - m$.

Let E consist of the right-hand halves of the codewords in Im ψ. Then E is a binary linear code of length 2^{m-1}, dimension $\geq 2^{m-1} - m$, and minimum weight 2. By Corollary 3.15, E is closed under componentwise multiplication.

Assume that the positions of codewords of E are numbered by 1, 2, ..., 2^{m-1} Let $\mathbf{x} = (x_1, \ldots, x_{2^m-1})$, $\mathbf{y} = (y_1, \ldots, y_{2^m-1})$ be any two codewords of E of weights 2 or 3. Then $\mathbf{x} + \mathbf{y} \in E$ and $\mathbf{x} * \mathbf{y} \in E$. Define

$$S_{\mathbf{x}} = \{i \,|\, 1 \leq i \leq 2^{m-1}, \quad x_i = 1\}.$$

Then $S_{\mathbf{x}} \cap S_{\mathbf{y}} = \emptyset$; otherwise, either $\mathbf{x} + \mathbf{y}$ or $\mathbf{x} * \mathbf{y}$ would be a codeword of weight 1 in E, a contradiction.

Denote the number of codewords of weight i in E by A_i. Define

$$J = \{j \in I \,|\, \exists\ \mathbf{x} \in E \quad \text{with} \quad w(\mathbf{x}) = 2 \text{ or } 3 \quad \text{and} \quad x_j = 1\}.$$

By the preceding paragraph, $|\,J\,| = 2A_2 + 3A_3$. Delete those components numbered by numbers in J from the codewords of E, we obtain a shortened binary linear code E^* of length $2^{m-1} - 2A_2 - 3A_3$, dimension $\geq 2^{m-1} - m - A_2 - A_3$. Let $\mathbf{z}$ be a codeword of weight 4 in E. Then $S_{\mathbf{z}} \cap J = \emptyset$; otherwise, as in the preceding paragraph it will lead to a contradiction. Hence the minimum weight of E^* is 4. But by Lemma 4.3, E^* cannot exist unless $A_2 = A_3 = 0$ and E is itself an extended binary Hamming code of length 2^{m-1}. Since $m \geq 5$ and E is closed under componentwise multiplication, we can use Lemma 4.2 to produce a codeword of weight 1, again a contradiction. $\square$

From the above discussion we conclude that when $m \leq 4$ all Reed–Muller codes $\mathrm{RM}(r, m), 0 \leq r \leq m$, are $\mathbb{Z}_4$-linear, that when $m = 5$, $\mathrm{RM}(r, 5)$, $r = 0, 1, 2, 4, 5$, are $\mathbb{Z}_4$-linear and $\mathrm{RM}(3, 5)$ is not, and when $m > 5$, $\mathrm{RM}(r, m), r = 0, 1, 2, m-1, m$, are $\mathbb{Z}_4$-linear and $\mathrm{RM}(m-2, m)$ is not. It was proved recently by X.-D. Hou *et al.* (1997) that when $m > 5$ and $2 < r < m-2$, $\mathrm{RM}(r, m)$ is not $\mathbb{Z}_4$-linear.

It is worthwhile to remark that $\mathrm{RM}(1, m)$ and $\mathrm{RM}(m-2, m) = H_{2^m}$ are dual to each other and that $\mathrm{RM}(1, m)$ is $\mathbb{Z}_4$-linear, but $\mathrm{RM}(m-2, m)$ is not.

Propositions 4.1 and 4.4 are due to Hammons *et al.* (1994).

CHAPTER 5

HENSEL'S LEMMA AND HENSEL LIFT

5.1. Hensel's Lemma

In studying $\mathbb{Z}_4$-codes it is convenient to introduce the Galois ring GR(4^m). Hensel's lemma is an important tool in studying Galois rings. In the following we restrict our study of Hensel's lemma to the simplest case, i.e. the case of polynomials over $\mathbb{Z}_4$, which is needed in studying GR(4^m). To extend it to the general case, i.e. the case of polynomials over $\mathbb{Z}_{p^e}$, where p is any prime and e is any integer > 1, is immediate.

Let $\mathbb{Z}_4[X]$ be the polynomial ring in an indeterminate X over $\mathbb{Z}_4$. We have defined a ring homomorphism

$$\alpha : \mathbb{Z}_4 \to \mathbb{Z}_2$$
$$0, 2 \mapsto 0$$
$$1, 3 \mapsto 1.$$

Henceforth we shall simply denote the map α by "$-$", i.e. $\overline{0} = \overline{2} = 0$ and $\overline{1} = \overline{3} = 1$. The map $-: \mathbb{Z}_4 \to \mathbb{Z}_2$ can be naturally extended to a map from $\mathbb{Z}_4[X]$ to $\mathbb{Z}_2[X]$ as follows:

$$\mathbb{Z}_4[X] \to \mathbb{Z}_2[X]$$
$$a_0 + a_1X + \cdots + a_nX^n \mapsto \overline{a}_0 + \overline{a}_1X + \cdots + \overline{a}_nX^n$$

It can be readily verified that this extended map is a ring homomorphism from $\mathbb{Z}_4[X]$ onto $\mathbb{Z}_2[X]$ with kernel $(2) = \mathbb{Z}_4[X]2 = \{2f(X) \mid f(X) \in \mathbb{Z}_4[X]\}$. This extended ring homomorphism will also be denoted by $-$ and the image of $f(X) \in \mathbb{Z}_4[X]$ under the map $-$ will be denoted by $\overline{f}(X)$.

For any $f(X) \in \mathbb{Z}_4[X]$ define

$$(f(X)) = \mathbb{Z}_4[X]\, f(X) = \{g(X)\, f(X) \mid g(X) \in \mathbb{Z}_4[X]\}$$

Let $f_1(X)$ and $f_2(X)$ be polynomials in $\mathbb{Z}_4[X]$. They are said to be *coprime* in $\mathbb{Z}_4[X]$ if there are polynomials $\lambda_1(X)$, $\lambda_2(X)$ in $\mathbb{Z}_4[X]$ such that

$$\lambda_1(X)f_1(X) + \lambda_2(X)f_2(X) = 1,$$

or, equivalently, if

$$\mathbb{Z}_4[X]\ f_1(X) + \mathbb{Z}_4[X]\ f_2(X) = \mathbb{Z}_4[X].$$

The coprimeness of polynomials in $\mathbb{Z}_2[X]$ can be defined in a similar way. It is well known that two polynomials $f_1(X)$ and $f_2(X)$ in $\mathbb{Z}_2[X]$ are coprime if and only if they have no common divisor of degree ≥ 1.

Lemma 5.1. *Let $f_1(X)$ and $f_2(X) \in \mathbb{Z}_4[X]$ and denote their images in $\mathbb{Z}_2[X]$ under $^-$ by $\overline{f}_1(X)$ and $\overline{f}_2(X)$, respectively. Then $f_1(X)$ and $f_2(X)$ are coprime in $\mathbb{Z}_4[X]$ if and only if $\overline{f}_1(X)$ and $\overline{f}_2(X)$ are coprime in $\mathbb{Z}_2[X]$.*

Proof. Assume that $\overline{f}_1(X)$ and $\overline{f}_2(X)$ are coprime in $\mathbb{Z}_2[X]$. Then there are polynomials $\lambda_1(X)$ and $\lambda_2(X)$ in $\mathbb{Z}_4[X]$ such that

$$\overline{\lambda}_1(X)\ \overline{f}_1(X) + \overline{\lambda}_2(X)\ \overline{f}_2(X) = 1.$$

Thus

$$\lambda_1(X)\ f_1(X) + \lambda_2(X)\ f_2(X) = 1 + 2k(X)\,, \tag{5.1}$$

where $k(X) \in \mathbb{Z}_4[X]$. Multiplying the above equation by $2k(X)$, we have

$$2k(X)\,\lambda_1(X)\ f_1(X) + 2k(X)\,\lambda_2(X)\ f_2(X) = 2k(X). \tag{5.2}$$

Substituting (5.2) into (5.1), we obtain

$$[1 - 2k(X)]\ \lambda_1(X)\ f_1(X) + (1 - 2k(X))\ \lambda_2(X)\ f_2(X) = 1.$$

Therefore $f_1(X)$ and $f_2(X)$ are coprime in $\mathbb{Z}_4[X]$. The converse part is easy. □

Lemma 5.2. (*Hensel's Lemma*) *Let $f(X)$ be a monic polynomial in $\mathbb{Z}_4[X]$ and assume that*

$$\overline{f}(X) = \overline{f}_1(X)\ \overline{f}_2(X),$$

where $\overline{f}_1(X)$ and $\overline{f}_2(X)$ are coprime polynomials in $\mathbb{Z}_2[X]$. Then there exist monic polynomials $g_1(X)$, $g_2(X) \in \mathbb{Z}_4[X]$ with the following properties:

(i) $f(X) = g_1(X)\ g_2(X)$,
(ii) $\overline{g}_1(X) = \overline{f}_1(X)$, $\overline{g}_2(X) = \overline{f}_2(X)$,
(iii) $\deg g_1(X) = \deg \overline{f}_1(X)$, $\deg g_2(X) = \deg \overline{f}_2(X)$,
(iv) $g_1(X)$ *and* $g_2(X)$ *are coprime in* $\mathbb{Z}_4[X]$.

Proof. Let $f_1(X) \in \mathbb{Z}_4[X]$ be an original of $\overline{f}_1(X)$ under the map $^-$ and $f_2(X) \in \mathbb{Z}_4[X]$ be one of $\overline{f}_2(X)$. We can choose both $f_1(X)$ and $f_2(X)$ to be monic, which implies that $\deg f_1(X) = \deg \overline{f}_1(X)$ and $\deg f_2(X) = \deg \overline{f}_2(X)$. Clearly, we have

$$f(X) - f_1(X)\ f_2(X) = 2k(X),$$

where $k(X) \in \mathbb{Z}_4[X]$ and $\deg k(X) < \deg f(X)$. Since $\overline{f}_1(X)$ and $\overline{f}_2(X)$ are coprime in $\mathbb{Z}_2[X]$, by Lemma 5.1 $f_1(X)$ and $f_2(X)$ are coprime in $\mathbb{Z}_4[X]$. Thus there exist $\lambda_1(X)$ and $\lambda_2(X) \in \mathbb{Z}_4[X]$ such that

$$\lambda_1(X)\ f_1(X) + \lambda_2(X)\ f_2(X) = k(X). \tag{5.3}$$

Dividing $\lambda_1(X)$ by $f_2(X)$, we obtain

$$\lambda_1(X) = q_1(X)\ f_2(X) + r_1(X), \tag{5.4}$$

where $q_1(X),\ r_1(X) \in \mathbb{Z}_4[X]$ and $\deg r_1(X) < \deg f_2(X)$. Similarly,

$$\lambda_2(X) = q_2(X)\ f_1(X) + r_2(X), \tag{5.5}$$

where $q_2(X), r_2(X) \in \mathbb{Z}_4[X]$ and $\deg r_2(X) < \deg f_1(X)$. Substituting (5.4) and (5.5) into (5.3), we obtain

$$[q_1(X)\ f_2(X) + r_1(X)]\ f_1(X) + [q_2(X)\ f_1(X) + r_2(X)]\ f_2(X) = k(X).$$

Thus

$$[q_1(X) + q_2(X)]\ f_1(X)\ f_2(X) = k(X) - r_1(X)\ f_1(X) - r_2(X)\ f_2(X).$$

The R.H.S. of the above equality is a polynomial of degree less than $\deg f(X)$ and its L.H.S. is a polynomial of degree $\geq \deg f_1(X) + \deg f_2(X) = \deg f(X)$, unless $q_1(X)+q_2(X) = 0$. Therefore we must have $q_1(X)+q_2(X) = 0$ and consequently

$$r_1(X)\ f_1(X) + r_2(X)\ f_2(X) = k(X).$$

Let

$$g_1(X) = f_1(X) + 2r_2(X),$$
$$g_2(X) = f_2(X) + 2r_1(X).$$

Then both $g_1(X)$ and $g_2(X)$ are monic polynomials in $\mathbb{Z}_4[X]$ and

$$\begin{aligned} g_1(X)\ g_2(X) &= f_1(X)\ f_2(X) + 2\,[r_1(X)\ f_1(X) + r_2(X)\ f_2(X)] \\ &= f_1(X)\ f_2(X) + 2k(X) \\ &= f(X). \end{aligned}$$

This proves (i). By the construction of $g_1(X)$ and $g_2(X)$, we have $\overline{g}_1(X) = \overline{f}_1(X)$, $\overline{g}_2(X) = \overline{f}_2(X)$, deg $g_1(X)$ = deg $f_1(X)$ = deg $\overline{f}_1(X)$, and deg $g_2(X)$ = deg $f_2(X)$ = deg $\overline{f}_2(X)$. Therefore (ii) and (iii) also hold. Since $\overline{f}_1(X)$ and $\overline{f}_2(X)$ are coprime in $\mathbb{Z}_2[X]$, by Lemma 5.1, $g_1(X)$ and $g_2(X)$ are coprime in $\mathbb{Z}_4[X]$. This proves (iv). □

By mathematical induction, Lemma 5.2 can be generalized as follows:

Lemma 5.3. (*Hensel's Lemma*) *Let $f(X)$ be a monic polynomial in $\mathbb{Z}_4[X]$ and assume that*

$$\overline{f}(X) = \overline{f}_1(X)\ \overline{f}_2(X) \cdots \overline{f}_r(X),$$

where $\overline{f}_1(X), \overline{f}_2(X), \ldots, \overline{f}_r(X)$ are pairwise coprime polynomials in $\mathbb{Z}_2[X]$. Then there exist monic polynomials $g_1(X), g_2(X), \ldots, g_r(X) \in \mathbb{Z}_4[X]$ with the following properties:

(i) $f(X) = g_1(X)\ g_2(X) \cdots g_r(X)$,
(ii) $\overline{g}_i(X) = \overline{f}_i(X), \quad i = 1, 2, \ldots, r$,
(iii) deg $g_i(X)$ = deg $\overline{f}_i(X), \quad i = 1, 2, \ldots, r$,
(iv) $g_1(X), g_2(X), \ldots, g_r(X)$ *are pairwise coprime in* $\mathbb{Z}_4[X]$. □

5.2. Basic Irreducible Polynomials

Let $f(X)$ be a monic polynomial of degree $m \geq 1$ in $\mathbb{Z}_4[X]$. If $\overline{f}(X)$ is irreducible over $\mathbb{Z}_2$, then $f(X)$ is called a *basic irreducible polynomial* of degree m in $\mathbb{Z}_4[X]$. If $\overline{f}(X)$ is primitive of degree m over $\mathbb{Z}_2$, then $f(X)$ is called a *basic primitive polynomial* of degree m in $\mathbb{Z}_4[X]$.

Now we shall use Hensel's lemma to prove the existence of basic irreducible polynomials of any degree over $\mathbb{Z}_4$.

Proposition 5.4. *For any positive integer m there exists a monic polynomial $f(X)$ of degree m in $\mathbb{Z}_4[X]$ such that $f(X) \mid (X^{2^m-1} - 1)$ in $\mathbb{Z}_4[X]$ and that $\overline{f}(X)$ is irreducible over $\mathbb{Z}_2$. Thus for any positive integer m, there exists a basic irreducible polynomial of degree m in $\mathbb{Z}_4[X]$.*

Proof. By the theory of Galois fields, (see Wan (1992), Chap. 3), for any positive integer m there exist irreducible polynomials of degree m in $\mathbb{Z}_2[X]$, each irreducible polynomial of degree m in $\mathbb{Z}_2[X]$ is a divisor of $X^{2^m-1} - 1$ in $\mathbb{Z}_2[X]$, and $X^{2^m-1} - 1$ has no multiple roots in any extension field of $\mathbb{Z}_2$. Let $f_2(X)$ be an irreducible polynomial of degree m in $\mathbb{Z}_2[X]$. Let

$$g_2(X) = \frac{X^{2^m-1} - 1}{f_2(X)},$$

then $f_2(X)$ and $g_2(X)$ are coprime in $\mathbb{Z}_2[X]$ and

$$X^{2^m-1} - 1 = f_2(X)\, g_2(X).$$

By Hensel's lemma these are monic polynomials $f(X)$ and $g(X)$ in $\mathbb{Z}_4[X]$ such that

$$X^{2^m-1} - 1 = f(X)\, g(X) \quad \text{in} \quad \mathbb{Z}_4[X],$$

$\overline{f}(X) = f_2(X)$, $\overline{g}(X) = g_2(X)$, $\deg f(X) = \deg f_2(X)$, $\deg g(X) = \deg g_2(X)$, furthermore $f(X)$ and $g(X)$ are coprime in $\mathbb{Z}_4[X]$. Then $f(X)$ is a monic polynomial of degree m in $\mathbb{Z}_4[X]$ such that $f(X) \mid (X^{2^m-1} - 1)$ in $\mathbb{Z}_4[X]$ and that $\overline{f}(X) = f_2(X)$ is irreducible over $\mathbb{Z}_2$. □

Corollary 5.5. (of the proof) *For any positive integer m there exists a monic polynomial $f(X)$ of degree m in $\mathbb{Z}_4[X]$ such that $f(X) \mid (X^{2^m-1} - 1)$ in $\mathbb{Z}_4[X]$ and that $\overline{f}(X)$ is a primitive polynomial of degree m over $\mathbb{Z}_2$. Thus, for any positive integer m there exists a basic primitive polynomial of degree m in $\mathbb{Z}_4[X]$.*

Proof. In the proof of Proposition 5.4, let $f_2(X)$ be a primitive polynomial of degree m over $\mathbb{Z}_2$. □

After some preparations in Secs. 5.3 and 5.4 we shall prove in Sec. 5.5 that if $f_2(X)$ is a polynomial over $\mathbb{Z}_2$ dividing $X^n - 1$ in $\mathbb{Z}_2[X]$ for some odd positive integer n, then there exists a unique monic polynomial $f(X)$ over $\mathbb{Z}_4$ dividing $X^n - 1$ in $\mathbb{Z}_4[X]$ and $\overline{f}(X) = f_2(X)$. Moreover, $f(X)$ is independent of the odd positive integer n. Such a polynomial $f(X)$ will be called the *Hensel lift* of $f_2(X)$.

5.3. Some Concepts from Commutative Ring Theory

In this section we recapitulate some concepts from commutative ring theory, which will be needed later. They can be found in Zariski and Samuel (1958).

Let R be a commutative ring. An element $z \in R$ is called a *zero divisor* if $z \neq 0$ and there is a nonzero element $y \in R$ such that $yz = 0$. An element $w \in R$ is called *nilpotent* if there is a positive integer n such that $w^n = 0$. An element $e \in R$ is called an *idempotent*, if $e^2 = e$; moreover, if $e^2 = e$ and $e \neq 0$ then e is called a *nonzero idempotent*. If e and e' are nonzero idempotents of R and $ee' = 0$, then they are said to be *orthogonal*.

Now assume that R has an identity element 1 and $1 \neq 0$. An element $u \in R$ is called an *invertible element* (or a *unit*) if there is an element $v \in R$ such that $uv = 1$. A nonzero element $p \in R$ is called an *irreducible element* if p is not a unit and if $p = ab$ where $a, b \in R$ then a is a unit or b is a unit.

For example, in $\mathbb{Z}_2$, 1 is the only unit, 0 is the only nilpotent element, and there is no zero divisor as well as irreducible element. In $\mathbb{Z}_4$, 1 and 3 are units, 0 and 2 are nilpotent, and 2 is a zero divisor as well as an irreducible element. In $\mathbb{Z}_2[X]$, 1 is the only unit, 0 is the only nilpotent element, there are no zero divisors, and irreducible elements are irreducible polynomials.

A nonempty set I of a commutative ring R is called an *ideal* if $a, b \in I$ and $r \in R$ imply $a + b \in I$ and $ra \in I$. Let $a \in R$, then the set $Ra = \{ra \mid r \in R\}$ is an ideal, called the *principal ideal* generated by a and denoted by (a).

For example, if R has an identity 1 then $R = (1)$. Every ideal of the ring $\mathbb{Z}_2[X]$ is principal.

An ideal M of R is called *maximal* if $M \neq R$ and there is no ideal not equal to R and containing M properly. An ideal P of R is called *prime* if $P \neq R$, and $ab \in P$ implies $a \in P$ or $b \in P$ An ideal Q of R is called *primary* if $Q \neq R$, and $ab \in Q$ implies $a \in Q$ or $b^n \in Q$ for some positive integer n.

Clearly, an ideal M of R is maximal if and only if the residue class ring R/M is a field, and an ideal P of R is prime if and only if $P \neq R$ and R/P has no zero divisors. Moreover, all maximal ideals are prime, but not conversely, and all prime ideals are primary, but not conversely.

For example, in $\mathbb{Z}_2[X]$ the ideal $(f(X))$ generated by an irreducible polynomial $f(X)$ is prime and also maximal. Conversely, every nonzero prime ideal of $\mathbb{Z}_2[X]$ is generated by an irreducible polynomial. The ideal $(f(X)^e)$ generated by a power of an irreducible polynomial $f(X)$ is primary. Conversely, every nonzero primary ideal is generated by a power of an irreducible polynomial.

An element $a \neq 0$ of a commutative ring R is called a *prime* or *primary* element, if the ideal (a) is a prime or primary ideal of R, respectively. If $R = \mathbb{Z}_2[X]$ or $\mathbb{Z}_4[X]$, prime elements and primary elements are also called *prime polynomials* and *primary polynomials*, respectively.

For example, in $\mathbb{Z}_2[X]$ prime polynomials are irreducible polynomials and conversely, primary polynomials are powers of irreducible polynomials and conversely.

Let I be an ideal of a commutative ring R. Define

$$\sqrt{I} = \{a \in R \,|\, a^n \in I \text{ for some positive integer } n\}.$$

It is easy to verify that $\sqrt{I}$ is an ideal of R. We call $\sqrt{I}$ the *radical* of I. Clearly, $I \subseteq \sqrt{I}$.

It is easy to prove that the radical of a primary ideal is prime and the radical of a prime ideal is itself.

Now we illustrate some of the foregoing concepts with the ring $\mathbb{Z}_4[X]$.

First, it is easy to prove that a polynomial $f(X)$ of $\mathbb{Z}_4[X]$ is a unit if and only if $\overline{f}(X) = 1$ in $\mathbb{Z}_2[X]$, and if and only if it can be expressed in the form $f(X) = \pm 1 + 2g(X)$, where $g(X) \in \mathbb{Z}_4[X]$. Moreover, a polynomial $f(X) \in \mathbb{Z}_4[X]$ is nilpotent if and only if $\overline{f}(X) = 0$ in $\mathbb{Z}_2[X]$, and if and only if it is a zero divisor or zero in $\mathbb{Z}_4[X]$.

The kernel of the ring homomorphism $-$: $\mathbb{Z}_4[X] \to \mathbb{Z}_2[X]$ defined in Sec. 5.1 is the principal ideal (2). (2) is a prime ideal, for $\mathbb{Z}_4[X]/(2) \simeq \mathbb{Z}_2[X]$, which has no zero divisors. Let P be a prime ideal of $\mathbb{Z}_4[X]$, then the image $\overline{P}$ of P under the ring homomorphism $-$: $\mathbb{Z}_4[X] \to \mathbb{Z}_2[X]$ is a prime ideal of $\mathbb{Z}_2[X]$. Moreover, we have

Lemma 5.6. *All prime ideals of* $\mathbb{Z}_4[X]$ *containing* (2) *properly are maximal.*

Proof. Let P be a prime ideal of $\mathbb{Z}_4[X]$ which contains (2) properly. Then $\overline{P}$ is a prime ideal of $\mathbb{Z}_4[X]/(2) \simeq \mathbb{Z}_2[X]$ and $\overline{P} \neq (0)$. Therefore $\overline{P}$ is a maximal ideal of $\mathbb{Z}_2[X]$ and $\mathbb{Z}_2[X]/\overline{P}$ is a field. By the second isomorphism theorem,

$$\begin{aligned} \mathbb{Z}_4[X]/P &\simeq (\mathbb{Z}_4[X]/(2))/(P/(2)) \\ &\simeq \mathbb{Z}_2[X]/\overline{P} \end{aligned}$$

Hence P is a maximal ideal of $\mathbb{Z}_4[X]$. $\square$

Lemma 5.7. *Let Q be an ideal of $\mathbb{Z}_4[X]$ containing (2) properly. Then Q is primary if and only if $\sqrt{Q}$ is prime.*

Proof. The "only if" part is immediate. We prove only the "if" part. Assume that $\sqrt{Q}$ is prime. Clearly, $(2) \subseteq \sqrt{Q}$. If $(2) = \sqrt{Q}$, then $Q \subseteq \sqrt{Q} = (2)$, which contradicts the hypothesis that Q contains (2) properly. Therefore $\sqrt{Q}$ contains (2) properly. By Lemma 5.6, $\sqrt{Q}$ is maximal. Since $\sqrt{Q}$ is prime, $\sqrt{Q} \neq \mathbb{Z}_4[X]$. Thus $Q \neq \mathbb{Z}_4[X]$. Let a, $b \in \mathbb{Z}_4[X]$ be such that $ab \in Q$. Assume that $b^n \notin Q$ for any positive integer n, i.e. $b \notin \sqrt{Q}$. Since $\sqrt{Q}$ is maximal, the ideal $(b, \sqrt{Q})$ generated by b and $\sqrt{Q}$ is $\mathbb{Z}_4[X]$. Then the identity 1 can be written as $1 = xb + r$, where $x \in \mathbb{Z}_4[X]$ and $r \in \sqrt{Q}$. There is a positive integer n such that $r^n \in Q$. Then

$$1 = 1^n = (xb + r)^n = yb + r^n, \text{ where } y \in \mathbb{Z}_4[X].$$

Multiplying by a, we obtain

$$a = yab + ar^n \in Q.$$

Hence Q is primary. $\square$

Lemma 5.8. *Let $f(X)$ be a polynomial in $\mathbb{Z}_4[X]$ and assume that $\overline{f}(X) = g(X)^e$, where $g(X)$ is an irreducible polynomial in $\mathbb{Z}_2[X]$ and e is a positive integer. Then $f(X)$ is a primary polynomial in $\mathbb{Z}_4[X]$.*

Proof. Let $(f(X))$ be the principal ideal generated by $f(X)$. By Lemma 5.7, it is enough to prove that $\sqrt{(f(X))}$ is a prime ideal. Since $1 \notin (f(X))$, we have also $1 \notin \sqrt{(f(X))}$. Thus $\sqrt{(f(X))} \neq (1) = \mathbb{Z}_4[X]$. Let $a(X)$, $b(X) \in \mathbb{Z}_4[X]$ and $a(X)\,b(X) \in \sqrt{(f(X))}$. Then there is a positive integer n such that $(a(X)\,b(X))^n \in (f(X))$. It follows that $(\overline{a}(X)\,\overline{b}(X))^n \in (\overline{f}(X)) = (g(X)^e)$. By the unique factorization theorem of $\mathbb{Z}_2[X]$, $g(X)\,|\,\overline{a}(X)$ or $g(X)\,|\,\overline{b}(X)$. If $g(X)\,|\,\overline{a}(X)$, then $\overline{f}(X)|\,\overline{a}(X)^e$. There are polynomials $c(X)$, $d(X) \in \mathbb{Z}_4[X]$ such that $a(X)^e = c(X)f(X) + 2d(X)$. Then $a(X)^{2e} = c(X)^2\, f(X)^2 \in (f(X))$ and, consequently, $a(X) \in \sqrt{(f(X))}$. If $g(X)\,|\,\overline{b}(X)$, then we can prove in a similar way that $b(X) \in \sqrt{(f(X))}$. Therefore $\sqrt{(f(X))}$ is prime. $\square$

Corollary 5.9. *Any basic irreducible polynomial in $\mathbb{Z}_4[X]$ is primary.* $\square$

5.4. Factorization of Monic Polynomials in $\mathbb{Z}_4[X]$

Theorem 5.10. *Let $f(X)$ be a monic polynomial of degree ≥ 1 in $\mathbb{Z}_4[X]$. Then*

(i) $f(X) = g_1(X) \cdots g_r(X)$, *where* $g_1(X), \ldots, g_r(X)$ *are pairwise coprime monic primary polynomials.*

(ii) *Let*

$$f(X) = g_1(X) \cdots g_r(X) = h_1(X) \cdots h_s(X) \tag{5.6}$$

be two factorization of $f(X)$ *into pairwise coprime monic primary polynomials, then* $r = s$ *and after renumbering,* $g_i(X) = h_i(X), i = 1, \ldots, r$.

Proof. (i) By the unique factorization theorem of polynomials in $\mathbb{Z}_2[X]$ we can assume that

$$\overline{f}(X) = f_1(X)^{e_1} \cdots f_r(X)^{e_r},$$

where $f_1(X), \ldots, f_r(X)$ are distinct irreducible polynomials in $\mathbb{Z}_2[X]$ and e_1, $\ldots, e_r$ are positive integers. By Lemma 5.3 there exist pairwise coprime monic polynomials $g_1(X), \ldots, g_r(X) \in \mathbb{Z}_4[X]$ such that

$$f(X) = g_1(X) \cdots g_r(X)$$

and

$$\overline{g}_i(X) = f_i(X)^{e_1}, \quad i = 1, \ldots, r.$$

By Lemma 5.8, all $g_i(X)$, $i = 1, \ldots, r$, are primary polynomials.

(ii) From $g_1(X) \cdots g_r(X) = h_1(X) \cdots h_s(X)$ we deduce that $g_1(X) \cdots g_r(X) \in (h_i(X))$ for all $i = 1, \ldots, s$. Since $(h_i(X))$ is primary, there is an integer k_i, $1 \leq k_i \leq r$ and a positive integer n_i such that $g_{k_i}(X)^{n_i} \in (h_i(X))$.

We assert that k_i is uniquely determined by $h_i(X)$. Assume that there is another k_i' and an n_i' such that $g_{k_i'}(X)^{n_i'} \in (h_i(X))$. Since $g_{k_i}(X)$ and $g_{k_i'}(X)$ are coprime, there are polynomials $a(X)$, $b(X) \in \mathbb{Z}_4[X]$ such that

$$1 = a(X)\, g_{k_i}(X) + b(X)\, g_{k_i'}(X).$$

Then

$$1 = 1^{n_i + n_i' - 1} = (a(X)\, g_{k_i}(X) + b(X)\, g_{k_i'}(X))^{n_i + n_i' - 1} \in (h_i(X)),$$

which is a contradiction. Our assertion is proved.

Similarly, for all $j = 1, \ldots, r$, there is a uniquely determined integer l_j, $1 \leq l_j \leq s$, and a positive integer m_j such that $h_{l_j}(X)^{m_j} \in (g_j(X))$. Then for any i, $1 \leq i \leq s$, we have

$$h_{l_{k_i}}(X)^{m_{k_i} n_i} \in (h_i(X)).$$

Since $h_i(X)$ and $h_j(X)$ are coprime for $i \neq j$, we must have $l_{k_i} = i$ for every $i = 1, \ldots, s$. It follows that the map

$$\{1, 2, \ldots, s\} \to \{1, 2, \ldots, r\}$$
$$i \mapsto k_i$$

is a well-defined injective map. Thus $r \geq s$. Similarly, $s \geq r$. Hence $r = s$. After renumbering, we can assume that $k_i = i$ for $i = 1, \ldots, r$. Then $l_i = i$ for $i = 1, 2, \ldots, r$. Thus $g_i(X)^{n_i} \in (h_i(X))$ and $h_i(X)^{m_i} \in (g_i(X))$ for $i = 1, 2, \ldots, r$.

For $j \neq 1$, $g_j(X)$ and $g_1(X)$ are coprime. By Lemma 5.1, $\overline{g}_j(X)$ and $\overline{g}_1(X)$ are coprime, which implies that $\overline{g}_j(X)$ and $\overline{g}_1(X)^{n_1}$ are coprime. Hence $\overline{g}_2(X) \cdots \overline{g}_r(X)$ and $\overline{g}_1(X)^{n_1}$ are coprime. By Lemma 5.1 again, $g_2(X) \cdots g_r(X)$ and $g_1(X)^{n_1}$ are coprime. Since $g_1(X)^{n_1} \in (h_1(X))$, $g_2(X) \cdots g_r(X)$ and $h_1(X)$ are coprime, i.e. there are polynomials $c(X), d(X) \in \mathbb{Z}_4[X]$ such that

$$c(X)\, g_2(X) \cdots g_r(X) + d(X)\, h_1(X) = 1.$$

Multiplying by $g_1(X)$, we obtain

$$c(X)\, g_1(X)\, g_2(X) \cdots g_r(X) + d(X)\, g_1(X)\, h_1(X) = g_1(X).$$

By (5.6), we have

$$c(X)\, h_1(X)\, h_2(X) \cdots h_r(X) + d(X)\, g_1(X)\, h_1(X) = g_1(X),$$

which implies $h_1(X) \mid g_1(X)$. Similarly, $g_1(X) \mid h_1(X)$. Since both $g_1(X)$ and $h_1(X)$ are monic polynomials, we must have $g_1(X) = h_1(X)$. Similarly, $g_i(X) = h_i(X)$, $i = 2, 3, \ldots, r$. □

From Theorem 5.10 we deduce

Proposition 5.11. *Let n be a positive odd integer. Then the polynomial $X^n - 1$ over $\mathbb{Z}_4$ can be factored into a product of finitely many pairwise coprime basic irreducible polynomials over $\mathbb{Z}_4$, say*

$$X^n - 1 = g_1(X)\; g_2(X) \cdots g_r(X). \tag{5.7}$$

Moreover, $g_1(X), g_2(X), \ldots, g_r(X)$ are uniquely determined up to a rearrangement.

Proof. Over $\mathbb{Z}_2$, we have the unique factorization

$$X^n - 1 = f_2^{(1)}(X)\; f_2^{(2)}(X) \cdots f_2^{(r)}(X),$$

where $f_2^{(1)}(X),\ f_2^{(2)}(X), \ldots, f_2^{(r)}(X)$ are irreducible polynomials over $\mathbb{Z}_2$. Since n is odd, $f_2^{(1)}(X),\ f_2^{(2)}(X), \ldots, f_2^{(r)}(X)$ are pairwise coprime. By Hensel's lemma, there are monic polynomials $g_1(X),\ g_2(X), \ldots, g_r(X)$ over $\mathbb{Z}_4$ such that $\overline{g}_i(X) = f_2^{(i)}(X)$ and deg $g_i(X) =$ deg $f_2^{(i)}(X)$ for $i = 1, 2, \ldots, r$, that $g_1(X),\ g_2(X), \ldots, g_r(X)$ are pairwise coprime, and that

$$x^n - 1 = g_1(X)\, g_2(X) \cdots g_r(X).$$

Since $\overline{g}_i(X) = f_2^{(i)}(X)$, $i = 1, 2, \ldots, r$, are irreducible over $\mathbb{Z}_2$, $g_1(X),\ g_2(X), \ldots, g_r(X)$ are basic irreducible. By Corollary 5.9, $g_i(X)$, $i = 1, 2, \ldots, r$, are primary. Then the uniqueness of (5.7) follows from Theorem 5.10. □

5.5. Hensel Lift

Proposition 5.4 can be generalized and strengthened as follows.

Proposition 5.12. *Let n be an odd positive integer and $f_2(X)$ be a polynomial in $\mathbb{Z}_2[X]$ dividing $X^n - 1$. Then there exists a unique monic polynomial $f(X)$ in $\mathbb{Z}_4[X]$ dividing $X^n - 1$ and $\overline{f}(X) = f_2(X)$.*

Proof. By Proposition 5.11 we have (5.7)

$$X^n - 1 = g_1(X)\, g_2(X) \cdots g_r(X)$$

over $\mathbb{Z}_4$, where $g_1(X),\ g_2(X), \ldots, g_r(X)$ are pairwise coprime basic irreducible polynomials over $\mathbb{Z}_4$. Then

$$X^n - 1 = \overline{g}_1(X)\, \overline{g}_2(X) \cdots \overline{g}_r(X)$$

over $\mathbb{Z}_2$, where $\overline{g}_1(X),\ \overline{g}_2(X), \cdots, \overline{g}_r(X)$ are distinct irreducible polynomials over $\mathbb{Z}_2$. By the unique factorization theorem in $\mathbb{Z}_2[X]$, we can assume that up to a rearrangement

$$f_2(X) = \overline{g}_1(X)\, \overline{g}_2(X) \cdots \overline{g}_s(X), \qquad \text{where} \quad 1 \le s \le r. \tag{5.8}$$

Let

$$f(X) = g_1(X)\, g_2(X) \cdots g_s(X),$$

then $f(X)$ is a monic polynomial over $\mathbb{Z}_4$ dividing $X^n - 1$ and $\overline{f}(X) = f_2(X)$.

Now let us come to the proof of the uniqueness of $f(X)$. Assume that $h(X)$ is any monic polynomial in $\mathbb{Z}_4[X]$ dividing $X^n - 1$ and $\overline{h}(X) = f_2(X)$. From the factorization (5.8) of $f_2(X)$ in $\mathbb{Z}_2[X]$ and by Hensel's lemma, we have pairwise coprime basic irreducible polynomials, $h_1(X)$, $h_2(X), \ldots, h_s(X)$ over $\mathbb{Z}_4$ such that $\overline{h}_i(X) = \overline{g}_i(X)$, $i = 1, 2, \ldots, s$, and

$$h(X) = h_1(X)h_2(X) \cdots h_s(X).$$

Since $h(X) \mid (X^n - 1)$ in $\mathbb{Z}_4[X]$, by Proposition 5.11 all $h_1(X), h_2(X), \ldots, h_s(X)$ appear in $\{g_1(X), g_2(X), \ldots, g_r(X)\}$. Since $\overline{h}_i(X) = \overline{g}_i(X)$, $i = 1, 2, \ldots, s$ and $\overline{g}_1(X), \overline{g}_2(X), \ldots, \overline{g}_r(X)$ are distinct, we must have $h_i(X) = g_i(X)$, $i = 1, 2, \ldots, r$. Consequently, $h(X) = g(X)$. □

Proposition 5.13. *Let n_1 and n_2 be odd positive integers and $f_2(X)$ be a polynomial in $\mathbb{Z}_2[X]$ dividing both $X^{n_1} - 1$ and $X^{n_2} - 1$. Let $f^{(1)}(X)$ and $f^{(2)}(X)$ be monic polynomials in $\mathbb{Z}_4[X]$ dividing $X^{n_1} - 1$ and $X^{n_2} - 1$, respectively, and $\overline{f}^{(1)}(X) = \overline{f}^{(2)}(X) = f_2(X)$. Then $f^{(1)}(X) = f^{(2)}(X)$.*

Proof. Let $n = (n_1, n_2)$, then n is also odd, $X^n - 1 = (X^{n_1} - 1, X^{n_2} - 1)$, and $f_2(X) \mid (X^n - 1)$. By Proposition 5.12 there is a unique monic polynomial $f(X)$ in $\mathbb{Z}_4[X]$ dividing $X^n - 1$ and $\overline{f}(X) = f_2(X)$. Since $X^n - 1$ divides $X^{n_1} - 1$, $f(X)$ also divides $X^{n_1} - 1$. By the uniqueness part of Proposition 5.12, $f(X) = f^{(1)}(X)$. Similarly, $f(X) = f^{(2)}(X)$. Therefore $f^{(1)}(X) = f^{(2)}(X)$. □

Corollary 5.14. *Let n be an odd positive integer and $f_2(X)$ be an irreducible polynomial in $\mathbb{Z}_2[X]$ dividing $X^n - 1$. Then there exists a unique basic irreducible polynomial $f(X)$ in $\mathbb{Z}_4[X]$ dividing $X^n - 1$ and $\overline{f}(X) = f_2(X)$. Moreover, $f(X)$ is independent of n.* □

Let $f_2(X)$ be a polynomial over $\mathbb{Z}_2$ without multiple roots and not divisible by X. It is well known that there is a positive odd integer n such that $f_2(X)$ divides $X^n - 1$, (see Wan (1992), Definition 7.2 and Theorem 7.8). For example, if $f_2(X)$ is an irreducible polynomial of degree m, then $n = 2^m - 1$ satisfies $f_2(X) \mid X^{2^m - 1} - 1$. By Proposition 5.12 there is a unique monic polynomial $f(X)$ over $\mathbb{Z}_4$ dividing $X^n - 1$ and $\overline{f}(X) = f_2(X)$. By Proposition 5.13, $f(X)$ is independent of the particular choice of n. This polynomial is called the *Hensel lift* of $f_2(X)$ and can be calculated by using Graeffe's method for finding a polynomial whose roots are the squares of the roots of $f_2(X)$, (see Uspensky (1948)), as the following proposition shows.

Proposition 5.15. *Let $f_2(X)$ be a polynomial over $\mathbb{Z}_2[X]$ without multiple roots and not divisible by X. Write $f_2(X) = e(X) - d(X)$, where $e(X)$ contains* *only even power terms and $d(X)$ only odd power terms. Then $e(X)^2 - d(X)^2$, computed in $\mathbb{Z}_4[X]$, is a polynomial having only even power terms and of degree $2\deg f_2(X)$. Let $f(X^2) = \pm(e(X)^2 - d(X)^2)$, where we take the $+$ or $-$ sign if $\deg e(X) > \deg d(X)$ or $\deg d(X) > \deg e(X)$, then $f(X)$ is the Hensel lift of $f_2(X)$.*

Proof. The first statement is clear. By the choice of $\pm$ sign, $f(X^2)$ is monic and, hence $f(X)$ is monic. We have

$$f(X^2) \equiv e(X^2) - d(X^2) = f_2(X^2) \pmod 2,$$

which implies $\overline{f}(X) = f_2(X)$. We also have

$$f(X^2) = \pm f_2(X)\, f_2(-X),$$

computed in $\mathbb{Z}_4[X]$. There is an odd positive integer n such that $f_2(X) \mid X^n - 1$ in $\mathbb{Z}_2[X]$. Computed in $\mathbb{Z}_4[X]$,

$$X^n - 1 = f_2(X)\, a(X) + 2b(X),$$

where $a(X)$, $b(X) \in \mathbb{Z}_4[X]$. Then

$$(-X)^n - 1 = f_2(-X)\, a(-X) + 2b(-X)$$

and

$$\begin{aligned} X^{2n} - 1 &= (X^n - 1)(X^n + 1) \\ &= -\, f_2(X)\, f_2(-X)\, a(X)\, a(-X) + 2\,[f_2(X)\, a(X)\, b(-X) \\ &\quad + f_2(-X)\, a(-X)\, b(X)]. \end{aligned}$$

Writing $f_2(X) = e(X) - d(X)$, $a(X) = e_a(X) - d_a(X)$, and $b(X) = e_b(X) - d_b(X)$, where $e(X)$, $e_a(X)$, $e_b(X)$ contain only even power terms and $d(X)$, $d_a(X)$, $d_b(X)$ only odd power terms, we can verify easily that

$$2\,[f_2(X)\, a(X)\, b(-X) + f_2(-X)\, a(-X)\, b(X)] = 0.$$

Therefore $f(X^2) \mid X^{2n} - 1$ in $\mathbb{Z}_4[X]$. Hence $f(X) \mid X^n - 1$ in $\mathbb{Z}_4[X]$. We conclude that $f(X)$ is the Hensel lift of $f_2(X)$. □

Example 5.1. Let $m = 2$ and $h_2(X) = X^2 + X + 1 = e(X) - d(X)$, where $e(X) = X^2 + 1$ and $d(X) = -X$. Then

$$e(X)^2 - d(X)^2 = X^4 + X^2 + 1.$$

Hence $h(X) = X^2 + X + 1$ is the Hensel lift of $X^2 + X + 1$. □

Example 5.2. Let $m = 3$ and $h_2(X) = X^3 + X + 1 = e(X) - d(X)$, where $e(X) = 1$ and $d(X) = -X^3 - X$. We have

$$-e(X)^2 + d(X)^2 = X^6 + 2X^4 + X^2 - 1.$$

Then $h(X) = X^3 + 2X^2 + X - 1$ is the Hensel lift of $X^3 + X + 1$. $\square$

Finally, the following example shows that not every monic polynomial $h(X) \in \mathbb{Z}_4[X]$ with the property that $\overline{h}(X)$ is irreducible over $\mathbb{Z}_2$ is the Hensel lift of $\overline{h}(X)$.

Example 5.3. Let $h(X) = X - 3 \in \mathbb{Z}_4[X]$. $h(X)$ is monic and $\overline{h}(X) = X + 1$ is irreducible over $\mathbb{Z}_2$. Clearly, $h(X) \nmid (X^n - 1)$ for any odd positive integer n. Therefore $h(X)$ is not the Hensel lift of $\overline{h}(X)$. $\square$

CHAPTER 6

GALOIS RINGS

This chapter introduces the main machinery, Galois rings, for the study of $\mathbb{Z}_4$-codes. The theory of Galois rings was developed by Krull in the twenties of this century, see Krull (1924). We do not intend to introduce general Galois rings but only the Galois ring $\mathrm{GR}(4^m)$ with 4^m elements instead. Extending to the general Galois rings is immediate. In preparing this chapter, Nechaev (1989) is helpful.

6.1. The Galois Ring $\mathrm{GR}(4^m)$

We recall that a basic irreducible polynomial $h(X)$ of degree m over $\mathbb{Z}_4$ is a monic polynomial of degree m over $\mathbb{Z}_4$ such that $\overline{h}(X)$ is irreducible over $\mathbb{Z}_2$ and that if $\overline{h}(X)$ is primitive, then $h(X)$ is called basic primitive over $\mathbb{Z}_4$.

For any given positive integer m the existence of a basic irreducible polynomial and a basic primitive polynomial of degree m over $\mathbb{Z}_4$ are guaranteed by Proposition 5.4 and Corollary 5.5, respectively. Let $h(X)$ be a basic irreducible polynomial of degree m over $\mathbb{Z}_4$. Consider the residue class ring

$$\mathbb{Z}_4[X]/(h(X)).$$

The residue classes

$$a_0 + a_1X + \cdots + a_{m-1}X^{m-1} + (h(X)),$$

where $a_0, a_1, \ldots, a_{m-1} \in \mathbb{Z}_4$, are all the distinct elements of $\mathbb{Z}[X]/(h(X))$. Therefore $|\,\mathbb{Z}_4[X]/(h(X))\,| = 4^m$. The ring $\mathbb{Z}_4[X]/(h(X))$ is called the *Galois ring* with 4^m elements and is denoted by $\mathrm{GR}(4^m)$.

Write $\xi = X + (h(X))$, then $h(\xi) = 0$, i.e., ξ is a root of $h(X)$, and the elements

$$a_0 + a_1\xi + \cdots + a_{m-1}\,\xi^{m-1},$$

where $a_0, a_1, \ldots, a_{m-1}$ runs through $\mathbb{Z}_4$ independently, exhaust all the distinct elements of $\mathrm{GR}(4^m)$. Therefore $\mathrm{GR}(4^m) = \mathbb{Z}_4[\xi]$.

For a commutative ring with identity 1 the order of 1 in the additive group of the ring is called the *characteristic* of the ring. Then $\mathrm{GR}(4^m)$ is of characteristic 4. We know that the kernel of the ring homomorphism

$$
\begin{aligned}
- : \mathbb{Z}_4[X] &\to \mathbb{Z}_2[X] \\
a_0 + a_1 X + \cdots + a_n X^n &\mapsto \overline{a}_0 + \overline{a}_1 X + \cdots + \overline{a}_n X^n
\end{aligned} \tag{6.1}
$$

is the ideal (2) and the image of $(h(X))$ under $-$ is $(\overline{h}(X))$. Therefore the ring homomorphism (6.1) induces a ring homomorphism

$$
\begin{aligned}
\mathbb{Z}_4[X]/(h(X)) &\to \mathbb{Z}_2[X]/(\overline{h}(X)) \\
a_0 + a_1 X + \cdots + a_{m-1} X^{m-1} + (h(X)) &\mapsto \overline{a}_0 + \overline{a}_1 X + \cdots \overline{a}_{m-1} X^{m-1} + (\overline{h}(X)),
\end{aligned} \tag{6.2}
$$

which will also be denoted by $-$. Denote the image of $\xi = X + (h(X))$ by $\overline{\xi}$, then $\overline{\xi} = X + (\overline{h}(X))$, $\overline{\xi}$ is a root of $\overline{h}(X)$,

$$\mathbb{Z}_2[X]/(\overline{h}(X)) = \mathbb{Z}_2[\overline{\xi}],$$

and (6.2) can be written as

$$
\begin{aligned}
- : \mathbb{Z}_4[\xi] &\to \mathbb{Z}_2[\overline{\xi}] \\
a_0 + a_1\, \xi + \cdots + a_{m-1}\, \xi^{m-1} &\mapsto \overline{a}_0 + \overline{a}_1\, \overline{\xi} + \cdots + \overline{a}_{m-1}\, \overline{\xi}^{m-1}
\end{aligned} \tag{6.3}
$$

Obviously, the following diagram is commutative.

$$
\begin{array}{ccc}
\mathbb{Z}_4[X] & \xrightarrow{-} & \mathbb{Z}_2[X] \\
\downarrow & & \downarrow \\
\mathbb{Z}_4[\xi] & \xrightarrow{-} & \mathbb{Z}_2[\overline{\xi}].
\end{array}
$$

Since $h(X)$ is assumed to be basic irreducible, $\overline{h}(X)$ is irreducible over $\mathbb{Z}_2$ and $\mathbb{Z}_2[\overline{\xi}]$ is the Galois field $\mathbb{F}_{2^m}$. Clearly, the kernel of (6.3) is the ideal (2), (2) is a maximal ideal of $\mathbb{Z}_4[\xi]$, and (2) consists of all the zero divisors of $\mathbb{Z}_4[\xi]$ together with the zero element 0. Since in a finite ring any nonzero element which is not a zero divisor is invertible, (2) is the unique maximal ideal of $\mathbb{Z}_4[\xi]$. We summarize the foregoing discussion into the following theorem.

Theorem 6.1. *Let $h(X)$ be a basic irreducible polynomial of degree m over $\mathbb{Z}_4$. Then the residue class ring $\mathrm{GR}(4^m) = \mathbb{Z}_4[X]/(h(X))$ is a finite ring of* *characteristic* 4 *with* 4^m *elements. Write* $\xi = X + (h(X))$, *then* $h(\xi) = 0$, *every element of* $\mathrm{GR}(4^m)$ *can be written uniquely in the following form*

$$a_0 + a_1\,\xi + \cdots + a_{m-1}\,\xi^{m-1}, \quad a_i \in \mathbb{Z}_4 \quad (0 \le i \le m-1) \tag{6.4}$$

and $\mathrm{GR}(4^m) = \mathbb{Z}_4[\xi]$. *Moreover, the ideal* (2) *of* $\mathbb{Z}_4[\xi]$ *is the unique maximal ideal which consists of all the zero divisors together with the zero element* 0. *Write* $\bar{\xi} = X + (\bar{h}(X))$, *then* $\bar{h}(\bar{\xi}) = 0$ *and* $\mathbb{Z}_4[\xi]/(2) \simeq \mathbb{Z}_2[\bar{\xi}]$ *is the Galois field* $\mathbb{F}_{2^m}$. □

The representation (6.4) is called the *additive representation* of the elements of the Galois ring $\mathrm{GR}(4^m) = \mathbb{Z}_4[X]/(h(X))$.

In general, a *Galois ring* is defined to be a finite commutative ring R with identity 1 such that the set of zero divisors of R with 0 added is a principal ideal (p) for some prime number p.

Proposition 6.2. *Let* R *be a Galois ring whose zero divisors together with* 0 *form a principal ideal* (p) *for some prime* p. *Then* (p) *is the only maximal ideal of* R, $R/(p)$ *is a Galois field* $\mathbb{F}_{p^m}$ *for some positive integer* m, *and the characteristic of* R *is a power of* p.

Proof. In a finite ring any nonzero element which is not a zero divisor is invertible. Therefore (p) is the only maximal ideal of R and $R/(p)$ is a finite field. Denote the natural homomorphism $R \to R/(p)$ by $^-$ and the image of $r \in R$ by $\bar{r}$. Let n be any positive integer and $a \in R$ or $R/(p)$, denote

$$\underbrace{a + a + \cdots + a}_{n}$$

by na. Then $p\bar{1} = \overline{p1} = 0$. Therefore $R/(p)$ is of characteristic p and $R/(p) \simeq \mathbb{F}_{p^m}$ for some positive integer m.

Let k be the characteristic of R. From $k1 = 0$ we deduce $k\bar{1} = \overline{k1} = 0$. Therefore $p\,|\,k$. Assume that $k = p^n l$, where n, l are positive integers and $(p, l) = 1$. If $l > 1$, then $a = p^n 1$ and $b = l\,1$ are nonzero elements of R and $ab = 0$. It follows that $l1 \in (p)$ and $l\bar{1} = \overline{l1} = 0$ in $R/(p)$. But $R/(p)$ is of characteristic p, so $p\,|\,l$, which contradicts $(p, l) = 1$. Therefore $l = 1$ and $k = p^n$ □

Proposition 6.3. *Let* R *be a Galois ring of characteristic* 4. *Then the set of zero divisors of* R *with* 0 *added is the principal ideal* (2), (2) *is the only* *maximal ideal of R, $R/(2) \simeq \mathbb{F}_{2^m}$ and $|R| = 4^m$ for some positive integer m.*

Proof. Since R is of characteristic 4, by Proposition 6.2 the set of zero divisors of R with 0 added is the principal ideal (2), (2) is the only maximal ideal of R, and $R/(2) \simeq \mathbb{F}_{2^m}$ for some positive integer m. Consider the map

$$\begin{aligned} R &\to (2) \\ r &\mapsto 2r. \end{aligned}$$

Clearly, this is a well-defined homomorphism from the additive group of R to that of (2), it is surjective, and its kernel includes (2). It is easy to see that the kernel is an ideal of R and 1 does not belong to the kernel. Since (2) is a maximal ideal of R, the kernel must be (2). By the fundamental theorem of homomorphism we have the additive group isomorphism

$$R/(2) \simeq (2).$$

It follows that $|(2)| = |R/(2)| = |\mathbb{F}_{2^m}| = 2^m$. Hence $|R| = |R/(2)|\,|(2)| = 4^m$. $\square$

Lemma 6.4. *Let R be a Galois ring of characteristic 4, $R/(2) \cong \mathbb{F}_{2^m}$ and $|R| = 4^m$ for some positive integer m. Let $f(X)$ be a polynomial over $\mathbb{Z}_4$ and assume that $\overline{f}(X)$ has a root $\overline{\beta}$ in $\mathbb{F}_{2^m}$ and $\overline{f}\,'(\overline{\beta}) \neq 0$. Then there exists a unique root $\alpha \in R$ of the polynomial $f(X)$ such that $\overline{\alpha} = \overline{\beta}$.*

Proof. Let $\overline{\beta} = \beta + (2)$, where $\beta \in R$. Since $\overline{f}\,'(\overline{\beta}) \neq 0$, $f'(\beta)$ is an invertible element of R. Let $\alpha = \beta - f'(\beta)^{-1} f(\beta) \in R$, then by Taylor's formula

$$f(\alpha) = f(\beta) + \frac{f'(\beta)}{1!}\,(-f'(\beta)^{-1} f(\beta)) + \frac{f''(\beta)}{2!}\,(-f'(\beta)^{-1} f(\beta))^2 + \cdots .$$

Since $\overline{f}(\overline{\beta}) = 0$, $f(\beta) \in (2)$ and $f(\beta)^2 = f(\beta)^3 = \cdots = 0$. Therefore $f(\alpha) = 0$ and $\overline{\alpha} = \alpha + (2) = \beta + (2) = \overline{\beta}$.

Let α' be any root of $f(X)$ such that $\overline{\alpha'} = \overline{\beta}$. Then $\overline{\alpha'} = \overline{\alpha}$ and $\alpha' = \alpha + 2\gamma$, where γ is an element of R. By Taylor's formula,

$$f(\alpha') = f(\alpha) + \frac{f'(\alpha)}{1!}\,(2\gamma) + \frac{f''(\alpha)}{2!}\,(2\gamma)^2 + \frac{f'''(\alpha)}{3!}\,(2\gamma)^3 + \cdots .$$

Since $f(\alpha) = f(\alpha') = 0$ and $\frac{f''(\alpha)}{2!}(2\gamma)^2 = \frac{f'''(\alpha)}{3!}(2\gamma)^3 = \cdots = 0$, we have $f'(\alpha)(2\gamma) = 0$. But $\overline{f'(\alpha)} = \overline{f'}(\overline{\alpha}) = \overline{f}'(\overline{\beta}) \neq 0$, so $f'(\alpha)$ is an invertible element of R which implies that $2\gamma = 0$. Therefore $\alpha' = \alpha$. □

Theorem 6.5. *Let R be a Galois ring of characteristic 4, $R/(2) \simeq \mathbb{F}_{2^m}$, and $|R| = 4^m$ for some positive integer m. Then R is ring isomorphic to $\mathbb{Z}_4[X]/(h(X))$ for any basic irreducible polynomial $h(X)$ of degree m over $\mathbb{Z}_4$.*

Proof. Let $h(X)$ be any basic irreducible polynomial of degree m over $\mathbb{Z}_4$. Then $\overline{h}(X)$ is irreducible over $\mathbb{Z}_2$ and $\deg \overline{h}(X) = m$. $\overline{h}(X)$ has a root in $R/(2) \simeq \mathbb{F}_{2^m}$, let it be $\overline{\beta}$. Then $\overline{h}(\overline{\beta}) = 0$. Since $\overline{h}(X)$ is irreducible, $\overline{h}(X)$ has no multiple root. Therefore $\overline{h}'(\overline{\beta}) \neq 0$. By Lemma 6.4 there exists a unique root $\alpha \in R$ of the polynomial $h(X)$ such that $\overline{\alpha} = \overline{\beta}$. Consider the map

$$\mathbb{Z}_4[X]/(h(X)) \to R$$

$$a_0 + a_1X + \cdots + a_{m-1}X^{m-1} + (h(X)) \mapsto a_0 + a_1\alpha + \cdots + a_{m-1}\alpha^{m-1}, \tag{6.5}$$

where $a_0, a_1, \ldots, a_{m-1} \in \mathbb{Z}_4$. Clearly, it is a well-defined ring homomorphism. Let us prove that it is injective. Assume that $a_0 + a_1\alpha + \cdots + a_{m-1}\alpha^{m-1} = 0$, then $\overline{a}_0 + \overline{a}_1\overline{\alpha} + \cdots + \overline{a}_{m-1}\,\overline{\alpha}^{m-1} = 0$. But $\overline{\alpha} = \overline{\beta}$ is a root of the irreducible polynomial $\overline{h}(X)$ of degree m over $\mathbb{Z}_2$, so $\overline{a}_0 = \overline{a}_1 = \cdots = \overline{a}_{m-1} = 0$, then we may write $a_i = 2b_i$, where $b_i = 0$ or 1 $(i = 0, 1, \ldots, m-1)$. Thus $2(b_0 + b_1\alpha + \cdots + b_{m-1}\alpha^{m-1}) = a_0 + a_1\alpha + \cdots + a_{m-1}\alpha^{m-1} = 0$, and $b_0 + b_1\alpha + \cdots + b_{m-1}\alpha^{m-1}$ is either a zero divisor or 0. That is, $b_0 + b_1\alpha + \cdots + b_{m-1}\alpha^{m-1} \in (2)$. Then $\overline{b}_0 + \overline{b}_1\overline{\alpha} + \cdots + \overline{b}_{m-1}\overline{\alpha}^{m-1} = 0$. Since $\overline{\alpha} = \overline{\beta}$ is a root of the irreducible polynomial $\overline{h}(X)$ of degree m over $\mathbb{Z}_2$, we have $\overline{b}_0 = \overline{b}_1 = \cdots = \overline{b}_{m-1} = 0$. Then $b_i = 2c_i$ where $c_i = 0$ or 1 $(i = 0, 1, \ldots, m-1)$. It follows that $a_i = 2b_i = 4c_i = 0\,(i = 0, 1, \ldots, m-1)$. Therefore the map (6.5) is injective. By Theorem 6.1 $|\mathbb{Z}_4[X]/(h(X))| = 4^m$ and by hypothesis $|R| = 4^m$. Therefore the map (6.5) is also surjective. Hence $R \simeq \mathbb{Z}_4[X]/(h(X))$. □

Corollary 6.6. *Any two Galois rings both of characteristic 4 and having the same number of elements are isomorphic.* □

This corollary justifies the notation $\mathrm{GR}(4^m)$.

6.2. The 2-Adic Representation

Theorem 6.7. (i) *In the Galois ring $\mathrm{GR}(4^m)$ there exists a nonzero element ξ of order $2^m - 1$, which is a root of a basic primitive polynomial $h(X)$ of* *degrees m over $\mathbb{Z}_4$ and* $\mathrm{GR}(4^m) = \mathbb{Z}_4[\xi]$. *Moreover, $h(X)$ is the unique monic polynomial of degree $\leq m$ over $\mathbb{Z}_4$ having ξ as a root.*

(ii) *Let $\mathcal{T} = \{0, 1, \xi, \ldots, \xi^{2^m-2}\}$, then any element $c \in \mathrm{GR}(4^m)$ can be written uniquely as*

$$c = a + 2b, \tag{6.6}$$

where $a, b \in \mathcal{T}$.

Proof. (i) By Proposition 6.3, $\mathrm{GR}(4^m)/(2) \simeq \mathbb{F}_{2^m}$. Let ξ_2 be a primitive element of $\mathbb{F}_{2^m}$, then $\xi_2^{2^m-1} = 1$ and $\xi_2^i \neq 1$ for $0 < i < 2^m - 1$. By Lemma 6.4 there exists a unique root $\xi \in \mathrm{GR}(4^m)$ of the polynomial $X^{2^m-1} - 1$ such that $\overline{\xi} = \xi_2$. Then $\xi^{2^m-1} = 1$. Since $\overline{\xi} = \xi_2$ is of order $2^m - 1$, ξ is also of order $2^m - 1$.

We know that the polynomial $X^{2^m} - 1$ can be factored into a product of distinct irreducible polynomials of degrees dividing m into $\mathbb{Z}_2[X]$, say

$$X^{2^m-1} - 1 = f_1(X)f_2(X)\cdots f_r(X).$$

We can assume that $f_1(X)$ is primitive of degree m over $\mathbb{Z}_2$ and $\overline{\xi}$ is a root of $f_1(X)$. Clearly $f_1'(\overline{\xi}) \neq 0$. By Hensel's lemma,

$$X^{2^m-1} - 1 = h_1(X)h_2(X)\cdots h_r(X) \quad \text{in} \quad \mathbb{Z}_4[X],$$

where $h_1(X), h_2(X), \ldots, h_r(X)$ are pairwise coprime monic polynomials and $\overline{h}_i(X) = f_i(X), i = 1, 2, \ldots, r$. Let $h(X) = h_1(X)$, then $h(X)$ is a basic primitive polynomial of degree m over $\mathbb{Z}_4$, $\overline{h}(\overline{\xi}) = \overline{h}_1(\overline{\xi}) = f_1(\overline{\xi}) = 0$ and $\overline{h}'(\overline{\xi}) = f_1'(\overline{\xi}) \neq 0$. By Lemma 6.4 the polynomial $h(X)$ has a unique root $\eta \in \mathrm{GR}(4^m)$ such that $\overline{\eta} = \overline{\xi}$. But η is also a root of $X^{2^m-1} - 1$. By the uniqueness of Lemma 6.4, $\eta = \xi$. Then $h(\xi) = 0$.

By the proof of Theorem 6.5, the map

$$\mathbb{Z}_4[X]/(h(X)) \to \mathrm{GR}(4^m)$$

$$a_0 + a_1X + \cdots + a_{m-1}X^{m-1} + (h(X)) \to a_0 + a_1\xi + \cdots + a_{m-1}\xi^{m-1} \tag{6.7}$$

is a ring isomorphism and $\mathrm{GR}(4^m) = \mathbb{Z}_4[\xi]$.

Let $g(X)$ be any monic polynomial of degree $\leq m$ over $\mathbb{Z}_4$ and assume that $g(\xi) = 0$. Let $f(X) = g(X) - h(X)$, then $\deg f(X) \leq m$ and $f(\xi) = 0$. Since (6.7) is a ring isomorphism, $f(X) \in (h(X))$. Thus $g(X) \in (h(X))$. Since both $h(X)$ and $g(X)$ are monic, $\deg g(X) \leq m$ and $\deg h(X) = m$, we must have $g(X) = h(X)$.

(ii) We know that $|\mathrm{GR}(4^m)| = 4^m$. If we can show that all the 4^m elements of the form (6.6) are distinct, then (ii) will be proved. Assume that

$$a + 2b = a' + 2b',$$

where $a, b, a', b' \in \mathcal{T}$. Mod 2, we obtain $\bar{a} = \bar{a}'$. Since both ξ and $\bar{\xi} = \xi_2$ are of order $2^m - 1$, the map $\xi^i \to \bar{\xi}^i (i = 0, 1, \ldots, 2^m - 2)$ is bijective. Therefore $a = a'$. It follows that $2b = 2b'$. If $b = 0$ and $b' = \xi^i\, (0 \le i \le 2^m - 2)$, then from $0 = 2\xi^i$ we deduce $0 = 0 \cdot \xi^{2^m-1-i} = 2\xi^i \cdot \xi^{2^m-1-i} = 2$, which contradicts $0 \ne 2$ in $\mathbb{Z}_4$. Therefore $b = 0$ if and only if $b' = 0$. Now assume that $b = \xi^i$ and $b' = \xi^{i'} (0 \le i, i' \le 2^m - 2)$. If $i \ne i'$, without loss of generality we can assume that $i > i'$, then $2\xi^{i-i'} = 2$. It follows that $\xi^{i-i'} - 1$ is a zero divisor or 0. Therefore $\xi^{i-i'} - 1 \in (2)$. Then $\bar{\xi}^{i-i'} = 1$, which contradicts that $\bar{\xi}$ is of order $2^m - 1$. □

The representation (6.6) is called the *2-adic representation* of the element $c \in \mathrm{GR}(4^m)$, which is a generalization of the multiplicative representation of the elements of $\mathbb{F}_{2^m}$.

Corollary 6.8. *Express any element $c \in \mathrm{GR}(4^m)$ in the form* (6.6)

$$c = a + 2b, \quad \textit{where} \quad a, b \in \mathcal{T}.$$

Then

(i) *all the elements c with $a \ne 0$ are invertible and form a multiplicative group of order $(2^m - 1)2^m$, which is a direct product $\langle\xi\rangle \times \mathcal{E}$ where $\langle\xi\rangle$ is a cyclic group of order $2^m - 1$ generated by ξ and $\mathcal{E} = \{1 + 2b | b \in \mathcal{T}\}$ has the structure of an abelian group of type 2^m and is isomorphic to the additive group of $\mathbb{F}_{2^m}$.*

(ii) *All the elements c with $a = 0$ are nilpotent* (*and are zero divisors or* 0), *and they form the ideal* (2) *of* $\mathrm{GR}(4^m)$.

(iii) *The order of c is a divisor of $2^m - 1$ if and only if $a \ne 0$ and $b = 0$.*

(iv) *Any element $\eta \in \mathrm{GR}(4^m)$ of order $2^m - 1$ is of the form ξ^i, where $(i, 2^m - 1) = 1$ and is a root of a basic primitive polynomial of degree m over $\mathbb{Z}_4$ and $\mathcal{T} = \{0, 1, \eta, \eta^2, \ldots, \eta^{2^m-2}\}$.* □

Example 6.1. Let $m = 3$, $h(X) = X^3 + 2X^2 + X - 1$, and $\xi = X + (h(X))$. Then $\mathbb{Z}_4[\xi] = \mathrm{GR}(4^3)$ and ξ is an element of order $2^3 - 1 = 7$. We have

$$\xi^0 = 1,\ \xi^1 = \xi,\ \xi^2 = \xi^2,$$
$$\xi^3 = 2\xi^2 + 3\xi + 1,$$
$$\xi^4 = 3\xi^2 + 3\xi + 2,$$
$$\xi^5 = \xi^2 + 3\xi + 3,$$
$$\xi^6 = \xi^2 + 2\xi + 1.$$

Therefore

$$\mathcal{T} = \{0,\ 1,\ \xi,\ \xi^2, 2\xi^2 + 3\xi + 1,\ 3\xi^2 + 3\xi + 2,\ \xi^2 + 3\xi + 3,\ \xi^2 + 2\xi + 1\}. \quad \square$$

The following formulas for adding elements of $\mathcal{T}$ are useful, (see Helleseth and Kumar (1995)).

Corollary 6.9. *Let c_1, $c_2 \in \mathcal{T}$, and express*

$$c_1 + c_2 = a + 2b, \quad a, b \in \mathcal{T}, \tag{6.8}$$

then

$$a = c_1 + c_2 + 2(c_1c_2)^{1/2}, \tag{6.9}$$
$$b = (c_1c_2)^{1/2}, \tag{6.10}$$

where $(c_1c_2)^{1/2}$ denotes the unique element in $\mathcal{T}$ such that $((c_1c_2)^{1/2})^2 = c_1c_2$.

Proof. Squaring (6.8), we have

$$(c_1 + c_2)^2 = a^2$$

Thus

$$(c_1 + c_2)^{2^m} = a^{2^m} = a. \tag{6.11}$$

On the other hand,

$$\begin{aligned}(c_1 + c_2)^{2^m} &= (c_1^2 + c_2^2 + 2c_1c_2)^{2^{m-1}} \\ &= (c_1^{2^2} + c_2^{2^2} + 2c_1^2c_2^2)^{2^{m-2}} \\ &= c_1^{2^m} + c_2^{2^m} + 2c_1^{2^{m-1}}c_2^{2^{m-1}} \\ &= c_1 + c_2 + 2(c_1c_2)^{1/2}. \end{aligned} \tag{6.12}$$

From (6.11) and (6.12) we deduce (6.9) and then (6.10). $\square$

More generally, we have

Corollary 6.10. *Let* $c_1, c_2, \ldots, c_k \in \mathcal{T}$, *and express*

$$\sum_{i=1}^{k} c_i = a + 2b, \quad a, b \in \mathcal{T},$$

then

$$a = \sum_{i=1}^{k} c_i + 2 \sum_{1 \le i < j \le k} (c_i c_j)^{1/2}$$

$$b = \sum_{1 \le i < j \le k} (c_i c_j)^{1/2}.$$

Proof. By induction. □

6.3. Automorphisms of GR(4^m)

The Frobenius map of the Galois field $\mathbb{F}_{2^m}$

$$\begin{aligned} f_2 : \mathbb{F}_{2^m} &\to \mathbb{F}_{2^m} \\ a &\to a^2 \end{aligned}$$

can be generalized to GR(4^m) as follows:

$$\begin{aligned} f : \mathrm{GR}(4^m) &\to \mathrm{GR}(4^m) \\ c = a + 2b &\to c^f = a^2 + 2b^2 . \end{aligned}$$

f is called the *generalized Frobenius map* of GR(4^m).

Theorem 6.11. *The generalized Frobenius map f of* GR(4^m) *is a ring automorphism of* GR(4^m), *the fixed elements of f are the elements of* $\mathbb{Z}_4$, *and f is of order m.*

Proof. First we prove that f is a ring automorphism of GR(4^m). Clearly, f is injective. Since $(2, 2^m - 1) = 1$, every element of the cyclic group $\langle \xi \rangle$ can be written as a square element of $\langle \xi \rangle$. It follows that f is surjective.

Let $c, c' \in$ GR(4^m) and

$$c = a + 2b, \quad c' = a' + 2b'$$

be their 2-adic representations. Then

$$c + c' = a + a' + 2(b + b').$$

By Corollary 6.9, $a + a'$ has the 2-adic representation

$$a + a' = (a + a' + 2(aa')^{1/2}) + 2(aa')^{1/2},$$

where $a + a' + 2(aa')^{1/2}$, $(aa')^{1/2} \in \mathcal{T}$. Then

$$c + c' = (a + a' + 2(aa')^{1/2}) + 2(b + b' + (aa')^{1/2}).$$

Let the 2-adic representation of $b + b' + (aa')^{1/2}$ be

$$b + b' + (aa')^{1/2} = a_1 + 2b_1,$$

then

$$c + c' = (a + a' + 2(aa')^{1/2}) + 2a_1$$

is the 2-adic representation of $c + c'$. Therefore

$$\begin{aligned}(c + c')^f &= (a + a' + 2(aa')^{1/2})^2 + 2a_1^2 \\ &= (a^2 + a'^2 + 2aa') + 2(b^2 + b'^2 + aa') \\ &= a^2 + a'^2 + 2(b^2 + b'^2) \\ &= (a^2 + 2b^2) + (a'^2 + 2b'^2) \\ &= c^f + c'^f\end{aligned}$$

This proves that f preserves the addition of $\mathrm{GR}(4^m)$. We also have

$$cc' = aa' + 2(ab' + a'b).$$

Let the 2-adic representation of $ab' + a'b$ be $ab' + a'b = a_2 + 2b_2$, then the 2-adic representation of cc' is $cc' = aa' + 2a_2$. Therefore

$$\begin{aligned}(cc')^f &= a^2a'^2 + 2a_2^2 \\ &= a^2a'^2 + 2(a^2b'^2 + a'^2b^2) \\ &= (a^2 + 2b^2)(a'^2 + 2b'^2) \\ &= c^f c^{f'}.\end{aligned}$$

This proves that f also preserves the multiplication of $\mathrm{GR}(4^m)$. Therefore f is a ring automorphism of $\mathrm{GR}(4^m)$.

We know that ξ is of order $2^m - 1$, from which it follows immediately that f is of order m. Clearly $a^2 = a$ implies $a = 0$ or 1. Therefore the fixed elements of f are 0, 1, 2, 3 and they form the ring $\mathbb{Z}_4$. □

Theorem 6.12. *Let σ be a ring automorphism of* $\mathrm{GR}(4^m)$, *then* $\sigma = f^i$ *for some* $i, 0 \le i \le m-1$.

Proof. By Theorem 6.7(i) there is an element $\xi \in \mathrm{GR}(4^m)$ such that ξ is of order $2^m - 1$, ξ is a root of a basic primitive polynomial $h(X)$ of degree m over $\mathbb{Z}_4$ and $\mathrm{GR}(4^m) = \mathbb{Z}_4[\xi]$. Then $\overline{h}(X)$ is a primitive polynomial of degree m over $\mathbb{Z}_2$. By Theorem 6.7 (ii), any element $c \in \mathrm{GR}(4^m)$ can be written uniquely as

$$c = a + 2b, \quad a, b \in \mathcal{T},$$

where $\mathcal{T} = \{0, 1, \xi, \xi^2, \ldots, \xi^{2^m-2}\}$. For any $a \in \mathcal{T}$, we have $a^{2^m} - a = 0$, i.e., the 2^m elements of $\mathcal{T}$ are roots of $X^{2^m} - X$. By Lemma 6.4, they are all the roots of $X^{2^m} - X$ in $\mathrm{GR}(4^m)$. It follows that $\mathcal{T}^\sigma = \mathcal{T}$ Clearly, $1^\sigma = 1$, so $2^\sigma = 2$. Thus $c^\sigma = a^\sigma + 2b^\sigma$. Therefore σ is determined by its action on $\mathcal{T}$

From $2^\sigma = 2$ we deduce $(2)^\sigma = (2)$. Therefore σ induces an automorphism $\overline{\sigma}$ of $\mathrm{GR}(4^m)/(2) \simeq \mathbb{F}_{2^m}$. That is, $\overline{c}^{\overline{\sigma}} = \overline{c^\sigma}$ for all $c \in \mathrm{GR}(4^m)$. Assume that $\overline{\xi}^{\overline{\sigma}} = \overline{\xi}^{2^i}$ for some $i, 0 \le i \le m-1$, and that $\xi^\sigma = \xi^j$, $1 \le j \le 2^m - 2$, then, $\overline{\xi}^j = \overline{\xi^j} = \overline{\xi^\sigma} = \overline{\xi}^{\overline{\sigma}} = \overline{\xi}^{2^i}$, which implies $j = 2^i$. Therefore $\sigma = f^i$. □

The cyclic group $\langle f \rangle$ generated by f is called the *Galois group* of $\mathrm{GR}(4^m)$ over $\mathbb{Z}_4$.

Recall that the trace map Tr from $\mathbb{F}_{2^m}$ to $\mathbb{F}_2$ is defined by

$$\mathrm{Tr}(a) = a + a^{f_2} + a^{f_2^2} + \cdots + a^{f_2^{m-1}} \quad \text{for all} \quad a \in \mathbb{F}_{2^m}.$$

Define the *generalized trace map* T from $\mathrm{GR}(4^m)$ to $\mathbb{Z}_4$ by

$$T(c) = c + c^f + c^{f^2} + \cdots + c^{f^{m-1}} \quad \text{for all} \quad c \in \mathrm{GR}(4^m).$$

Proposition 6.13. *We have*

(i) $T(c + c') = T(c) + T(c')$ *for all* $c, c' \in \mathrm{GR}(4^m)$,
(ii) $T(ac) = aT(c)$ *for all* $a \in \mathbb{Z}_4$ *and* $c \in \mathrm{GR}(4^m)$,
(iii) $\overline{} \circ f = f_2 \circ \overline{}$, *i.e.*, $\overline{c^f} = \overline{c}^{f_2}$ *for all* $c \in \mathrm{GR}(4^m)$,
(iv) $\overline{} \circ T = \mathrm{Tr} \circ \overline{}$, *i.e.*, $\overline{T(c)} = \mathrm{Tr}(\overline{c})$ *for all* $c \in \mathrm{GR}(4^m)$.

Moreover, T is a surjective map from $\mathrm{GR}(4^m)$ *to* $\mathbb{Z}_4$.

Proof. The four formulas in the proposition are easy to verify. The last assertion follows from the fact that Tr is a surjective map from $\mathbb{F}_{2^m}$ to $\mathbb{F}_2$, (iv) and (ii). $\square$

Proposition 6.14. *Let $h(X)$ be a basic irreducible polynomial of degree m over $\mathbb{Z}_4$ and η be a root of $h(X)$ in $\mathrm{GR}(4^m)$. Then $\eta, \eta^f, \eta^{f^2}, \ldots, \eta^{f^{m-1}}$ are all the distinct roots of $h(X)$ in $\mathrm{GR}(4^m)$ and $h(X)$ has the following unique factorization into linear factors in $\mathrm{GR}(4^m)[X]$:*

$$h(X) = (X-\eta)(X-\eta^f)\cdots(X-\eta^{f^{m-1}}). \tag{6.13}$$

In particular, if $h(X)$ is a basic primitive polynomial of degree m, $h(X) \mid (X^{2^m-1}-1)$ in $\mathbb{Z}_4[X]$, and ξ is a root of $h(X)$ in $\mathrm{GR}(4^m)$, then $\xi, \xi^2, \xi^{2^2}, \ldots, \xi^{2^{m-1}}$ are all the distinct roots of $h(X)$ in $\mathrm{GR}(4^m)$ and $h(X)$ has the following unique factorization:

$$h(X) = (X-\xi)(X-\xi^2)\cdots(X-\xi^{2^{m-1}}).$$

Proof. Let $h(X)$ be a basic irreducible polynomial of degree m over $\mathbb{Z}_4$ and η be a root of $h(X)$ in $\mathrm{GR}(4^m)$. From $h(\eta)=0$ we deduce that $h(\eta^{f^i}) = h(\eta)^{f^i} = 0$ for $i = 0, 1, 2, \ldots$, i.e., $\eta^{f^i}, i = 0, 1, 2, \ldots$, are roots of $h(X)$ in $\mathrm{GR}(4^m)$. By Proposition 6.13, $\overline{h}(\overline{\eta}^{f_2^i}) = \overline{h}(\overline{\eta})^{f_2^i} = \overline{h(\eta)^{f^i}} = 0$, that is, $\overline{\eta}^{f_2^i}, i = 0, 1, 2, \ldots$, are roots of $\overline{h}(X)$ in $\mathbb{F}_{2^m} = \mathrm{GR}(4^m)/(2)$. Since $\overline{h}(X)$ is irreducible of degree m over $\mathbb{Z}_2, \overline{\eta}, \overline{\eta}^{f_2}, \ldots, \overline{\eta}^{f_2^{m-1}}$ are distinct in pairs and are all the m roots of $\overline{h}(X)$, and $\overline{\eta}^{f_2^m} = \overline{\eta}$. It follows that for $i \neq j$, $0 \le i, j \le m-1$, $\overline{\eta^{f^i} - \eta^{f^j}} = \overline{\eta}^{f_2^i} - \overline{\eta}^{f_2^j} \neq 0$, so $\eta, \eta^f, \ldots, \eta^{f^{m-1}}$ are distinct in pairs.

Let η' be a root of $h(X)$ in $\mathrm{GR}(4^m)$. Then $\overline{\eta'}$ is a root of $\overline{h}(X)$ in $\mathbb{F}_{2^m}$. Therefore $\overline{\eta'} = \overline{\eta}^{f_2^i}$ for some $i, 0 \le i \le m-1$. But $\eta^{f^i} \in \mathrm{GR}(4^m)$ and $\overline{\eta^{f^i}} = \overline{\eta}^{f_2^i}$ By Lemma 6.4, $\eta' = \eta^{f^i}$ This proves that $\eta, \eta^f, \ldots, \eta^{f^{m-1}}$ are all the distinct roots of $h(X)$ in $\mathrm{GR}(4^m)$.

We have the unique factorization of $\overline{h}(X)$ into linear factors $\overline{h}(X) = (X - \overline{\eta})(X - \overline{\eta}^{f_2}) \ldots (X - \overline{\eta}^{f_2^{m-1}})$ in $\mathbb{F}_{2^m}[X]$. Then the unique factorization (6.13) of $h(X)$ into linear factors in $\mathrm{GR}(4^m)$ follows from Hensel's lemma and Lemma 6.4. $\square$

6.4. Basic Primitive Polynomials Which Are Hensel Lifts

The proof of Theorem 6.7 (i) shows that the basic primitive polynomial $h(X) \in \mathbb{Z}_4[X]$ of degree m in that proposition is the Hensel lift of the binary

Table 6.1. Basic primitive polynomials of degree ≤ 10 over $\mathbb{Z}_4$ which are Hensel lifts of binary primitive polynomials.

Degree 3	1213	1323			
Degree 4	10231	13201			
Degree 5	100323	113013	113123	121003	123133
	130133				
Degree 6	1002031	1110231	1211031	1301121	1302001
	1320111				
Degree 7	10020013	10030203	10201003	10221133	10233123
	11122323	11131123	11321133	11332133	11332203
	12122333	12303213	12311203	12331333	13002003
	13210123	13212213	13223213		
Degree 8	100103121	100301231	102231321	111002031	111021311
	111310321	113120111	121102121	121201121	121301001
	121320031	123013111	123132201	130023121	130200111
	132103001				
Degree 9	1000030203	1001011333	1001233203	1002231013	1020100003
	1020332213	1021123003	1021301133	1021331123	1022121323
	1023112133	1110220323	1111300013	1111311013	1112201133
	1113303003	1130312123	1131003213	1131003323	1131030123
	1132331203	1133013203	1133022333	1210032123	1210220333
	1211003133	1211213013	1213232203	1230103133	1230313123
	1231310123	1232100323	1232310133	1232322013	1233113203
	1300013333	1301110213	1301301213	1301323323	1302210213
	1302212123	1303122003	1303313333	1320322013	1320333013
	1321003133	1322110203	1323013013		
Degree 10	10000203001	10002102111	10002123121	10020213031	10030023231
	10030200001	10203103311	10203122121	10211131111	10213010311
	10213330231	10231100111	10233222121	11100113201	11111110231
	11113111201	11120120001	11120232311	11122031321	11131011031
	11301031201	11301210321	11301320031	11312010231	11321001121
	11323133321	11323202111	11330130201	11330223121	12100122031
	12102023121	12110012311	12120311321	12122130201	12122233201
	12132020121	12132120001	12132203311	12301210311	12311302121
	12313022111	12321103231	12321222031	12331133031	12333132311
	13002310311	13011013111	13011232231	13020010231	13022100121
	13022212321	13031202001	13033113321	13201002031	13201021311
	13201111111	13203331201	13223211031	13230112321	13232003001

Note. For degree 3, the entry 1213 represents the polynomial $X^3 + 2X^2 + X + 3$.

primitive polynomial $\overline{h}(X)$. But Example 5.3 points out that a basic primitive polynomial $h(X)$ over $\mathbb{Z}_4$ is not necessarily the Hensel lift of the binary primitive polynomial $\overline{h}(X)$. Now we give a necessary and sufficient condition when a basic primitive polynomial $h(X)$ over $\mathbb{Z}_4$ is the Hensel lift of the binary primitive polynomial $\overline{h}(X)$.

Proposition 6.15. *Let $h(X)$ be a basic primitive polynomial of degree m over $\mathbb{Z}_4$. Then $h(X)$ is the Hensel lift of the binary primitive polynomial $\overline{h}(X)$ if and only if $h(X)$ has a root ξ of order $2^m - 1$ in* $\mathrm{GR}(4^m)$.

Proof. First assume that $h(X)$ has a root ξ of order $2^m - 1$ in $\mathrm{GR}(4^m)$. By Theorem 6.7 (i) ξ is a root of a basic primitive polynomial of degree m over $\mathbb{Z}_4$ and this polynomial is the unique monic polynomial of degree m over $\mathbb{Z}_4$ having ξ as a root. Then this basic primitive polynomial must be $h(X)$. By the proof of Theorem 6.7 (i) this polynomial divides $X^{2^m-1} - 1$ in $\mathbb{Z}_4[X]$, i.e., $h(X)|(X^{2^m-1} - 1)$ in $\mathbb{Z}_4[X]$. Since $\overline{h}(X)$ is a binary primitive polynomial of degree m, we also have $\overline{h}(X)|(X^{2^m-1} - 1)$ in $\mathbb{Z}_2[X]$. Therefore $h(X)$ is the Hensel lift of $\overline{h}(X)$.

Conversely, assume that $h(X)$ is the Hensel lift of $\overline{h}(X)$. Since $\overline{h}(X)$ is a binary primitive polynomial of degree m, we have $\overline{h}(X)|(X^{2^m-1} - 1)$ in $\mathbb{Z}_2[X]$. By Proposition 5.12 there exists a unique monic polynomial $f(X)$ in $\mathbb{Z}_4[X]$ dividing $X^{2^m-1} - 1$ in $\mathbb{Z}_4[X]$ and $\overline{f}(X) = \overline{h}(X)$. Then $f(X)$ is the Hensel lift of $\overline{h}(X)$. Therefore $f(X) = h(X)$ and $h(X)|(X^{2^m-1} - 1)$ in $\mathbb{Z}_4[X]$. Let ξ be a root of $h(X)$ in $\mathrm{GR}(4^m)$, then $\xi^{2^m-1} = 1$ and hence $\overline{\xi}^{2^m-1} = 1$. From $h(\xi) = 0$ we deduce that $\overline{h}(\overline{\xi}) = 0$. Since $\overline{h}(X)$ is primitive of degree m over $\mathbb{Z}_2$, $\overline{\xi}$ is of order $2^m - 1$. Therefore ξ is also of order $2^m - 1$. $\square$

All basic primitive polynomials of degree ≤ 10 over $\mathbb{Z}_4$ which are Hensel lifts of binary primitive polynomials were listed by Boztaş *et al.* (1992). Their list is reproduced in Table 6.1.

6.5. Dependencies among ξ^j

For later use we prove the following proposition, (see Hammons *et al.* (1994)), which contains some results about dependencies among the powers ξ^j.

Proposition 6.16. *Let $\xi \in \mathrm{GR}(4^m)$ be such that both ξ and $\overline{\xi}$ are of order $2^m - 1$. Then*

(i) $\pm\xi^j \pm \xi^k$ *is invertible for* $0 \le j < k < 2^m - 1$, *where* $m \ge 2$.

(ii) $\xi^j - \xi^k \ne \pm\xi^l$ *for distinct* j, k, l *in the range* $[0, 2^m - 2]$, *where* $m \ge 2$.

(iii) *Assume that* $m \ge 3$ *and* i, j, k, l *are in the range* $[0, 2^m - 2]$ *and* $i \ne j, k \ne l$. *Then*

$$\xi^i - \xi^j = \xi^k - \xi^l \Leftrightarrow i = k \text{ and } j = l.$$

(iv) *For odd* $m \ge 3$,

$$\xi^i + \xi^j + \xi^k + \xi^l = 0 \Rightarrow i = j = k = l.$$

Proof. (i) Assume on the contrary that $\pm\xi^j \pm \xi^k = 2\lambda$, where $\lambda \in \mathrm{GR}(4^m)$, then applying $^-$ we obtain $\overline{\xi}^j + \overline{\xi}^k = 0$, which contradicts the fact that $\overline{\xi}$ is of order $2^m - 1$ in $\mathbb{F}_{2^m}$.

(ii) Assume that $\xi^j - \xi^k = \xi^l$, then $\xi^k + \xi^l = \xi^j$ and $1 + \xi^{l-k} = \xi^{j-k}$. Let $l - k = a$ and $j - k = b$, then $1 + \xi^a = \xi^b$ and $a \ne b$. Squaring gives $1 + 2\xi^a + \xi^{2a} = \xi^{2b}$, but applying the Frobenius map gives $1 + \xi^{2a} = \xi^{2b}$, so $2\xi^a = 0$, a contradiction. Similarly, if $\xi^j - \xi^k = -\xi^l$, then $\xi^j + \xi^l = \xi^k$ which leads also to a contradiction.

(iii) From $\xi^i - \xi^j = \xi^k - \xi^l$ we deduce $1 + \xi^a = \xi^b + \xi^c$, where $a = l - i$, $b = j - i$, and $c = k - i$. Squaring and subtracting the result of applying the Frobenius map, we obtain $2\xi^a = 2\xi^{b+c}$. By the uniqueness of 2-adic representation (Theorem 6.7(ii)), $\xi^a = \xi^{b+c}$ Then $1 + \xi^{b+c} = \xi^b + \xi^c$ and $(1 - \xi^b)(1 - \xi^c) = 0$. By (i), $b = 0$ or $c = 0$. By assumption $b \ne 0$, therefore $c = 0$, i.e., $i = k$. Then $j = l$.

(iv) We have $1 + \xi^a = -\xi^b - \xi^c$, where $a = j - i, b = k - i$, and $c = l - i$. Squaring and subtracting the result of applying the Frobenius map gives $2\xi^a = 2(\xi^{2b} + \xi^{2c} + \xi^{b+c})$. Substituting $\xi^a = -1 - \xi^b - \xi^c$ into it, we obtain $2(-1 - \xi^b - \xi^c) = 2(\xi^{2b} + \xi^{2c} + \xi^{b+c})$. Therefore $1 + \overline{\xi}^b + \overline{\xi}^c = \overline{\xi}^{2b} + \overline{\xi}^{2c} + \overline{\xi}^{b+c}$. Let $\overline{\xi}^b = x + 1$ and $\overline{\xi}^c = y + 1$, then $x^2 + y^2 = xy$. Substituting $y = tx$, we find $x^2(1 + t + t^2) = 0$. As m is odd, $1 + t + t^2 \ne 0$ in $\mathbb{F}_{2^m}$. Therefore $x = 0$ and $y = 0$. It follows that $b = c = 0$. From $1 + \xi^a = -\xi^b - \xi^c$ we deduce that $\xi^a = 1$ and $a = 0$. Hence $i = j = k = l$. $\square$

CHAPTER 7

CYCLIC CODES

7.1. A Review of Binary Cyclic Codes

A binary linear code C of length n is called a *binary cyclic code* if

$$(c_0, c_1, \ldots, c_{n-1}) \in C \Rightarrow (c_{n-1}, c_0, c_1, \ldots, c_{n-1}) \in C. \tag{7.1}$$

We represent any word $(a_0, a_1, \ldots, a_{n-1}) \in \mathbb{F}_2^n$ by the residue class of the polynomial $a_0 + a_1X + \cdots + a_{n-1}X^{n-1}$ over $\mathbb{F}_2$ mod $X^n - 1$. Then we have a bijection

$$\begin{aligned} \mathbb{F}_2^n &\to \mathbb{F}_2[X]/(X^n - 1) \\ (a_0, a_1, \ldots, a_{n-1}) &\mapsto a_0 + a_1X + \cdots + a_{n-1}X^{n-1} + (X^n - 1), \end{aligned} \tag{7.2}$$

where $(X^n - 1)$ is the principal ideal generated by $X^n - 1$ in the polynomial ring $\mathbb{F}_2[X]$. For simplicity, we write $a_0 + a_1X + \cdots + a_{n-1}X^{n-1}$ for $a_0 + a_1X + \cdots + a_{n-1}X^{n-1} + (X^n - 1)$, namely, we use the unique residue class representative of degree $< n$ $a_0 + a_1X + \cdots + a_{n-1}X^{n-1}$ in the residue class $a_0 + a_1X + \cdots + a_{n-1}X^{n-1} + (X^n - 1)$ to represent the residue class, and we call the residue class simply the polynomial $a_0 + a_1X + \cdots + a_{n-1}X^{n-1}$. Denote the image of a binary code C under the map (7.2) also by C. Clearly, (7.1) is equivalent to

$$c_0 + c_1X + \cdots + c_{n-1}X^{n-1} \in C \Rightarrow X(c_0 + c_1X + \cdots + c_{n-1}X^{n-1}) \in C, \tag{7.3}$$

from which we deduce immediately:

Proposition 7.1. *A nonempty subset of $\mathbb{F}_2^n$ is a binary cyclic code if and only if its image under the map* (7.2) *is an ideal of the residue class ring* $\mathbb{F}_2[X]/(X^n - 1)$. □

It follows from Proposition 7.1 that binary cyclic codes of length n are precisely the ideals of the residue class ring $\mathbb{F}_2[X]/(X^n - 1)$.

We recall the following well-known facts on binary cyclic codes and the ideals of $\mathbb{F}_2[X]/(X^n - 1)$, which are stated as Propositions 7.2–7.5, the proofs of which can be found in MacWilliams and Sloane (1977), Chap. 5.

Proposition 7.2. *Every ideal I of $\mathbb{F}_2[X]/(X^n-1)$ is principal. More precisely, I is generated by the polynomial of least degree $g(X) \in I$. Moreover, if $g(X) \neq 0$, then $g(X)$ is a divisor of $X^n - 1$ in $\mathbb{F}_2[X]$.* □

The polynomial $g(X)$ in Proposition 7.2 is called the *generator polynomial* of I. Let

$$h(X) = h_0 + h_1 X + \cdots + h_m X^m \in \mathbb{F}_2[X], \quad \text{where } h_0 = h_m = 1\,.$$

Define

$$\tilde{h}(X) = h_m + \cdots + h_1 X^{m-1} + h_0 X^m$$

and call $\tilde{h}(X)$ the *reciprocal polynomial* to $h(X)$. It is easy to verify that $\tilde{h}(X) = X^m h(1/X)$.

Proposition 7.3. *Let C be a binary cyclic code of length n, then the dual code $C^\perp$ of C is also cyclic. Moreover, assume that both C and $C^\perp$ are nonzero, let I and I' be the ideals corrresponding to C and $C^\perp$, respectively, under the bijection* (7.2), *and let $g(X)$ and $\tilde{h}(X)$ be the generator polynomials of I and I', respectively, then $\tilde{h}(X)$ is the reciprocal polynomial to $h(X) = (X^n - 1)/g(X)$.* □

Proposition 7.4. *Let $g(X)$ be a divisor of $X^n - 1$ and $g(X) \neq 1$. Then $(g(X))$ is a prime ideal of $\mathbb{F}_2[X]/(X^n - 1)$ if and only if $g(X)$ is an irreducible factor of $X^n - 1$ in $\mathbb{F}_2[X]$. Moreover, every prime ideal of $\mathbb{F}_2[X]/(X^n - 1)$ is maximal.* □

Proposition 7.5. *Assume that $2 \nmid n$. Every nonzero ideal I of $\mathbb{F}_2[X]/(X^n - 1)$ is generated by a unique idempotent polynomial $e(X)$, i.e., there exist a* *unique polynomial* $e(X) \in \mathbb{F}_2[X]/(X^n - 1)$ *with the properties:* $I = (e(X))$ *and* $e(X)^2 = e(X) \neq 0$. □

The unique polynomial $e(X) \in \mathbb{F}_2[X]/(X^n - 1)$ in Proposition 7.5 such that $I = (e(X))$ and $e(X)^2 = e(X) \neq 0$ is called the *generating idempotent* of I. $e(X)$ is also called the generating idempotent of the binary cyclic code of length n corresponding to I under the bijection (7.2). Moreover, let I_1 and I_2 be ideals of $\mathbb{F}_2[X]/(X^n - 1)$. Define

$$I_1 \cap I_2 = \{a(X) \in \mathbb{F}_2[X]/(X^n - 1) | a(X) \in I_1 \text{ and } a(X) \in I_2\},$$
$$I_1 + I_2 = \{a_1(X) + a_2(X) | a_1(X) \in I_1 \text{ and } a_2(X) \in I_2\}.$$

It is easy to see that both $I_1 \cap I_2$ and $I_1 + I_2$ are ideals of $\mathbb{F}_2[X]/(X^n - 1)$. We call $I_1 \cap I_2$ and $I_1 + I_2$ the intersection and sum of I_1 and I_2, respectively. We have

Proposition 7.6. *Let* $2 \nmid n$, I_1 *and* I_2 *be two nonzero ideals of* $\mathbb{F}_2[X]/(X^n-1)$, *and* $e_1(X)$ *and* $e_2(X)$ *be the generating idempotents of* I_1 *and* I_2, *respectively. Then* $e_1(X)\,e_2(X)$ *is the generating idempotent of* $I_1 \cap I_2$ *and* $e_1(X) + e_2(X) - e_1(X)e_2(X)$ *is the generating idempotent of* $I_1 + I_2$. *In particular, when* $e_1(X)$ *and* $e_2(X)$ *are orthogonal,* $e_1(X) + e_2(X)$ *is the generating idempotent of* $I_1 + I_2$.

Proof. For any $f(X) \in I_1 \cap I_2$, we have $f(X) = f_1(X)e_1(X) = f_2(X)e_2(X)$, where $f_1(X), f_2(X) \in \mathbb{F}_2[X]/(X^n - 1)$. Then

$$\begin{aligned} f(X)e_1(X)e_2(X) &= f_1(X)e_1(X)e_1(X)e_2(X) = f_1(X)e_1(X)e_2(X) \\ &= f_2(X)e_2(X)e_2(X) = f_2(X)e_2(X) = f(X), \end{aligned}$$

i.e., $e_1(X)e_2(X)$ is the identity of $I_1 \cap I_2$ and hence, is the generating idempotent of $I_1 \cap I_2$.

Similarly, we can show that $e_1(X) + e_2(X) - e_1(X)e_2(X)$ is the identity of $I_1 + I_2$. □

Since X and $X^n - 1$ are coprime, there are polynomials $a(X)$ and $b(X) \in \mathbb{F}_2[X]$ such that

$$a(X)\,X + b(X)(X^n - 1) = 1.$$

Performing reduction modulo $X^n - 1$, we obtain

$$a(X)\,X \equiv 1 \pmod{X^n - 1}.$$

Hence $a(X)$ is the inverse of X in $\mathbb{F}_2[X]/(X^n - 1)$ and we denote $a(X)$ by X^{-1}

Proposition 7.7. *Let $2 \nmid n$ and C be a nonzero binary cyclic code of length n with the generating idempotent $e(X)$, then $C^\perp$ has the generating idempotent $1 - e(X^{-1})$.*

Proof. Let I be the ideal corresponding to C under the bijection (7.2) and $g(X)$ be the generator polynomial of I. By Proposition 7.2, $g(X) \mid X^n - 1$. Let $X^n - 1 = g(X)\,h(X)$, where $h(X) \in \mathbb{F}_2[X]$. Since $2 \nmid n$, $g(X)$ and $h(X)$ are coprime. Then $\mathbb{F}_2[X] = (g(X)) + (h(X))$. Thus

$$\mathbb{F}_2[X]/(X^n - 1) = (g(X)) + (h(X)) \quad \text{and} \quad (g(X))\,(h(X)) = (0)\,.$$

Let

$$1 = e_1(X) + e_2(X), \qquad \text{where} \qquad e_1(X) \in (g(X)),\, e_2(X) \in (h(X))\,.$$

It is easy to verify that $e_1(X)$ and $e_2(X)$ are the generating idempotents of $(g(X))$ and $(h(X))$, respectively. By hypothesis, $e_1(X) = e(X)$. Therefore $e_2(X) = 1 - e(X)$. We have $e_2(X) = r(X)\,h(X)$ and $h(X) = t(X)\,e_2(X)$, where $r(X)$, $t(X) \in \mathbb{F}_2[X]$. Then we deduce $e_2(X^{-1}) = r(X^{-1})\,h(X^{-1})$ and $h(X^{-1}) = t(X^{-1})\,e_2(X^{-1})$. Therefore $e_2(X^{-1})$ is the generating idempotent of $(h(X^{-1}))$. But $(h(X^{-1})) = (\tilde{h}(X))$ and $(\tilde{h}(X))$ is the ideal corresponding to the dual code $C^\perp$ under the bijection (7.2). Therefore $e_2(X^{-1}) = 1 - e(X^{-1})$ is the generating idempotent of $C^\perp$ □

Propositions 7.6 and 7.7 are well-known, see MacWilliams and Sloane (1977), Chap. 8.

7.2. Quaternary Cyclic Codes

A *quaternary cyclic code* C of length n is a quaternary linear code C of length n with the property

$$(c_0, c_1, \ldots, c_{n-1}) \in C \Rightarrow (c_{n-1}, c_0, c_1, \ldots, c_{n-2}) \in C\,. \tag{7.4}$$

We call quaternary cyclic code simply $\mathbb{Z}_4$-*cyclic code* in the following. As in the binary case, we have a bijection

$$\begin{aligned} \mathbb{Z}_4^n &\to \mathbb{Z}_4[X]/(X^n - 1) \\ (a_0, a_1, \ldots, a_{n-1}) &\mapsto a_0 + a_1 X + \cdots + a_{n-1}X^{n-1} + (X^n - 1)\,. \end{aligned} \tag{7.5}$$

For simplicity we write $a_0 + a_1X + \cdots + a_{n-1}X^{n-1}$ for $a_0 + a_1X + \cdots + a_{n-1}X^{n-1} + (X^n - 1)$. Denote the image of a $\mathbb{Z}_4$-code C under (7.5) also by C, then under the bijection (7.5) the codeword $\mathbf{c} = (c_0, c_1, \ldots, c_{n-1}) \in C$ is mapped into $c(X) = c_0 + c_1X + \cdots + c_{n-1}X^{n-1}$, which will also be called a codeword of C. The property (7.4) is equivalent to

$$c_0 + c_1X + \cdots + c_{n-1}X^{n-1} \in C \Rightarrow X(c_0 + c_1X + \cdots + c_{n-1}X^{n-1}) \in C\,. \quad (7.6)$$

As in the binary case we have

Proposition 7.8. *A nonempty set of $\mathbb{Z}_4^n$ is a $\mathbb{Z}_4$-cyclic code if and only if its image under* (7.5) *is an ideal of the residue class ring $\mathbb{Z}_4[X]/(X^n - 1)$.* □

Thus $\mathbb{Z}_4$-cyclic codes of length n are precisely the ideals in the residue class ring $\mathbb{Z}_4[X]/(X^n - 1)$.

Let $g(X)$ be a monic polynomial over $\mathbb{Z}_4$ dividing $X^n - 1$ and let $C = (g(X))$ be the principal ideal of $\mathbb{Z}_4[X]/(X^n - 1)$ generated by $g(X)$. Then C is called the $\mathbb{Z}_4$-cyclic code with *generator polynomial* $g(X)$. Let $h(X) = (X^n - 1)/g(X)$, then $h(X)g(X) \equiv 0 \pmod{X^n - 1}$. Let $\deg g(X) = m$, then $\deg h(X) = n - m$. Write

$$g(X) = g_0 + g_1X + \cdots + g_mX^m$$

and

$$h(X) = h_0 + h_1X + \cdots + h_{n-m}X^{n-m}\,,$$

then $g_m = h_{n-m} = 1$ and $g_0 = h_0 = \pm 1$. Since $h(X)g(X) \equiv 0 \pmod{X^n - 1}$, $X^{n-m}g(X)$ can be expressed as a linear combination of $g(X)$, $Xg(X)$, …, $X^{n-m-1}g(X)$. Therefore the codewords $g(X)$, $Xg(X)$, …, $X^{n-m-1}g(X)$ of C form a basis of the code C. That is, the $(n - m) \times m$ matrix[a]

$$G = \begin{pmatrix} g_0 & g_1 & \cdots & g_m & & & \\ & g_0 & g_1 & \cdots & g_m & & \\ & & \ddots & & & \ddots & \\ & & & g_0 & g_1 & \cdots & g_m \end{pmatrix}$$

is a *generator matrix* of C and C is of type 4^{n-m}.

Clearly, a word $c(X) = c_0 + c_1X + \cdots + c_{n-1}X^{n-1}$ is a codeword of C if and only if $c(X)h(X) = 0$. $h(X)$ is called the *check polynomial* of C. Define an $m \times n$ matrix H by

[a] We agree that the zeros in matrices are sometimes omitted, i.e., the blanks in matrices represent the omitted zeros if it is clear from the context.

$$H = \begin{pmatrix} h_{n-m} & \cdots & h_1 & h_0 & & & \\ 0 & h_{n-m} & \cdots & h_1 & h_0 & & \\ & & \ddots & & & \ddots & \\ & & & h_{n-m} & \cdots & h_1 & h_0 \end{pmatrix}$$

It is easy to verify that the codewords $g(X)$, $Xg(X)$, $\ldots$, $X^{n-m-1}g(X)$ are orthogonal to every row of H. It follows that each codeword of C is orthogonal to every row of H. Clearly, the system of linear equations

$$H^t(X_0, X_1, \ldots, X_{n-1}) = 0$$

has 4^m solutions. Therefore a word orthogonal to each row of H if and only if it is a codeword of C. Thus H is a *parity check matrix* of C. Define the *reciprocal polynomial* $\tilde{h}(X)$ to $h(X)$ to be

$$\tilde{h}(X) = h_{n-m} + \cdots + h_1 X^{n-m-1} + h_0 X^{n-m}\,.$$

Then the $\mathbb{Z}_4$-cyclic code with $\tilde{h}(X)$ as its generator matrix is the dual code of C. We conclude

Proposition 7.9. *Let $g(X)$ be a monic polynomial over $\mathbb{Z}_4$ dividing $X^n - 1$ and $h(X) = (X^n - 1)/g(X)$. Let $C = (g(X))$ be the $\mathbb{Z}_4$-cyclic code with generator polynomial $g(X)$, then $C^\perp$ is a $\mathbb{Z}_4$-cyclic code whose generator polynomial $\tilde{h}(X)$ is the reciprocal polynomial to $h(X)$.* □

7.3. Sun Zi Theorem

Sun Zi Theorem appeared first in *The Arithmetic of Sun Zi*, 3–5AD. It is one of the important achievements of ancient Chinese mathematics and is often called the Chinese Remainder Theorem in the western literature of mathematics. It can be regarded as a theorem of simultaneous congruences modulo finitely many pairwise coprime integers and can be interpreted as a theorem of the direct sum decomposition of the ring of integers modulo the product of these pairwise coprime integers, see Wan (1992), Chap. 4. For the present purpose we generalize it to a theorem on simultaneous congruences modulo finitely many pairwise coprime polynomials in $\mathbb{Z}_4[X]$ and interpret it as a theorem of the direct sum decomposition of the ring $\mathbb{Z}_4[X]/(f(X))$, where $f(X)$ is the product of these pairwise coprime polynomials in $\mathbb{Z}_4[X]$.

We need more concepts from commutative ring theory, which will be sketched below.

Let R be a commutative ring and $I_1, I_2, \ldots, I_r$ be ideals of R. Define

$$I_1 \cap I_2 \cap \cdots \cap I_r = \{a \in R | a \in I_i,\ i = 1, 2, \ldots, r\},$$
$$I_1 + I_2 + \cdots + I_r = \{a_1 + a_2 + \cdots + a_r | a_i \in I_i,\ i = 1, 2, \ldots, r\},$$
$$I_1 I_2 \cdots I_r = \left\{ \sum a_1 a_2 \cdots a_r | a_i \in I_i,\ i = 1, 2, \ldots, r \text{ and the sum is finite} \right\}$$

Then $I_1 \cap I_2 \cap \cdots \cap I_r$, $I_1 + I_2 + \cdots + I_r$, and $I_1 I_2 \cdots I_r$ are ideals of R, and they are called the *intersection, sum*, and *product* of $I_1, I_2, \ldots, I_r$ respectively. Clearly

$$I_1 I_2 \cdots I_r \subset I_1 \cap I_2 \cap \cdots \cap I_r .$$

Let I be an ideal of R with identity. If there are finitely many elements $a_1, \ldots, a_m \in I$ such that

$$I = \{r_1 a_1 + \cdots + r_m a_m | r_1, \ldots, r_m \in R\},$$

then I is said to have a *finite basis* $\{a_1, \ldots, a_m\}$ and we write $I = (a_1, a_2, \ldots, a_m)$.

For example, the principal ideal (a) generated by $a \in R$ has a finite basis $\{a\}$ consisting of a single element a.

Let R be a commutative ring and $R_1, R_2, \ldots, R_r$ be r nonzero ideals of R. If every element $a \in R$ can be expressed uniquely as

$$a = a_1 + a_2 + \cdots + a_r, \qquad a_i \in R_i,$$

then we say that R is decomposed into a *direct sum* of its ideals $R_1, R_2, \ldots, R_r$, which is denoted by

$$R = R_1 \dot{+} R_2 \dot{+} \cdots \dot{+} R_r .$$

We have the following theorem, the proof of which can be found in Wan (1992), Chap. 4.

Theorem 7.10. *Let R be a commutative ring with identity* 1. *Assume that R is decomposed into a direct sum of r nonzero ideals $R_1, R_2, \ldots, R_r$*

$$R = R_1 \dot{+} R_2 \dot{+} \cdots \dot{+} R_r \tag{7.7}$$

and that 1 *has the decomposition*

$$1 = e_1 + e_2 + \cdots + e_r \tag{7.8}$$

in this direct sum decomposition. Then

(i) $e_1, e_2, \ldots, e_r$ *are* r *mutually orthogonal nonzero idempotents of* R, *i.e.*, $e_i \neq 0$ *and* $e_i e_j = \delta_{ij} e_i$ *for* $i, j = 1, 2, \ldots, n$.

(ii) $R_i = Re_i$ *with* e_i *as its identity and* $R_i R_j = \{0\}$.

Conversely, if 1 *is decomposed into a sum of* r *mutually orthogonal nonzero idempotents as in* (7.8) *and let* $R_i = Re_i$, *then* R_i *is a nonzero ideal of* R *with* e_i *as its identity,* $R_i R_j = \{0\}$, *and* R *is the direct sum of* $R_1, R_2, \ldots, R_r$, *i.e., we have* (7.7). $\square$

Lemma 7.11. *Let* $f_1(X), f_2(X), \ldots, f_r(X)$ *be* r *pairwise coprime polynomials over* $\mathbb{Z}_4$ *and let* $\hat{f}_i(X)$ *denote the product of all* $f_j(X)$ *except* $f_i(X)$. *Then* $\hat{f}_i(X)$ *and* $f_i(X)$ *are coprime for* $i = 1, 2, \ldots, r$.

Proof. By Lemma 5.1 the coprimeness of $f_i(X)$ and $f_j(X)$ for $i \neq j$ implies the coprimeness of $\overline{f_i}(X)$ and $\overline{f_j}(X)$. But $\overline{f}_1(X), \overline{f}_2(X), \ldots, \overline{f}_r(X)$ are polynomials over $\mathbb{Z}_2$. So $\overline{\hat{f}_i}(X) = \overline{f_1}(X) \cdots \overline{f}_{i-1}(X)\,\overline{f}_{i+1}(X) \cdots \overline{f}_r(X)$ and $\overline{f_i}(X)$ are coprime. Again by Lemma 5.1, $\hat{f}_i(X)$ and $f_i(X)$ are coprime. $\square$

Lemma 7.12. *Let* $f_1(X), f_2(X), \ldots, f_r(X)$ *be* r *pairwise coprime polynomials in* $\mathbb{Z}_4[X]$, *then*

$$(f_1(X)\, f_2(X) \cdots f_r(X)) = (f_1(X)) \cap (f_2(X)) \cap \cdots \cap (f_r(X)) . \tag{7.9}$$

Proof. Clearly, $f_1(X) f_2(X) \cdots f_r(X) \in (f_i(X))$ for every i. Therefore L.H.S. of (7.9) $\subseteq$ R.H.S. of (7.9). It remains to prove that R.H.S. of (7.9) $\subseteq$ L.H.S. of (7.9). We apply induction on r.

The case $r = 1$ is trivial. Let $r > 1$ and assume that (7.9) holds for $r - 1$. That is, we have

$$(f_1(X)\, f_2(X) \cdots f_{r-1}(X)) = (f_1(X)) \cap (f_2(X)) \cap \cdots \cap (f_{r-1}(X)) .$$

Let $g(X) \in (f_1(X)) \cap (f_2(X)) \cap \cdots \cap (f_r(X))$, then $g(X) \in (f_1(X)\, f_2(X) \cdots f_{r-1}(X)) \cap (f_r(X))$. Thus there are polynomials $g_1(X), g_r(X) \in \mathbb{Z}_4[X]$ such that

$$g(X) = g_1(X)\, f_1(X)\, f_2(X) \cdots f_{r-1}(X) = g_r(X)\, f_r(X) .$$

By Lemma 7.11, $f_1(X)\, f_2(X) \cdots f_{r-1}(X)$ and $f_r(X)$ are coprime. Then there are polynomials $h_1(X), h_r(X) \in \mathbb{Z}_4[X]$ such that

$$h_1(X)\, f_1(X)\, f_2(X) \cdots f_{r-1}(X) + h_r(X)\, f_r(X) = 1 .$$

Multiplying the above equation by $g(X)$, we obtain

$$\begin{aligned} g(X) &= g(X)\,h_1(X)\,f_1(X)\,f_2(X)\cdots f_{r-1}(X) + g(X)\,h_r(X)\,f_r(X) \\ &= (g_r(X)\,h_1(X) + g_1(X)\,h_r(X))\,f_1(X)\,f_2(X)\cdots f_r(X) \\ &\in (f_1(X)\,f_2(X)\cdots f_r(X)) . \end{aligned}$$

Therefore R.H.S. of (7.9) $\subseteq$ L.H.S. of (7.9). Hence (7.9) holds also for r. □

Let $a(X)$, $b(X)$, $f(X) \in \mathbb{Z}_4[X]$. We say that $a(X)$ is *congruent* to $b(X)$ mod $f(X)$ if there exists a polynomial $q(X) \in \mathbb{Z}_4[X]$ such that

$$a(X) - b(X) = q(X)\,f(X) .$$

Then we write

$$a(X) \equiv b(X) \pmod{f(X)} .$$

Theorem 7.13. (Sun Zi Theorem) *Let $f_1(X)$, $f_2(X)$, ..., $f_r(X)$ be r pairwise coprime polynomials of degree ≥ 1 over $\mathbb{Z}_4$ and $a_1(X), a_2(X), \ldots, a_r(X)$ be any r polynomials over $\mathbb{Z}_4$. Then the simultaneous congruences*

$$\begin{aligned} x &\equiv a_1(X) \pmod{f_1(X)} \\ x &\equiv a_2(X) \pmod{f_2(X)} \\ &\cdots \\ x &\equiv a_r(X) \pmod{f_r(X)} \end{aligned} \tag{7.10}$$

has a solution in $\mathbb{Z}_4[X]$. Moreover, the solution of (7.10) is unique mod $f_1(X)\,f_2(X)\cdots f_r(X)$, i.e., if $g(X)$ and $h(X)$ are two solutions of (7.8), then $g(X) \equiv h(X) \pmod{f_1(X)\,f_2(X)\cdots f_r(X)}$.

Proof. Let $\hat{f}_i(X)$ be the product of all $f_j(X)$ except $f_i(X)$. By Lemma 7.11 $\hat{f}_i(X)$ and $f_i(X)$ are coprime, $i = 1, 2, \ldots, r$. Then there are polynomials $b_i(X)$ and $c_i(X)$ over $\mathbb{Z}_4$ such that

$$b_i(X)\,\hat{f}_i(X) + c_i(X)\,f_i(X) = 1 . \tag{7.11}$$

It is easy to verify that

$$a_1(X)\,b_1(X)\,\hat{f}_1(X) + a_2(X)\,b_2(X)\,\hat{f}_2(X) + \cdots + a_r(X)\,b_r(X)\,\hat{f}_r(X)$$

is a solution of (7.10).

Now let $g(X)$ and $h(X)$ be two solutions of (7.10). Then

$$g(X) \equiv h(X) \pmod{f_i(X)}, \qquad i = 1, 2, \ldots, r$$

That is, $g(X) - h(X) \in (f_i(X))$, $i = 1, 2, \ldots, r$. By Lemma 7.12

$$g(X) - h(X) \in (f_1(X)\, f_2(X) \cdots f_r(X)).$$

That is,

$$g(X) \equiv h(X) \pmod{f_1(X)\, f_2(X) \cdots f_r(X)}.$$

□

Sun Zi Theorem can also be interpreted as a theorem on the direct sum decomposition of the residue class ring $\mathbb{Z}_4[X]/(f_1(X)\, f_2(X) \cdots f_r(X))$ as follows:

Theorem 7.14. (Sun Zi Theorem) *Let $f_1(X)$, $f_2(X)$, $\ldots$, $f_r(X)$ be r pairwise coprime polynomials of degree ≥ 1 over $\mathbb{Z}_4$ and $f(X) = f_1(X)\, f_2(X) \cdots f_r(X)$. Denote the residue class ring $\mathbb{Z}_4[X]/(f(X))$ by R. For $i = 1, 2, \ldots, r$, let*

$$e_i = b_i(X)\, \hat{f}_i(X) + (f(X)),$$

where $b_i(X)$ is the polynomial $b_i(X)$ appearing in (7.11) and $\hat{f}_i(X)$ is the product of all $f_j(X)$ except $f_i(X)$. Then

(i) *$e_1, e_2, \ldots, e_r$ are r mutually orthogonal nonzero idempotents of R, i.e., $e_i \neq 0$ for $i = 1, 2, \ldots, r$ and $e_i e_j = \delta_{ij} e_i$ for $i, j = 1, 2, \ldots, r$.*

(ii) *$1 = e_1 + e_2 + \cdots + e_r$.*

(iii) *$R_i = Re_i$ is an ideal of R, and e_i is the identity of R_i, $i = 1, 2, \ldots, r$.*

(iv) *$R = R_1 \dotplus R_2 \dotplus \cdots \dotplus R_r$.*

Proof. First prove (i). From (7.11) we deduce

$$b_i(X)\, \hat{f}_i(X) \equiv 1 \pmod{f_i(X)}. \tag{7.12}$$

Clearly

$$b_i(X)\, \hat{f}_i(X) \equiv 0 \pmod{f_j(X)}, \quad \text{if } j \neq i. \tag{7.13}$$

Squaring both sides of (7.12) and (7.13), we obtain

$$(b_i(X)\, \hat{f}_i(X))^2 \equiv 1 \pmod{f_i(X)},$$

$$(b_i(X)\, \hat{f}_i(X))^2 \equiv 0 \pmod{f_j(X)}, \quad \text{if } j \neq i.$$

By the uniqueness part of Theorem 7.13,

$$(b_i(X)\,\hat{f}_i(X))^2 \equiv b_i(X)\,\hat{f}_i(X) \pmod{f(X)},$$

which implies $e_i^2 = e_i$, $i = 1, 2, \ldots, r$ When $i \neq j$, we have

$$b_i(X)\,\hat{f}_i(X)\,b_j(X)\,\hat{f}_j(X) \equiv 0 \pmod{f(X)},$$

i.e., $e_ie_j = 0$.

If $e_i = 0$ for some i, from (7.11) we deduce that $c_i(X)\,f_i(X) \equiv 1 \pmod{f(X)}$, which implies $0 \equiv 1 \pmod{f_i(X)}$, a contradiction.

Therefore $e_1, e_2, \ldots, e_r$ are r mutually orthogonal nonzero idempotents of R.

Next we prove (ii). From (7.12) and (7.13) we deduce

$$b_1(X)\,\hat{f}_1(X) + b_2(X)\,\hat{f}_2(X) + \cdots + b_r(X)\,\hat{f}_r(X) \equiv 1 \pmod{f_i(X)},$$

$$i = 1, 2, \ldots, r.$$

By Lemma 7.12,

$$b_1(X)\,\hat{f}_1(X) + b_2(X)\,\hat{f}_2(X) + \cdots + b_r(X)\,\hat{f}_r(X) \equiv 1 \pmod{f(X)}.$$

Therefore $e_1 + e_2 + \cdots + e_r = 1$. (iii) and (iv) are immediate consequences of the converse part of Theorem 7.10. □

Corollary 7.15. *Let $f_1(X), f_2(X), \ldots, f_r(X)$ be r pairwise coprime monic polynomials of degree ≥ 1 over $\mathbb{Z}_4$ and $f(X) = f_1(X)\,f_2(X)\cdots f_r(X)$. Then for any $i = 1, 2, \ldots, r$, the map*

$$\begin{aligned} \mathbb{Z}_4[X]/(f_i(X)) &\to (\mathbb{Z}_4[X]/(f(X))\,e_i = Re_i \\ k(X) + (f_i(X)) &\mapsto (k(X) + (f(X)))\,e_i \end{aligned} \tag{7.14}$$

is an isomorphism of rings.

Proof. Clearly, (7.14) is a homomorphism of rings. Let us prove that (7.14) is injective. Let $k(X) + (f_i(X)) \in \mathbb{Z}_4[X]/(f_i(X))$ be such that $(k(X) + (f(X)))\,e_i = 0$. Then

$$k(X)\,b_i(X)\,\hat{f}_i(X) \equiv 0 \pmod{f(X)}.$$

It follows that

$$k(X)\,b_i(X)\,\hat{f}_i(X) \equiv 0 \pmod{f_i(X)}.$$

Multiplying both sides of (7.11) by $k(X)$ and taking modulo $f_i(X)$, we obtain

$$k(X)\, b_i(X)\, \hat{f}_i(X) \equiv k(X) \pmod{f_i(X)}.$$

Therefore $k(X) \equiv 0 \pmod{f_i(X)}$ and $k(X) + (f_i(X)) = (f_i(X))$. This proves that (7.14) is injective.

Finally, let us prove that (7.14) is surjective. Let $(l(X) + (f(X)))\, e_i$ be any element of Re_i. Since $f_i(X)$ is monic, we can divide $l(X)$ by $f_i(X)$ and obtain

$$l(X) = q(X)\, f_i(X) + r(X),$$

where $q(X)$, $r(X) \in \mathbb{Z}_4[X]$ and $\deg r(X) < \deg f_i(X)$. Then $r(X) + (f_i(X)) = l(X) + (f_i(X)) \in \mathbb{Z}_4[X]/(f_i(X))$ and under (7.14), $r(X) + (f_i(X))$ is mapped into $(l(X) + (f(X)))\, e_i$. Therefore (7.14) is surjective. □

Corollary 7.16. *Let $f_1(X)$, $f_2(X)$, ..., $f_r(X)$ be r pairwise coprime monic polynomials of degree ≥ 1 over $\mathbb{Z}_4$ and $f(X) = f_1(X)\, f_2(X) \cdots f_r(X)$. Then*

$$\mathbb{Z}_4[X]/(f(X)) \simeq \mathbb{Z}_4[X]/(f_1(X)) \dot{+} \mathbb{Z}_4[X]/(f_2(X)) \dot{+} \cdots \dot{+} \mathbb{Z}_4[X]/(f_r(X)). \quad \square$$

7.4. Ideals in $\mathbb{Z}_4[X]/(X^n-1)$

By Proposition 7.8 $\mathbb{Z}_4$-cyclic codes of length n are precisely the ideals in the residue class ring $\mathbb{Z}_4[X]/(X^n - 1)$. Now we are going to study the ideals in $\mathbb{Z}_4[X]/(X^n - 1)$. We write $\mathcal{R}$ simply for $\mathbb{Z}_4[X]/(X^n - 1)$. It should be noticed that the unique factorization theorem does not hold in $\mathcal{R}$. For example, in $\mathbb{Z}_4[X]/(X^4 - 1)$ the polynomial $X^4 - 1$ has two distinct factorizations into irreducible polynomials:

$$\begin{aligned} X^4 - 1 &= (X - 1)\,(X + 1)\,(X^2 + 1) \\ &= (X + 1)^2\,(X^2 + 2X - 1). \end{aligned}$$

It should also be noticed that the number of distinct roots of a polynomial of degree m over $\mathbb{Z}_4$ in an extension ring of $\mathbb{Z}_4$, for instance $\mathrm{GR}(4^m)$, may be greater than m. For example, every element $1 + 2\alpha$, where $\alpha \in \mathrm{GR}(4^m)$, is a root of $X^2 - 1$. Therefore we must be careful when working with $\mathcal{R}$.

From now on we assume that n is odd. Then we have

Proposition 7.17. *Let n be an odd positive integer. Then*

(i) *$X^n - 1$ can be factored uniquely into a product of pairwise coprime basic irreducible polynomials $f_1(X)$, $f_2(X)$, ..., $f_r(X)$:*

$$X^n - 1 = f_1(X)\, f_2(X) \cdots f_r(X)\,.$$

(ii) *Let $\hat{f}_i(X)$ be the product of all $f_j(X)$ except $f_i(X)$. Then $\hat{f}_i(X)$ and $f_i(X)$ are coprime for $i = 1, 2, \ldots, r$ and there exist polynomials $b_i(X)$ and $c_i(X)$ over $\mathbb{Z}_4$ such that*

$$b_i(X)\, \hat{f}_i(X) + c_i(X)\, f_i(X) = 1\,. \tag{7.15}$$

(iii) *Let*

$$e_i = b_i(X)\, \hat{f}_i(X) + (X^n - 1), \qquad i = 1, 2, \ldots, r\,, \tag{7.16}$$

then e_1, e_2, ..., e_r are mutually orthogonal nonzero idempotents of $\mathcal{R}$, $1 = e_1 + e_2 + \cdots + e_r$ in $\mathcal{R}$, $\mathcal{R}_i = \mathcal{R}e_i$ is an ideal of $\mathcal{R}$ with e_i as its identity, $i = 1, 2, \ldots, r$, and $\mathcal{R}$ has the direct sum decomposition

$$\mathcal{R} = \mathcal{R}_1 \dot{+} \mathcal{R}_2 \dot{+} \cdots \dot{+} \mathcal{R}_r\,.$$

(iv) *For any $i = 1, 2, \ldots, r$, the map*

$$\begin{aligned} \mathbb{Z}_4[X]/(f_i(X)) &\to \mathcal{R}_i = \mathcal{R}e_i \\ k(X) + (f_i(X)) &\mapsto (k(X) + (X^n - 1))\, e_i \end{aligned} \tag{7.17}$$

is an isomorphism of rings.

Proof. (i) is Proposition 5.11. (ii) follows from Lemma 7.11. (iii) follows from Theorem 7.14 (Sun Zi Theorem). (iv) follows from Corollary 7.15. □

We need the following general result.

Proposition 7.18. *Let R be a commutative ring and*

$$R = R_1 \dot{+} R_2 \dot{+} \cdots \dot{+} R_r$$

be a direct sum decomposition of R. Then

(i) *For each $i = 1, 2, \ldots, r$, let I_i be an ideal of R_i, then $I_1 + I_2 + \cdots + I_r$ is an ideal of R.*

(ii) *For any ideal I of R, let $I_i = I \cap R_i$, $i = 1, 2, \ldots, r$, then I_i is an ideal of R_i and $I = I_1 + I_2 + \cdots + I_r$.* □

The proof is immediate and is omitted.

Let us determine the ideals of $\mathbb{Z}_4[X]/(f_i(X))$ first.

Lemma 7.19. *Let $f(X)$ be a basic irreducible polynomial of degree m over $\mathbb{Z}_4$. Then the only ideals of $\mathbb{Z}_4[X]/(f(X))$ are (0), $(1 + (f(X)))$ and $(2 + (f(X)))$.*

Proof. We have the ring homomorphism

$$- : \mathbb{Z}_4[X]/(f(X)) \to \mathbb{Z}_2[X]/(\overline{f}(X))$$

$$a_0+a_1X+\cdots+a_{m-1}X^{m-1}+(f(X)) \mapsto \overline{a}_0+\overline{a}_1X+\cdots+\overline{a}_{m-1}X^{m-1}+(\overline{f}(X)),$$

(cf. (6.2)). Let I be a nonzero ideal of $\mathbb{Z}_4[X]/(f(X))$ and $g(X) + (f(X)) \in I$ for some $g(X) \notin (f(X))$ and $\deg g(X) < m$. Since $\overline{f}(X)$ is irreducible over $\mathbb{Z}_2$, the greatest common divisor

$$(\overline{g}(X), \overline{f}(X)) = 1 \quad \text{or} \quad \overline{f}(X) .$$

If $(\overline{g}(X), \overline{f}(X)) = 1$, then $\overline{g}(X)$ and $\overline{f}(X)$ are coprime in $\mathbb{Z}_2[X]$. By Lemma 5.1, $g(X)$ and $f(X)$ are coprime in $\mathbb{Z}_4[X]$. Thus there are polynomials $b(X)$ and $c(X) \in \mathbb{Z}_4[X]$ such that

$$b(X)\,g(X) + c(X)\,f(X) = 1 .$$

It follows that

$$b(X)\,g(X) \equiv 1 \pmod{f(X)} ,$$

which implies $1+(f(X)) \in I$. Consequently, $I = (1+(f(X)))$. If $(\overline{g}(X), \overline{f}(X)) = \overline{f}(X)$, then $\overline{g}(X) = 0$ and $g(X) = 2g_1(X)$, where $g_1(X) \in \mathbb{Z}_2[X]$. Clearly, $g_1(X)$ and $\overline{f}(X)$ are coprime in $\mathbb{Z}_2[X]$. By Lemma 5.1 $g_1(X)$ and $f(X)$ are coprime in $\mathbb{Z}_4[X]$. There are polynomials $b_1(X)$ and $c_1(X) \in \mathbb{Z}_4[X]$ such that

$$b_1(X)\,g_1(X) + c_1(X)\,f(X) = 1 .$$

It follows that

$$b_1(X)\,g_1(X) \equiv 1 \pmod{f(X)} .$$

Hence

$$b_1(X)\,g(X) \equiv 2 \pmod{f(X)},$$

which implies $2 + (f(X)) \in I$. Therefore $(2 + (f(X))) \subseteq I$. Because

$$(\mathbb{Z}_4[X]/(f(X)))/(2 + (f(X))) \simeq \mathbb{Z}_2[X]/(\overline{f}(X)),$$

which is a field, $(2 + (f(X)))$ is a maximal ideal of $\mathbb{Z}_4[X]/(f(X))$. Hence $I = (2 + (f(X)))$.

Lemma 7.20. *Let n be an odd positive integer and $X^n - 1 = f_1(X)\,f_2(X) \cdots f_r(X)$ be the unique factorization of X^n-1 into basic irreducible polynomials over $\mathbb{Z}_4$. Then under the isomorphism* (7.17), *the ideals* (0), $(1+(f_i(X)))$, *and* $(2 + (f_i(X)))$ *of* $\mathbb{Z}_4[X]/(f_i(X))$ *are mapped into* (0), $(\hat{f}_i(X) + (X^n - 1))$ *and* $(2\hat{f}_i(X) + (X^n - 1))$ *of* $\mathcal{R}_i = \mathcal{R}\,e_i$, *respectively.*

Proof. Under the isomorphism (7.17), we have

$$1 + (f_i(X)) \mapsto (1 + (X^n - 1))\,e_i\,.$$

By (7.16), $e_i = b_i(X)\,\hat{f}_i(X) + (X^n - 1)$. Therefore

$$1 + (f_i(X)) \mapsto b_i(X)\,\hat{f}_i(X) + (X^n - 1)\,.$$

Clearly, $b_i(X)\,\hat{f}_i(X) + (X^n - 1) \in (\hat{f}_i(X) + (X^n - 1))$. Multiplying both sides of (7.15) by $\hat{f}_i(X)$, we obtain

$$b_i(X)\,\hat{f}_i(X)\,\hat{f}_i(X) + c_i(X)\,(X^n - 1) = \hat{f}_i(X)\,.$$

Then

$$b_i(X)\,\hat{f}_i(X)\,\hat{f}_i(X) + (X^n - 1) = \hat{f}_i(X) + (X^n - 1)\,,$$

which implies $\hat{f}_i(X) + (X^n - 1) \in (b_i(X)\,\hat{f}_i(X) + (X^n - 1))$. Therefore $(b_i(X)\,\hat{f}_i(X) + (X^n - 1)) = (\hat{f}_i(X) + (X^n - 1))$ and the image of $(1 + (f_i(X)))$ under (7.17) is $(\hat{f}_i(X) + (X^n - 1))$.

Similarly, we can prove that the image of $(2 + (f_i(X)))$ under (7.17) is $(2\hat{f}_i(X) + (X^n - 1))$. □

At the beginning of Sec. 7.2 we adopted the convention that we write $a_0 + a_1X + \cdots + a_{n-1}X^{n-1}$ simply for $a_0 + a_1X + \cdots + a_{n-1}X^n + (X^n - 1)$. Then the ideal $(\hat{f}_i(X) + (X^n - 1))$ and $(2\hat{f}_i(X) + (X^n - 1))$ of $\mathcal{R}$ will be written simply as $(\hat{f}_i(X))$ and $(2\hat{f}_i(X))$ respectively.

From Propositions 7.17, 7.18 and Lemmas 7.19, 7.20 we deduce immediately

Proposition 7.21. *Let n be an odd positive integer, $X^n - 1 = f_1(X)\,f_2(X) \cdots f_r(X)$ be the unique factorization of $X^n - 1$ into basic irreducible polynomials, and $\hat{f}_i(X)$ be the product of all $f_j(X)$ except $f_i(X)$. Then any ideal of the ring $\mathcal{R}$ is a sum of some $(\hat{f}_i(X))$ and $(2\hat{f}_i(X))$.* □

Corollary 7.22. *The number of $\mathbb{Z}_4$-cyclic codes of odd length n is 3^r, where r is the number of basic irreducible polynomial factors in $X^n - 1$.* □

Theorem 7.23. *Let $2 \nmid n$ and I be an ideal of $\mathcal{R}$. Then these are unique monic polynomials $f(X)$, $g(X)$, and $h(X)$ over $\mathbb{Z}_4$ such that $I = (f(X)\,h(X),\, 2f(X)\, g(X))$, where $f(X)\,g(X)\,h(X) = X^n - 1$ and*

$$|I| = 4^{\deg g(X)}\, 2^{\deg h(X)}\,. \tag{7.18}$$

Proof. By Proposition 5.11, $X^n - 1$ has a unique factorization into basic irreducible polynomials: $X^n - 1 = f_1(X)\,f_2(X) \cdots f_r(X)$. By Proposition 7.21, I is a sum of some $(\hat{f}_i(X))$ and $(2\hat{f}_i(X))$. We abbreviate $f_i(X)$ and $\hat{f}_i(X)$ as f_i and $\hat{f}_i$, respectively. By rearranging $f_1, \ldots, f_r$, we can assume that

$$I = (\hat{f}_{k+1}) + (\hat{f}_{k+2}) + \cdots + (\hat{f}_{k+l}) + (2\hat{f}_{k+l+1}) + (2\hat{f}_{k+l+2}) + \cdots + (2\hat{f}_r)\,.$$

Then

$$I = (f_1 f_2 \cdots f_k f_{k+l+1}\, f_{k+l+2} \cdots f_r, \quad 2 f_1 f_2 \cdots f_k f_{k+1} f_{k+2} \cdots f_{k+l})\,.$$

Let

$$\begin{aligned} f(X) &= f_1 f_2 \cdots f_k\,, \\ g(X) &= f_{k+1} k_{k+2} \cdots f_{k+l}\,, \\ h(X) &= f_{k+l+1} f_{k+l+2} \cdots f_r\,, \end{aligned}$$

where we understand that

$$\begin{aligned} f(X) &= 1 \qquad \text{if } \ k = 0\,, \\ g(X) &= 1 \qquad \text{if } \ l = 0\,, \\ h(X) &= 1 \qquad \text{if } \ k + l = r\,. \end{aligned}$$

Then $I = (f(X)\,h(X),\, 2f(X)g(X))$ and $f(X)\,g(X)\,h(X) = X^n - 1$.

When $h(X) \neq 1$, we have $(f(X)h(X)) \cap (2f(X)g(X)) = (0)$ and then $I = (f(X)h(X)) \dot{+} (2f(X)g(X))$. Therefore

$$\begin{aligned}|I| &= |f(X)h(X)|\;|2f(X)g(X)| \\ &= 4^{n-\deg f(X)-\deg h(X)}\,2^{n-\deg f(X)-\deg g(X)} \\ &= 4^{\deg g(X)}\,2^{\deg h(X)}\,.\end{aligned}$$

When $h(X) = 1$, $I = (f(X),\, 2f(X)g(X)) = f(X)$. Then

$$|I| = 4^{n-\deg f(X)}\,.$$

Since $\deg h(X) = 0$, we also have (7.18). □

Corollary 7.24. *Let C be a $\mathbb{Z}_4$-cyclic code of odd length n and assume that $C = (f(X)h(X),\, 2f(X)g(X))$, where $f(X)$, $g(X)$ and $h(X)$ are monic polynomials over $\mathbb{Z}_4$ such that $f(X)g(X)h(X) = X^n - 1$. Then $C^\perp$ is also a $\mathbb{Z}_4$-cyclic code, $C^\perp = (\tilde{g}(X)\tilde{h}(X),\, 2\tilde{g}(X)\tilde{f}(X))$, and $|C^\perp| = 4^{\deg f(X)}2^{\deg h(X)}$.*

Proof. By the definition of $\mathbb{Z}_4$-cyclic codes it is easy to verify that $C^\perp$ is a $\mathbb{Z}_4$-cyclic code. By Proposition 7.9 $(f(X))^\perp = (\tilde{g}(X)\tilde{h}(X))$. Clearly, $C = (f(X)h(X),\ 2f(X)g(X)) \subseteq f(X)$. This implies $(f(X))^\perp \subseteq C^\perp$. Hence $(\tilde{g}(X)\tilde{h}(X)) \subseteq C^\perp$. Similarly $(2\tilde{g}(X)\tilde{f}(X)) \subseteq (\tilde{g}(X)) = (f(X)h(X))^\perp$ Clearly, $(2\tilde{g}(X)\tilde{f}(X)) \subseteq (2f(X)g(X))^\perp$. Thus $(2\tilde{g}(X)\tilde{f}(X)) \subseteq (f(X)h(X))^\perp \cap (2f(X)g(X))^\perp = C^\perp$. Therefore $(\tilde{g}(X)\tilde{h}(X),\, 2\tilde{g}(X)\tilde{f}(X)) \subseteq C^\perp$.

By Theorem 7.23,

$$|C| = 4^{\deg g(X)}2^{\deg h(X)}\,,$$

$$|(\tilde{g}(X)\tilde{h}(X),\, 2\tilde{g}(X)\tilde{f}(X))| = 4^{\deg\ \tilde{f}(X)}2^{\deg \tilde{h}(X)} = 4^{\deg f(X)}2^{\deg h(X)}\,,$$

and by Proposition 1.2,

$$\begin{aligned}|C^\perp| &= 4^{n-\deg g(X)-\deg h(X)}2^{\deg h(X)} \\ &= 4^{\deg\ f(X)}2^{\deg h(X)}\,.\end{aligned}$$

Therefore $C^\perp = (\tilde{g}(X)\tilde{h}(X),\, 2\tilde{g}(X)\tilde{f}(X))$. □

Theorem 7.25. *Let $2 \nmid n$. Then every ideal of $\mathcal{R}$ is of the form $(f_0(X),\, 2f_1(X))$, where $f_0(X)$ and $f_1(X)$ are monic divisors of $X^n - 1$ over $\mathbb{Z}_4$ and $f_1(X)|f_0(X)$.*

Proof. Let I be an ideal of $\mathcal{R}$. By Theorem 7.23 $I = (f(X)h(X), 2f(X)g(X))$ where $f(X)g(X)h(X) = X^n - 1$. $g(X)$ and $h(X)$ are coprime, from which we deduce easily that $I = (f(X)h(X), 2f(X))$. Let $f_0(X) = f(X)h(X)$ and $f_1(X) = f(X)$, then $I = (f_0(X), 2f_1(X))$ and $f_1(X)|f_0(X)$. □

Theorem 7.26. *Let $2 \nmid n$. Then every ideal of $\mathcal{R}$ is principal.*

Proof. Let I be an ideal of $\mathcal{R}$. By Theorem 7.25, $I = (f_0(X), 2f_1(X))$, where $f_0(X)$ and $f_1(X)$ are monic divisors of $X^n - 1$ over $\mathbb{Z}_4$ and $f_1(X)|f_0(X)$. Let $g(X) = f_0(X) + 2f_1(X)$. We assert that $I = (g(X))$. Clearly $(g(X)) \subseteq I$. Let $\hat{f}_0(X) = (X^n - 1)/f_0(X)$ and $\hat{f}_1(X) = f_0(X)/f_1(X)$. Then $\hat{f}_0(X)$, $\hat{f}_1(X)$ are coprime over $\mathbb{Z}_4$. We have $2f_1(X)\hat{f}_1(X) = 2f_0(X) = 2g(X) \in (g(X))$ and $2f_1(X)\hat{f}_0(X) = 2g(X)\hat{f}_0(X) \in (g(X))$. It follows that $2f_1(X) \in (g(X))$ and, hence, $f_0(X) \in (g(X))$. Therefore $I \subseteq (g(X))$. We conclude that $I = (g(X))$. □

We remark that the generating polynomial $g(X)$ of the ideal I in the proof of Theorem 7.26 is not necessarily a divisor of $X^n - 1$ in $\mathbb{Z}_4[X]$. For example, let $n = 3$, $f_0(X) = X - 1$, $f_1(X) = 1$, then $g(X) = X + 1$ and $g(X) \nmid X^3 - 1$.

The ring homomorphism

$$\begin{aligned} - : \mathbb{Z}_4 &\to \mathbb{Z}_2 \\ 0, 2 &\mapsto 0 \\ 1, 3 &\mapsto 1 \end{aligned}$$

can be extended to a ring homomorphism

$$\begin{aligned} &\mathcal{R} = \mathbb{Z}_4[X]/(X^n - 1) \to \mathbb{Z}_2[X]/(X^n - 1) \\ &a_0 + a_1X + \cdots + a_{n-1}X^{n-1} + (X^n - 1) \\ &\qquad \mapsto \overline{a}_0 + \overline{a}_1X + \cdots + \overline{a}_{n-1}X^{n-1} + (X^n - 1)\,, \end{aligned} \tag{7.19}$$

which will be denoted also by $-$ and the image of $f(X) \in \mathcal{R}$ will be denoted by $\overline{f}(X)$.

Proposition 7.27. *Let $2 \nmid n$ and $f(X)$ be a monic divisor of $X^n - 1$ in $\mathbb{Z}_4[X]$, then the principal ideal $(f(X))$ of $\mathbb{Z}_4[X]$ has a unique generating idempotent. Moreover, let $e_2(X)$ be the unique generating idempotent of the principal ideal*

$(\bar{f}(X))$ of $\mathbb{Z}_2[X]$ and $\theta(X) \in \mathbb{Z}_4[X]$ be such that $\bar{\theta}(X) = e_2(X)$, then $\theta(X)^2$ is the unique generating idempotent of $(f(X))$.

Proof. Let $g(X) = (X^n - 1)/f(X)$, then $g(X)$ is also a monic polynomial in $\mathbb{Z}_4[X]$ and $f(X)$ and $g(X)$ are coprime in $\mathbb{Z}_4[X]$. There are $u(X)$ and $v(X)$ in $\mathbb{Z}_4[X]$ such that

$$f(X)u(X) + g(X)v(X) = 1\,.$$

Set $\nu(X) = f(X)u(X)$, then $\nu(X) = 1 - g(X)v(X)$ and $\nu(X)^2 = \nu(X) - g(X)\nu(X)v(X) \equiv \nu(X) \pmod{X^n - 1}$. Thus $\nu(X)$ is an idempotent in $(f(X))$. But

$$\begin{aligned} f(X)\nu(X) &= f(X)\,(1 - g(X)\nu(X)) \\ &= f(X) - f(X)g(X)\nu(X) \\ &\equiv f(X) \pmod{X^n - 1}\,. \end{aligned}$$

Therefore $\nu(X)$ is the identity of $(f(X))$, i.e., $\nu(X)$ is the unique generating idempotent of $(f(X))$. Then $\bar{\nu}(X)$ is the unique generating idempotent of $(\bar{f}(X))$. Thus $\bar{\nu}(X) = e_2(X) = \bar{\theta}(X)$. We may write $\theta(X) = \nu(X) + 2b(X)$, where $b(X) \in \mathbb{Z}_4[X]$. Then $\theta(X)^2 = \nu(X)^2 = \nu(X)$. That is, $\theta(X)^2$ is the unique generating idempotent of $f(X)$. □

Let $f(X)$ be a monic divisor of $X^n - 1$ in $\mathbb{Z}_4[X]$. By Proposition 7.8 we may regard $(f(X))$ as a quaternary cyclic code. Then the generating idempotent of the ideal $(f(X))$ is also called the *generating idempotent* of the code.

Finally, we have

Proposition 7.28. *Let $2 \nmid n$, I_1 and I_2 be nonzero ideals of $\mathcal{R}$ with generating idempotents $e_1(X)$ and $e_2(X)$ respectively. Then $e_1(X)e_2(X)$ is the generating idempotent of $I_1 \cap I_2$ and $e_1(X) + e_2(X) - e_1(X)e_2(X)$ is that of $I_1 + I_2$. In particular, if $e_1(X)$ and $e_2(X)$ are orthogonal, then $e_1(X) + e_2(X)$ is the generating idempotent of $I_1 + I_2$.*

Proof. Same as Proposition 7.6. □

Proposition 7.29. *Let $2 \nmid n$, $f(X)$ be a monic divisor of $X^n - 1$ in $\mathbb{Z}_4[X]$ and the cyclic code C with generator polynomial $f(X)$ have the generating idempotent $e(X)$. Then $C^\perp$ has the generating idempotent $1 - e(X^{-1})$.*

Proof. Same as Proposition 7.7. □

Theorems 7.25 and 7.26 are due to Calderbank and Sloane (1995) and their proofs rest on the Lasker–Noether decomposition theorem of ideals in Noetherian rings. Theorem 7.23 is an equivalent form of Theorem 7.25 and is due to Pless and Qian (1996), and their proof is elementary and is adopted in this chapter. Proposition 7.27 is due to Bonnecaze *et al.* (1995), but the present proof is simpler. Corollary 7.24 and Propositions 7.28 and 7.29 are due to Pless and Qian (1996).

CHAPTER 8

KERDOCK CODES

8.1. The Quaternary Kerdock Codes

Let m be any integer ≥ 2 and $h(X)$ be a basic primitive polynomial of degree m over $\mathbb{Z}_4$ such that $h(X)|(X^{2^m-1}-1)$. The existence of such a polynomial $h(X)$ is guaranteed by Corollary 5.5. Clearly, $h(X)$ is the Hensel lift of the binary primitive polynomial $\overline{h}(X)$ of degree m.

Let $n = 2^m - 1$ and $g(X)$ be the reciprocal polynomial to the polynomial $(X^n - 1)/(X-1)h(X)$.

Definition 8.1. The *shortened quaternary Kerdock code* $\mathcal{K}(m)^-$ is the quaternary cyclic code of length $2^m - 1$ with generator polynomial $g(X)$. The positions of the coordinates of codewords of $\mathcal{K}(m)^-$ are numbered as 0, 1, $2, \dots, 2^m - 2$. The *quaternary Kerdock code* $\mathcal{K}(m)$ is the code obtained from $\mathcal{K}(m)^-$ by adding a zero-sum check symbol to each codeword of $\mathcal{K}(m)^-$ at position ∞, which is situated in front of the position 0. $\square$

In Sec. 8.3 we shall prove that when m is an odd integer ≥ 3, the binary image of $\mathcal{K}(m)$ is the Kerdock code K_{m+1} of length 2^{m+1} First, clearly we have

Proposition 8.1. *Let* $\deg g(X) = \delta$, *then* $\delta = 2^m - m - 2$. *Let*

$$g(X) = g_0 + g_1 X + \cdots + g_\delta X^\delta,$$

where $g_i \in \mathbb{Z}_4$, *and let* $g_\infty = -(g_0+g_1+\cdots+g_\infty)$, *then the following* $(m+1)\times 2^m$ *matrix*

$$\begin{pmatrix} g_\infty & g_0 & g_1 & \cdots & g_\delta & & & \\ g_\infty & & g_0 & g_1 & \cdots & g_\delta & & \\ \vdots & & & \ddots & & & \ddots & \\ g_\infty & & & & g_0 & g_1 & \cdots & g_\delta \end{pmatrix} \tag{8.1}$$

is a generator matrix of $\mathcal{K}(m)$. □

Proposition 8.2. *Let* ξ *be a root of* $h(X)$ *in some extension ring of* $\mathbb{Z}_4$, *for instance, in* $\mathrm{GR}(4^m)$. *Then the* $(m+1) \times 2^m$ *matrix*

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \xi & \xi^2 & \cdots & \xi^{n-1} \end{pmatrix} \tag{8.2}$$

is also a generator matrix of $\mathcal{K}(m)$, *where the entries* $\xi^j (0 \leq j \leq n-1)$ *in the second row of* (8.2) *are to be replaced by the corresponding* m *tuples* ${}^t(b_{ij}, b_{2j}, \ldots, b_{mj})$ *if* $\xi^j = b_{1j} + b_{2j}\xi + \cdots + b_{mj}\xi^{m-1}$.

Proof. Let C_1 be the $\mathbb{Z}_4$-linear codes of length $n = 2^m - 1$ with generator matrix

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \xi & \xi^2 & \cdots & \xi^{n-1} \end{pmatrix} \tag{8.3}$$

For $\mathbf{a} = (a_0, a_1, \ldots, a_{n-1}) \in \mathbb{Z}_4^n$, let $a(X) = a_0 + a_1 X + \cdots + a_{n-1}X^{n-1}$ Then $\mathbf{a} \in C_1^\perp$ if and only if $\sum_{i=0}^{n-1} a_i = 0$ and $a(\xi) = 0$. $\sum_{i=0}^{n-1} a_i = 0$ is equivalent to $(X-1)|a(X)$. Dividing $a(X)$ by $h(X)$, we obtain $a(X) = q(X)h(X) + r(X)$, where $\deg r(X) < \deg h(X) = m$. Substituting $X = \xi$ into this equation, we obtain that $a(\xi) = 0$ if and only if $r(\xi) = 0$. But $r(\xi) = 0$ implies $r(X) = 0$ by the uniqueness of the additive representation of elements of $\mathrm{GR}(4^m) = \mathbb{Z}_4[\xi]$. Therefore $a(\xi) = 0$ is equivalent to $h(X)|a(X)$. Hence $\mathbf{a} \in C_1^\perp$ if and only if $(X-1)|a(X)$ and $h(X)|a(X)$.

We assert further that $(X-1)|a(X)$ and $h(X)|a(X)$ if and only if $(X-1)h(X)|a(X)$. "If part" is trivial. Assume that $(X-1)|a(X)$ and $h(X)|a(X)$. Then $a(X) = q(X)h(X)$. Substituting $X = 1$ into this equation, we obtain $q(1)h(1) = a(1) = 0$. We assumed $m \geq 2$, so $\overline{h}(1) \neq 0$, thus $h(1)$ is an invertible element of $\mathrm{GR}(4^m)$. It follows that $q(1) = 0$ and $(X-1)|q(X)$. Therefore $(X-1)h(X)|a(X)$. Our assertion is proved.

From our assertion we deduce that $\mathbf{a} \in C_1^\perp$ if and only if $(X-1)h(X)|a(X)$. That is, $C_1^\perp$ is a cyclic code of length n with generator polynomial $(X-1)h(X)$. By Proposition 7.9, C_1 is a cyclic code of length n with the reciprocal polynomial to $(X^n-1)/(X-1)h(X)$ as the generator polynomial. But the reciprocal polynomial to $(X^n - 1)/(X - 1)h(X)$ is $g(X)$ and $\mathcal{K}(m)^-$ is defined to be the cyclic code of length n over $\mathbb{Z}_4$ with the generator polynomial $g(X)$. Therefore $\mathcal{C}_1 = \mathcal{K}(m)^-$ and (8.3) is a generator matrix of $\mathcal{K}(m)^-$. $\mathcal{K}(m)$ is the code obtained from $\mathcal{K}(m)^-$ by adding a zero-sum check symbol, hence $\mathcal{K}(m)$ has generator matrix (8.2). □

It follows immediately from Proposition 8.2 and Corollary 6.8 (iii), (iv) that different basic primitive polynomials of the same degree m over $\mathbb{Z}_4$ define permutation-equivalent quaternary Kerdock codes.

Example 8.1. Let $m = 2$ and $h(X) = X^2 + X + 1$ be the unique basic primitive polynomial of degree 2. Then $\tilde{g}(X) = (X^3 - 1)/(X - 1)h(X) = 1$ and $g(X) = 1$. By Proposition 8.1 $\mathcal{K}(2)$ has generator matrix

$$\begin{pmatrix} 3 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \end{pmatrix}.$$

By Proposition 8.2, $\mathcal{K}(2)$ has generator matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 3 \end{pmatrix}.$$

It is easy to verify that these two generator matrices generate the same linear $\mathbb{Z}_4$-code.

Example 8.2. Let $m = 3$ and $h(X) = X^3 + 2X^2 + X - 1$ be the basic primitive polynomial of degree 3. We find $g(X) = X^3 + 2X^2 + X - 1 = h(X)$. So $\mathcal{K}(3)^-$ is self-dual. It follows that $\mathcal{K}(3)$ is also self-dual. The generator matrices of $\mathcal{K}(3)$ given by Propositions 8.1 and 8.2 are

$$\begin{pmatrix} 1 & 3 & 1 & 2 & 1 & 0 & 0 & 0 \\ 1 & 0 & 3 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 3 & 1 & 2 & 1 & 0 \\ 1 & 0 & 0 & 0 & 3 & 1 & 2 & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 2 & 3 & 1 \\ 0 & 0 & 1 & 0 & 3 & 3 & 3 & 2 \\ 0 & 0 & 0 & 1 & 2 & 3 & 1 & 1 \end{pmatrix},$$

respectively. It is easy to prove that the above two matrices and the matrix (1.6) generate the same code. Therefore $\mathcal{K}(3)$ is the octacode $\mathcal{O}_8$. □

Corollary 8.3. *Both $\mathcal{K}(m)^-$ and $\mathcal{K}(m)$ are $\mathbb{Z}_4$-linear codes of types 4^{m+1}.* □

Corollary 8.4. *The binary Linear code $K^{(1)}$ associated with $\mathcal{K}(m)$ is equivalent to* RM(1, m).

Proof. If $\xi^j = b_{1j} + b_{2j}\xi + \cdots + b_{mj}\xi^{m-1}$, where $\xi_{ij} \in \mathbb{Z}_4$ then $\overline{\xi}^j = \overline{b}_{1j} + \overline{b}_{2j}\overline{\xi} + \cdots + \overline{b}_{mj}\overline{\xi}^{m-1}$ Therefore $K^{(1)}$ has

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \overline{\xi} & \overline{\xi}^2 & \cdots & \overline{\xi}^{n-1} \end{pmatrix} \tag{8.4}$$

as its generator matrix, where $\overline{\xi}^j (0 \leq j \leq n-1)$ should be replaced by the column ${}^t(\overline{b}_{1j}, \overline{b}_{2j}, \ldots, \overline{b}_{mj})$. Since $\overline{\xi}$ is a root of the primitive polynomial $\overline{h}(X)$, $\overline{\xi}$ is of order $n = 2^m - 1$. So, the columns, 0, 1, $\overline{\xi}, \ldots, \overline{\xi}^{n-1}$ are distinct in pairs and they are some rearrangement of all the 2^m m-dimensional column vectors over $\mathbb{F}_2$. Hence $K^{(1)}$ is equivalent to RM(1, m). □

8.2. Trace Descriptions of $\mathcal{K}(m)$

Proposition 8.5. *The codes $\mathcal{K}(m)^-$ and $\mathcal{K}(m)$ have the following trace descriptions over the ring* GR(4^m) $= \mathbb{Z}_4[\xi]$, *where ξ is a root of the basic primitive polynomial $h(X)$ in* GR(4^m).

(i) $\mathcal{K}(m)^- = \{\varepsilon 1^n + \mathbf{v}^{(\lambda)} | \varepsilon \in \mathbb{Z}_4,\ \lambda \in \mathbb{Z}_4[\xi]\}$, *where 1^n is the all 1 n-tuple and*

$$\mathbf{v}^{(\lambda)} = (T(\lambda\xi^0),\ T(\lambda\xi),\ T(\lambda\xi^2), \ldots, T(\lambda\xi^{n-1})). \tag{8.5}$$

(ii) $\mathcal{K}(m) = \{\varepsilon 1^{n+1} + \mathbf{u}^{(\lambda)} | \varepsilon \in \mathbb{Z}_4,\ \lambda \in \mathbb{Z}_4[\xi]\}$, *where 1^{n+1} is the all 1 $(n+1)$-tuple and*

$$\mathbf{u}^{(\lambda)} = (T(\lambda\xi^\infty),\ T(\lambda\xi^0),\ T(\lambda\xi), \ldots, T(\lambda\xi^{n-1})) \tag{8.6}$$

with the convention that $\xi^\infty = 0$.

Proof. (i) Let

$$C_2 = \{\varepsilon 1^n + \mathbf{v}^{(\lambda)} | \varepsilon \in \mathbb{Z}_4,\ \lambda \in \mathbb{Z}_4[\xi]\},$$

where $\mathbf{v}^{(\lambda)}$ is the vector (8.5). Under the correspondence (7.5), the vector $\varepsilon 1^n + \mathbf{v}^{(\lambda)}$ can be expressed as the polynomial

$$\varepsilon \sum_{i=0}^{n-1} X^i + \sum_{i=0}^{n-1} T(\lambda\xi^i)X^i.$$

First we prove that

$$\left(\sum_{i=0}^{n-1} X^i\right) (\widetilde{X-1}) \equiv 0 \,(\mathrm{mod}\, X^n - 1) \tag{8.7}$$

and

$$\left(\sum_{i=0}^{n-1} T(\lambda\xi^i)X^i\right) \tilde{h}(X) \equiv 0 \,(\mathrm{mod}\, X^n - 1)\,. \tag{8.8}$$

The first formula is clear, since

$$\begin{aligned}\left(\sum_{i=0}^{n-1} X^i\right) (\widetilde{X-1}) &= \left(\sum_{i=0}^{n-1} X^i\right)(1-X) \\ &= 1 - X^n \\ &\equiv 0 \,(\mathrm{mod}\, X^n - 1)\,.\end{aligned}$$

For the second formula, by the definition of generalized trace map from $\mathrm{GR}(4^m)$ to $\mathbb{Z}_4$ given in Sec. 6.3, we have

$$\begin{aligned}\sum_{i=0}^{n-1} T(\lambda\xi^i)X^i &= \sum_{i=0}^{n-1} \sum_{k=0}^{m-1} (\lambda\xi^i)^{f^k} X^{n-1} \\ &= \sum_{k=0}^{m-1} \lambda^{f^k} \sum_{i=0}^{n-1} (\xi^{f^k})^i X^i,\end{aligned}$$

and by Proposition 6.14 we have

$$h(X) = (X - \xi)(X - \xi^f)(X - \xi^{f^2}) \cdots (X - \xi^{f^{m-1}}).$$

Then

$$\tilde{h}(X) = (1 - \xi X)(1 - \xi^f X)(1 - \xi^{f^2} X) \cdots (1 - \xi^{f^{m-1}} X).$$

Since $(\xi^{f^k})^n = (\xi^n)^{f^k} = 1$, we have

$$\begin{aligned}\left(\sum_{i=0}^{n-1} (\xi^{f^k})^i X^i\right)(1 - \xi^{f^k} X) &= 1 - (\xi^{f^k})^n X^n \\ &= 1 - X^n \\ &\equiv 0 \,(\operatorname{mod} X^n - 1).\end{aligned}$$

Therefore we have (8.8). Consequently,

$$\left(\varepsilon \sum_{i=0}^{n-1} X^i + \sum_{i=0}^{n-1} T(\lambda\xi^i)X^i\right)\widetilde{(X-1)h}(X) \equiv 0 \,(\operatorname{mod} X^n - 1).$$

$\mathcal{K}(m)^-$ is the cyclic code with generator polynomial $g(X)$, which is the reciprocal polynomial to $(X^n - 1)/(X-1)h(X)$. It follows that the check polynomial of $\mathcal{K}(m)^-$ is $\widetilde{(X-1)h}(X)$. Therefore we have proved $\mathcal{C}_2 \subseteq \mathcal{K}(m)^-$. By Corollary 8.3, $|\mathcal{K}(m)^-| = 4^{m+1}$. If we can show that $|\mathcal{C}_2| = 4^{m+1}$, then $\mathcal{C}_2 = \mathcal{K}(m)^-$.

Suppose that $\varepsilon 1^n + \mathbf{v}^{(\lambda)} = \varepsilon' 1^n + \mathbf{v}^{(\lambda')}$, where $\varepsilon, \varepsilon' \in \mathbb{Z}_4$ and $\lambda, \lambda' \in \mathrm{GR}(4^m)$. Then $(\varepsilon - \varepsilon')1^n + \mathbf{v}^{(\lambda-\lambda')} = \mathbf{0}$. Thus

$$(\varepsilon - \varepsilon') \sum_{i=0}^{n-1} X^i + \sum_{i=0}^{n-1} T((\lambda - \lambda')\xi^i)X^i = 0.$$

By (8.8), we have $(\sum_{i=0}^{n-1} T((\lambda-\lambda')\xi^i)X^i)\,\tilde{h}(X) = 0$. Multiplying the above equation by $\tilde{h}(X)$, we obtain

$$(\varepsilon - \varepsilon')\left(\sum_{i=0}^{n-1} X^i\right)\tilde{h}(X) = 0. \tag{8.9}$$

Dividing $\tilde{h}(X)$ by $X - 1$, we have

$$\tilde{h}(X) = q(X)(X-1) + \tilde{h}(1), \tag{8.10}$$

where $q(X) \in \mathbb{Z}_4[X]$. By (8.7), $(\sum_{i=0}^{n-1} X^i)(X-1) = 0$. Substituting (8.10) into (8.9), we obtain

$$(\varepsilon - \varepsilon')\left(\sum_{i=0}^{n-1} X^i\right)\tilde{h}(1) = 0.$$

Since $\overline{\tilde{h}}(1) \neq 0$, $\tilde{h}(1)$ is an invertible element of $\mathbb{Z}_4$. It follows that $\varepsilon = \varepsilon'$ Then we have $\mathbf{v}^{(\lambda-\lambda')} = 0$. In particular,

$$T(\lambda - \lambda') = (\lambda - \lambda') + (\lambda - \lambda')^f + \cdots + (\lambda - \lambda')^{f^{m-1}} = 0,$$

$$T((\lambda - \lambda')\xi) = (\lambda - \lambda')\xi + (\lambda - \lambda')^f \xi^f + \cdots + (\lambda - \lambda')^{f^{m-1}} \xi^{f^{m-1}} = 0,$$

$$\vdots$$

$$\begin{aligned} T((\lambda - \lambda')\xi^{m-1}) &= (\lambda - \lambda')\xi^{m-1} + (\lambda - \lambda')^f (\xi^f)^{m-1} + \cdots \\ &+ (\lambda - \lambda')^{f^{m-1}} (\xi^{f^{m-1}})^{m-1} = 0. \end{aligned}$$

By definition of the generalized Frobenius map f of GR(4^m), $\xi^{f^i} = \xi^{2^i}$ for $i = 0, 1, \ldots, m-1$. By Proposition 6.16 (i), all $\xi^{f^i} - \xi^{f^j} = \xi^{2^i} - \xi^{2^j}$ ($0 \leq i$, $j \leq m-1$ and $i \neq j$) are invertible elements of GR(4^m). So, the van der Monde determinant

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \xi & \xi^f & \cdots & \xi^{f^{m-1}} \\ \vdots & \vdots & & \vdots \\ \xi^{m-1} & (\xi^f)^{m-1} & \cdots & (\xi^{f^{m-1}})^{m-1} \end{vmatrix}$$

is an invertible element of GR(4^m). It follows that $\lambda - \lambda' = 0$, i.e., $\lambda = \lambda'$. Therefore $|\mathcal{C}_2| = 4^{m+1}$.

(ii) follows from (i), since the zero-check sum for $\varepsilon 1^n$ is ε and for $\mathbf{v}^{(\lambda)}$ is 0. □

Furthermore, we have

Proposition 8.6. *Let m be an integer ≥ 2. Let $\mathbf{c} = (c_\infty, c_0, c_1, \ldots, c_{n-1})$ be an arbitrary codeword of $\mathcal{K}(m)$, then the 2-adic representation of c_t*

$$c_t = a_t + 2b_t, \quad t \in \{\infty, 0, 1, \ldots, n-1\}, \tag{8.11}$$

is given by

$$a_t = A + \mathrm{Tr}(\pi\overline{\xi}^t), \tag{8.12}$$

$$b_t = B + \mathrm{Tr}(\eta\overline{\xi}^t) + \sum_{0 \leq j < k \leq m-1} (\pi\overline{\xi}^t)^{2^j + 2^k}, \tag{8.13}$$

where the elements A, $B \in \mathbb{Z}_2$ and $\pi, \eta \in \mathbb{F}_{2^m}$ are arbitrary and we adopt the convention that $\bar{\xi}^{\infty} = 0$. When m is odd, let

$$Q(x) = \sum_{j=1}^{(m-1)/2} \mathrm{Tr}(x^{1+2^j}) \quad \text{for all} \quad x \in \mathbb{F}_{2^m},$$

then b_t can be written as

$$b_t = B + \mathrm{Tr}(\eta\bar{\xi}^t) + Q(\pi\bar{\xi}^t). \tag{8.14}$$

Proof. By Proposition 8.5, there is a unique $\varepsilon \in \mathbb{Z}_4$ and a unique $\lambda \in \mathbb{Z}_4[\xi]$ such that

$$c_t = \varepsilon + T(\lambda\xi^t), \quad t \in \{\infty, 0, 1, \dots, n-1\}$$

with the convention that $\xi^{\infty} = 0$. Let the 2-adic representation of λ be $\lambda = \xi^r + 2\xi^s$, $r, s \in \{\infty, 0, 1, \dots, n-1\}$, then

$$c_t = \varepsilon + T(\xi^{r+t}) + 2T(\xi^{s+t}) = a_t + 2b_t.$$

Since $a_t, b_t = 0$ or 1, applying the map $-$, we obtain

$$a_t = A + \mathrm{Tr}(\pi\bar{\xi}^t),$$

where $A = \bar{\varepsilon}$ and $\pi = \bar{\xi}^r$ There remains to compute b_t. Clearly, $c_t^2 = a_t^2 = a_t$. Therefore

$$\begin{aligned} 2b_t &= c_t - c_t^2 \\ &= (\varepsilon - \varepsilon^2) + (T(\xi^{r+t}) - (T(\xi^{r+t}))^2) + 2\varepsilon T(\xi^{r+t}) + 2T(\xi^{s+t}). \end{aligned}$$

It is clear that

$$\varepsilon - \varepsilon^2 = 2\beta(\varepsilon).$$

We compute

$$\begin{aligned} T(\xi^{r+t}) - (T(\xi^{r+t}))^2 &= T(\xi^{r+t})(1 - T(\xi^{r+t})) \\ &= (\xi^{r+t} + (\xi^{r+t})^2 + (\xi^{r+t})^{2^2} + \cdots + (\xi^{r+t})^{2^{m-1}}) \\ &\quad \times (1 - \xi^{r+t} - (\xi^{r+t})^2 - (\xi^{r+t})^{2^2} - \cdots - (\xi^{r+t})^{2^{m-1}}) \\ &= 2 \sum_{0 \le j < k \le m-1} (\xi^{r+t})^{2^j + 2^k}, \end{aligned}$$

$$2\varepsilon T(\xi^{r+t}) + 2T(\xi^{s+t}) = 2T((\varepsilon\xi^r + \xi^s)\,\xi^t).$$

Then

$$2b_t = 2\beta(\varepsilon) + 2T((\varepsilon\xi^r + \xi^s)\xi^t) + 2 \sum_{0\le j<k\le m-1} (\xi^{r+t})^{2^j+2^k}$$

Thus

$$b_t = \pm \left(\beta(\varepsilon) + T((\varepsilon\xi^r + \xi^s)\xi^t) + \sum_{0\le j<k\le m-1} (\xi^{r+t})^{2^j+2^k} \right)$$

Since $b_t = 0$ or 1, we have $b_t = \overline{b}_t$. Therefore we have (8.13)

$$b_t = B + \mathrm{Tr}(\eta\overline{\xi}^{\,t}) + \sum_{0\le j<k\le m-1} (\pi\overline{\xi}^{\,t})^{2^j+2^k},$$

where $B = \beta(\varepsilon)$, $\eta = \overline{\varepsilon}\overline{\xi}^{\,r} + \overline{\xi}^{\,s}$, and $\pi = \overline{\xi}^{\,r}$. When m is odd, we have

$$\begin{aligned} Q(\pi\overline{\xi}^{\,t}) &= \sum_{j=1}^{(m-1)/2} \mathrm{Tr}(\pi\overline{\xi}^{\,t})^{1+2^j} \\ &= \sum_{0\le j<k\le m-1} (\pi\overline{\xi}^{\,t})^{2^j+2^k}. \end{aligned}$$

Therefore we have (8.14)

$$b_t = B + \mathrm{Tr}(\eta\overline{\xi}^{\,t}) + Q(\pi\overline{\xi}^{\,t}). \qquad \square$$

8.3. The Kerdock Codes

Let m be an integer ≥ 2. Denote the binary image of the quaternary Kerdock code $\mathcal{K}(m)$ by $K(m)$, i.e., $K(m) = \phi(\mathcal{K}(m))$. First we have

Theorem 8.7. *Let m be an integer ≥ 2. Then $K(m)$ is a nonlinear binary code of length 2^{m+1} and with 4^{m+1} codewords. This code is distance invariant and all its codewords are of even weight.*

Proof. It is clear that $K(m)$ is of length 2^{m+1}. Since $|K(m)| = |\mathcal{K}(m)|$ and $\mathcal{K}(m)$ is of type 4^{m+1}, $|K(m)| = 4^{m+1}$. The distance invariance of $K(m)$ follows from Theorem 3.6.

Since $\mathcal{K}(m)$ is obtained from $\mathcal{K}(m)^-$ by adding a zero-sum check symbol to each codeword of $\mathcal{K}(m)^-$, by Proposition 3.4 all codewords of $K(m)$ are of even weight.

There remains to prove that $K(m)$ is nonlinear. By Proposition 8.5 for any $\lambda,\ \mu \in \mathbb{Z}_4[\xi]$,

$$\mathbf{u}^{(\lambda)} = (T(\lambda\xi^\infty),\ T(\lambda\xi^0), T(\lambda\xi), \ldots, T(\lambda\xi^{n-1}))$$

and

$$\mathbf{u}^{(\mu)} = (T(\mu\xi^\infty),\ T(\mu\xi^0), T(\mu\xi), \ldots, T(\mu\xi^{n-1}))$$

are codewords of $\mathcal{K}(m)$. If we can show that $2\alpha(\mathbf{u}^{(\lambda)}) * \alpha(\mathbf{u}^{(\mu)}) \notin \mathcal{K}(m)$, for some $\lambda, \mu \in \mathbb{Z}_4[\xi]$, where $*$ denotes the componentwise product, then the nonlinearity of $K(m)$ will follow from Proposition 3.16.

First we give the following remark. We know that the map

$$\begin{aligned}\mathrm{Tr} : \mathbb{F}_{2^m} &\to \mathbb{F}_2\\ \pi &\to \mathrm{Tr}\pi = \pi + \pi^2 + \cdots + \pi^{2^{m-1}}\end{aligned}$$

is a surjective homomorphism from the additive group of $\mathbb{F}_{2^m}$ to $\mathbb{F}_2$ and that for any $\pi \in \mathbb{F}_{2^m}^*$, $\pi\bar{\xi}^\infty$, $\pi\bar{\xi}^0$, $\pi\bar{\xi}, \ldots, \pi\bar{\xi}^{n-1}$ are all the 2^m elements of $\mathbb{F}_{2^m}$. Therefore the number of 1's and the number of 0's in the binary vector

$$(\mathrm{Tr}(\pi\bar{\xi}^\infty),\ \mathrm{Tr}(\pi\bar{\xi}^0),\ \mathrm{Tr}(\pi\bar{\xi}), \ldots, \mathrm{Tr}(\pi\bar{\xi}^{n-1}))$$

are all equal to 2^{m-1}, so are the number of 1's and the number of 0's in the binary vector

$$1^{n+1} + (\mathrm{Tr}(\pi\bar{\xi}^\infty),\ \mathrm{Tr}(\pi\bar{\xi}^0),\ \mathrm{Tr}(\pi\bar{\xi}), \ldots, \mathrm{Tr}(\pi\bar{\xi}^{n-1})).$$

Since $\mathrm{Tr} : \mathbb{F}_{2^m} \to \mathbb{F}_2$ is a surjective homomorphism, there are k and $l\,(0 \leq k,\ l \leq n-1)$ such that $\mathrm{Tr}(\bar{\xi}^k) = 1$ and $\mathrm{Tr}(\bar{\xi}^l) = 0$. Let us choose $\lambda = \xi^k$ and $\mu = \xi^l$. Suppose that $\mathbf{c} = 2\alpha(\mathbf{u}^{(\lambda)}) * \alpha(\mathbf{u}^{(\mu)}) \in \mathcal{K}(m)$.

Let $\mathbf{c} = (c_\infty, c_0, c_1, \ldots, c_{n-1})$ and $c_t = a_t + 2b_t$ be the 2-adic representation of c_t, $t \in \{\infty, 0, 1, \ldots, n-1\}$. by Proposition 8.6 these exist $A, B \in \mathbb{F}_2$ and $\pi, \eta \in \mathbb{F}_{2^m}$ such that (8.12) and (8.13) hold, i.e.,

$$a_t = A + \mathrm{Tr}(\pi\bar{\xi}^t),$$

$$b_t = B + \mathrm{Tr}(\eta\bar{\xi}^t) + \sum_{0 \leq j < k \leq m-1} (\pi\bar{\xi}^t)^{2^j + 2^k},\ t \in \{\infty, 0, 1, \ldots, n-1\}.$$

Clearly, $a_t = 0$ for all t. From the above remark we deduce $A = 0$ and $\pi = 0$. Then $c_t = 2b_t$ and

$$b_t = B + \mathrm{Tr}(\eta\bar{\xi}^t),\ t \in \{\infty, 0, 1, \ldots, n-1\}. \tag{8.15}$$

Let $\mathbf{b} = (b_\infty, b_0, b_1, \ldots, b_{n-1})$, then $\mathbf{c} = 2\mathbf{b}$. But $\mathbf{c} = 2\alpha(\mathbf{u}^{(\lambda)}) * \alpha(\mathbf{u}^{(\mu)})$. Therefore $\mathbf{b} = \alpha(\mathbf{u}^{(\lambda)}) * \alpha(\mathbf{u}^{(\mu)})$, we have

$$\alpha(\mathbf{u}^{(\lambda)}) = (\mathrm{Tr}(\bar{\xi}^k\bar{\xi}^\infty),\ \mathrm{Tr}(\bar{\xi}^k\bar{\xi}^0),\ \mathrm{Tr}(\bar{\xi}^k\bar{\xi}), \ldots, \mathrm{Tr}(\bar{\xi}^k\bar{\xi}^{n-1})),$$

$$\alpha(\mathbf{u}^{(\mu)}) = (\mathrm{Tr}(\bar{\xi}^l\bar{\xi}^\infty),\ \mathrm{Tr}(\bar{\xi}^l\bar{\xi}^0),\ \mathrm{Tr}(\bar{\xi}^l\bar{\xi}), \ldots, \mathrm{Tr}(\bar{\xi}^l\bar{\xi}^{n-1})).$$

By the above remark, the number of 1's and the number of 0's in both $\alpha(\mathbf{u}^{(\lambda)})$ and $\alpha(\mathbf{u}^{(\mu)})$ are equal to 2^{m-1}. We have $\mathrm{Tr}(\bar{\xi}^k\bar{\xi}^\infty) = \mathrm{Tr}(\bar{\xi}^l\bar{\xi}^\infty) = 0$, which implies $\mathbf{b} = \alpha(\mathbf{u}^{(\lambda)}) * \alpha(\mathbf{u}^{(\mu)}) \neq \mathbf{0}$. We also have $\mathrm{Tr}(\bar{\xi}^k\bar{\xi}^0) = \mathrm{Tr}(\bar{\xi}^k) = 1$ and $\mathrm{Tr}(\bar{\xi}^l\bar{\xi}^0) = \mathrm{Tr}(\bar{\xi}^l) = 0$, which implies the number of 1's in $\mathbf{b} = \alpha(\mathbf{u}^{(\lambda)}) * \alpha(\mathbf{u}^{(\mu)})$ is less than 2^{m-1} Again by the above remark, from (8.15) it follows that $B = 0$ and $\eta = 0$. Thus $\mathbf{b} = \mathbf{0}$. We get a contradiction. □

Proposition 8.8. *Let m be an integer ≥ 2. Then the binary image of the linear subcode of $\mathcal{K}(m)$ with generator matrix*

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \ldots & 1 \\ 0 & 2 & 2\xi & 2\xi^2 & \ldots & 2\xi^{n-1} \end{pmatrix} \tag{8.16}$$

is the first-order Reed–Muller code RM(1, $m+1$) *contained in $K(m)$. This linear subcode of $\mathcal{K}(m)$ consists of those codewords* $\mathbf{c}$ *for which $\lambda \in 2\mathrm{GR}(4^m)$ in the trace description* (8.6) *and for which $\pi = 0$ in the 2-adic representation* (8.11)–(8.13).

Proof. Denote the linear subcode of $\mathcal{K}(m)$ with generator matrix (8.16) by C_3. We have

$$\varphi(1^{2^m}) = (0^{2^m}, 1^{2^m})$$

$$\varphi(2^{2^m}) = (1^{2^m}, 1^{2^m})$$

and

$$\varphi(0, 2, 2\xi, 2\xi^2, \ldots, 2\xi^{n-1}) = (0, 1, \bar{\xi}, \bar{\xi}^2, \ldots, \bar{\xi}^{n-1}, 0, 1, \bar{\xi}, \bar{\xi}^2, \ldots, \bar{\xi}^{n-1}).$$

Then $\varphi(\mathcal{C}_3)$ has generator matrix

$$\begin{pmatrix} \varphi(2^{2^m}) \\ \varphi(0,2,2\xi,\dots,2\xi^{n-1}) \\ \varphi(1^{2^m}) \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \overline{\xi} & \overline{\xi}^2 & \cdots & \overline{\xi}^{n-1} & 0 & 1 & \overline{\xi} & \overline{\xi}^2 & \cdots & \overline{\xi}^{n-1} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 1 & 1 & 1 & \cdots & 1 \end{pmatrix}$$

Therefore $\varphi(\mathcal{C}_3) = \mathrm{RM}(1, m+1)$.

Next let us prove the second statement. By Proposition 8.2 any codeword $\mathbf{c}$ of $\mathcal{K}(m)$ can be expressed uniquely in the form

$$\varepsilon 1^{2^m} + (a_1, a_2, \dots, a_m)\,(0 \;\, 1 \;\, \xi \;\, \xi^2 \dots \xi^{n-1}),$$

where $\varepsilon, a_1, a_2, \dots, a_m \in \mathbb{Z}_4$. By Proposition 8.5, $\mathbf{c}$ can also be expressed uniquely in the form

$$\varepsilon 1^{2^m} + \mathbf{u}^{(\lambda)},$$

where $\varepsilon \in \mathbb{Z}_4$, $\lambda \in \mathrm{GR}(4^m)$, and $\mathbf{u}^{(\lambda)}$ is (8.6). Thus there is a bijective map

$$\mathbb{Z}_4^{m+1} \to \mathbb{Z}_4 \times \mathrm{GR}(4^m)$$
$$(\varepsilon, a_1, a_2, \dots, a_m) \to (\varepsilon, \lambda).$$

It is easy to verify that this map is an additive group isomorphism. A codeword $\mathbf{c}$ of $\mathcal{C}_3$ can be expressed uniquely as

$$\begin{aligned} &\varepsilon 1^{2^m} + (a_1, a_2, \dots, a_m)\,(0 \;\, 2 \;\, 2\xi \;\, 2\xi^2 \cdots \, 2\xi^{n-1}) \\ &\quad = \varepsilon 1^{2^m} + (2a_1, 2a_2, \dots, 2a_m)\,(0 \;\, 1 \;\, \xi \;\, \xi^2 \cdots \, \xi^{n-1}), \end{aligned}$$

and hence, can be expressed uniquely as

$$\varepsilon 1^{2^m} + \mathbf{u}^{(2\lambda)}.$$

That is, $\mathcal{C}_3$ consists of those codewords $\mathbf{c}$ for which $\lambda \in 2\mathrm{GR}(4^m)$ in the trace description (8.6). As in the proof of Proposition 8.6, let the 2-adic representation of λ be $\lambda = \xi^r + 2\xi^s$, then $r = \infty$ and hence $\pi = \overline{\xi}^r = \overline{\xi}^\infty = 0$. Therefore $\mathcal{C}_3$ consists of those codewords $\mathbf{c}$ for which $\pi = 0$ in the 2-adic representation (8.11)–(8.13). □

From now on we assume that m is an odd integer and ≥ 3. We recall that the Kerdock code K_{m+1} of length 2^{m+1} is the binary code which consists of

RM(1, $m+1$) together with $2^m - 1$ cosets of RM(2, $m+1$) relative to RM(1, $m+1$) with coset representatives

$$(L_\pi(\bar{\xi}^\infty),\, L_\pi(\bar{\xi}^0), \ldots, L_\pi(\bar{\xi}^{n-1}),\, R_\pi(\bar{\xi}^\infty),\, R_\pi(\bar{\xi}^0), \cdots, R_\pi(\bar{\xi}^{n-1})),$$

where π runs through $\mathbb{F}_{2^m}^*$,

$$L_\pi(\bar{\xi}^j) = \sum_{i=1}^{(m-1)/2} \mathrm{Tr}(\pi\bar{\xi}^j)^{1+2^i},$$

$$R_\pi(\bar{\xi}^j) = \sum_{i=1}^{(m-1)/2} \mathrm{Tr}(\pi\bar{\xi}^j)^{1+2^i} + \mathrm{Tr}(\pi\bar{\xi}^j), \quad j \in \{0, 1, \ldots, 2^m - 1\},$$

(cf. MacWilliams and Sloane (1977), Chap. 15, §5). Then we have

Theorem 8.9. *Let m be odd and ≥ 3. Then $K(m) = K_{m+1}$.*

Proof. Let $\mathbf{c}$ be any codeword of $\mathcal{K}(m)$ and $\mathbf{c} = \mathbf{a} + 2\mathbf{b}$ be its 2-adic representation. By Proposition 8.6 these are elements $A, B \in \mathbb{Z}_2$ and $\pi, \eta \in \mathbb{F}_{2^m}$ such that

$$a_t = A + \mathrm{Tr}(\pi\bar{\xi}^t),$$
$$b_t = B + \mathrm{Tr}(\eta\bar{\xi}^t) + Q(\pi\bar{\xi}^t), \quad t \in \{\infty, 0, 1, \ldots, n-1\},$$

where

$$Q(x) = \sum_{j=1}^{(m-1)/2} \mathrm{Tr}(x^{1+2^j}), \quad \text{for all} \quad x \in \mathbb{F}_{2^m}.$$

Then

$$\phi(\mathbf{c}) = (\beta(\mathbf{c}),\ \gamma(\mathbf{c})) = (\mathbf{b}, \mathbf{a} + \mathbf{b}).$$

Let

$$\mathbf{u} = B1^{2^m} + (\mathrm{Tr}(\eta\bar{\xi}^\infty),\, \mathrm{Tr}(\eta\bar{\xi}^0), \ldots, \mathrm{Tr}(\eta\bar{\xi}^{n-1})),$$
$$\mathbf{v} = A1^{2^m},$$

then $\mathbf{u} \in \mathrm{RM}(1, m)$, $\mathbf{v} \in \mathrm{RM}(0, m)$. By the $|u|u+v|$ construction,

$$(\mathbf{u}, \mathbf{u} + \mathbf{v}) \in \mathrm{RM}(1, m+1).$$

Therefore the codeword $\phi(\mathbf{c})$ and

$$(Q(\pi\bar{\xi}^\infty),\, Q(\pi\bar{\xi}^0), \ldots, Q(\pi\bar{\xi}^{n-1}),$$
$$\mathrm{Tr}(\pi\bar{\xi}^\infty) + Q(\pi\bar{\xi}^\infty),\ \mathrm{Tr}(\pi\bar{\xi}^0) + Q(\pi\bar{\xi}^0), \ldots, \mathrm{Tr}(\pi\bar{\xi}^{n-1}) + Q(\pi\bar{\xi}^{n-1}))$$

belong to the same coset of RM $(2, m+1)$ relative to RM $(1, m+1)$. Clearly,

$$Q(\pi\bar{\xi}^j) = L_\pi(\bar{\xi}^j),$$

$$\mathrm{Tr}(\pi\bar{\xi}^j) + Q(\pi\bar{\xi}^j) = R_\pi(\bar{\xi}^j).$$

Therefore $K(m) \subset K_{m+1}$. But the number of codewords of $K(m)$ and K_{m+1} are both equal to 4^{m+1}. Therefore $K(m) = K_{m+1}$. □

By Examples 3.4 and 8.2 the Nordstrom–Robinson code is the binary image of the quaternary Kerdock code $\mathcal{K}(3)$. The Kerdock codes K_{m+1} ($m \geq 3$ and m is odd) were introduced by Kerdock (1972). They are binary nonlinear codes which contains at least twice as many codewords as the best binary linear code with the same length and minimum distance. Nechaev (1989) used Galois rings and trace descriptions of some $\mathbb{Z}_4$-sequences to study the Kerdock codes. He proved that the Kerdock code punctured in two coordinates may be constructed as a family of segments of highest binary coordinates of some linear recursive sequences family over $\mathbb{Z}_4$ and that this code has the cyclic form, see also Hammons *et al.* (1994).

In preparing this chapter, Nechaev (1989) and Hammons *et al.* (1994) are helpful.

8.4. Weight Distributions of the Kerdock Codes

The weight distribution of the Kerdock code K_{m+1}, where m is an odd integer ≥ 3, was computed by Kerdock (1972) and can be found in MacWilliams and Sloane (1977), Table 15.7. Regarding K_{m+1} as the binary image of the quaternary Kerdock code $\mathcal{K}(m)$, Hammons *et al.* (1994) computed its weight distribution as follows.

Proposition 8.10. *The binary Kerdock code K_{m+1} of length 2^{m+1} (m odd $\geq$ 3) has the following weigth distribution*

Table 8.1. Weight distribution of K_{m+1} (m odd ≥ 3).

Weight	No. of codewords
0	1
$2^m - 2^{(m-1)/2}$	$2^{m+1}(2^m - 1)$
2^m	$2^{m+2} - 2$
$2^m + 2^{(m+1)/2}$	$2^{m+1}(2^m - 1)$
2^{m+1}	1

We need the following lemmas.

Lemma 8.11. *Denote the Galois ring* $\mathrm{GR}(4^m)$ *simply by* R *and the set of invertible elements of* R *by* R^*. *Then*

$$\sum_{\nu\in R} i^{T(\nu)} = \sum_{i\in R\backslash R^*} i^{T(\nu)} = \sum_{R^*} i^{T(\nu)} = 0.$$

Proof. Consider the generalized trace map

$$T : \mathrm{GR}(4^m) \to \mathbb{Z}_4$$
$$\lambda \to T(\lambda)$$

defined in Sec. 6.3. By Proposition 6.13, T is a surjective additive group homomorphism. It follows that as λ runs through $\mathrm{GR}(4^m)$, $T(\lambda)$ takes the values 0, 1, 2, 3 equally often. But $i^0 + i^1 + i^2 + i^3 = 0$. Therefore

$$\sum_{\nu\in R} i^{T(\nu)} = 0\,. \tag{8.17}$$

If we restrict T to the ideal $(2) = R\backslash R^*$, we get a surjective group homorphism $T : (2) \to \{0,2\}$ and, hence, $T(\lambda)$ takes the values 0 and 2 equally often. But $i^0 + i^2 = 0$. So,

$$\sum_{\nu\in R\backslash R^*} i^{T(\nu)} = 0. \tag{8.18}$$

From (8.17) and (8.18) we deduce

$$\sum_{\nu\in R^*} i^{T(\nu)} = 0. \tag{8.19}$$

Lemma 8.12. *The diophantine equation* $X^2 + Y^2 = 2^m$ *(*m *odd and* ≥ 3*) has a unique solution* $(2^{(m-1)/2}, 2^{(m-2)/2})$.

Proof. Let (x, y) be a solution, where x and y are non-negative integers. Write $x = 2^{r_1}(2x_1 + 1)$ and $y = 2^{r_2}(2y_1 + 1)$, where x_1 and y_1 are non-negative integers. Then

$$2^{2r_1}(2^2x_1^2 + 2^2x_1 + 1) + 2^{2r_2}(2^2y_1^2 + 2^2y_1 + 1) = 2^m.$$

If $r_1 > r_2$, then

$$2^{2(r_1-r_2)}(2^2x_1^2 + 2^2x_1 + 1) + 2^2y_1^2 + 2^2y_1 + 1 = 2^{m-2r_2},$$

which is impossible. So, $r_1 = r_2$ and then

$$2^2(x_1^2 + x_1 + y_1^2 + y_1) + 2 = 2^{m-2r_1},$$

which implies $m - 2r_1 = 1$ and $x_1 = y_1 = 0$. Therefore $x = y = 2^{(m-1)/2}$ □

Proof of Proposition 8.10. By Proposition 8.8, the codewords $\mathbf{u}^{(\lambda)} \in \mathcal{K}(m)$ for which $\lambda \in 2R$ form a first-order Reed–Muller code RM$(1, m+1)$. We know that the weight distribution of RM$(1, m+1)$ is

Table 8.2. Weight distribution of RM(1, $m+1$).

Weight	No. of codewords
0	1
2^m	$2^{m+2} - 2$
2^{m+1}	1

Now consider the codeword $\mathbf{v}^{(\lambda)} = T(\lambda\xi^0), T(\lambda\xi), \ldots, T(\lambda\xi^{n-1}) \in \mathcal{K}(m)^-$, where $\lambda \in R^*$. Let $w_a = w_a(\mathbf{v}^{(\lambda)})$, where $a \in \mathbb{Z}_4$. We claim that there exist $\delta_1, \delta_2 = \pm 1$ such that

$$\begin{aligned} w_0 &= 2^{m-2} - 1 + \delta_1 2^{(m-3)/2}, \quad w_1 = 2^{m-2} + \delta_2 2^{(m-3)/2}, \\ w_2 &= 2^{m-2} - \delta_1 2^{(m-3)/2}, \quad w_3 = 2^{m-2} - \delta_2 2^{(m-3)/2}. \end{aligned} \tag{8.20}$$

Let

$$S = \sum_{j=0}^{2^m-2} i^{T(\lambda\xi^j)},$$

then

$$S = w_0 - w_2 + i(w_1 - w_3) \tag{8.21}$$

and

$$|S|^2 = 2^m - 1 + \sum_{j \neq k} i^{T(\lambda(\xi^j - \xi^k))}.$$

By Proposition 6.16 (i) and (iii), $\xi^j - \xi^k$ $(0 \leq j, k \leq 2^m - 2, j \neq k)$ are distinct invertible elements of GR(4^m). They are $(2^m - 1)(2^m - 2) =$

$4^m - 3\cdot 2^m + 2$ in number. By Proposition 6.16 (ii) the other invertible elements of $\mathrm{GR}(4^m)$ are $\pm\xi^j$ $(0 \le j \le 2^m - 2)$. Therefore

$$\sum_{j\neq k} i^{T(\lambda(\xi^j - \xi^k))} = \sum_{\nu\in R^*} i^{T(\nu)} - \sum_{j=0}^{2^m-2} i^{T(\lambda\xi^j)} - \sum_{k=0}^{2^m-2} i^{T(-\lambda\xi^k)}.$$

By Lemma 8.11,

$$\sum_{\nu\in R^*} i^{T(\nu)} = 0.$$

Obviously,

$$\sum_{j=0}^{2^m-2} i^{T(\lambda\xi^j)} = S$$

and

$$\sum_{k=0}^{2^m-2} i^{T(-\lambda\xi^k)} = \overline{S}.$$

Therefore

$$|S|^2 = 2^m - 1 - S - \overline{S}.$$

It follows that

$$(S+1)(\overline{S}+1) = 2^m.$$

Substituting (8.21) into the above equation, we get

$$(w_0 - w_2 + 1)^2 + (w_1 - w_3)^2 = 2^m.$$

By Lemma 8.12, we must have

$$w_0 - w_2 = -1 \pm 2^{(m-1)/2}, \tag{8.22}$$

$$w_1 - w_3 = \pm 2^{(m-1)/2}. \tag{8.23}$$

On the other hand, $\overline{\lambda} = 1$ for $\lambda \in R^*$ and then $\overline{\mathbf{v}^{(\lambda)}} = (\mathrm{Tr}(1), \mathrm{Tr}(\overline{\xi}), \dots, \mathrm{Tr}(\overline{\xi}^n))$. Since $\mathrm{Tr} : \mathbb{F}_{2^m} \to \mathbb{F}_2$ is a surjective homomorphism, we have

$$w_1 + w_3 = 2^{m-1}, \tag{8.24}$$

$$w_0 + w_2 = 2^{m-1} - 1. \tag{8.25}$$

Then (8.20) follows from (8.22)-(8.25).

Now we consider the four codewords of $\mathcal{K}(m)$ obtained from $\varepsilon 1^n + \mathbf{v}^{(\lambda)} \in \mathcal{K}(m)^-$ ($\varepsilon = 0, 1, 2, 3$), where $\lambda \in R^*$, by appending the zero-sum check symbol ε. For the weights of the codeword $1^{n+1} + \mathbf{u}^{(\lambda)} \in \mathcal{K}(m)$, we have

$$w_1 = 2^{m-2} + \delta_1 2^{(m-3)/2}, \quad w_2 = 2^{m-2} + \delta_2 2^{(m-3)/2},$$

$$w_3 = 2^{m-2} - \delta_1 2^{(m-3)/2}, \quad w_0 = 2^{m-2} - \delta_2 2^{(m-3)/2}.$$

Thus $1^{n+1} + \mathbf{u}^{(\lambda)}$ is a codeword of Lee weight

$$w_1 + w_3 + 2w_2 = 2^m + \delta_2\, 2^{(m-1)/2}.$$

Similarly, $2^{n+1} + \mathbf{u}^{(\lambda)}$, $3^{n+1} + \mathbf{u}^{(\lambda)}$, and $\mathbf{u}^{(\lambda)}$ are codewords of Lee weights

$$2^m + \delta_1 2^{(m-1)/2},\ 2^m - \delta_2 2^{(m-1)/2}, \quad \text{and} \quad 2^m - \delta_1 2^{(m-1)/2},$$

respectively. Of these four codewords obtained from $\mathbf{v}^{(\lambda)}, \lambda \in R^*$, two have Lee weight $2^m + 2^{(m-1)/2}$ and two have Lee weight $2^m - 2^{(m-1)/2}$. This holds for all $2^m(2^m - 1)$ codewords $\mathbf{v}^{(\lambda)}$, $\lambda \in R^*$ Therefore Table 8.1 is established. $\square$

Remark 8.1. When m is even, $m \geq 2$, a similar argument shows that $\phi(\mathcal{K}(m))$ has the following weight distribution.

Table 8.3. Weight distribution of $\phi(\mathcal{K}(m))$, (m even ≥ 2).

Weight	No. of codewords
0	1
$2^m - 2^{m/2}$	$2^m(2^m - 1)$
2^m	$2^{m+1}(2^m + 1) - 2$
$2^m + 2^{m/2}$	$2^m(2^m - 1)$
2^{m+1}	1

This code is not as good as a double-error-correction BCH code.

8.5. Soft-Decision Decoding of Quaternary Kerdock Codes

For simplicity we write $\Delta = \{\infty, 0, 1, \ldots, n-1\}$, where $n = 2^m - 1$. Let $\mathbf{r} = (r_t,\ t \in \Delta)$ be a received word, where $r_t \in \mathbb{Z}_4$. The brute-force decoding of $\mathbf{r}$ requires the computation of its correlation with all codewords $\varepsilon 1^{2^m} + \mathbf{u}^{(\lambda)}$, where

$$\mathbf{u}^{(\lambda)} = (T(\lambda\xi^\infty), T(\lambda\xi^0), T(\lambda\xi), T(\lambda\xi^2), \ldots, T(\lambda\xi^{n-1})),$$

$\varepsilon \in \mathbb{Z}_4$, and $\lambda \in \mathrm{GR}(4^m)$. That is, we have to compute the correlation

$$\zeta(\varepsilon, \lambda) = \sum_{t \in \Delta} i^{r_t} i^{-(\varepsilon + T(\lambda \xi^t))} \tag{8.26}$$

for all $\varepsilon \in \mathbb{Z}_4$ and $\lambda \in \mathbb{Z}_4[\xi]$. If Real$\{\zeta(\varepsilon_0, \lambda_0)\}$ is a maximum for the pair $(\varepsilon_0, \lambda_0)$, we decode $\mathbf{r}$ into the codeword $\varepsilon_0 1^{2^m} + \mathbf{u}^{(\lambda_0)}$. If we compute (8.26) directly, it requires $4^{m+1}2^m$ multiplications and $4^{m+1}(2^m - 1)$ additions.

Let $\lambda = \xi^r + 2\xi^s$ be the 2-adic representation of λ, where $r, s \in \Delta$. Then we can write (8.26) as follows:

$$\zeta(\varepsilon, \xi^r + 2\xi^s) = i^{-\varepsilon} \sum_{t \in \Delta} i^{r_t - T(\xi^{r+t})}(-1)^{\mathrm{Tr}(\bar{\xi}^{s+t})}, \tag{8.27}$$

where we adopt the convention that for $l \in \Delta$, $l + \infty = \infty$. If use (8.27) to compute $\zeta(\varepsilon, \lambda)$, the computational complexity is reduced. Furthermore, the correlation sums $\zeta(\varepsilon, \xi^r + 2\xi^s)$ may be viewed (after some reordering of indexes) as $i^{-\varepsilon}$ times the Hadamard transform of the 2^m complex vectors $(i^{r_t - T(\xi^{r+t})}, t \in \Delta)$ of length 2^m. Using the FHT, each of these can be computed using $m2^m$ additions/subtractions. Thus the overall requirement is for about 4^m multiplications and $m4^m$ additions/subtractions.

The above soft-decision decoding algorithm is suggested by Hammons *et al.* (1994) and can be regarded as an extension of the fast Hadamard transform soft-decision decoding algorithm for the binary first-order Reed–Muller code to the quaternary Kerdock code.

For another decoding algorithm of the Kerdock codes, see Adoul (1987).

CHAPTER 9

PREPARATA CODES

9.1. The Quaternary Preparata Codes

We follow the notation of the previous chapter. That is, m is an integer ≥ 2, $h(X)$ is a basic primitive polynomial of degree m dividing $X^n - 1$ in $\mathbb{Z}_4[X]$, where $n = 2^m - 1$, ξ is a root of $h(X)$ in $\mathrm{GR}(4^m)$, and $g(X)$ is the reciprocal polynomial to the polynomial $(X^n - 1)/(X - 1)h(X)$.

Definition 9.1. The $\mathbb{Z}_4$-cyclic code of length n with generator polynomial $h(X)$ is called the *shortened quaternary Preparata code* and denoted by $\mathcal{P}(m)^-$. The $\mathbb{Z}_4$-linear code obtained from $\mathcal{P}(m)^-$ by adding a zero-sum check symbol to each codeword of $\mathcal{P}(m)^-$ is called the *quaternary Preparata code* and denoted by $\mathcal{P}(m)$. □

Proposition 9.1. *$\mathcal{P}(m)^-$ has parity check matrix*

$$(1 \ \xi \ \xi^2 \ \cdots \ \xi^{n-1}). \tag{9.1}$$

$\mathcal{P}(m)$ is the dual code of $\mathcal{K}(m)$ and has parity check matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \xi & \xi^2 & \cdots & \xi^{n-1} \end{pmatrix} \tag{9.2}$$

Proof. By definition

$$\mathcal{P}(m)^- = \{a(X)h(X) \bmod X^n - 1 \mid a(X) \in \mathbb{Z}_4[X]\}.$$

For any codeword $\mathbf{c} = (c_0, c_1, c_2, \ldots, c_{n-1}) \in \mathcal{P}(m)^-$, we have

$$c(X) = c_0 + c_1X + c_2X_2 + \cdots + c_{n-1}X^{n-1} \equiv b(X)h(X) (\bmod X^n - 1)$$

for some $b(X) \in \mathbb{Z}_4[X]$. Therefore $c(\xi) = 0$, i.e.,

$$(c_0, c_1, c_2, \ldots, c_{n-1}) \cdot (1\ \xi\ \xi^2\ \cdots\ \xi^{n-1}) = 0. \tag{9.3}$$

Conversely, assume that $\mathbf{c} = (c_0, c_1, c_2, \ldots, c_{n-1}) \in \mathbb{Z}_4^n$ has the property (9.3). Let $c(X) = c_0 + c_1 X + c_2 X^2 + \cdots + c_{n-1} X^{n-1}$. Then (9.3) is equivalent to $c(\xi) = 0$. We know that $h(X)$ is monic. Dividing $c(X)$ by $h(X)$, we have

$$c(X) = q(X)h(X) + r(X),$$

where $q(X)$, $r(X) \in \mathbb{Z}_4[X]$ and $\deg r(X) < \deg h(X) = m$. Substituting ξ in the above equation, we obtain $r(\xi) = 0$. By Theorem 6.1 the additive representation of every element in $\mathrm{GR}(4^m)$ is unique. It follows that $r(X) = 0$. Therefore $c(X) = q(X)h(X) \in \mathcal{P}(m)^-$ We conclude that $\mathcal{P}(m)^-$ has parity check matrix (9.1).

Now let $(c_\infty, c_0, c_1, c_2, \ldots, c_{n-1})$ be a codeword of $\mathcal{P}(m)$. By definition, $c_\infty = -\sum_{i=0}^{n-1} c_i$ and $c_0 + c_1 X + c_2 X^2 + \cdots + c_{n-1} X^{n-1}$ is a multiple of $h(X)$. Therefore

$$(c_\infty, c_0, c_1, c_2, \ldots, c_{n-1}) \quad (1, 1, 1, 1, \ldots, 1) = 0$$

and

$$(c_\infty, c_0, c_1, c_2, \ldots, c_{n-1}) \quad (0, 1, \xi, \xi^2, \ldots, \xi^{n-1}) = 0.$$

But (9.2) is the generator matrix of $\mathcal{K}(m)$, so $\mathcal{P}(m) \subset \mathcal{K}(m)^\perp$ By Corollary 8.3, $\mathcal{K}(m)$ is of type 4^{m+1} Then by Proposition 1.2, $\mathcal{K}(m)^\perp$ is of type $4^{2^m - m - 1}$. Since $\mathcal{P}(m)^- = (h(X))$, $\mathcal{P}(m)^-$ is of type $4^{2^m - m - 1}$ Thus $\mathcal{P}(m)$ is also of type $4^{2^m - m - 1}$. Therefore $\mathcal{P}(m) = \mathcal{K}(m)^\perp$ and (9.2) is a parity check matrix of $\mathcal{P}(m)$. $\square$

Corollary 9.2. *Both $\mathcal{P}(m)^-$ and $\mathcal{P}(m)$ are $\mathbb{Z}_4$-linear codes of type* $4^{2^m - m - 1}$ $\square$

Corollary 9.3. *The binary linear code $P^{(1)}$ associated with $\mathcal{P}(m)$ is* $\mathrm{RM}(m-2, m)$.

Proof. By definition

$$P^{(1)} = \{\overline{\mathbf{c}} \mid \mathbf{c} \in \mathcal{P}(m)\}.$$

For any $\mathbf{c} = (c_\infty, c_0, c_1, \ldots, c_{n-1}) \in \mathcal{P}(m)$, we have

$$c_\infty + \sum_{i=0}^{n-1} c_i = 0$$

and

$$\sum_{i=0}^{n-1} c_i\xi^i = 0.$$

By reduction modulo 2, we obtain

$$\overline{c_\infty} + \sum_{i=0}^{n-1} \overline{c_i} = 0$$

and

$$\sum_{i=0}^{n-1} \overline{c_i}\,\overline{\xi}_i^i = 0.$$

Thus $\overline{\mathbf{c}}$ is a codeword of binary linear code with parity check matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \overline{\xi} & \overline{\xi}^2 & \cdots & \overline{\xi}^{n-1} \end{pmatrix} \tag{9.4}$$

So $\overline{\mathbf{c}} \in \mathrm{RM}(m-2, m)$. Hence $P^{(1)} \subset \mathrm{RM}(m-2, m)$. $\mathcal{P}(m)$ is of type 4^{2^m-m-1}, so dim $P^{(1)} = 2^m - m - 1$. But dim $\mathrm{RM}(m-2, m) = 2^m - m - 1$. Therefore $P^{(1)} = \mathrm{RM}(m-2, m)$. □

Digression. Let us study the number of distinct zeros of some polynomials over $\mathbb{F}_{2^m}$. We begin with the following well-known result in finite group theory.

Lemma 9.4. *Let G be a finite cyclic group of order $n > 1$ and d be a positive integer relatively prime to n, then the equation $X^d = e$ has a unique solution $X = e$ in G, where e denotes the identity element of G.*

Proof. Let a be a generator of the cyclic group G, i.e., $a^n = 1$ and $a^k \neq e$ for $0 < k < n$. We know that $a^l = e$ if and only if $n|l$. Let a^i be a solution of the equation $X^d = e$, i.e., $a^{id} = e$. Then $n|id$. Since $(d, n) = 1$, we have $n|i$. Therefore $a^i = e$. □

From Lemma 9.4, we deduce

Lemma 9.5. *Let m is an odd integer ≥ 3, then the polynomial $X^3 + a$, where $a \in \mathbb{F}_{2^m}$, has at most one root in $\mathbb{F}_{2^m}$.*

Proof. For $a = 0$, the polynomial X^3 has only 0 as a triple root. Now consider the case $a \neq 0$. Let x_1 and x_2 be two roots of $X^3 + a$. Then $x_1^3 = x_2^3 = a$ and $x_1 \neq 0$, $x_2 \neq 0$. If follows that $(x_1/x_2)^3 = 1$. Since $\mathbb{F}_{2^m}^*$ is a cyclic group of order $2^m - 1$ and m is odd and ≥ 3, $(3, 2^m - 1) = 1$. By Lemma 9.4, $x_1/x_2 = 1$. Therefore $x_1 = x_2$. □

Corollary 9.6. *Let m be an odd integer ≥ 3, then the polynomial $p(X) = X^4 + aX + b \in \mathbb{F}_{2^m}[X]$ has at most two distinct roots in $\mathbb{F}_{2^m}$.*

Proof. Let x be a root of $p(X)$ in $\mathbb{F}_{2^m}$. Then

$$\begin{aligned} p(X + x) &= (X + x)^4 + a(X + x) + b \\ &= X^4 + aX + x^4 + ax + b \\ &= X(X^3 + a). \end{aligned}$$

By Lemma 9.5, $p(X + x)$ has at most two distinct roots in $\mathbb{F}_{2^m}$, so does $p(X)$. □

Corollary 9.7. *Let m be an odd integer ≥ 3, then the polynomial $q(X) = X^5 + aX^4 + dX + e \in \mathbb{F}_{2^m}[X]$ has at most three distinct roots in $\mathbb{F}_{2^m}$.*

Proof. Let x be a root of $q(X)$ in $\mathbb{F}_{2^m}$. Then

$$\begin{aligned} q(X + x) &= (X + x)^5 + a(X + x)^4 + d(X + x) + e \\ &= X(X^4 + (x + a)X^3 + (x^4 + d)). \end{aligned}$$

If $x^4 + d = 0$, $q(X + x) = X^4(X + (x + a))$, which clearly has at most two distinct roots. If $x^4 + d \neq 0$, then by Corollary 9.6,

$$X^4 + \frac{x + a}{x^4 + d} X + \frac{1}{x^4 + d}$$

has at most two distinct roots in $\mathbb{F}_{2^m}$, and so does its reciprocal polynomial

$$\frac{1}{x^4 + d} X^4 + \frac{x + a}{x^4 + d} X^3 + 1.$$

Thus the polynomial

$$X^4 + (x + a)X^3 + (x^4 + d)$$

has at most two distinct roots in $\mathbb{F}_{2^m}$. It follows that $q(X + x)$ has at most three distinct roots in $\mathbb{F}_{2^m}$, and so does $q(X)$. □

Now we return to the study of quaternary Preparata codes.

Proposition 9.8. *Let m be an integer ≥ 2, then all codewords of $\mathcal{P}(m)$ are of even Lee weight. Moreover, when m is even and ≥ 2, $\mathcal{P}(m)$ has minimum Lee distance 4 and when m is odd and ≥ 3, $\mathcal{P}(m)$ has minimum Lee distance 6.*

Proof. The first assertion follows from Proposition 3.4. Let us prove the second assertion. Since $\mathcal{P}(m)$ is a quaternary linear code, it is enough to show that $\mathcal{P}(m)$ has minimum Lee weight 4 or 6, when m is even and ≥ 2 or odd and $n \geq 3$, respectively.

First we assert that $\mathcal{P}(m)$ has no codeword of Lee weight 2. Let $\mathbf{c} = (c_\infty, c_0, c_1, \ldots, c_{n-1})$ be a codeword of $\mathcal{P}(m)$. Since (9.2) is the parity matrix of $\mathcal{P}(m)$, we have

$$c_\infty + \sum_{i=0}^{n-1} c_i = 0 \tag{9.5}$$

and

$$\sum_{i=0}^{n-1} c_i \xi^i = 0. \tag{9.6}$$

Assume that $w_L(\mathbf{c}) = 2$. Denote by $\mathbf{e}_i$ the 2^m-tuple whose ith component is 1 and all other components are 0's. By (9.5), $\mathbf{c}$ must be of the form $\mathbf{c} = \mathbf{e}_i - \mathbf{e}_j$ ($i, j = \infty, 0, 1, \ldots, n-1, i \neq j$). If $i = \infty$, by (9.6) we have $-\xi^j = 0$, a contradiction. Similarly, $j = \infty$ is also impossible. Assume that both i and $j \neq \infty$, by (9.6) we have $\xi^i - \xi^j = 0$. Since ξ is of order $n = 2^m - 1$, this is also impossible. Our assertion is proved.

Then we distinguish the following two cases.

(a) m is even and ≥ 2. We have $3 | 2^m - 1$. Let $t = (2^m - 1)/3$, then $\xi^{3t} = 1$ and $\xi^{3t} - 1 = (\xi^t - 1)(\xi^{2t} + \xi^t + 1) = 0$. By Proposition 6.16 (i), $\xi^t - 1$ is an invertible element of $\mathrm{GR}(4^m)$. Therefore $\xi^{2t} + \xi^t + 1 = 0$, which yields a codeword of Lee weight 3 in $\mathcal{P}(m)^-$. Adjoining a zero-sum check symbol to this codeword, we get a codeword of Lee weight 4 in $\mathcal{P}(m)$.

(b) m is odd and ≥ 3. Assume that $\mathbf{c}$ is a codeword of Lee weight 4 in $\mathcal{P}(m)$. By (9.5), $\mathbf{c}$ must be one of the following forms:

$$\pm(\mathbf{e}_i + \mathbf{e}_j + \mathbf{e}_k + \mathbf{e}_l),$$

$$\mathbf{e}_i + \mathbf{e}_j - \mathbf{e}_k - \mathbf{e}_l,$$

$$\pm(2\mathbf{e}_i + \mathbf{e}_j + \mathbf{e}_k),$$

$$2\mathbf{e}_i + 2\mathbf{e}_j.$$

where i, j, $k, l \in \{\infty, 0, 1, \ldots, n-1\}$ and are distinct in pairs. Following Helleseth (1996), we treat these four cases one by one in the following way.

(b.1) $\mathbf{c} = 2\mathbf{e}_i + 2\mathbf{e}_j$. Then (9.6) leads to $2\xi^i + 2\xi^j = 0$. Thus $\xi^i + \xi^j \equiv 0 \pmod 2$. Consequently, $\overline{\xi}^i + \overline{\xi}^j = 0$, which is impossible for $i \neq j$.

(b.2) $\mathbf{c} = \pm(2\mathbf{e}_i + \mathbf{e}_j + \mathbf{e}_k)$. Then (9.6) leads to $2\xi^i + \xi^j + \xi^k = 0$. Thus we also have $\overline{\xi}^i + \overline{\xi}^k = 0$. As in case (b.1) this is also impossible.

(b.3) $\mathbf{c} = \mathbf{e}_i + \mathbf{e}_j - \mathbf{e}_k - \mathbf{e}_l$. Then (9.6) gives

$$\xi^i + \xi^j = \xi^k + \xi^l. \tag{9.7}$$

By Corollary 6.9,

$$\xi^i + \xi^j = (\xi^i + \xi^j + 2(\xi^i\xi^j)^{1/2}) + 2(\xi^i\xi^j)^{1/2}, \tag{9.8}$$

$$\xi^k + \xi^l = (\xi^k + \xi^l + 2(\xi^k\xi^l)^{1/2}) + 2(\xi^k\xi^l)^{1/2}, \tag{9.9}$$

where $\xi^i + \xi^j + 2(\xi^i\xi^j)^{1/2}$, $(\xi^i\xi^j)^{1/2}$, $\xi^k + \xi^l + 2(\xi^k\xi^l)^{1/2}$, $(\xi^k\xi^l)^{1/2} \in \mathcal{T}$ By (9.7)–(9.9) and Theorem 6.7 (ii), we have $(\xi^i\xi^j)^{1/2} = (\xi^k\xi^l)^{1/2}$. Consequently

$$\xi^i\xi^k = \xi^k\xi^l. \tag{9.10}$$

By reduction mod 2, (9.7) and (9.10) give

$$\overline{\xi}^i + \overline{\xi}^j = \overline{\xi}^k + \overline{\xi}^l \quad \text{and} \quad \overline{\xi}^i\overline{\xi}^j = \overline{\xi}^k\overline{\xi}^l,$$

respectively. Then

$$f(X) = (X - \overline{\xi}^i)(X - \overline{\xi}^j) = (X - \overline{\xi}^k)(X - \overline{\xi}^l)$$

has four distinct roots, a contradiction.

(b.4) $\mathbf{c} = \pm(\mathbf{e}_i + \mathbf{e}_j + \mathbf{e}_k + \mathbf{e}_l)$. Then (9.6) gives

$$\xi^i + \xi^j + \xi^k + \xi^l = 0. \tag{9.11}$$

Let

$$\xi^i + \xi^j + \xi^k + \xi^l = a + 2b, \tag{9.12}$$

where $a, b \in \mathcal{T}$. By Corollary 6.10

$$b = (\xi^i\xi^j)^{1/2} + (\xi^i\xi^k)^{1/2} + (\xi^i\xi^l)^{1/2} + (\xi^j\xi^k)^{1/2} + (\xi^j\xi^l)^{1/2} + (\xi^k\xi^l)^{1/2}.$$

By (9.11) and (9.12), $b = 0$. It follows that

$$(\xi^i\xi^j)^{1/2} + (\xi^i\xi^k)^{1/2} + (\xi^i\xi^l)^{1/2} + (\xi^j\xi^k)^{1/2} + (\xi^j\xi^l)^{1/2} + (\xi^k\xi^l)^{1/2} = 0.$$

Squaring, we obtain

$$\xi^i\xi^j + \xi^i\xi^k + \xi^i\xi^l + \xi^j\xi^k + \xi^j\xi^l + \xi^k\xi^l \equiv 0 \pmod 2. \tag{9.13}$$

Applying the map $\bar{}: \mathbb{Z}_4[\xi] \to \mathbb{F}_2[\bar{\xi}]$ to (9.11) and (9.13), we get

$$\bar{\xi}^i + \bar{\xi}^j + \bar{\xi}^k + \bar{\xi}^l = 0$$

and

$$\bar{\xi}^i\bar{\xi}^j + \bar{\xi}^i\bar{\xi}^k + \bar{\xi}^i\bar{\xi}^l + \bar{\xi}^j\bar{\xi}^k + \bar{\xi}^j\bar{\xi}^l + \bar{\xi}^k\bar{\xi}^l = 0.$$

Then

$$\begin{aligned} f(X) &= (X - \bar{\xi}^i)(X - \bar{\xi}^j)(X - \bar{\xi}^k)(X - \bar{\xi}^l) \\ &= X^4 + aX + b \end{aligned}$$

has four distinct roots in $\mathbb{F}_{2^m}$, which contradicts Corollary 9.6.

We proved that $\mathcal{P}(m)$ has no codewords of Lee weight 4.

Finally we have to prove that $\mathcal{P}(m)$ contains a codeword of Lee weight 6. Consider the matrix (9.4). Since $m \geq 3$, ${}^t(1, 1, 0, 0, 0^{m-3})$, ${}^t(1, 0, 1, 0, 0^{m-3})$, and ${}^t(1, 0, 0, 1, 0^{m-3})$ are the zeroth, first, and second columns of (9.4), respectively. But ${}^t(1, 1, 1, 1, 0^{m-3})$ must be a column of (9.4), and let it be the ith column, where $3 \leq i \leq n-1$. Then $1 + \bar{\xi} + \bar{\xi}^2 + \bar{\xi}^i = 0$. We distinguish the following two cases:

(α) $-1 + \xi + \xi^2 + \xi^i = 0$. Multiplying by $1 - \xi$, we obtain $1 + 2\xi + \xi^3 - \xi^i + \xi^{i+1} = 0$. If $i = 3$, we have $1 + \xi^4 = 2\xi$. By Proposition 6.16 (i) $1 + \xi^4$ is invertible, but 2ξ is a zero divisor, which is a contradiction. If $3 < i < n-1$, $\mathbf{e}_0 + 2\mathbf{e}_1 + \mathbf{e}_3 - \mathbf{e}_i + \mathbf{e}_{i+1}$ is a codeword of Lee weight 6. If $i = n-1$, $2\mathbf{e}_0 + 2\mathbf{e}_1 + \mathbf{e}_3 - \mathbf{e}_{n-1}$ is a codeword of Lee weight 6.

(β) $-1 + \xi + \xi^2 + \xi^i \neq 0$. Then $-1 + \xi + \xi^2 + \xi^i = 2\xi^j$. If $j = 0$, we have $1 + \xi + \xi^2 + \xi^i = 0$, which contradicts Proposition 6.16 (iv). If $j = 1$, we have $-1 - \xi + \xi^2 + \xi^i = 0$, which contradicts Proposition 6.16 (iii). Similarly, $j = 2$ and $j = i$ are also impossible. Therefore $j \neq 0, 1, 2, i$ and $-\mathbf{e}_0 + \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_i + 2\mathbf{e}_j$ is a codeword of Lee weight 6. □

9.2. The "Preparata" Codes

Denote the binary image of the quaternary Preparata code $\mathcal{P}(m)$ by $P(m)$, i.e., $P(m) = \phi(\mathcal{P}(m))$. First, we have

Theorem 9.9. *Let m be an integer ≥ 2. $P(m)$ is a binary code of length 2^{m+1} and has $2^{2^{m+1}-2m-2}$ codewords. It is distance invariant, all its codewords have even weight and is the formal dual of $K(m)$. Its weight enumerator is*

$$W_{P(m)}(X,Y) = \frac{1}{4^{m+1}}\, W_{K(m)}(X+Y,\, X-Y). \tag{9.14}$$

When $m \geq 3$, $P(m)$ is nonlinear. When m is even and ≥ 2, the minimum distance of $P(m)$ is 4, and when m is odd and ≥ 3, the minimum distance of $P(m)$ is 6.

Proof. Clearly, $P(m)$ is a binary code of length 2^{m+1}. Since $\mathcal{P}(m)$ is of type 4^{2^m-m-1}, $|\mathcal{P}(m)| = 4^{2^m-m-1}$. But $|P(m)| = |\mathcal{P}(m)|$, so $|P(m)| = 2^{2^{m+1}-2m-2}$.

By Theorem 3.6, $P(m)$ is distance invariant, and by Proposition 3.4 all codewords of $P(m)$ have even weight. Since $\mathcal{P}(m) = \mathcal{K}(m)^{\perp}$, $P(m) = K(m)_{\perp}$ is the formal dual of $K(m)$ and by Theorem 3.7 we have (9.14).

Now let us prove that $P(m)$ is nonlinear when $m \geq 3$. Let $h(X) = h_0 + h_1X + \cdots + h_mX^m$, then $h_0 = \pm 1$ and $h_m = 1$. Since $m \geq 3$, we have $2^m \geq 2m+2$. Thus both

$$\mathbf{c} = \left(-\sum_{i=0}^{m} h_i,\, h_0,\, h_1, \ldots, h_{m-1},\, h_m,\, \underbrace{0, \ldots, 0}_{m},\, 0, \ldots, 0\right)$$

and

$$\mathbf{c}' = \left(-\sum_{i=0}^{m} h_i,\, \underbrace{0,\, 0, \ldots, 0}_{m},\, h_0,\, h_1, \ldots, h_m,\, 0, \ldots, 0\right)$$

are codewords of $\mathcal{P}(m)$. But

$$2\alpha(\mathbf{c}) * \alpha(\mathbf{c}') = (2,\, \underbrace{0, \ldots, 0}_{m},\, 2,\, 0, \ldots, 0)$$

is not a codeword of $\mathcal{P}(m)$. By Proposition 3.16, $P(m)$ is nonlinear.

The last assertion follows from Propositions 3.3 and 9.8. □

Remark 9.1. When $m = 2$, $h(X) = 1 + X + X^2$ is the unique basic primitive polynomial of degree 2. Then $\mathcal{P}(2) = \{\varepsilon 1^4 | \varepsilon \in \mathbb{Z}_4\}$ and condition (3.20) of Proposition 3.16 trivially holds. Therefore $P(2)$ is linear. □

Remark 9.2. The decoding algorithm given in the next section gives an alternate proof that $P(m)$ has minimum distance 6 when m is odd and ≥ 3. We can give a third proof by using the Krawtchouk polynomials as follows. Let A_i and A'_i be the number of codewords of weight i in $K(m)$ and $P(m)$, respectively. By (9.14) and Propositions 2.6 and 8.10, we have

$$4^{m+1}A'_k = K_k(0) + 2^{m+1}(2^m - 1)\left(K_k(2^m - 2^{(m-1)/2}) + K_k(2^m + 2^{(m-1)/2})\right)$$
$$+(2^{m+2} - 2)K_k(2^m) + K_k(2^{m+1}),$$

where $K_k(x)$'s are the Krawtchouk polynomials for $q = 2$. Using the formulas of $K_k(x)$, where $k = 2, 4, 6$, given in Proposition 2.16, it can be readily checked that $A'_2 = A'_4 = 0$ and $A'_6 \neq 0$. Therefore the minimum distance of $P(m)$ is 6.

That the minimum distance of $P(m)$ is 4 when m is even and ≥ 2 can be proved in a similar way. □

When m is an odd integer ≥ 3, $P(m)$ is called the "Preparata" code; here we use the quotation mark to distinguish it from the Preparata's original code P_{m+1} which will be introduced in Sec. 9.4. We will see that they have the same code length, the same number of codewords, the same minimum distance, and the same weight enumerator. But there is an essential difference between $P(m)$ and P_{m+1}. The latter is contained in the extended binary Hamming code of length 2^{m+1} (see Proposition 9.15), whose minimum weight is 4. For $P(m)$, we have

Proposition 9.10. *For odd $m \geq 5$, $P(m)$ is contained in a nonlinear code with the same weight distribution as the extended binary Hamming code of the same length, and the linear code spanned by the codewords of $P(m)$ has minimum weight* 2.

Proof. We recall that the $\mathbb{Z}_4$-linear code ZRM(1, m) is of length 2^m and generated by RM(0, m) and 2RM (1, m). Hence ZRM(1, m) has generator matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 2 & 2\xi & 2\xi^2 & \cdots & 2\xi^{n-1} \end{pmatrix}.$$

Therefore

$$\mathrm{ZRM}(1, m) \subseteq \mathcal{K}(m).$$

It follows that

$$\mathcal{P}(m) \subseteq \mathrm{ZRM}(1, m)^{\perp}$$

and

$$P(m) \subseteq \phi(\mathrm{ZRM}(1, m)^{\perp}).$$

By Proposition 4.1, $\phi(\mathrm{ZRM}(1, m)) = \mathrm{RM}(1, m + 1)$. Therefore ϕ(ZRM $(1, m)^{\perp}) = \mathrm{RM}(1, m+1)_{\perp}$ and

$$P(m) \subseteq \mathrm{RM}(1, m + 1)_{\perp}$$

$\mathrm{RM}(1, m+1)_\perp$ is $\mathbb{Z}_4$-linear of length 2^{m+1} and its weight enumerator is the same as $\mathrm{RM}(1,m+1)^\perp$. But $\mathrm{RM}(1,m+1)^\perp = \mathrm{RM}(m-1,m+1)$, which is the extended binary Hamming code of length 2^{m+1} Hence the weight enumerator of $\mathrm{RM}(1,m+1)_\perp$ is the same as the extended binary Hamming code of the same length. If $\mathrm{RM}(1,m+1)_\perp$ is linear, then by the uniqueness of the extended binary Hamming code, $\mathrm{RM}(1,m+1)_\perp = \mathrm{RM}(m-1,m+1)$. Since $m \geq 5$, by Proposition 4.4 $\mathrm{RM}(m-1,m+1)$ is not $\mathbb{Z}_4$-linear, which is a contradiction. Therefore $\mathrm{RM}(1,m+1)_\perp$ is nonlinear.

Let us come to the proof of the second assertion. By definition, $\mathcal{P}(m)^-$ is the cyclic code of length $n = 2^m - 1$ generated by a basic primitive polynomial $h(X)$ dividing $X^n - 1$. Write $h(X) = \sum_{j=0}^m h_j X^j$, where $h_0 \neq 0, 2$ and $h_m = 1$. Let $h_\infty = -h(1)$. Since $\overline{h}(1) \neq 0$, $h_\infty = \pm 1$. When m is odd and ≥ 5, $2^m - 1 > 2m + 3$. Therefore

$$\mathbf{a} = (h_\infty, h_0, h_1, \ldots, h_m, 0, 0, \ldots, 0, 0, \ldots, 0)$$

and

$$\mathbf{b} = (h_\infty, \underbrace{0, 0, \ldots, 0}_{m+1}, h_0, h_1, \ldots, h_m, 0, \ldots, 0)$$

are codewords of $\mathcal{P}(m)$, and so is $\mathbf{a}+\mathbf{b}$. Then $\phi(\mathbf{a})$, $\phi(\mathbf{b})$, and $\phi(\mathbf{a}+\mathbf{b}) \in P(m)$. By (3.18),

$$\phi(2(\alpha(\mathbf{a}) * \alpha(\mathbf{b}))) = \phi(\mathbf{a}) + \phi(\mathbf{b}) + \phi(\mathbf{a}+\mathbf{b}).$$

Thus $\phi(2(\alpha(\mathbf{a}) * \alpha(\mathbf{b})))$ belongs to the linear code spanned by the codewords of $P(m)$. Clearly,

$$\begin{aligned} \phi(2(\alpha(\mathbf{a}) * \alpha(\mathbf{b}))) &= \phi(2, 0, \ldots, 0) \\ &= (1, 0, \ldots, 0, 1, 0, \ldots, 0) \end{aligned}$$

is a codeword of weight 2. By Theorem 9.9 all codewords of $P(m)$ are of even weight, so the linear code spanned by the codewords of $P(m)$ has minimum weight 2. □

9.3. Decoding $\mathcal{P}(m)$ in the $\mathbb{Z}_4$-Domain

Hammons *et al.* (1994) also suggested a simple decoding algorithm for the "Preparata" code $P(m)$, when m is odd and ≥ 3, by working in the $\mathbb{Z}_4$-domain. This is an optimal syndrome decoder: it corrects all error patterns of Lee weight at most 2, detects all errors of Lee weight 3, and detects some errors of Lee weight 4.

Let H be the parity check matrix (9.2)

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \xi & \xi^2 & \cdots & \xi^{n-1} \end{pmatrix}$$

of the code $\mathcal{P}(m)$. Let $\mathbf{c} = (c_\infty, c_0, c_1, \ldots, c_{n-1}) \in \mathcal{P}(m)$ be the transmitted codeword and $\mathbf{r} = (r_\infty, r_0, r_1, \ldots, r_{n-1})$ be the received word. Then $\mathbf{e} = \mathbf{r} - \mathbf{c} = (e_\infty, e_0, e_1, \ldots, e_{n-1})$ is the *error pattern*. We compute the *syndrome* $H^t\mathbf{r}$, which has two components $r_\infty + \sum_{j=0}^{n-1} r_j$ and $\sum_{j=0}^{n-1} r_j\xi^j$. Let

$$r_\infty + \sum_{j=0}^{n-1} r_j = t$$

and

$$\sum_{j=0}^{\infty} r_j\xi^j = a + 2b,$$

where $t \in \mathbb{Z}_4$ and $a, b \in \mathcal{T}$.

Since $P(m) = K(m)_\perp$, they have the same weight distribution $\{A'_0, A'_1, \ldots, A'_{n+1}\}$ and the same weight enumerator

$$\begin{aligned} W_{P(m)}(X,Y) &= W_{K(m)_\perp}(X,Y) \\ &= \sum_{i=0}^{n+1} A'_i X^{n+1-i} Y^i. \end{aligned}$$

By Theorem 9.9 $W_{P(m)}(X,Y)$ is the MacWilliams transform of $W_{K(m)}(X, Y)$. Denote the weight distribution of $K(m)$ by $\{A_0, A_1, \ldots, A_{n+1}\}$. By Proposition 3.8 the MacWilliams transform of $\{A'_0, A'_1, \ldots, A'_{n+1}\}$ is $\{A_0, A_1, \ldots, A_{n+1}\}$. By Proposition 8.10 the number of nonzero weights of $K(m)$, i.e., the number of nonzero A_i where $0 < i \leq n+1$, is equal to 4. That is, 4 is the external distance of $P(m)$. By Corollary 2.23, for any vector $\mathbf{v} \in \mathbb{F}_2^{2^{m+1}}$ there is a codeword $\mathbf{c} \in P(m)$ such that $d(\mathbf{v}, \mathbf{c}) \leq 4$.

In other words, for any $\mathbf{u} \in \mathbb{Z}_4^{n+1}$ we have $d_{\mathrm{L}}(\mathbf{u}, \mathcal{P}(m)) \leq 4$. In particular, for the received word $\mathbf{r}$ we have $d_{\mathrm{L}}(\mathbf{r}, \mathcal{P}(m)) \leq 4$. It is not difficult to prove that $t = \pm 1$ if and only if $d_{\mathrm{L}}(\mathbf{r}, \mathcal{P}(m)) = 1$ or 3.

First consider the case $t = 1$. Then $d_{\mathrm{L}}(\mathbf{r}, \mathcal{P}(m)) = 1$ or 3. If $b = 0$, we decide that there is a unique single error pattern

$$\mathbf{e} = \mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$$

of Lee weight 1 if $a = \xi^i$, where $i = \infty, 0, 1, 2, \ldots$ or $n-1$. If $b \neq 0$, then $d_L(\mathbf{r}, \mathcal{P}(m)) = 3$, and the error pattern is of Lee weight 3 and is detected.

Then consider the case $t = -1$. We also have $d_L(\mathbf{r}, \mathcal{P}(m)) = 1$ or 3. If $a = b$, we decide that there is a unique single error pattern $-\mathbf{e}_i$ of Lee weight 1 if $a = b = \xi^i$. If $a \neq b$, then $d_L(\mathbf{r}, \mathcal{P}(m)) = 3$, and the error pattern is of Lee weight 3 and is detected.

Now consider the case $t = 0$. Then $d_L(\mathbf{r}, \mathcal{P}(m)) = 0, 2$ or 4. If $a = b = 0$, then $\mathbf{r}$ is a codeword of $\mathcal{P}(m)$ and $d_L(\mathbf{r}, \mathcal{P}(m)) = 0$. If $a = 0$ but $b \neq 0$, the error pattern must be of the form $2\mathbf{e}_i + 2\mathbf{e}_k$, where $i \neq k$, which is of Lee weight 4 and can be detected. If $a \neq 0$, assume that the error pattern is of Lee weight 2, then it must be of the form $\mathbf{e}_i - \mathbf{e}_k$, where $i \neq k$. Thus

$$a + 2b = \xi^i - \xi^k$$

Raising the above equation to 2^m-th power, we obtain

$$a = \xi^i + \xi^k + 2\xi^{i \cdot 2^{m-1}} \xi^{k \cdot 2^{m-1}}.$$

It follows that

$$b = -\xi^k - \xi^{i \cdot 2^{m-1}} \xi^{k \cdot 2^{m-1}}$$

Applying the map $-$ to the above two equations, we obtain

$$\bar{a} = \bar{\xi}^i + \bar{\xi}^k, \quad \bar{b} = \bar{\xi}^k + \bar{\xi}^{i \cdot 2^{m-1}} \bar{\xi}^{k \cdot 2^{m-1}}$$

which can be written as

$$\bar{a} = \bar{\xi}^i + \bar{\xi}^k, \quad (\bar{b} + \bar{\xi}^k)^2 = \bar{\xi}^i \bar{\xi}^k$$

The unique solution of the above simultaneous equations is $\bar{\xi}^k = \bar{b}^2/\bar{a}$, $\bar{\xi}^i = \bar{a} + \bar{b}^2/\bar{a}$, Therefore the error positions i and k can be determined. Note that when $\bar{b} = 0$ or $\bar{b} = \bar{a}$, the double error involves the ∞-position.

Finally, consider the case $t = 2$. If $a = 0$, then $b = \xi^i$ where $i = \infty, 0, 1, \ldots, n-1$ and we assume that $\xi^\infty = 0$. Thus the error pattern is of the form $2\mathbf{e}_i$, where i is uniquely determined by b. If $a \neq 0$, then the error pattern is either of the form $\mathbf{e}_i + \mathbf{e}_k$ $(i \neq k)$ or of the form $-\mathbf{e}_i - \mathbf{e}_k$ $(i \neq k)$. For the first case we have

$$a + 2b = \xi^i + \xi^k.$$

Proceeding as above, we obtain

$$\bar{a} = \bar{\xi}^i + \bar{\xi}^k, \; \bar{b}^2 = \bar{\xi}^i \bar{\xi}^k$$

So $\bar{\xi}^i$ and $\bar{\xi}^k$ are distinct roots of the equation

$$X^2 + \bar{a}X + \bar{b}^2 = 0. \tag{9.15}$$

A necessary and sufficient condition for this equation to have distinct roots is that

$$\mathrm{Tr}(\bar{b}^2/\bar{a}^2) = \mathrm{Tr}(\bar{b}/\bar{a}) = 0,$$

(cf. MacWilliams and Sloane (1977), Chap. 9). Therefore if the condition $\mathrm{Tr}(\bar{b}, \bar{a}) = 0$ is fulfilled, then the error positions i and k can be determined by solving Eq. (9.15).

Then consider the second case. We have

$$a + 2b = -\xi^i - \xi^k,$$

where $a \neq 0$. Proceeding as above, we find

$$\bar{a} = \bar{\xi}^i + \bar{\xi}^k, \quad (\bar{b} + \bar{a})^2 = \bar{\xi}^j \bar{\xi}^k.$$

So $\bar{\xi}^i$, $\bar{\xi}^k$ are distinct roots of the equation

$$X^2 + \bar{a}X + (\bar{a}^2 + \bar{b}^2) = 0. \tag{9.16}$$

A necessary and sufficient condition for this equation to have distinct roots is that

$$\mathrm{Tr}\left(\frac{\bar{a}^2 + \bar{b}^2}{\bar{a}^2}\right) = \mathrm{Tr}\left(1 + \frac{\bar{b}}{\bar{a}}\right) = 1 + \mathrm{Tr}\left(\frac{\bar{b}}{\bar{a}}\right) = 0.$$

Thus if the condition $1 + \mathrm{Tr}(\bar{b}/\bar{a}) = 0$ is fulfilled, then the error positions i and k can be determined by solving Eq. (9.16).

A decision tree for the alogrithm is shown in Fig. 9.1.

9.4. The Preparata Codes

Let m be an odd integer ≥ 3 and $n = 2^m - 1$. We are going to construct a binary code of length 2^{m+1}. The vectors in $\mathbb{F}_2^{2^{m+1}}$ are written in the form $(\mathbf{x}, \mathbf{y})$, where $\mathbf{x}$, $\mathbf{y} \in \mathbb{F}_2^{2^m}$, the positions of $\mathbf{x}$ and $\mathbf{y}$ are both numbered by the 2^m elements of $\mathbb{F}_{2^m}$, the zero element of $\mathbb{F}_{2^m}$ corresponds to the first position in $\mathbf{x}$ and $\mathbf{y}$, and the components at the αth positions in $\mathbf{x}$ and $\mathbf{y}$ are denoted by x_α and y_α, respectively, where $\alpha \in \mathbb{F}_{2^m}$.

Definition 9.2. The *Preparata code* P_{m+1} of length 2^{m+1} consists of all codewords $(\mathbf{x}, \mathbf{y})$, where $\mathbf{x}, \mathbf{y} \in \mathbb{F}_2^{2^m}$, satisfying

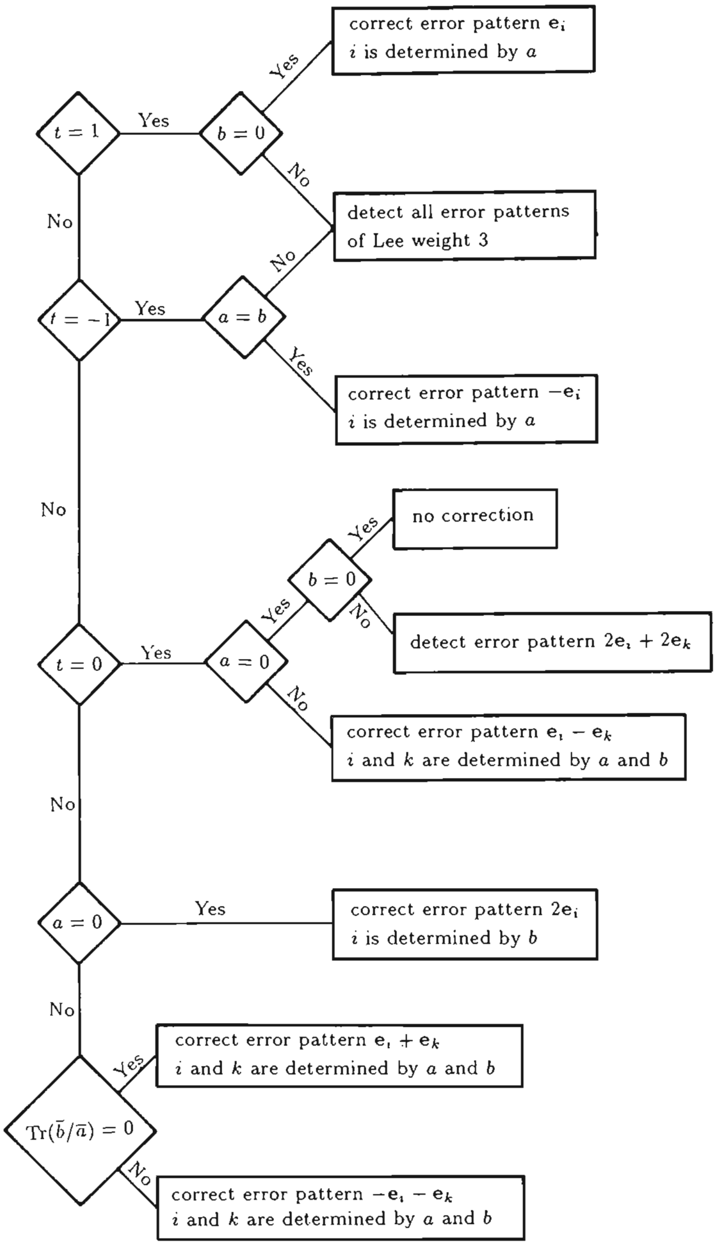

Fig. 9.1. A decoding algorithm for $\mathcal{P}(m)$.

1° Both $w(\mathbf{x})$ and $w(\mathbf{y})$ are even.
2° $\sum_{x_\alpha=1}\alpha=\sum_{y_\alpha=1}\alpha$.
3° $\sum_{x_\alpha=1}\alpha^3+\left(\sum_{x_\alpha=1}\alpha\right)^3=\sum_{y_\alpha=1}\alpha^3$.

The code obtained by deleting the first coordinates is called the *shortened Preparata code* of length 2^{m+1}, denoted by $P(m)^-$. □

The following identity will be used quite often in the following.

$$(a+b)^3=a^3+a^2b+ab^2+b^3 \quad \text{for all} \quad a,b\in\mathbb{F}_{2^m}. \tag{9.17}$$

The study of the properties of the Kerdock code P_{m+1} becomes easier if we find some automorphisms of the code first.

Lemma 9.11. *The group* Aut P_{m+1} *contains the permutations*

(i) $(\mathbf{x},\mathbf{y})\to(\mathbf{x}',\mathbf{y}')$, where $x'_\alpha=x_{\alpha+c}$, $y'=y_{\alpha+c}$ for any $c\in\mathbb{F}_q$,
(ii) $(\mathbf{x},\mathbf{y})\to(\mathbf{y},\mathbf{x})$,
(iii) $(\mathbf{x},\mathbf{y})\to(\mathbf{x}',\mathbf{y}')$, where $x'_\alpha=x_{\beta\alpha}$, $y'_\alpha=y_{\beta\alpha}$ for any $\beta\in\mathbb{F}_q^*$,
(iv) $(\mathbf{x},\mathbf{y})\to(\mathbf{x}',\mathbf{y}')$, where $x'_\alpha=x_{\alpha^2}$, $y'_\alpha=y_{\alpha^2}$.

Proof. We check only condition 3° for the map (i) since all other properties are trivially true. We have

$$\begin{aligned}
&\sum_{x'_\alpha=1}\alpha^3+\left(\sum_{x'_\alpha=1}\alpha\right)^3\\
&=\sum_{x_\alpha=1}(\alpha+c)^3+\left(\sum_{x_\alpha=1}(\alpha+c)\right)^3\\
&=\sum_{x_\alpha=1}(\alpha+c)^3+\left(\sum_{x_\alpha=1}\alpha\right)^3 && (w(\mathbf{x})\text{ is even})\\
&=\sum_{x_\alpha=1}(\alpha^3+\alpha^2c+\alpha c^2+c^3)+\left(\sum_{x_\alpha=1}\alpha\right)^3 && (\text{By }(9.17))\\
&=\sum_{x_\alpha=1}\alpha^3+\left(\sum_{x_\alpha=1}\alpha^2\right)c+\left(\sum_{x_\alpha=1}\alpha\right)c^2+\left(\sum_{x_\alpha=1}\alpha\right)^3 && (w(\mathbf{x})\text{ is even})
\end{aligned}$$

$$= \sum_{y_\alpha=1} \alpha^3 + \left(\sum_{y_\alpha=1} \alpha^2\right) c + \left(\sum_{y_\alpha=1} \alpha\right) c^2 \qquad \text{(Conditions } 2^\circ \text{ and } 3^\circ)$$

$$= \sum_{y_\alpha=1} (\alpha^3 + \alpha^2 c + \alpha c^2 + c^3) \qquad (w(\mathbf{y}) \text{ is even})$$

$$= \sum_{y_\alpha=1} (\alpha + c)^3 \qquad \text{(By (9.17))}$$

$$= \sum_{y_{\alpha'}=1} \alpha^3. \qquad \square$$

Proposition 9.12. *The binary code P_{m+1} is distance invariant, has minimum distance* 6, *and has 2^k codewords, where $k = 2^{m+1} - 2m - 2$.*

Proof. First we prove that P_{m+1} is distance invariant. Let $(\mathbf{u}, \mathbf{v})$ be any codeword of P_{m+1}. Let $\alpha_0 = \sum_{u_\alpha=1} \alpha$. Consider the map

$$P_{m+1} \to \mathbb{F}_2^{2^{m+1}}$$

$$(\mathbf{x}, \mathbf{y}) \to (\mathbf{x}', \mathbf{y}'),$$

where

$$x'_\alpha = x_{\alpha+\alpha_0} + u_\alpha,$$

$$y'_\alpha = y_{\alpha+\alpha_0} + v_\alpha.$$

Assume that $(\mathbf{x}, \mathbf{y}) \in P_{m+1}$, we want to show that $(\mathbf{x}', \mathbf{y}') \in P_{m+1}$ also. Conditions 1° and 2° are easily checked. For condition 3°, we compute

$$\sum_{x'_\alpha=1} \alpha^3 + \left(\sum_{x'_\alpha=1} \alpha\right)^3 = \sum_{x_{\alpha+\alpha_0}=1} \alpha^3 + \sum_{u_\alpha=1} \alpha^3 + \left(\sum_{x_{\alpha+\alpha_0}=1} \alpha\right)^3 + \left(\sum_{u_\alpha=1} \alpha\right)^3$$

We have

$$\sum_{x_{\alpha+\alpha_0}=1} \alpha^3 = \sum_{x_\alpha=1} (\alpha + \alpha_0)^3$$

$$= \sum_{x_\alpha=1} \alpha^3 + \sum_{x_\alpha=1} \alpha^2 \alpha_0 + \sum_{x_\alpha=1} \alpha \alpha_0^2 + \sum_{x_\alpha=1} \alpha_0^3$$

$$= \sum_{x_\alpha=1} \alpha^3 + \left(\sum_{x_\alpha=1} \alpha\right)^2 \alpha_0 + \left(\sum_{x_\alpha=1} \alpha\right) \alpha_0^2,$$

since $\sum_{x_\alpha=1} \alpha_0^3 = 0$. We also have

$$\sum_{x_{\alpha+\alpha_0}=1} \alpha = \sum_{x_\alpha=1} (\alpha+\alpha_0) = \sum_{x_\alpha=1} \alpha.$$

So,

$$\begin{aligned}
\sum_{x'_\alpha=1} \alpha^3 + \left(\sum_{x'_\alpha=1} \alpha\right)^3 &= \sum_{x_\alpha=1} \alpha^3 + \left(\sum_{x_\alpha=1} \alpha\right)^2 \alpha_0 + \left(\sum_{x_\alpha=1} \alpha\right) \alpha_0^2 \\
&\quad + \sum_{u_\alpha=1} \alpha^3 + \left(\sum_{x_\alpha=1} \alpha\right)^3 + \left(\sum_{u_\alpha=1} \alpha\right)^3 \\
&= \sum_{y_\alpha=1} \alpha^3 + \left(\sum_{y_\alpha=1} \alpha\right)^2 \alpha_0 \\
&\quad + \left(\sum_{y_\alpha=1} \alpha\right) \alpha_0^2 + \sum_{v_\alpha=1} \alpha^3.
\end{aligned}$$

But

$$\begin{aligned}
\sum_{y'_\alpha=1} \alpha^3 &= \sum_{y_{\alpha+\alpha_0}=1} \alpha^3 + \sum_{v_\alpha=1} \alpha^3 \\
&= \sum_{y_\alpha=1} (\alpha+\alpha_0)^3 + \sum_{v_\alpha=1} \alpha^3 \\
&= \sum_{y_\alpha=1} \alpha^3 + \left(\sum_{y_\alpha=1} \alpha\right)^2 \alpha_0 + \left(\sum_{y_\alpha=1} \alpha\right) \alpha_0^2 + \sum_{v_\alpha=1} \alpha^3,
\end{aligned}$$

since $\sum_{y_\alpha=1} \alpha_0^3 = 0$. Therefore

$$\sum_{x'_\alpha=1} \alpha^3 + \left(\sum_{x'_\alpha=1} \alpha\right)^3 = \sum_{y'_\alpha=1} \alpha^3.$$

Hence $(\mathbf{x}', \mathbf{y}') \in P_{m+1}$. Clearly, the map $(\mathbf{x}, \mathbf{y}) \to (\mathbf{x}', \mathbf{y}')$ is an injection from P_{m+1} to P_{m+1}. Since P_{m+1} is a finite set, it is a bijection. It is clear that for all $(\mathbf{x}, \mathbf{y}) \in P_{m+1}$

$$\begin{aligned}
d((\mathbf{x}', \mathbf{y}'), (\mathbf{u}, \mathbf{v})) &= w((\mathbf{x}' - \mathbf{u},\ \mathbf{y}' - \mathbf{v})) \\
&= w((\mathbf{x}, \mathbf{y})) \\
&= d((\mathbf{x}, \mathbf{y}), (\mathbf{0}, \mathbf{0})).
\end{aligned}$$

Therefore P_{m+1} is distance invariant.

Next we prove that P_{m+1} has minimum distance 6. It is enough to show that the minimum weight is 6. Obviously, there are no codewords of weight 2.

Let us prove that $P(m)$ has no codewords of weight 4. First, assume that there is a codeword $(\mathbf{x}, \mathbf{y})$, where $x_\alpha = x_\beta = y_\gamma = y_\delta = 1$, $\alpha \neq \beta$, $\gamma \neq \delta$, and all other components of $\mathbf{x}$ and $\mathbf{y}$ are zeros. By the distance invariance of P_{m+1} and Lemma 9.11 (i) we can assume that $\alpha = 0$. Then condition 3° of Definition 9.2 yields $\gamma^3 + \delta^3 = 0$, which implies $(\gamma\delta^{-1})^3 = 1$. Since m is odd, $(3, 2^m - 1) = 1$. Thus we get a contradiction. Then assume that there is a codeword $(\mathbf{x}, \mathbf{y})$ with $w(\mathbf{x}) = 4$ and $\mathbf{y} = \mathbf{0}$. Let $x_0 = x_\alpha = x_\beta = x_\gamma = 1$ where 0, α, β, γ are four distinct elements of $\mathbb{F}_{2^m}$. Then conditions 2° and 3° of Definition 9.2 imply

$$\alpha + \beta + \gamma = 0,$$

$$\alpha^3 + \beta^3 + \gamma^3 = 0.$$

Substituting the first equation into the second and then using (9.17), we obtain $\alpha\beta(\alpha + \beta) = 0$, whence $\alpha = \beta$, a contradiction. Similarly, there is no codeword $(\mathbf{x}, \mathbf{y})$ with $\mathbf{x} = \mathbf{0}$ and $w(\mathbf{y}) = 4$.

Now we prove that there are indeed codewords of weight 6 in P_{m+1}. Let α, β, γ be three distinct elements of $\mathbb{F}_{2^m}$. Without loss of generality we can assume that $\beta \neq 0$ and $\gamma \neq 0$. Define λ by $\lambda^3 = \alpha^3 + \beta^3 + \gamma^3$ We assert that $\lambda \neq \alpha, \beta, \gamma$; otherwise, assume that $\lambda = \alpha$, then $\beta^3 + \gamma^3 = 0$, which leads to a contradiction as before. Then define μ by $\mu = \alpha + \beta + \gamma + \lambda$. We assert that $\mu \neq 0$; otherwise we have both equations $\alpha + \beta + \gamma + \lambda = 0$ and $\alpha^3 + \beta^3 + \gamma^3 + \lambda^3 = 0$. Then $\alpha + \beta = \gamma + \lambda$ and $\alpha^3 + \beta^3 = \gamma^3 + \lambda^3$ Factorizing the second equation and then using the first equation, we obtain $\alpha\beta = \gamma\lambda$. Similarly, $\alpha\gamma = \beta\lambda$. Thus $\alpha^2\beta\gamma = \lambda^2\beta\gamma$. Since $\beta\gamma \neq 0$, we have $\alpha^2 = \lambda^2$ and $\alpha = \lambda$, a contradiction. Then $(\mathbf{x}, \mathbf{y})$ with $x_0 = x_\mu = y_\alpha = y_\beta = y_\gamma = y_\lambda = 1$ and all other components zero is a codeword of weight 6 in P_{m+1}.

Finally, let us compute the number of codewords of P_{m+1}. A vector $\mathbf{x} \in \mathbb{F}_2^{2^m}$ satisfying condition 1° of Definition 9.2 can be chosen in $2^{2^m - 1}$ ways. We now count for a given $\mathbf{x} \in \mathbb{F}_2^{2^m}$ satisfying condition 1°, how many $(y_\alpha; \alpha \in \mathbb{F}_{2^m}^*)$'s in $\mathbb{F}_2^{2^m - 1}$ satisfy conditions 2° and 3° For such a $(y_\alpha; \alpha \in \mathbb{F}_{2^m}^*)$ define $y_0 = \sum_{\alpha \in \mathbb{F}_{2^m}^*} y_\alpha$, then we get a codeword $(\mathbf{x}, \mathbf{y}) \in P_{m+1}$. Let $\bar{\xi}$ be a primitive element of $\mathbb{F}_{2^m}$, then conditions 2° and 3° can be regarded as two equations in $2^m - 1$ unknowns $a_0, a_1, a_2, \ldots, a_{n-1}$:

$$\left.\begin{aligned} \sum_{x_\alpha = 1} \alpha &= a_0 + a_1\bar{\xi} + a_2\bar{\xi}^2 + \cdots + a_{n-1}\bar{\xi}^{n-1}, \\ \sum_{x_\alpha = 1} \alpha^3 + \left(\sum_{x_\alpha = 1} \alpha\right)^3 &= a_0 + a_1\bar{\xi}^3 + a_2\bar{\xi}^6 + \cdots + a_{n-1}\bar{\xi}^{3(n-1)}. \end{aligned}\right\} \quad (9.18)$$

Express each $\bar{\xi}^j$, and also $\sum_{x_\alpha=1}\alpha$ and $\sum_{x_\alpha=1}\alpha^3 + \left(\sum_{x_\alpha=1}\alpha\right)^3$, as linear combinations in $1, \bar{\xi}, \bar{\xi}^2, \ldots, \bar{\xi}^{m-1}$ with coefficients in $\mathbb{F}_2$, these two equations becomes $2m$ linear equations in $a_0, a_1, a_2, \ldots, a_{n-1}$ with coefficients in $\mathbb{F}_2$. We claim that these $2m$ linear equations are linearly independent. Denote by $m_i(X)$ the minimum polynomial of $\bar{\xi}^i$. Clearly, $m_1(X)$ is a degree m. Since m is odd, $(3, 2^m - 1) = 1$. Thus $m_3(X)$ is also of degree m. Then the binary cyclic code C of length $n = 2^m - 1$ with generator polynomial $m_1(X)m_3(X)$ has dimension $n - 2m = 2^m - 2m - 1$. A word $c(X) = c_0 + c_1X + c_2X^2 + \cdots + c_{n-1}X^{n-1}$ is a codeword of C if and only if $c(\bar{\xi}) = c(\bar{\xi}^3) = 0$, i.e., if and only if $(c_0, c_1, c_2, \ldots, c_{n-1})$ is a solution of the linear homogeneous equations corresponding to (9.18). This proves the linear independence of (9.18). Therefore for each choice of $\mathbf{x} \in \mathbb{F}_2^{2^m}$ with $w(\mathbf{x})$ being even, there are 2^{2^m-2m-1} choices of $(y_\alpha, \alpha \in \mathbb{F}_{2^m}^*)$ such that (9.18) holds. Hence $|P_{m+1}| = 2^{2^m-1} \cdot 2^{2^m-2m-1} = 2^{2^{m+1}-2m-2}$ □

Corollary 9.13. *The shortened Preparata code $P(m)^-$ is a binary nonlinear code of length $2^m - 1$ and has minimum distance* 5. □

The Preparata codes were introduced by Preparata (1968) and their weight distribution were obtained by Semankov and Zinovév (1969), see also Chap. 5 of MacWilliams and Sloane (1977). After the Kerdock codes were introduced by Kerdock (1972) and their weight distributions were computed, it was found that the weight enumerator of the Preparata code P_{m+1} is the MacWilliams transform of the weigth enumerator of the Kerdock code K_{m+1}, (see Theorem 24, Chap. 5 of MacWilliams and Sloane (1977)), which was regarded as a mystery in coding theory. Hammons *et al.* (1994) explains this conundrum by showing that a variant of P_{m+1}, i.e., $P(m)$, is the formal dual of K_{m+1}. By Theorem 9.9, the weight enumerator of $P(m)$ is the MacWilliams transform of the weight enumerator of K_{m+1}. Therefore we have

Proposition 9.14. *The Preparata code P_{m+1} and the "Preparata" code $P(m)$ have the same length, the same number of codewords, the same minimum distance, and the same weight enumerator.* □

However, in contrast to Proposition 9.10 we have

Proposition 9.15. *The Preparata code P_{m+1} of length 2^{m+1} is a subcode of the extended binary Hamming code of the same length.*

Proof. Denote the code P_{m+1} by C_0. To each $\beta \in \mathbb{F}_{2^m}^*$ we associate a word $(\mathbf{u}^{(\beta)}, \mathbf{v}^{(\beta)})$, where $u_0^{(\beta)} = u_\beta^{(\beta)} = v_0^{(\beta)} = v_\beta^{(\beta)} = 1$ and all other components of $\mathbf{u}^{(\beta)}$ and $\mathbf{v}^{(\beta)}$ are zeros. Then we define the code

$$C_\beta = \{(\mathbf{x}, \mathbf{y}) + (\mathbf{u}^{(\beta)}, \mathbf{v}^{(\beta)}) | (\mathbf{x}, \mathbf{y}) \in C_0\}.$$

We assert that C_β has minimum weight 4. Clearly, $w(\mathbf{u}^{(\beta)}, \mathbf{v}^{(\beta)}) = 4$. We have to show that $w((\mathbf{x}, \mathbf{y}) + (\mathbf{u}^{(\beta)}, \mathbf{v}^{(\beta)})) \geq 4$ for all $(\mathbf{x}, \mathbf{y}) \in C_0$. If $w(\mathbf{x}, \mathbf{y}) \geq 8$, this is obvious. Now assume that $w(\mathbf{x}, \mathbf{y}) = 6$ and $w((\mathbf{x}, \mathbf{y}) + (\mathbf{u}^{(\beta)}, \mathbf{v}^{(\beta)})) < 4$. Then $w((\mathbf{x}, \mathbf{y}) + (\mathbf{u}^{(\beta)}, \mathbf{v}^{(\beta)})) = 2$. We can assume that

$$x_0 = x_\beta = y_0 = y_\beta = y_\gamma = y_\delta = 1$$

or

$$x_0 = x_\beta = x_\gamma = x_\delta = y_0 = y_\beta = 1,$$

where 0, β, γ, δ are distinct elements of $\mathbb{F}_{2^m}$, while all the other components are zeros. For both cases, by 2° we have $\beta = \beta + \gamma + \delta$, which implies $\gamma = \delta$, a contradiction. Our assertion is proved.

Next we assert that the codes $C_\beta (\beta \in \mathbb{F}_{2^m})$ are pairwise disjoint. Assume that $C_\beta \cap C_\gamma \neq \emptyset$ for a pair of distinct elements β, $\gamma \in \mathbb{F}_{2^m}^*$. Then there are two distinct codewords $(\mathbf{x}, \mathbf{y})$ and $(\mathbf{x}', \mathbf{y}')$ of C_0 such that

$$(\mathbf{x}, \mathbf{y}) + (\mathbf{u}^{(\beta)}, \mathbf{v}^{(\beta)}) = (\mathbf{x}', \mathbf{y}') + (\mathbf{u}^{(\gamma)}, \mathbf{v}^{(\gamma)}).$$

Transposing, we get

$$(\mathbf{x}, \mathbf{y}) - (\mathbf{x}', \mathbf{y}') = (\mathbf{u}^{(\gamma)}, \mathbf{v}^{(\gamma)}) - (\mathbf{u}^{(\beta)}, \mathbf{v}^{(\beta)}).$$

Thus

$$d((\mathbf{x}, \mathbf{y}), (\mathbf{x}', \mathbf{y}')) = d((\mathbf{u}^{(\gamma)}, \mathbf{v}^{(\gamma)}), (\mathbf{u}^{(\beta)}, \mathbf{v}^{(\beta)})) = 4,$$

a contradiction. Similarly, $C_\beta \cap C_0 = \emptyset$ for all $\beta = \mathbb{F}_{2^m}^*$.

Now let

$$C = \bigcup_{\beta \in \mathbb{F}_{2^m}} C_\beta.$$

We claim that C is a linear code. Let $(\mathbf{x}, \mathbf{y}) + (\mathbf{u}^{(\beta)}, \mathbf{v}^{(\beta)})$ and $(\mathbf{x}', \mathbf{y}') + (\mathbf{u}^{(\gamma)}, \mathbf{v}^{(\gamma)})$ be any two codewords of C, where $(\mathbf{x}, \mathbf{y})$ and $(\mathbf{x}', \mathbf{y}')$ are codewords of C_0, and β, $\gamma \in \mathbb{F}_{2^m}$. If $\beta = 0$, we agree that $(\mathbf{u}^{(\beta)}, \mathbf{v}^{(\beta)}) = (\mathbf{0}, \mathbf{0})$, where $\mathbf{0} = 0^{2^m}$. Similarly, if $\gamma = 0$, we agree that $(\mathbf{u}^{(\gamma)}, \mathbf{v}^{(\gamma)}) = (\mathbf{0}, \mathbf{0})$. We assert that there is a codeword $(\mathbf{x}'', \mathbf{y}'') \in C_0$ and a $\delta \in \mathbb{F}_{2^m}$ such that

$$(\mathbf{x}, \mathbf{y}) + (\mathbf{u}^{(\beta)}, \mathbf{v}^{(\beta)}) + (\mathbf{x}', \mathbf{y}') + (\mathbf{u}^{(\gamma)}, \mathbf{v}^{(\gamma)}) = (\mathbf{x}'', \mathbf{y}'') + (\mathbf{u}^{(\delta)}, \mathbf{v}^{(\delta)}),$$

i.e.,

$$\mathbf{x}'' = \mathbf{x} + \mathbf{x}' + \mathbf{u}^{(\beta)} + \mathbf{u}^{(\gamma)} + \mathbf{u}^{(\delta)} \tag{9.19}$$

and

$$\mathbf{y}'' = \mathbf{y} + \mathbf{y}' + \mathbf{v}^{(\beta)} + \mathbf{v}^{(\gamma)} + \mathbf{v}^{(\delta)}. \tag{9.20}$$

For any $\delta \in \mathbb{F}_{2^m}$, define $\mathbf{x}''$ and $\mathbf{y}''$ by (9.19) and (9.20), respectively, then clearly $(\mathbf{x}'', \mathbf{y}'')$ satisfies conditions 1° and 2° of Definition 9.2. Let us examine when condition 3° is also satisfied. We have

$$\begin{aligned}
\sum_{x''_\alpha=1} \alpha^3 + \left(\sum_{x''_\alpha=1} \alpha\right)^3 &= \sum_{x_\alpha=1} \alpha^3 + \sum_{x'_\alpha=1} \alpha^3 \\
&\quad + \beta^3 + \gamma^3 + \delta^3 + \left(\sum_{x_\alpha=1} \alpha + \sum_{x'_\alpha=1} \alpha + \beta + \gamma + \delta\right)^3, \\
\sum_{y''_\alpha=1} \alpha^3 &= \sum_{y_\alpha=1} \alpha^3 + \sum_{y'_\alpha=1} \alpha^3 + \beta^3 + \gamma^3 + \delta^3 \\
&= \sum_{x_\alpha=1} \alpha^3 + \left(\sum_{x_\alpha=1} \alpha\right)^3 + \sum_{x'_\alpha=1} \alpha^3 \\
&\quad + \left(\sum_{x'_\alpha=1} \alpha\right)^3 + \beta^3 + \gamma^3 + \delta^3.
\end{aligned}$$

Thus condition 3° for $(\mathbf{x}'', \mathbf{y}'')$ is equivalent to

$$\left(\sum_{x_\alpha=1} \alpha\right)^3 + \left(\sum_{x'_\alpha=1} \alpha\right)^3 = \left(\sum_{x_\alpha=1} \alpha + \sum_{x'_\alpha=1} \alpha + \beta + \gamma + \delta\right)^3,$$

which has a unique solution δ. With this δ, we can define $\mathbf{x}''$ and $\mathbf{y}''$ by (9.19) and (9.20), respectively. Then $(\mathbf{x}'', \mathbf{y}'') \in C_0$. Therefore we conclude that C is a binary linear code of length 2^{m+1}, with cardinality

$$|C| = |\mathbb{F}_{2^m}|\,|C_0| = 2^{2^{m+1}-m-2},$$

and has minimum distance 4. Therefore C must be the extended binary Hamming code of length 2^{m+1}. Clearly, $P_{m+1} \subseteq C$. □

The above description of the Preparata codes is due to Baker *et al.* (1983), but the proof of the distance invariance of the code P_{m+1} is different from theirs, which the author could not verify.

Clearly, both P_{m+1} and "$P(m)$" have the same length and minimum distance as the $[2^{m+1}, 2^{m+1}-2m-3, 6]$ extended BCH code, but contain twice as many codewords. It is also known that P_{m+1} has the greatest possible number of codewords for this minimum distance, (see Chap. 17 of MacWilliams and Sloane (1977)). So does "$P(m)$".

CHAPTER 10

GENERALIZATIONS OF QUATERNARY KERDOCK AND PREPARATA CODES

10.1. Quaternary Reed–Muller Codes

From the definitions of the quaternary codes $\mathcal{K}(m)$ and $\mathcal{P}(m)$ we see that they can be regarded as the $\mathbb{Z}_4$-analogs of the binary first-order Reed–Muller code $\mathrm{RM}(1, m)$ and the $(m-2)$th-order Reed–Muller code $\mathrm{RM}(m-2, m) = \mathrm{RM}(1, m)^{\perp}$, respectively. This suggests us to define the quaternary Reed–Muller codes $\mathrm{QRM}(r, m)$ of any order $r, 0 \leq r \leq m$, which are $\mathbb{Z}_4$-analogs of the binary Reed–Muller codes $\mathrm{RM}(r, m)$ of order r and includes the codes $\mathcal{K}(m)$ and $\mathcal{P}(m)$ as special cases.

Let m be an integer ≥ 2 and $n = 2^m - 1$. Let $h(X)$ be a basic primitive polynomial of degree m dividing $X^n - 1$ and ξ be one of its roots. Then the m distinct roots of $h(X)$ are ξ, $\xi^2, \ldots, \xi^{2^{m-1}}$, and ξ is of order $2^m - 1$. Consider the $(m+1) \times 2^m$ matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \xi & \xi^2 & \cdots & \xi^{n-1} \end{pmatrix}, \tag{10.1}$$

whose rows are numbered by $0, 1, 2, \ldots, m$ and columns by $\infty, 0, 1, 2, \ldots, n-1$, where ξ^j should be replaced by ${}^t(b_{1j}, b_{2j}, \ldots, b_{mj})$ if $\xi^j = b_{1j} + b_{2j}\xi + \cdots + b_{mj}\xi^{m-1}$ ($j = \infty, 0, 1, \ldots, n-1$) and we agree that $\xi^{\infty} = 0$. Denote the ith row of the matrix (10.1) by $\mathbf{v}_i$. Then $\mathbf{v}_i$ ($i = 0, 1, 2, \ldots, m$) are 2^m-tuples over $\mathbb{Z}_4$ and $\mathbf{v}_0$ is the all 1 2^m-tuple 1^{2^m}. Define a componentwise multiplication of 2^m-tuples in $\mathbb{Z}_4^{2^m}$ as follows:

$$(x_\infty, x_0, x_1, \ldots, x_{n-1})(y_\infty, y_0, y_1, \ldots, y_{n-1})$$
$$= (x_\infty y_\infty, x_0y_0, x_1y_1, \ldots, x_{n-1}y_{n-1}),$$

where, for simplicity, we use the concatenation to denote the componentwise multiplication instead of the symbol $*$ used previously.

Definition 10.1. Let m be an integer ≥ 2, $n = 2^m - 1$, and r be an integer such that $0 \leq r \leq m$. The *quaternary rth-order Reed–Muller code* QRM(r, m) is the code generated by all 2^m-tuples of the form

$$\mathbf{v}_{i_1}\mathbf{v}_{i_2}\cdots\mathbf{v}_{i_s}, \quad 1 \leq i_1 < i_2 < \cdots < i_s \leq m, \quad 0 \leq s \leq r.$$

We agree that $\mathbf{v}_{i_1}\mathbf{v}_{i_2}\cdots\mathbf{v}_{i_s} = 1^{2^m}$, when $s = 0$. □

From this definition the following propositions follow immediately.

Proposition 10.1. QRM$(1, m) = \mathcal{K}(m)$.

Proof. By Proposition 8.2 and Definition 10.1. □

Proposition 10.2. $\alpha(\text{QRM}(r, m)) = \text{RM}(r, m)$.

Proof. By Definition 10.1 and the definition of binary Reed–Muller codes. □

Now let us prove the following lemma.

Lemma 10.3. *The following* 2^m 2^m*-tuples over* $\mathbb{Z}_4$

$$\mathbf{v}_{i_1}\mathbf{v}_{i_2}\cdots\mathbf{v}_{i_s}, \quad 1 \leq i_1 < i_2 < \cdots < i_s \leq m, \quad 0 \leq s \leq m, \tag{10.2}$$

form a basis of the free $\mathbb{Z}_4$*-module* $\mathbb{Z}_4^{2^m}$

Proof. From the theory of binary Reed–Muller codes of length 2^m, it is well-known that the following 2^m 2^m-tuples over $\mathbb{Z}_2$

$$\overline{\mathbf{v}}_{i_1}\overline{\mathbf{v}}_{i_2}\cdots\overline{\mathbf{v}}_{i_s}, \quad 1 \leq i_1 < i_2 < \cdots < i_s \leq m, \quad 0 \leq s \leq m, \tag{10.3}$$

form a basis of the vector space $\mathbb{Z}_2^{2^m}$ over $\mathbb{Z}_2$. For any $\mathbf{v} \in \mathbb{Z}_4^{2^m}$, we have $\overline{\mathbf{v}} \in \mathbb{Z}_2^{2^m}$. Then there are elements $a_{i_1 i_2 \cdots i_s} \in \mathbb{Z}_2$ $(1 \leq i_1 < i_2 < \cdots < i_s \leq m, 0 \leq s \leq m)$ such that

$$\overline{\mathbf{v}} = \sum_{s=0}^{m} \sum_{1 \le i_1 < \cdots < i_s \le m} a_{i_1 i_2 \cdots i_s} \overline{\mathbf{v}}_{i_1} \overline{\mathbf{v}}_{i_2} \cdots \overline{\mathbf{v}}_{i_s}.$$

Thus

$$\mathbf{v} = \sum_{s=0}^{m} \sum_{1 \le i_1 < \cdots < i_s \le m} a_{i_1 i_2 \cdots i_s} \mathbf{v}_{i_1} \mathbf{v}_{i_2} \cdots \mathbf{v}_{i_s} + 2\mathbf{u}, \tag{10.4}$$

where $\mathbf{u} \in \mathbb{Z}_4^{2^m}$. Similarly, there are elements $b_{i_1 i_2 \cdots i_s} \in \mathbb{Z}_2$ such that

$$\mathbf{u} = \sum_{s=0}^{m} \sum_{1 \le i_1 < \cdots < i_s \le m} b_{i_1 i_2 \cdots i_s} \mathbf{v}_{i_1} \mathbf{v}_{i_2} \cdots \mathbf{v}_{i_s} + 2\mathbf{w}, \tag{10.5}$$

where $\mathbf{w} \in \mathbb{Z}_4^{2^m}$. Substituting (10.5) into (10.4) we obtain

$$\mathbf{v} = \sum_{s=0}^{m} \sum_{1 \le i_1 < \cdots < i_s \le m} (a_{i_1 i_2 \cdots i_s} + 2b_{i_1 i_2 \cdots i_s}) \mathbf{v}_{i_1} \mathbf{v}_{i_2} \cdots \mathbf{v}_{i_s}.$$

It follows that the 2^m 2^m-tuples over $\mathbb{Z}_4$ (10.2) form a basis of the free $\mathbb{Z}_4$-module $\mathbb{Z}_4^{2^m}$. □

Corollary 10.4. *For* $0 \le r \le m$, QRM(r, m) *is of type* $4^{K_{r,m}}$, *where*

$$K_{r,m} = 1 + \binom{m}{1} + \binom{m}{2} + \cdots + \binom{m}{r}$$ □

Digression. Let j be a positive integer and let the dyadic expansion of j be

$$j = a_0 2^0 + a_1 2^1 + a_2 2^2 + \cdots + a_l 2^l, \tag{10.6}$$

where

$$a_0, a_1, a_2, \ldots, a_{l-1} = 0 \text{ or } 1, \quad \text{and} \quad a_l = 1.$$

For example,

$$3 = 1 \cdot 2^0 + 1 \cdot 2^1, \quad 9 = 1 \cdot 2^0 + 0 \cdot 2^1 + 0 \cdot 2^2 + 1 \cdot 2^3,$$

$$26 = 0 \cdot 2^0 + 1 \cdot 2^1 + 0 \cdot 2^2 + 1 \cdot 2^3 + 1 \cdot 2^4, \text{etc.}$$

The number of 1's among the coefficients $a_0, a_1, a_2, \ldots, a_l$ in the dyadic expansion (10.6) of j will be called the 2-weight of j and denoted by $w_2(j)$, i.e.,

$$w_2(j) = a_0 + a_1 + a_2 + \cdots + a_l.$$

For example,

$$w_2(3) = 2, \quad w_2(9) = 2, \quad w_2(26) = 3, \quad \text{etc.}$$

We define the 2-weight of 0, denoted by $w_2(0)$, to be 0, i.e., $w_2(0) = 0$.

Let m be a fixed positive integer. Let r and s be integers such that $0 \le r,\ s \le 2^m - 2$. We define r and s to be equivalent, if there is a non-negative integer i such that $2^i r \equiv s \pmod{2^m - 1}$. Clearly, this defines an equivalence relation in the set of integers $\{0,\ 1,\ 2, \ldots, 2^m - 2\}$. The equivalence classes are called the *cyclotomic cosets* mod $2^m - 1$. For example, when $m = 4$, the cyclotomic cosets mod $2^4 - 1$ are

$$\{0\},\ \{1,2,4,8\},\ \{3,6,12,9\},\ \{5,10\},\ \{7,14,13,11\}.$$

Clearly, if r and s belong to the same cyclotomic coset, then $w_2(r) = w_2(s)$. A number in a cyclotomic coset is called a representative of the cyclotomic coset.

The following proposition gives an equivalent definition of the quaternary Reed–Muller code $\mathrm{QRM}(r, m)$, $0 \le r \le m$.

Proposition 10.5. *Let m be an integer ≥ 2. Then $\mathrm{QRM}(0, m)$ is the $\mathbb{Z}_4$-repetition code $\{\varepsilon 1^{2^m} | \varepsilon \in \mathbb{Z}_4\}$ of length 2^m, and for $1 \le r \le m$ $\mathrm{QRM}(r, m)$ is generated by $\mathrm{QRM}(0, m)$ together with all 2^m-tuples of the form*

$$(T(\lambda_j\xi^\infty),\ T(\lambda_j\xi^0),\ T(\lambda_j\xi^j),\ T(\lambda_j\xi^{j\cdot 2}), \ldots, T(\lambda_j\xi^{j(n-1)})), \tag{10.7}$$

where j runs through a system of representatives of those cyclotomic cosets mod $2^m - 1$ *for which $w_2(j) \le r$ and λ_j runs through* $\mathrm{GR}(4^m)$.

Proof. By Definition 10.1, $\mathrm{QRM}(0, m) = \{\varepsilon 1^{2^m} | \varepsilon \in \mathbb{Z}_4\}$.

Now let $1 \le r \le m$ and denote the $\mathbb{Z}_4$-linear code generated by $\mathrm{QRM}(0, m)$ together with all 2^m-tuples of the form (10.7) by C_r. By Proposition 10.1, $\mathrm{QRM}(1, m) = \mathcal{K}(m)$ and by Proposition 8.6

$$\mathcal{K}(m) = \{\varepsilon 1^{2^m} + \mathbf{u}^{(\lambda)} | \varepsilon \in \mathbb{Z}_4,\ \lambda \in \mathrm{GR}(4^m)\},$$

where

$$\mathbf{u}^{(\lambda)} = (T(\lambda\xi^\infty),\ T(\lambda\xi^0),\ T(\lambda\xi),\ T(\lambda\xi^2), \ldots, T(\lambda\xi^{n-1})).$$

Therefore $\mathrm{QRM}(1, m) = C_1$. In particular, for each $\mathbf{v}_i (i = 1,\ 2, \ldots, m)$ there exists a unique $\mu_i \in \mathrm{GR}(4^m)$ such that

$$\mathbf{v}_i = (T(\mu_i\xi^\infty),\ T(\mu_i\xi^0),\ T(\mu_i\xi),\ T(\mu_i\xi^2), \ldots, T(\mu_i\xi^{n-1})).$$

Now let us consider the case $r = 2$. By Definition 10.1,

$$\mathrm{QRM}(2,m) = \left\{ \varepsilon 1^{2^m} + \sum_{i=1}^{m} a_i \mathbf{v}_i + \sum_{1 \le i < j \le m} b_{ij} \mathbf{v}_i \mathbf{v}_j | \varepsilon, a_i, b_{ij} \in \mathbb{Z}_4 \right\}.$$

We want to prove that $\mathrm{QRM}(2,m) = \mathcal{C}_2$. By the case $r = 1$, $\mathrm{QRM}(1,m) = \mathcal{C}_1 \subseteq \mathcal{C}_2$. Thus

$$\varepsilon 1^{2^m} + \sum_{i=1}^{m} a_i \mathbf{v}_i \in \mathcal{C}_2.$$

Let us prove that $\mathbf{v}_i \mathbf{v}_j \in \mathcal{C}_2$ for $1 \le i < j \le m$. Recall that $\mathbf{v}_i \mathbf{v}_j$ is a componentwise product:

$$\begin{aligned} \mathbf{v}_i \mathbf{v}_j = (&T(\mu_i \xi^{\infty}) T(\mu_j \xi^{\infty}),\, T(\mu_i \xi^0) T(\mu_j \xi^0),\, T(\mu_i \xi) T(\mu_j \xi), \\ &\ldots, T(\mu_i \xi^{n-1}) T(\mu_j \xi^{n-1})). \end{aligned}$$

For $k \in \{\infty, 0, 1, 2, \ldots, n-1\}$, we compute

$$\begin{aligned} T(\mu_i \xi^k) T(\mu_j \xi^k) &= \sum_{s=0}^{m-1} (\mu_i \xi^k)^{2^s} \sum_{t=0}^{m-1} (\mu_j \xi^k)^{2^t} \\ &= T(\mu_i \mu_j \xi^{2 \cdot k}) + T(\mu_i \mu_j^2 \xi^{(1+2)k}) \\ &\quad + \cdots + T(\mu_i \mu_j^{2^{m-1}} \xi^{(1+2^{m-1})k}). \end{aligned}$$

Clearly,

$$\begin{aligned} &(T(\mu_i \mu_j \xi^{\infty}),\, T(\mu_i \mu_j \xi^0),\, T(\mu_i \mu_j \xi^2), \\ &T(\mu_i \mu_j \xi^{2 \cdot 2}), \ldots, T(\mu_i \mu_j \xi^{2(n-1)})) \in \mathcal{C}_1 \subseteq \mathcal{C}_2 \end{aligned}$$

and for $1 \le l \le m-1$,

$$\begin{aligned} &(T(\mu_i \mu_j^{2^l} \xi^{\infty}),\, T(\mu_i \mu_j^{2^l} \xi^0),\, T(\mu_i \mu_j^{2^l} \xi^{1+2^l}), \\ &T(\mu_i \mu_j^{2^l} \xi^{(1+2^l)2}), \ldots, T(\mu_i \mu_j^{2^l} \xi^{(1+2^l)(n-1)})) \in \mathcal{C}_2. \end{aligned}$$

Therefore $\mathbf{v}_i \mathbf{v}_j \in \mathcal{C}_2$ for $1 \le i < j \le m$. It follows that $\mathrm{QRM}(2,m) \subseteq \mathcal{C}_2$.

On the other hand, denote the $\mathbb{Z}_4$-code obtained by deleting the components at position ∞ of codewords of $\mathcal{C}_2$ by $\mathcal{C}_2^-$. Clearly, $\mathcal{C}_2^-$ is a $\mathbb{Z}_4$-linear code generated by 1^{2^m-1} together with all (2^m-1)-tuples

$$(T(\lambda_j \xi^0),\, T(\lambda_j \xi^j),\, T(\lambda_j \xi^{j \cdot 2}), \ldots, T(\lambda_j \xi^{j(n-1)})),$$

where j runs through a system of representatives of those cyclotomic cosets mod $2^m - 1$ of which $w_2(j) \leq 2$ and λ_j runs through $\mathrm{GR}(4^m)$. As in the proof of Proposition 8.5, it is easy to verify that all these generators are annihilated by the polynomial

$$\tilde{h}_2(X) = (1 - X) \prod_{\substack{1 \leq j \leq 2^m - 2 \\ w_2(j) \leq 2}} (1 - \xi^j X).$$

$\tilde{h}_2(X)$ is the reciprocal polynomial to the polynomial

$$h_2(X) = (X - 1) \prod_{\substack{1 \leq j \leq 2^m - 2 \\ w_2(j) \leq 2}} (X - \xi^j).$$

Let $g_2(X)$ be the reciprocal polynomial to the polynomial

$$\frac{X^{2^m - 1} - 1}{h_2(X)} = \prod_{\substack{1 \leq j \leq 2^m - 2 \\ w_2(j) > 2}} (X - \xi^j)$$

and denote the $\mathbb{Z}_4$-cyclic code generated by $g_2(X)$ by C. Then $\tilde{h}_2(X)$ is the check polynomial of C. Therefore $C_2 \subseteq C$. Then $\mathrm{QRM}(2, m) \subseteq C_2 \subseteq C$. Clearly,

$$g_2(X) = \prod_{\substack{1 \leq j \leq 2^m - 2 \\ w_2(j) > 2}} (1 - \xi^j X).$$

It is known that

$$\{j | 1 \leq j \leq 2^m - 1,\ w_2(j) = r\} = \binom{m}{r}.$$

Thus

$$\deg g_2(X) = \binom{m}{3} + \binom{m}{4} + \cdots + \binom{m}{m-1}.$$

It follows that C is of type $4^{K_{2,m}}$, where

$$K_{2,m} = 2^m - 1 - \deg g_2(X) = 1 + \binom{m}{1} + \binom{m}{2}.$$

By Corollary 10.4, $\mathrm{QRM}(r, m)$ is of type $4^{K_{2,m}}$. Hence $\mathrm{QRM}(2, m) = C_2 = C$.

The cases $r \geq 3$ can be proved in the same way as $r = 2$. $\square$

Corollary 10.6. *Denote the $\mathbb{Z}_4$-code obtained by deleting the components at position ∞ of the codewords of* $\mathrm{QRM}(r, m)$ *by* $\mathrm{QRM}(r, m)^-$. *Then* $\mathrm{QRM}(r, m)^-$ *is a $\mathbb{Z}_4$-cyclic code with generator polynomial*

$$g_r(X) = \prod_{\substack{1 \le j \le 2^m - 2 \\ w_2(j) > r}} (1 - \xi^j X) = \varepsilon_r \prod_{\substack{1 \le j \le 2^m - 2 \\ w_2(j) < m - r}} (X - \xi^j),$$

where $\varepsilon_r = \pm 1$.

Corollary 10.7. *Let* m *be an integer* ≥ 2 *and* $0 \le r \le m-1$. *Then for any word* $\mathbf{c} = (c_\infty, c_0, c_1, \dots, c_{n-1}) \in \mathrm{QRM}(r, m)$, *we have* $c_\infty + c_0 + c_1 + \cdots + c_{n-1} = 0$ *in* $\mathbb{Z}_4$.

Proof. It is enough to prove our corollary for all generators of $\mathrm{QRM}(r,m)$ given in Proposition 10.5. First, since $m \ge 2$, for 1^{2^m} we have

$$\underbrace{1 + 1 + 1 + \cdots + 1}_{2^m} = 2^m = 0.$$

Second, for the 2^m-tuple (10.7) we have

$$\begin{aligned} T(\lambda_j \xi^\infty) &+ T(\lambda_j \xi^0) + T(\lambda_j \xi^j) + T(\lambda_j \xi^{j\cdot 2}) + \cdots + T(\lambda_j \xi^{j(n-1)}) \\ &= \sum_{i=0}^{n-1} T(\lambda_j \xi^{ji}) \\ &= \sum_{i=0}^{n-1} \sum_{k=0}^{m-1} (\lambda_j \xi^{ji})^{2^k} \\ &= \sum_{k=0}^{m-1} \lambda_j^{2^k} \sum_{i=0}^{n-1} \xi^{j\cdot 2^k \cdot i} \\ &= \sum_{k=0}^{m-1} \lambda_j^{2^k} \frac{1 - \xi^{j\cdot 2^k \cdot n}}{1 - \xi^{j \cdot 2^k}} \\ &= 0, \end{aligned}$$

since $\xi^n = \xi^{2^m - 1} = 1$ and $\xi^{j\cdot 2^k} \neq 1$ for $w_2(j) \le r \le m-1$. □

Proposition 10.8. *Let* m *be an integer* ≥ 2 *and* $0 \le r \le m-1$. *Then*

$$\mathrm{QRM}(r,m)^\perp = \mathrm{QRM}(m-r-1, m).$$

Proof. First, we prove that the all 1 2^m-tuple $1^{2^m} \in \mathrm{QRM}(m-r-1,m)$ belongs to $\mathrm{QRM}(r,m)^\perp$. Since $m \ge 2$, $1^{2^m} \cdot 1^{2^m} = 0$. Moreover, we have to prove that 1^{2^m} is orthogonal to all 2^m-tuples of the form (10.7), where $w_2(j) \le r$. By the proof of Corollary 10.7, we have

$$\sum_{i=0}^{n-1} T(\lambda_j \xi^{ji}) = 0.$$

Therefore $1^{2^m} \in \mathrm{QRM}(r,m)^\perp$

Next, we prove that any $\mathbf{c} = (c_\infty, c_0, c_1, c_2, \dots, c_{n-1}) \in \mathrm{QRM}(m-r-1,m)$ belongs to $\mathrm{QRM}(r,m)^\perp$ Clearly, $\mathbf{c} \in \mathrm{QRM}(m-r-1,m)$ if and only if $\mathbf{c}-c_\infty 1^{2^m} \in \mathrm{QRM}(m-r-1,m)$, and $\mathbf{c} \in \mathrm{QRM}(r,m)^\perp$ if and only if $\mathbf{c}-c_\infty 1^{2^m} \in \mathrm{QRM}(r,m)^\perp$. Therefore it is sufficient to show that for $\mathbf{c}$ with $c_\infty = 0$, $\mathbf{c} \in \mathrm{QRM}(m-r-1,m)$ implies $\mathbf{c} \in \mathrm{QRM}(r,m)^\perp$ Let $\mathbf{c} = (0, \mathbf{c}') \in \mathrm{QRM}(m-r-1,m)$, where $\mathbf{c}' = (c_0, c_1, \dots, c_{n-1})$. Then $\mathbf{c}' \in \mathrm{QRM}(m-r-1,m)^-$ By Corollary 10.6, $\mathrm{QRM}(m-r-1,m)^-$ is a $\mathbb{Z}_4$-cyclic code with generator polynomial

$$g_{m-r-1}(X) = \varepsilon_{m-r-1} \prod_{\substack{1 \le j \le 2^m-2 \\ w_2(j) < r+1}} (X - \xi^j).$$

So, $c(X) = c_0 + c_1 X + \cdots + c_{n-1}X^{n-1}$ is a multiple of $g_{m-r-1}(X)$. By Corollary 10.7, $c(1) = \sum_{i=0}^{n-1} c_i = 0$. Then $c(X)$ is also a multiple of $X-1$. Since $\overline{g}_{m-r-1}(1) \neq 0$, $g_{m-r-1}(1)$ is an invertible element of $\mathbb{Z}_4$. It follows that $c(X)$ is a multiple of $(X-1)g_{m-r-1}(X)$. Then $c(X)$ is annihilated by the polynomial

$$f_{m-r-1}(X) = \frac{X^n - 1}{(X-1)g_{m-r-1}(X)} = \varepsilon_{m-r-1} \prod_{\substack{1 \le j \le 2^m-2 \\ w_2(j) \ge r+1}} (X - \xi^j),$$

i.e., $c(X) f_{m-r-1}(X) = 0$. Therefore $c(X)$ belongs to the dual code of the $\mathbb{Z}_4$-cyclic code with generator polynomial

$$\begin{aligned} \tilde{f}_{m-r-1}(X) &= \varepsilon_{m-r-1} \prod_{\substack{1 \le j \le 2^m-2 \\ w_2(j) > r}} (1 - \xi^j X) \\ &= \prod_{\substack{1 \le j \le 2^m-2 \\ w_2(j) < m-r}} (X - \xi^j) = g_r(X). \end{aligned}$$

By Corollary 10.6, the $\mathbb{Z}_4$-cyclic code with generator polynomial $g_r(X)$ is $\mathrm{QRM}(r,m)^-$. Hence $c(X) \in (\mathrm{QRM}(r,m)^-)^\perp$. Since $c_\infty = 0$, $\mathbf{c} \in \mathrm{QRM}(r,m)^\perp$.

Therefore we have proved $\mathrm{QRM}(m-r-1,m) \subseteq \mathrm{QRM}(r,m)^\perp$. Since $\mathrm{QRM}(r,m)$ is of type $4^{K_{r,m}}$, where

$$K_{r,m} = 1 + \binom{m}{1} + \cdots + \binom{m}{r},$$

by Proposition 1.2, $\mathrm{QRM}(r,m)^{\perp}$ is of type $4^{2^m - K_{r,m}}$. But

$$\begin{aligned} 2^m - K_{r,m} &= \binom{m}{r+1} + \binom{m}{r+2} + \cdots + \binom{m}{m-1} + 1 \\ &= 1 + \binom{m}{1} + \cdots + \binom{m}{m-r-1}. \end{aligned}$$

Thus $4^{2^m - K_{r,m}}$ is also the type of $\mathrm{QRM}(m-r-1,m)$. Therefore $\mathrm{QRM}(m-r-1,m) = \mathrm{QRM}(r,m)^{\perp}$. □

Corollary 10.9. $\mathrm{QRM}(m-2,m) = \mathcal{P}(m)$.

Proof. We have

$$\begin{aligned} \mathrm{QRM}(m-2,m) &= \mathrm{QRM}(1,m)^{\perp} && \text{(Proposition 10.8)} \\ &= \mathcal{K}(m)^{\perp} && \text{(Proposition 10.1)} \\ &= \mathcal{P}(m). && \text{(Proposition 9.1)} \end{aligned}$$ □

The quaternary Reed–Muller codes were first studied by Hammons *et al.* (1994).

10.2. Quaternary Goethals Codes

As another generalization of quaternary Preparata codes, Hammons *et al.* (1994) introduce the quaternary Goethals codes as follows.

Definition 10.2. Let m be an odd integer ≥ 3 and ξ be an element of order $2^m - 1$ in the Galois ring $\mathrm{GR}(4^m)$. The *quaternary Goethals code* $\mathcal{G}(m)$ of length 2^m is defined to be the $\mathbb{Z}_4$-linear code with parity check matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \xi & \xi^2 & \cdots & \xi^{n-1} \\ 0 & 2 & 2\xi^3 & 2\xi^6 & \cdots & 2\xi^{3(n-1)} \end{pmatrix}, \tag{10.8}$$

where $n = 2^m - 1$ and each $\xi^j (j \geq 0)$ should be replaced by ${}^t(b_{1j}, b_{2j}, \ldots, b_{mj})$ if $\xi^j = b_{1j} + b_{2j}\xi + \cdots + b_{mj}\xi^{m-1}$. The columns of the matrix (10.8) are numbered by ∞, 0, 1, 2, …, $n-1$.

If we delete the ∞-components of the codewords of $\mathcal{G}(m)$, the code thus obtained is called the *shortened quaternary Goethals code* and denoted by $\mathcal{G}(m)^{-}$. □

Denote the binary image of $\mathcal{G}(m)$ by $\phi(\mathcal{G}(m))$, and call it the "Goethals" code. The "Goethals" code $\phi(\mathcal{G}(m))$ is a binary nonlinear code and has the same length, the same number of codewords, the same minimum distance, and the same weight (and distance) enumerator as the original Goethals code G_{m+1} introduced by Goethals (1974, 1976), when m is odd and ≥ 5. First let us study $\mathcal{G}(m)$. We have the following proposition which is due to Hammons *et al.* (1994).

Proposition 10.10. *The quaternary Goethals code* $\mathcal{G}(m)$ *of length* 2^m, m *odd* ≥ 3, *is of type* $4^{2^m-2m-1}2^m$ *and of minimal Lee distance* 8.

Proof. $\mathcal{G}(m)$ is $\mathbb{Z}_4$-linear. Its dual code has generator matrix (10.8) and, hence, has type $4^{m+1}2^m$. Therefore by Proposition 1.2 $\mathcal{G}(m)$ is of type $4^{2^m-2m-1}2^m$. The first two rows of (10.8) form a parity check matrix of the quaternary Preparata code $\mathcal{P}(m)$. Therefore $\mathcal{G}(m) \subseteq \mathcal{P}(m)$. Since the minimal Lee distance of $\mathcal{P}(m)$ is 6, the minimal Lee distance of $\mathcal{G}(m)$ is at least 6. By Proposition 3.4 the minimal Lee weight of $\mathcal{G}(m)$ is even. To show that $\mathcal{G}(m)$ has minimal Lee distance 8 we have to show that $\mathcal{G}(m)$ has no codewords of Lee weight 6 and that $\mathcal{G}(m)$ has a codeword of Lee weight 8.

First we prove that $\mathcal{G}(m)$ has a codeword of Lee weight 8. By reduction mod 2 from (10.8) we obtain

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \bar{\xi} & \bar{\xi}^2 & \cdots & \bar{\xi}^{n-1} \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

the first two rows of which is a parity check matrix of the extended binary Hamming code H_{2^m} of length 2^m. H_{2^m} is of minimal Hamming weight 4. Let $\mathbf{e}_i + \mathbf{e}_j + \mathbf{e}_k + \mathbf{e}_l$ be a codeword of H_{2^m}, where i, j, k, l are distinct. Then $2\mathbf{e}_i + 2\mathbf{e}_j + 2\mathbf{e}_k + 2\mathbf{e}_l$ is a codeword of Lee weight 8 of $\mathcal{G}(m)$.

Then we prove that $\mathcal{G}(m)$ has no codeword of Lee weight 6. We prove by contradiction. Let $\mathbf{c}$ be a codeword of Lee weight 6 of $\mathcal{G}(m)$. Since $\mathbf{c}$ is orthogonal to the first row of (10.8), it must be one of the following forms:

$$2\mathbf{e}_i + 2\mathbf{e}_j + \mathbf{e}_k - \mathbf{e}_l,$$

$$\mathbf{e}_i + \mathbf{e}_j + \mathbf{e}_k - \mathbf{e}_l - \mathbf{e}_g - \mathbf{e}_h,$$

$$\pm(\mathbf{e}_i + \mathbf{e}_j + \mathbf{e}_k + \mathbf{e}_l + \mathbf{e}_g - \mathbf{e}_h),$$

where i, j, k, l, g, h are distinct. We treat these cases one by one following Helleseth (1996).

(a) $\mathbf{c} = 2\mathbf{e}_i + 2\mathbf{e}_j + \mathbf{e}_k - \mathbf{e}_l$. Since $\mathbf{c}$ is orthogonal to every row of (10.8), it is also orthogonal to every row of

$$2\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \xi & \xi^2 & \cdots & \xi^{n-1} \\ 0 & 1 & \xi^3 & \xi^6 & \cdots & \xi^{3(n-1)} \end{pmatrix}. \tag{10.9}$$

Clearly, $2\mathbf{e}_i + 2\mathbf{e}_j$ is orthogonal to every row of (10.9). It follows that $\mathbf{e}_k - \mathbf{e}_l$ is also orthogonal to every row of (10.9). So, $\overline{\mathbf{e}_k - \mathbf{e}_l}$ is orthogonal to every row of

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \overline{\xi} & \overline{\xi}^2 & \cdots & \overline{\xi}^{n-1} \\ 0 & 1 & \overline{\xi}^3 & \overline{\xi}^6 & \cdots & \overline{\xi}^{3(n-1)} \end{pmatrix},$$

i.e., $\overline{\mathbf{e}_k - \mathbf{e}_l}$ is in the extended doubly-error-correcting BCH code of length 2^m. But $w(\overline{\mathbf{e}_k - \mathbf{e}_l}) = 2$, which is a contradiction.

(b) $\mathbf{c} = \mathbf{e}_i + \mathbf{e}_j + \mathbf{e}_k - \mathbf{e}_l - \mathbf{e}_g - \mathbf{e}_h$. Since $\mathbf{c}$ is orthogonal to the last two rows of (10.8), we have

$$\xi^i + \xi^j + \xi^k = \xi^l + \xi^g + \xi^h, \tag{10.10}$$

$$2\xi^{3i} + 2\xi^{3j} + 2\xi^{3k} = 2\xi^{3l} + 2\xi^{3g} + 2\xi^{3h}. \tag{10.11}$$

Let

$$\xi^i + \xi^j + \xi^k = \xi^l + \xi^g + \xi^h = a + 2b,$$

where $a, b \in \mathcal{T}$. By Corollary 6.10,

$$b = (\xi^i\xi^j)^{1/2} + (\xi^j\xi^k)^{1/2} + (\xi^k\xi^i)^{1/2} = (\xi^l\xi^g)^{1/2} + (\xi^g\xi^h)^{1/2} + (\xi^h\xi^l)^{1/2}.$$

Squaring, we obtain

$$\xi^i\xi^j + \xi^j\xi^k + \xi^k\xi^i \equiv \xi^l\xi^g + \xi^g\xi^h + \xi^h\xi^l \pmod 2. \tag{10.12}$$

From (10.11) we deduce

$$\xi^{3i} + \xi^{3j} + \xi^{3k} \equiv \xi^{3l} + \xi^{3g} + \xi^{3h} \pmod 2. \tag{10.13}$$

Applying the map $^-\colon \mathrm{GR}(4^m) \to \mathbb{F}_{2^m}$ to (10.10), (10.12) and (10.13), we obtain

$$\overline{\xi}^i + \overline{\xi}^j + \overline{\xi}^k = \overline{\xi}^l + \overline{\xi}^g + \overline{\xi}^h, \tag{10.14}$$

$$\overline{\xi}^i\overline{\xi}^j + \overline{\xi}^j\overline{\xi}^k + \overline{\xi}^k\overline{\xi}^i = \overline{\xi}^l\overline{\xi}^g + \overline{\xi}^g\overline{\xi}^h + \overline{\xi}^h\overline{\xi}^l, \tag{10.15}$$

$$(\bar{\xi}^i)^3 + (\bar{\xi}^j)^3 + (\bar{\xi}^k)^3 = (\bar{\xi}^l)^3 + (\bar{\xi}^g)^3 + (\bar{\xi}^h)^3. \tag{10.16}$$

From (10.14)–(10.16) we deduce

$$\begin{aligned}
\bar{\xi}^i\bar{\xi}^j\bar{\xi}^k &= (\bar{\xi}^i + \bar{\xi}^j + \bar{\xi}^k)^3 + ((\bar{\xi}^i)^3 + (\bar{\xi}^j)^3 + (\bar{\xi}^k)^3) \\
&\quad + (\bar{\xi}^i + \bar{\xi}^j + \bar{\xi}^k)(\bar{\xi}^i\bar{\xi}^j + \bar{\xi}^j\bar{\xi}^k + \bar{\xi}^k\bar{\xi}^i) \\
&= (\bar{\xi}^l + \bar{\xi}^g + \bar{\xi}^h)^3 + ((\bar{\xi}^l)^3 + (\bar{\xi}^g)^3 + (\bar{\xi}^h)^3) \\
&\quad + (\bar{\xi}^l + \bar{\xi}^g + \bar{\xi}^h)(\bar{\xi}^l\bar{\xi}^g + \bar{\xi}^g\bar{\xi}^h + \bar{\xi}^h\bar{\xi}^i) \\
&= \bar{\xi}^l\bar{\xi}^g\bar{\xi}^h
\end{aligned}$$

Therefore

$$f(X) = (X - \bar{\xi}^i)(X - \bar{\xi}^j)(X - \bar{\xi}^k) = (X - \bar{\xi}^l)(X - \bar{\xi}^g)(X - \bar{\xi}^h)$$

has six distinct roots in $\mathbb{F}_{2^m}$, which is a contradiction.

(c) $\mathbf{c} = \pm(\mathbf{e}_i + \mathbf{e}_j + \mathbf{e}_k + \mathbf{e}_l + \mathbf{e}_g - \mathbf{e}_h)$. Since $\mathbf{c}$ is orthogonal to the last two rows of (10.8), we have

$$\xi^i + \xi^j + \xi^k + \xi^l + \xi^g = \xi^h, \tag{10.17}$$

$$2\xi^{3i} + 2\xi^{3j} + 2\xi^{3k} + 2\xi^{3l} + 2\xi^{3g} = 2\xi^{3h} \tag{10.18}$$

By Corollary 6.10, from (10.17) we deduce

$$(\xi^i\xi^j)^{1/2} + (\xi^i\xi^k)^{1/2} + (\xi^i\xi^l)^{1/2} + (\xi^i\xi^g)^{1/2} + \cdots + (\xi^l\xi^g)^{1/2} = 0.$$

Squaring, we obtain

$$\xi^i\xi^j + \xi^i\xi^k + \xi^i\xi^l + \xi^i\xi^g + \cdots + \xi^l\xi^g \equiv 0 \pmod 2). \tag{10.19}$$

From (10.18) we deduce

$$\xi^{3i} + \xi^{3j} + \xi^{3k} + \xi^{3l} + \xi^{3g} \equiv \xi^{3h} \pmod 2). \tag{10.20}$$

Applying the map $^-$: $\mathrm{GR}(4^m) \to \mathbb{F}_{2^m}$ to (10.17), (10.19) and (10.20), we obtain

$$\bar{\xi}^i + \bar{\xi}^j + \bar{\xi}^k + \bar{\xi}^l + \bar{\xi}^g = \bar{\xi}^h, \tag{10.21}$$

$$\bar{\xi}^i\bar{\xi}^j + \bar{\xi}^i\bar{\xi}^k + \bar{\xi}^i\bar{\xi}^l + \bar{\xi}^i\bar{\xi}^g + \cdots + \bar{\xi}^l\bar{\xi}^g = 0, \tag{10.22}$$

$$(\bar{\xi}^i)^3 + (\bar{\xi}^j)^3 + (\bar{\xi}^k)^3 + (\bar{\xi}^l)^3 + (\bar{\xi}^g)^3 = (\bar{\xi}^h)^3 \tag{10.23}$$

From (10.21)-(10.23) we deduce

$$\bar{\xi}^i\bar{\xi}^j\bar{\xi}^k + \cdots + \bar{\xi}^k\bar{\xi}^l\bar{\xi}^g = 0.$$

Therefore

$$\begin{aligned} f(X) &= (X - \bar{\xi}^i)(X - \bar{\xi}^j)(X - \bar{\xi}^k)(X - \bar{\xi}^l)(X - \bar{\xi}^g) \\ &= X^5 + \sigma_1 X^4 + \sigma_4 X + \sigma_5, \end{aligned}$$

where

$$\begin{aligned} \sigma_1 &= \bar{\xi}^i + \bar{\xi}^j + \bar{\xi}^k + \bar{\xi}^l + \bar{\xi}^g = \bar{\xi}^h, \\ \sigma_4 &= \bar{\xi}^i\bar{\xi}^j\bar{\xi}^k\bar{\xi}^l + \cdots + \bar{\xi}^j\bar{\xi}^k\bar{\xi}^l\bar{\xi}^g, \\ \sigma_5 &= \bar{\xi}^i\bar{\xi}^j\bar{\xi}^k\bar{\xi}^l\bar{\xi}^g. \end{aligned}$$

Then $f(X)$ has five distinct roots in $\mathbb{F}_{2^m}$, which contradicts Corollary 9.7. □

A complete decoding algorithm for $\mathcal{G}(m)$, i.e., an algorithm that for any received word to find the closest codeword, can be found in Helleseth and Kumar (1995). This is an algebraic decoding algorithm that corrects all errors of Lee weight ≤ 3. We will not reproduce this algorithm here.

Corollary 10.11. *The shortened quaternary Goethals code $\mathcal{G}(m)^-$ has parity check matrix*

$$\begin{pmatrix} 1 & \xi & \xi^2 & \cdots & \xi^{n-1} \\ 2 & 2\xi^3 & 2\xi^6 & \cdots & 2\xi^{3(n-1)} \end{pmatrix} \tag{10.24}$$

and is a $\mathbb{Z}_4$-cyclic code of length $2^m - 1$, of type $4^{2^m-2m-1}2^m$, and of minimal Lee distance 7. □

Now let us study the "Goethals" code. We have

Proposition 10.12. *Let m be an odd integer ≥ 3. The "Goethals" code $\phi(\mathcal{G}(m))$ is a binary code of length 2^{m+1}. It is distance invariant, and has $2^{2^{m+1}-3m-2}$ codewords and minimal Hamming distance 8. If $m \geq 5$, it is nonlinear; but $\phi(\mathcal{G}(3))$ is linear.*

Proof. Clearly $\phi(\mathcal{G}(m))$ is a binary code of length 2^{m+1}. By Proposition 10.10, $\mathcal{G}(m)$ is of type $4^{2^m-2m-1}2^m$. Therefore $|\phi(\mathcal{G}(m))| = |\mathcal{G}(m)| = 2^{2^{m+1}-3m-2}$. By Theorem 3.6, $\phi(\mathcal{G}(m))$ is distance invariant. By Propositions 3.3 and 10.10 $\phi(\mathcal{G}(m))$ has minimal Hamming distance 8

Assume that $m \geq 5$. Let us prove that $\phi(\mathcal{G}(m))$ is nonlinear. Let $h(X)$ be the basic primitive polynomial of degree m with ξ as one of its roots. We know that ξ is of order $2^m - 1$. By hypothesis m is odd, so $3 \nmid 2^m - 1$ and ξ^3 is also of order $2^m - 1$. Let $h_3(X)$ be the basic primitive polynomial of degree m with ξ^3 as one of its roots. Then $(h(X), h_3(X)) = 1$ and $h(X)h_3(X)$ is of degree $2m$. Let

$$h(X)h_3(X) = a_0 + a_1 X + a_2 X^2 + \cdots + a_{2m} X^{2m},$$

then $a_0 = \pm 1$ and $a_{2m} = 1$. Since $m \geq 5$, we have $2^m \geq 4m + 2$. Parallel to the proof of the nonlinearity of $P(m)$ in Theorem 9.9, both

$$\mathbf{c} = \left(-\sum_{i=0}^{2m} a_i, a_0, a_1, \ldots, a_{2m-1}, a_{2m}, \underbrace{0, \ldots, 0}_{2m}, 0, \ldots, 0 \right)$$

and

$$\mathbf{c} = \left(-\sum_{i=0}^{2m} a_i, \underbrace{0, 0, \ldots, 0}_{2m}, a_0, a_1, \ldots, a_{2m}, 0, \ldots, 0 \right)$$

are codewords of $\mathcal{G}(m)$. Clearly,

$$2\alpha(\mathbf{c}) * \alpha(\mathbf{c}') = (2, \underbrace{0, \ldots, 0}_{2m}, 2, 0, \ldots, 0)$$

is of Lee weight 4 and, hence, is not a codeword of $\mathcal{G}(m)$. By Proposition 3.16, $\phi(\mathcal{G}(m))$ is nonlinear when $m \geq 5$.

Now consider the case $m = 3$. It is enough to show that the binary image $\phi(\mathcal{G}(3)^-)$ of $\mathcal{G}(3)^-$ is linear. For $m = 3$ we have $n = 2^3 - 1 = 7$. We can assume that ξ is a root of the basic primitive polynomial $h(X) = X^3 + 2X^2 + X + 3$ and that ξ^3 is a root of the basic primitive polynomial $h_3(X) = X^3 + 3X^2 + 2X + 3$. We have the complete factorization

$$X^7 - 1 = (X - 1)h(X)h_3(X).$$

By Theorem 7.23,

$$\mathcal{G}(3)^- = (h(X)h_3(X), 2(X - 1)h(X)).$$

We have

$$h(X)h_3(X) = 1 + X + X^2 + X^3 + X^4 + X^5 + X^6$$

and

$$2(X-1)h(X) = 2 + 2X^2 + 2X^3 + 2X^4$$

Let

$$\mathbf{1} = (1,1,1,1,1,1,1),$$
$$\mathbf{c}_1 = (2,0,2,2,2,0,0),$$
$$\mathbf{c}_2 = (0,2,0,2,2,2,0),$$
$$\mathbf{c}_3 = (0,0,2,0,2,2,2).$$

Then

$$\mathcal{G}(3)^- = \{\varepsilon\mathbf{1} + a_1\mathbf{c}_1 + a_2\mathbf{c}_2 + a_3\mathbf{c}_3 | \varepsilon \in \mathbb{Z}_4,\ a_1, a_2, a_3 \in \mathbb{Z}_2\}.$$

It is not difficult to verify that the condition in Corollary 3.17 is fulfilled. Therefore by Corollary 3.17, $\phi(\mathcal{G}(3)^-)$ is linear. □

The Goethals codes G_{m+1}, where m is any odd integer ≥ 5, and the formal dual $\varphi(\mathcal{G}(m)^\perp)$ of $\varphi(\mathcal{G}(m))$, were introduced by Goethals (1974, 1976). Both of them are distance invariant binary nonlinear codes of length 2^{m+1}. (A simple description of G_{m+1}, similar to the one of P_{m+1} given in Sec. 9.4, can also be found in Baker *et al.* (1983).) G_{m+1} contains $2^{2^{m+1}-3m-2}$ codewords and has minimum distance 8. Thus G_{m+1} and $\phi(\mathcal{G}(m))$ have the same length, the same number of codewords, and the same minimum distance. Goethals also computed the weight distributions of both G_{m+1} and $\varphi(\mathcal{G}(m)^\perp)$ and observed that the weight enumerator of G_{m+1} is the MacWilliams transform of that of $\varphi(\mathcal{G}(m)^\perp)$. By Theorem 3.7 the weight enumerator of $\varphi(\mathcal{G}(m))$ is the MacWilliam transform of that of $\varphi(\mathcal{G}(m)^\perp)$. Therefore G_{m+1} and $\varphi(\mathcal{G}(m))$ also have the same weight enumerator. Finally both G_{m+1} and $\varphi(\mathcal{G}(m))$ contain four times as many codewords as the extended triple-error-correcting BCH code of the same length.

Table 10.1. Weight distribution of $\varphi(\mathcal{G}(m)^\perp)$, $m = 2t+1$.

Weight	No. of codewords
0 or 2^{2t+2}	1
$2^{2t+1} \pm 2^{t+1}$	$2^{2t}(2^{2t+1}-1)(2^{2t+2}-1)/3$
$2^{2t+1} \pm 2^t$	$2^{2t+2}(2^{2t+1}-1)(2^{2t+1}+4)/3$
2^{2t+1}	$2(2^{2t+2}-1)(2^{4t+1}-2^{2t}+1)$

10.3. Quaternary Delsarte–Goethals and Goethals–Delsarte Codes

The quaternary Goethals codes and its $\mathbb{Z}_4$-duals can be further generalized as follows, (see Hammons *et al.* (1994)).

Definition 10.3. Let m be an odd integer ≥ 3, $m = 2t+1$, $1 \leq r \leq t$, and ξ be an element of order $2^m - 1$ in $\mathrm{GR}(4^m)$. The *quaternary Delsarte-Goethals code* $\mathcal{DG}(m, \delta)$, where $\delta = (m+1)/2 - r$, is the $\mathbb{Z}_4$-linear code with generator matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \xi & \xi^2 & \cdots & \xi^{n-1} \\ 0 & 2 & 2\xi^3 & 2\xi^6 & \cdots & 2\xi^{2(n-1)} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 2 & 2\xi^{1+2^j} & 2\xi^{(1+2^j)2} & \cdots & 2\xi^{(1+2^j)(n-1)} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 2 & 2\xi^{1+2^r} & 2\xi^{(1+2^r)2} & \cdots & 2\xi^{(1+2^r)(n-1)} \end{pmatrix} \tag{10.25}$$

The *quaternary Goethals-Delsarte code* $\mathcal{GD}(m, \delta)$ is the $\mathbb{Z}_4$-linear code with the matrix (10.25) as its parity check matrix. □

Clearly, when $r = 1$, $\mathcal{GD}(m, (m-1)/2)$ is the quaternary Goethals code $\mathcal{G}(m)$ studied in the previous section.

Denote the binary images of $\mathcal{DG}(m, \delta)$ and $\mathcal{GD}(m, \delta)$ by $\phi(\mathcal{DG}(m, \delta))$ and $\phi(\mathcal{GD}(m, \delta))$, respectively. Then we have

Proposition 10.13. *Let m be an odd integer ≥ 3, $m = 2t+1$, $1 \leq r \leq t$, and $\delta = (m+1)/2 - r$. Then the quaternary Delsarte-Goethals code $\mathcal{DG}(m, \delta)$ is of length 2^m and has type $4^{m+1}2^{rm}$ and minimum Lee weight $2^m - 2^{m-\delta}$ Its binary image $\phi(\mathcal{DG}(m, \delta))$ is the Delsarte-Goethals code $DG(m+1, \delta)$, which is a binary code of length 2^{m+1}, is distance invariant, and has $2^{2(m+1)+rm}$ codewords and minimum Hamming distance $2^m - 2^{m-\delta}$. When $m \geq 5$, $DG(m+1, \delta)$ is nonlinear.*

Proof. That $\mathcal{DG}(m, \delta)$ is of length 2^m and has type $4^{m+2}2^{rm}$ is clear from Definition 10.3. If we can show that its binary image is the binary Delsarte–Goethals code $DG(m+1, \delta)$, then its minimum Lee weight equals $2^m - 2^{m-\delta}$ follows from the minimal Hamming distance of $DG(m+1, \delta)$ equals $2^m - 2^{m-\delta}$

Comparing Eqs. (37) and (34) of Chap. 15 of MacWilliams and Sloane (1977), we see that the difference between the Kerdock code K_{m+1} and $DG(m+1,\delta)$ comes from the words $(\mathbf{c},\mathbf{c})$, where $\mathbf{c}$ belongs to the code defined by Eq. (31) of that chapter. We already know from Proposition 8.2 that the first two rows of (10.25) produce the Kerdock code, and it is easily seen that the remaining rows produce the required codewords $(\mathbf{c},\mathbf{c})$. Therefore $\phi(\mathcal{DG}(m,\delta)) = DG(m+1,\delta)$. It follows that $|DG(m+1,\delta)| = |\mathcal{DG}(m,\delta)| = 2^{2(m+1)+rm}$. The distance invariance of $DG(m+1,\delta)$ follows from Theorem 3.6.

The proof of the minimum Hamming distance of $DG(m+1,\delta)$ being equal to $2^m - 2^{m-\delta}$ and when $m \geq 5$, the proof of the nonlinearity of $DG(m+1,\delta)$ can be found in §5, Chap. 15 of MacWilliams and Sloane (1977). □

The Delsarte–Goethals codes were introduced and studied by Delsarte and Goethals (1975).

Moreover, we have

Proposition 10.14. *Let m be an odd integer ≥ 3, $m = 2t+1$, $1 \leq r \leq t$, and $\delta = (m+1)/2 - r$. Then the Goethals-Delsarte $\mathbb{Z}_4$-code $\mathcal{GD}(m,\delta)$ is of length 2^m and has type $4^{2^m-(r+1)m-1}2^{rm}$ and minimum Lee weight 8. Its binary image $\phi(\mathcal{GD}(m,\delta))$ is a binary code of length 2^{m+1}, it has $2^{2^{m+1}-(r+2)m-2}$ codewords, and is distance invariant. It has the same weight distribution as the binary Goethals-Delsarte code $GD(m+1,\delta)$. When $m \geq 5$, $\phi(\mathcal{GD}(m,\delta))$ is nonlinear.* □

The proof of this proposition is omitted.

The binary Goethals–Delsarte codes were introduced and studied by Hergert (1990). In particular, he proved that the weight enumerator of $GD(m+1,\delta)$ is the MacWillams transform of that of $DG(m+1,\delta)$.

10.4. Automorphism Groups

Let C be a $\mathbb{Z}_4$-codes of length n and the coordinate positions of the codewords of C be indexed by $1, 2, \ldots, n$. Let σ be a permutation of $1, 2, \ldots, n$. For any codeword $\mathbf{c} = (c_1, c_2, \ldots, c_n)$ define

$$\sigma(\mathbf{c}) = (c_{\sigma(1)}, c_{\sigma(2)}, \ldots, c_{\sigma(n)}).$$

If $\sigma(\mathbf{c}) \in C$ for all $\mathbf{c} \in C$, σ is called a permutation automorphism of C. Recall that the automorphism group $\mathrm{Aut}(C)$ of C is group generated by all permutation automorphisms and the sign-changes of certain coordinates that preserve the set of codewords of $\mathcal{C}$.

Let us now study the automorphism groups of the quaternary Kerdock, Preparata, Delsarte–Goethals and Goethals–Delsarte codes.

As before, let m be an integer ≥ 2, $n = 2^m - 1, \xi$ be a root of a basic primitive polynomial of degree m over $\mathbb{Z}_4$ and dividing $X^n - 1$, and $\mathcal{T} = \{\xi^\infty = 0, \xi^0 = 1, \xi^1, \ldots, \xi^{n-1}\}$. Let $\mathcal{C}$ be a $\mathbb{Z}_4$-code of length 2^m and the positions of coordinates be indexed by $\mathcal{T}$. Assume that $\mathcal{C}$ consists of all codewords $\mathbf{c} = (c_0, c_1, c_\xi, \ldots, c_{\xi^{n-1}})$ which satisfy the following system of linear equations over $\mathbb{Z}_4$

$$\sum_{x \in \mathcal{T}} c_x = 0, \tag{10.26}$$

$$\sum_{x \in \mathcal{T}} c_x x = 0, \tag{10.27}$$

and linear equations of the form

$$2 \sum_{x \in \mathcal{T}} c_x x^{2^j+1} = 0 \tag{10.28}$$

where j's are integers ≥ 1. Clearly, $\mathcal{C}$ is $\mathbb{Z}_4$-linear.

Lemma 10.15. *For any $a, b \in \mathcal{T}$ and $a \neq 0$, the map*

$$x \to \tau(ax + b) = (ax + b)^{2^m} \tag{10.29}$$

is a bijection on the set $\mathcal{T}$. The set of maps of the form (10.29) *forms a doubly transitive permutation group G on $\mathcal{T}$ and G is of order $2^m(2^m - 1)$.*

Proof. Since τ is a map from $\mathrm{GR}(4^m)$ to $\mathcal{T}$, (10.29) is a map from $\mathcal{T}$ to $\mathcal{T}$. To prove bijective it is enough to show that it is injective. Assume that for $x, x_1 \in \mathcal{T}$, $\tau(ax + b) = \tau(ax_1 + b)$. We have

$$\begin{aligned}
\tau(ax + b) &= (ax + b)^{2^m} \\
&= a^{2^m} x^{2^m} + b^{2^m} + 2(ax)^{2^{m-1}} b^{2^{m-1}} \\
&= ax + b + 2(ax)^{2^{m-1}} b^{2^{m-1}}
\end{aligned}$$

and similarly

$$\tau(ax_1 + b) = ax_1 + b + 2(ax_1)^{2^{m-1}} b^{2^{m-1}}.$$

It follows that $ax + b \equiv ax_1 + b \pmod 2$. Therefore $x \equiv x_1 \pmod 2$ and, hence $x = x_1$. The injectivity of (10.29) is proved.

Denote the set of maps of the form (10.29) by G. Let

$$x \to \tau(a_1 x + b_1) \tag{10.30}$$

be another map of the form (10.29) where $a_1, b_1 \in \mathcal{T}$ and $a_1 \neq 0$. The composite of (10.29) and (10.30) is

$$\begin{aligned} x \to \tau(a_1\tau(ax+b)+b_1) &= \tau(a_1(ax+b)^{2^m}+b_1) \\ &= \tau(a_1(ax+b+2(axb)^{2^{m-1}})+b_1) \\ &= \tau(a_1ax+a_1b+b_1) \\ &= \tau(a_1ax+a_2), \end{aligned}$$

which is also of the form (10.29), where $a_2 + 2b_2$ is the 2-adic representation of $a_1b + b_1$. Therefore G is closed under the composition of maps. It is easy to verify that the map

$$x \to \tau(a^{-1}x - a^{-1}b)$$

is the inverse of (10.29). Hence G is a group.

Finally, let us prove the double transitivity of G. Let x_1 and x_2 be two distinct elements of $\mathcal{T}$. We are going to prove that there is an element of G which carries x_1 and x_2 into 0 and 1, respectively. If $x_1 = 0$, then $x_2 \neq 0$ and the map

$$x \to \tau(x_2^{-1}x)$$

leaves $x_1 = 0$ fixed and carries x_2 to 1. If $x_1 \neq 0$, the map

$$x \to \tau(x - x_1)$$

carries x_1 to 0, which is reduced to the previous case. □

For any $\mathbf{x} = (x_0, x_1, x_\xi, \ldots, x_{\xi^{n-1}}) \in \mathbb{Z}_4^{n+1}$ and $\sigma \in G$, define

$$\sigma(\mathbf{x}) = (x_{\sigma(0)}, x_{\sigma(1)}, \ldots, x_{(\sigma(\xi^{n-1}))}).$$

Then we have

Lemma 10.16. *For any* $\mathbf{c} \in C$ *and* $\sigma \in G$, $\sigma(\mathbf{c}) \in C$.

Proof. Let $\mathbf{c} = (c_0, c_1, c_\xi, \ldots, c_{\xi^{n-1}})$. Then $\mathbf{c}$ satisfies (10.26)–(10.28). Clearly, $\sigma(\mathbf{c})$ satisfies (10.26). Repeated applications of the generalized Frobenius map f to (10.27) gives

$$\sum_{x\in\mathcal{T}} c_x x^{2^k} = 0, \quad k = 0, 1, 2, \ldots. \tag{10.31}$$

Assume that $\sigma^{-1}(x) = \tau(ax+b)$ for all $x \in \mathcal{T}$, where $a, b \in \mathcal{T}$ and $a \neq 0$. From (10.26), (10.27) and (10.31) we deduce that

$$\begin{aligned}
\sum_{x\in\mathcal{T}} c_{\sigma(x)} x &= \sum_{x\in\mathcal{T}} c_x \sigma^{-1}(x) \\
&= \sum_{x\in\mathcal{T}} c_x (ax+b)^{2^m} \\
&= \sum_{x\in\mathcal{T}} c_x (ax+b+2(axb)^{2^{m-1}}) \\
&= a\sum_{x\in\mathcal{T}} c_x x + b\sum_{x\in\mathcal{T}} c_x + 2(ab)^{2^{m-1}} \sum_{x\in\mathcal{T}} c_x x^{2^{m-1}} \\
&= 0.
\end{aligned}$$

Finally,

$$\begin{aligned}
2\sum_{x\in\mathcal{T}} c_{\sigma(x)} x^{2^j+1} &= 2\sum_{x\in\mathcal{T}} c_x (\sigma^{-1}(x))^{2^j+1} \\
&= 2\sum_{x\in\mathcal{T}} c_x ((ax+b)^{2^m})^{2^j+1} \\
&= 2\sum_{x\in\mathcal{T}} c_x (ax+b+2(axb)^{2^{m-1}})^{2^j+1} \\
&= 2\sum_{x\in\mathcal{T}} c_x (ax+b)^{2^j+1} \\
&= 2\sum_{x\in\mathcal{T}} c_x (a^{2^j+1}x^{2^j+1} + a^{2^j}x^{2^j}b + axb^{2^j} + b^{2^j+1}) \\
&= 0.
\end{aligned}$$

Therefore $\sigma(\mathbf{c}) \in \mathcal{C}$. □

Proposition 10.17. *The automorphism group* Aut $\mathcal{P}(m)$ *of the quaternary Preparata code* $\mathcal{P}(m)$*, where* $m \geq 2$*, contains a subgroup generated by* G*, the negation, and the generalized Frobenius map* f *acting on* $\mathcal{T}$*, which is a doubly transitive group of order* $2^{m+1}(2^m-1)m$*. So is the automorphism group* Aut $\mathcal{K}(m)$ *of the quaternary Kerdock code* $\mathcal{K}(m)$*, where* $m \geq 2$.

Proof. It follows from Lemma 10.16 that Aut $\mathcal{P}(m)$ contains G. Since $\mathcal{P}(m)$ is $\mathbb{Z}_4$-linear, the negation belongs to Aut $\mathcal{P}(m)$. Repeated applications of the generalized Frobenius map f to (10.28) gives

$$\sum_{x\in\mathcal{T}} c_x x^{2^k(2^j+1)} = 0, \quad k = 0, 1, 2, \dots. \tag{10.32}$$

From (10.26), (10.31) and (10.32) we deduce that Aut $\mathcal{P}(m)$ contains the generalized Frobenius map f acting on $\mathcal{T}$. The first assertion is proved.

Since $\mathcal{K}(m)$ is the dual code of $\mathcal{P}(m)$, Aut $\mathcal{K}(m) =$ Aut $\mathcal{P}(m)$. □

For a binary code C, a permutation of coordinate positions of the codewords leaving the code invariant is called an *automorphism* of C. The set of automorphisms of C forms a group, called the *group of automorphisms* of C and denoted by Aut C. Clearly, we have

Lemma 10.18. *Let $\mathcal{C}$ be a $\mathbb{Z}_4$-code and $C = \phi(\mathcal{C})$ be its binary image. Then an automorphism of $\mathcal{C}$ induces an automorphism of C and different automorphisms of $\mathcal{C}$ induce different automorphisms of C. Therefore* Aut $\mathcal{C}$ *can be regarded as a subgroup of* Aut C. □

For odd m, the automorphism groups of the binary Kerdock codes K_{m+1} are determined by Carlet (1991) and that of the binary Preparata codes P_{m+1} by Kantor (1982, 1983). For $m \geq 5$ both groups are of order $2^{m+1}(2^m - 1)m$. Therefore we have

Theorem 10.19. *For m odd and ≥ 5 the subgroup mentioned in Proposition* 10.17 *is the full automorphism group of the quaternary Kerdock code $\mathcal{K}(m)$ and Preparata code $\mathcal{P}(m)$.*

Proof. For the Kerdock code $\mathcal{K}(m)$ it follows directly from Proposition 10.17, Lemma 10.18, and the foregoing result of Carlet (1991). For the Preparata code $\mathcal{P}(m)$ we use the fact that it has the same automorphism group as its dual. □

The case $m = 3$ is exceptional. The quaternary Kerdock code $\mathcal{K}(3)$ coincides with the quaternary Preparata code $\mathcal{P}(3)$ and also coincides with the octacode. It is known that the octacode has an automorphism group of order 1344, see Conway and Sloane (1993a), but its binary image, the

Nordstrom–Robinson code, has an automorphism group of order 80640, see Berlekamp (1971) and also Conway and Sloane (1990).

Similarly, we have

Proposition 10.20. *The automorphism group* Aut $\mathcal{DG}(m,\delta)$ *and the automorphism group* Aut $\mathcal{GD}(m,\delta)$, *where* m *is an odd integer* ≥ 3, $\delta = (m+1)/2 - r$, *and* $1 \leq r \leq (m-1)/2$, *coincide and each one of them contains a subgroup generated by the group* G *defined in Lemma* 10.15, *the negation, and the generalized Frobenius map* f *acting on* $\mathcal{T}$, *which is a doubly transitive group of order* $2^{m+1}(2^m-1)m$. □

The automorphism group of $DG(m+1,\delta)$, where $m \geq 5$, is determined by Carlet (1993), which is of order $2^{m+1}(2^m-1)m$. Hence, by Lemma 10.18 and Proposition 10.20 we have

Proposition 10.21. *For* m *odd and* ≥ 5, *the subgroup mentioned in Proposition* 10.20 *is the full automorphism group of the quaternary Delsarte–Goethals code* $\mathcal{DG}(m,\delta)$ *and of the quaternary Goethal–Delsarte code* $\mathcal{GD}(m,\delta)$. □

CHAPTER 11

QUATERNARY QUADRATIC RESIDUE CODES

11.1. A Review of Binary Quadratic Residue Codes

Throughout this chapter we assume that p is an odd prime and $p \equiv \pm 1 \pmod 8$. Then 2 is a quadratic residue mod p and $2^{(p-1)/2} \equiv 1 \pmod p$, (see Serre (1973)). Let m be the least positive integer such that $p \mid 2^m - 1$. Then there is a primitive pth root of unity ω in $\mathbb{F}_{2^m}$ and $X^p - 1$ has the complete factorization

$$X^p - 1 = \prod_{i=0}^{p-1} (X - \omega^i) \tag{11.1}$$

in $\mathbb{F}_{2^m}[X]$.

Denote by $\mathbb{F}_p^{*2}$ the set of square elements of $\mathbb{F}_p^*$, i.e.,

$$\mathbb{F}_p^{*2} = \{a^2 | a \in \mathbb{F}_p^*\}.$$

It is known that $|\mathbb{F}_p^{*2}| = |\mathbb{F}_p^* \backslash \mathbb{F}_p^{*2}| = (p-1)/2$. For simplicity write $Q = \mathbb{F}_p^{*2}$ and $N = \mathbb{F}_p^* \backslash \mathbb{F}_p^{*2}$. Let

$$g_2(X) = \prod_{r \in Q} (X - \omega^r)$$

and

$$h_2(X) = \prod_{s \in N} (X - \omega^s)$$

so that by (11.1), we have

$$X^p - 1 = (X - 1)\, g_2(X)\, h_2(X). \tag{11.2}$$

For any $f(X) = a_0 + a_1X + \cdots + a_nX^n \in \mathbb{F}_{2^m}[X]$ define $f(X)^2 = a_0^2 + a_1^2X + \cdots + a_n^2X^n$. Since $2 \in Q$, we have

$$g_2(X)^2 = \prod_{r \in Q} (X - \omega^{2r}) = \prod_{r' \in Q} (X - \omega^{r'}) = g_2(X),$$

$$h_2(X)^2 = \prod_{s \in N} (X - \omega^{2s}) = \prod_{s' \in N} (X - \omega^{s'}) = h_2(X).$$

Therefore (11.2) is a factorization in $\mathbb{F}_2[X]$.

Consider the binary cyclic codes of length p:

$$\begin{aligned} Q_2(p) &= (g_2(X)), \\ Q_2'(p) &= ((X-1)\,g_2(X)), \\ N_2(p) &= (h_2(X)), \\ N_2'(p) &= ((X-1)\,h_2(X)). \end{aligned}$$

$Q_2(p)$ and $N_2(p)$ are binary $[p, \frac{p+1}{2}]$-codes and called the *binary augmented quadratic residue codes.* $Q_2'(p)$ and $N_2'(p)$ are binary $[p, \frac{p-1}{2}]$-codes and called the *binary expurgated quadratic residue codes.* Clearly, $Q_2'(p)$ is the even weight subcode of $Q_2(p)$, i.e., the subcode consisting of the even weight codewords of $Q_2(p)$. Similarly, $N_2'(p)$ is the even weight subcode of $N_2(p)$. If G is a generator matrix of $Q_2'(p)$ (or $N_2'(p)$), then

$$\begin{pmatrix} 1\ 1\ \cdots\ 1 \\ G \end{pmatrix}$$

is a generator matrix of $Q_2(p)$ (or $N_2(p)$, respectively). In the following, for simplicity, the codes $Q_2(p)$, $Q_2'(p)$, $N_2(p)$ and $N_2'(p)$ will be denoted by Q_2, Q_2', N_2 and N_2', respectively, if the code length p is clear from the context.

Let j be an integer and $0 < j < p$. Then $(j, p) = 1$. Denote by j^{-1} the integer such that $0 < j^{-1} < p$ and $j \cdot j^{-1} \equiv 1 \pmod{p}$. Define a permutation π_j on the places of coordinates as follows:

$$\pi_j : i \mapsto ji \qquad \text{for} \qquad i = 0, 1, \ldots, p-1.$$

π_j induces a transformation on $\mathbb{F}_2[X]/(X^p - 1)$ in the following way:

$$\pi_j : \mathbb{F}_2[X]/(X^p - 1) \to \mathbb{F}_2[X]/(X^p - 1)$$

$$f(X) = a_0 + \sum_{i=1}^{p-1} a_i X^j \mapsto \pi_j(f(X)) = a_0 + \sum_{i=1}^{p-1} a_{ji} X^i.$$

Clearly,

$$\pi_j(f(X)) = a_0 + \sum_{i=1}^{p-1} a_i X^{j^{-1}i} = f(X^{j^{-1}}).$$

In particular,

$$\pi_j(g_2(X)) = g_2(X^{j^{-1}}) \qquad \text{and} \qquad \pi_j(h_2(X)) = h_2(X^{j^{-1}}).$$

Let α be an element in some extension field of $\mathbb{F}_2$. Then α is a root of $g_2(X)$ if and only if α^j is a root of $g_2(X^{j^{-1}})$. Therefore deg $g_2(X^{j^{-1}}) =$ deg $g_2(X)$. Similarly, deg $h_2(X^{j^{-1}}) =$ deg $h_2(X)$.

Moreover, we have

Proposition 11.1. *Let j be an integer and $0 < j < p$. If $j \in Q$, then π_j leaves every one of Q_2, Q_2', N_2 and N_2' invariant. If $j \in N$, then π_j carries Q_2, Q_2', N_2 and N_2' into N_2, N_2', Q_2 and Q_2', respectively.* □

For proofs of this and the following propositions, corollaries, and lemmas see MacWilliams and Sloane (1977), Chap. 16.

It is known that when p is an odd prime, -1 is a quadratic residue mod p if and only if $p \equiv 1 \pmod 4$, (see Serre (1973)). Therefore when $p \equiv -1 \pmod 8$, π_{-1} interchanges Q_2 and N_2, and also Q_2' and N_2', and when $p \equiv 1 \pmod 8$, π_{-1} leaves each of Q_2, Q_2', N_2 and N_2' invariant.

Proposition 11.2. *The polynomials*

$$\theta_1(X) = \sum_{i=0}^{p-1} X^i, \qquad \theta_Q(X) = \sum_{r\in Q} X^r \qquad \textit{and} \qquad \theta_N(X) = \sum_{s\in N} X^s$$

are idempotents in $\mathbb{Z}_2[X]/(X^n - 1)$ and $\theta_1(X) + \theta_Q(X) + \theta_N(X) = 1$. We can choose a primitive pth root of unity α so that $\theta_Q(\alpha) = 0$. Then when $p \equiv -1 \pmod 8$, the generating idempotents of Q_2, Q_2', N_2 and N_2' are $\theta_Q(X)$, $1 + \theta_N(X)$, $\theta_N(X)$ and $1 + \theta_Q(X)$, respectively, and when $p \equiv 1 \pmod 8$, $\theta_1(X)$, $\theta_Q(X)$ and $\theta_N(X)$ are mutually orthogonal, and the generating idempotents of Q_2, Q_2', N_2 and N_2' are $1 + \theta_N(X)$, $\theta_Q(X)$, $1 + \theta_Q(X)$ and $\theta_N(X)$, respectively. □

Lemma 11.3. *When $p \equiv -1$ (mod 8), we have the following identities over $\mathbb{Z}$:*

$$\theta_Q(X)^2 = \frac{1}{4}(p-3)\,\theta_Q(X) + \frac{1}{4}(p+1)\,\theta_N(X)\,,$$

$$\theta_N(X)^2 = \frac{1}{4}(p+1)\,\theta_Q(X) + \frac{1}{4}(p-3)\,\theta_N(X)\,,$$

$$\theta_Q(X)\,\theta_N(X) = \frac{1}{4}(p+1) + \frac{1}{4}(p-3)\,\theta_1(X)\,.$$

When $p \equiv 1 \pmod 8$, we have the following identities over $\mathbb{Z}$:

$$\theta_Q(X)^2 = \frac{1}{2}(p-1) + \frac{1}{4}(p-5)\,\theta_Q(X) + \frac{1}{4}(p-1)\,\theta_N(X)\,,$$

$$\theta_N(X)^2 = \frac{1}{4}(p-1)\,\theta_Q(X) + \frac{1}{4}(p-5)\,\theta_N(X)\,,$$

$$\theta_Q(X)\,\theta_N(X) = 0\,. \qquad \square$$

Denote by $Q_2^\perp$ and $N_2^\perp$ the dual codes of Q_2 and N_2, respectively. From Propositions 11.2 and 7.7, we deduce immediately:

Proposition 11.4. *If $p \equiv -1 \pmod 8$, then $Q_2^\perp = Q_2'$ and $N_2^\perp = N_2'$. If $p \equiv 1 \pmod 8$, then $Q_2^\perp = N_2'$ and $N_2^\perp = Q_2'$.* $\square$

Denote by $\bar{Q}_2$ and $\bar{N}_2$ the *extended binary quadratic residue codes.* They are binary codes obtained by adjoining the zero-sum check symbol $c_\infty = \sum_{i=0}^{p-1} c_i$ to every codeword $(c_0, c_1, \ldots, c_{p-1})$ of Q_2 and N_2, respectively, at the position ∞. Thus they are linear code of length $p+1$ and have the same dimension $(p+1)/2$ of Q_2 and N_2. If G is a generator matrix of Q_2' (or N_2'), then

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & & & & \\ \vdots & & G & & \\ 0 & & & & \end{pmatrix}$$

is a generator matrix of $\bar{Q}_2$ (or $\bar{N}_2$, respectively).

Proposition 11.5. *For $p \equiv -1 \pmod 8$, both $\bar{Q}_2$ and $\bar{N}_2$ are self-dual $[p+1, (p+1)/2]$-codes. For $p \equiv 1 \pmod 8$,* $\bar{Q}_2^\perp = \bar{N}_2$ and $\bar{N}_2^\perp = \bar{Q}_2$. $\square$

By Proposition 11.1, Q_2 and N_2 are equivalent. Therefore it is sufficient to consider Q_2.

Let C be the $p \times p$ circulant matrix with the coefficients of X^0, X^1, ..., X^{p-1} of $\theta_Q(X)$ as its first row. Define

$$\mathbf{c} = \begin{cases} 0^p & \text{if } p \equiv 1 \pmod 8, \\ 1^p & \text{if } p \equiv -1 \pmod 8, \end{cases}$$

where 0^p and 1^p are the all 0 p-tuple and all 1 p-tuple, respectively, and

$$G = \begin{pmatrix} 1 & 1^p \\ {}^t\mathbf{c} & C \end{pmatrix} .$$

Then the rows of G generate $\bar{Q}_2$ though they are not linearly-independent. Both the columns and the rows are numbered by ∞, 0, 1, ..., $p-1$.

We may regard the coordinate places ∞, 0, 1, ..., $p-1$ as the nonhomogeneous coordinates of the $p+1$ points of the projective line $\mathrm{PG}(1, \mathbb{F}_p)$. It is known that an element

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{PSL}_2(\mathbb{F}_2)$$

induces a permutation on the points of $\mathrm{PG}(1, \mathbb{F}_p)$ in the following way

$$z \mapsto \frac{az+b}{cz+d} \qquad \text{for all} \qquad z \in \mathrm{PG}(1, \mathbb{F}_p),$$

where we agree that if $c = 0$, the point ∞ is left fixed and if $c \neq 0$, the point ∞ goes to a/c and the point $-d/c$ goes to ∞. It is also known that $\mathrm{PSL}_2(\mathbb{F}_p)$ acts triply transitively on $\mathrm{PSL}_2(\mathbb{F}_p)$.

Proposition 11.6. $\mathrm{PSL}_2(\mathbb{F}_p) \subseteq \mathrm{Aut}\,\bar{Q}_2$. □

Proposition 11.7. *The minimum weight of the quadratic residue code Q_2 of length a prime number p, which is* $\equiv \pm 1 \pmod 8$, *and with generator polynomial $g_2(X)$ is an odd number d for which*

(i) $d^2 > p$ *if* $p \equiv 1 \pmod 8$,
(ii) $d^2 - d + 1 \geq p$ *if* $p \equiv -1 \pmod 8$. □

Example 11.1. Let $p = 7$. Clearly $7 \equiv -1 \pmod 8$ and 3 is the least positive integer m such that $p \mid 2^m - 1$. Let ω be a primitive element of $\mathbb{F}_{2^3}$, and assume that ω is a root of the primitive polynomial $X^3 + X + 1$. We have $R = \{1, 2, 4\}$ and $N = \{3, 5, 6\}$. Then

$$g_2(X) = (X - \omega)(X - \omega^2)(X - \omega^4) = X^3 + X + 1\,,$$
$$h_2(X) = (X - \omega^3)(X - \omega^5)(X - \omega^6) = X^3 + X^2 + 1\,,$$

and we have the factorization of $X^7 - 1$ in $\mathbb{F}_2[X]$

$$X^7 - 1 = (X - 1)\, g_2(X)\, h_2(X)\,.$$

The binary quadratic residue code $Q_2(7)$ is the cyclic code generated by $g_2(X) = X^3 + X + 1$. Hence $Q_2(7)$ is a [7,4]-code and $|Q_2(7)| = 2^4 = 16$. By Proposition 11.7 (ii) the minimum distance d of $Q_2(7)$ satisfies $d^2 - d + 1 \geq 7$, from which we deduce $d \geq 3$. By the sphere-packing bound

$$|Q_2(7)| \left(\binom{7}{0} + \binom{7}{1} + \cdots + \binom{7}{\left[\frac{d-1}{2}\right]} \right) \leq 2^7\,,$$

where $[\frac{d-1}{2}]$ is the integral part of $\frac{d-1}{2}$. But

$$|Q_2(7)| \left(\binom{7}{0} + \binom{7}{1} \right) = 2^4(1 + 7) = 2^7\,.$$

Therefore $d = 3$ and the code $Q_2(7)$ is perfect. $Q_2(7)$ is known as the *binary Hamming code of length* 7, which is denoted by H_7.

By adding a zero-sum check symbol to every codeword of the code H_7, we obtain the *extended binary Hamming code of length* $8 = 2^3$, which is denoted by H_8. It is easy to verify that H_8 is doubly even and self-dual. □

Example 11.2. Let $p = 23$. Clearly $23 \equiv -1 \pmod 8$ and 11 is the least positive integer m such that $p \,|\, 2^m - 1$. Let ω be a primitive 23rd root of unity in $\mathbb{F}_{2^{11}}$, and assume that ω is a root of the irreducible polynomial $X^{11} + X^9 + X^7 + X^6 + X^5 + X + 1$ of period 23. We have

$$R = \{1, 2, 3, 4, 6, 8, 9, 12, 13, 16, 18\}$$

and

$$N = \{5, 7, 10, 11, 14, 15, 17, 19, 20, 21, 22\}\,.$$

Then

$$g_2(X) = \prod_{r \in R} (X - \omega^r) = X^{11} + X^9 + X^7 + X^6 + X^5 + X + 1\,,$$

$$h_2(X) = \prod_{s \in N} (X - \omega^s) = X^{11} + X^{10} + X^6 + X^5 + X^4 + X^2 + 1$$

and we have the factorization of $X^{23} - 1$ over $\mathbb{F}_2$

$$X^{23} - 1 = (X - 1)\, g_2(X)\, h_2(X)\,.$$

The binary quadratic residue code $Q_2(23)$ is the cyclic code generated by $g_2(X)$. Hence $Q_2(23)$ is a $[23, 12]$-code and $|Q_2(23)| = 2^{12}$. By Proposition 11.7(ii) the minimum distance d satisfies $d^2 - d + 1 \geq 23$. Since d is odd we have $d \geq 7$. But

$$|Q_2(23)| \left(\binom{23}{0} + \binom{23}{1} + \binom{23}{2} + \binom{23}{3} \right) = 2^{23}$$

It follows that $d = 7$ and $Q_2(23)$ is a perfect code. The code $Q_2(23)$ is known as the *binary Golay code*, which is usually denoted by G_{23}.

By adding a zero-sum check symbol to every codeword of the code G_{23}, we obtain the *extended binary Golay code* G_{24}, which is a doubly even self-dual $[24, 12, 8]$-code. □

Example 11.3. Let $p = 17$. Then $p \equiv 1 \pmod 8$ and 8 is the least positive integer m such that $p \mid 2^m - 1$. Let ω be a primitive 17th root of unity in $\mathbb{F}_{2^8}$ and assume that ω is a root of the irreducible polynomial $X^8 + X^5 + X^4 + X^3 + 1$ of period 17. We have

$$R = \{1,\ 2,\ 4,\ 8,\ 9,\ 13,\ 15,\ 16\}$$

and

$$N = \{3,\ 5,\ 6,\ 7,\ 10,\ 11,\ 12,\ 14\}\,.$$

Then

$$g_2(X) = \prod_{r \in R} (X - \omega^r) = X^8 + X^5 + X^4 + X^3 + 1\,,$$

$$h_2(X) = \prod_{s \in N} (X - \omega^s) = X^8 + X^7 + X^6 + X^4 + X^2 + X + 1$$

and

$$X^{17} - 1 = (X - 1)\, g_2(X)\, h_2(X)\,,$$

which is a factorization over $\mathbb{F}_2$.

The binary augmented quadratic residue codes $Q_2(17)$ and $N_2(17)$ are cyclic binary $[17, 9]$-codes generated by $g_2(X)$ and $h_2(X)$, respectively, and the binary expurgated quadratic residue codes $Q_2'(17)$ and $N_2'(17)$ are cyclic binary $[17,8]$-codes generated by $(X-1)\,g_2(X)$ and $(X-1)\,h_2(X)$, respectively. By Proposition 11.5, the extended binary quadratic residue codes $\bar{Q}_2(17)$ and $\bar{N}_2(17)$ are dual to each other. $\square$

11.2. Quaternary Quadratic Residue Codes

We follow the notation of the preceding section. We have the factorization of X^p-1 in $\mathbb{F}_2[X]$ (11.2)

$$X^p - 1 = (X-1)\,g_2(X)\,h_2(X)\,,$$

where

$$g_2(X) = \prod_{r\in R} (X-\omega^r) \qquad \text{and} \qquad h_2(X) = \prod_{s\in N} (X-\omega^s)\,.$$

By Hensel's lemma, there are monic polynomials $X-a$, $g(X), h(X) \in \mathbb{Z}_4[X]$ such that they are pairwise coprime, $X-\bar{a} = X-1$, $\bar{g}(X) = g_2(X)$, $\bar{h}(X) = h_2(X)$, and

$$X^p - 1 = (X-a)\,g(X)\,h(X) \qquad \text{in} \qquad \mathbb{Z}_4[X].$$

Substituting $X=1$ into the above equation, we obtain $(1-a)\,g(1)\,h(1) = 0$. Since $\bar{g}(1) = g_2(1) \neq 0$ and $\bar{h}(1) = h_2(1) \neq 0$, $g(1)$ and $h(1)$ are both invertible elements of $\mathbb{Z}_4$. Therefore $a=1$ and

$$X^p - 1 = (X-1)\,g(X)\,h(X) \tag{11.3}$$

in $\mathbb{Z}_4[X]$. Moreover, $g(X)$ and $h(X)$ are the Hensel lifts of $g_2(X)$ and $h_2(X)$, respectively, and hence, they are uniquely determined by $g_2(X)$ and $h_2(X)$.

Definition 11.1. The *quaternary quadratic residue codes* $Q_4(p)$, $Q_4'(p)$, $N_4(p)$, $N_4'(p)$ are defined to be the $\mathbb{Z}_4$-cyclic codes of length p generated by $g(X)$, $(X-1)\,g(X)$, $h(X)$, $(X-1)\,h(X)$, respectively. $\square$

Clearly, both $Q_4(p)$ and $N_4(p)$ are of type $4^{(p+1)/2}$, and both $Q_4'(p)$ and $N_4'(p)$ are of type $4^{(p-1)/2}$. In the following, for simplicity, we denote $Q_4(p)$, $Q_4'(p)$, $N_4(p)$ and $N_4'(p)$ by Q_4, Q_4', N_4 and N_4', respectively, if the code length p is clear from the context.

Let us compute the generating idempotents of the quaternary quadratic residue codes.

Proposition 11.8. *Let* $p \equiv \pm 1 \pmod 8$ *and write* $p \pm 1 = 8r$, *where* r *is a positive integer. The generating idempotents of* Q_4, Q_4', N_4 *and* N_4' *are given by the following table:*

Table 11.1. Generating idempotents of quaternary quadratic residue codes.

	$p+1=8r$		$p-1=8r$	
	r odd	r even	r odd	r even
Q_4	$\theta_Q(X)+2\theta_N(X)$	$3\theta_Q(X)$	$1+2\theta_Q(X)+3\theta_N(X)$	$1+\theta_N(X)$
Q_4'	$1+2\theta_Q(X)+3\theta_N(X)$	$1+\theta_N(X)$	$\theta_Q(X)+2\theta_N(X)$	$3\theta_Q(X)$
N_4	$2\theta_Q(X)+\theta_N(X)$	$3\theta_N(X)$	$1+3\theta_Q(X)+2\theta_N(X)$	$1+\theta_Q(X)$
N_4'	$1+3\theta_Q(X)+2\theta_N(X)$	$1+\theta_Q(X)$	$2\theta_Q(X)+\theta_N(X)$	$3\theta_N(X)$

Proof. It follows from Proposition 7.27 and Lemma 11.3. □

Parallel to Proposition 11.1, we have

Proposition 11.9. *Let* j *be an integer and* $0 < j < p$. *If* $j \in Q$, *then* π_j *leaves every one of* Q_4, Q_4', N_4 *and* N_4' *invariant. If* $j \in N$, *then* π_j *carries* Q_4, Q_4', N_4 *and* N_4' *into* N_4, N_4', Q_4 *and* Q_4', *respectively.*

Proof. By Proposition 11.8. □

Parallel to Proposition 11.4 we have

Proposition 11.10. *If* $p \equiv -1 \pmod 8$, *then* $Q_4^\perp = Q_4'$ *and* $N_4^\perp = N_4'$. *If* $p \equiv 1 \pmod 8$, *then* $Q_4^\perp = N_4'$ *and* $N_4^\perp = Q_4'$.

Proof. By Propositions 7.29 and 11.8. □

Corollary 11.11. *If* $p \equiv -1 \pmod 8$, Q_4' *and* N_4' *are self-orthogonal codes.* □

Definition 11.2. The *extended quaternary quadratic residue codes* $\bar{Q}_4$ and $\bar{N}_4$ are defined to be the $\mathbb{Z}_4$-codes obtained from Q_4 and N_4, respectively, by adjoining the zero-sum check symbol $c_\infty = -\sum_{i=0}^{p-1} c_i$ to every codeword $(c_0, c_1, \ldots, c_{p-1})$ of Q_4 and N_4 at coordinate position ∞. □

We have the following lemma:

Lemma 11.12. *Let G be a generator matrix of Q_4' (or N_4'). Then*

$$\begin{pmatrix} 1\ 1 \cdots 1 \\ G \end{pmatrix} \tag{11.4}$$

is a generator matrix of Q_4 (or N_4, respectively). Moreover, when $p \equiv -1 \pmod 8$,

$$\left(\begin{array}{c|ccc} 1 & 1 & 1 \cdots & 1 \\ 0 & & & \\ \vdots & & G & \\ 0 & & & \end{array}\right) \tag{11.5}$$

is a generator matrix of $\bar{Q}_4$ (or $\bar{N}_4$, respectively), and when $p \equiv 1 \pmod 8$,

$$\left(\begin{array}{c|ccc} 3 & 1 & 1 \cdots & 1 \\ 0 & & & \\ \vdots & & G & \\ 0 & & & \end{array}\right) \tag{11.6}$$

is a generator matrix of $\bar{Q}_4$ (or $\bar{N}_4$, respectively).

Proof. By Definition 11.1, Q_4 and Q_4' are cyclic $\mathbb{Z}_4$-codes of length p with generator polynomials $g(X)$ and $(X-1)\,g(X)$, respectively. Over $\mathbb{Z}_2$, we have

$$X^p - 1 = (X-1)\, g_2(X)\, h_2(X)\,,$$

where $X-1, g_2(X)$ and $h_2(X)$ are pairwise coprime. Over $\mathbb{Z}_4$, we have

$$X^p - 1 = (X-1)\, g(X)\, h(X)\,,$$

where $\bar{g}(X) = g_2(X)$ and $\bar{h}(X) = h_2(X)$. By Lemma 5.1, $X-1$, $g(X)$ and $h(X)$ are pairwise coprime over $\mathbb{Z}_4$. But

$$X^{p-1} + X^{p-2} + \cdots + X + 1 = g(X)\, h(X)\,.$$

Therefore there are polynomials $a(X)$ and $b(X)$ such that

$$a(X)(X^{p-1} + X^{p-2} + \cdots + X + 1) + b(X)\,(X - 1)\,g(X) = g(X)\,.$$

It follows that (11.4), with G a generator matrix of Q'_4, is a generator matrix of Q_4. Similarly, (11.4), with G a generator matrix of N'_4, is a generator matrix of N_4.

The second and third assertions follow immediately from the first one. □

Definition 11.3. When $p \equiv 1 \pmod 8$ we define $\tilde{Q}_4$ (or $\tilde{N}_4$) to be the $\mathbb{Z}_4$-codes generated by the following matrix

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & & & & \\ \vdots & & G & & \\ 0 & & & & \end{pmatrix}, \tag{11.7}$$

where G is a generator matrix of Q'_4 (or N'_4, respectively). □

From Proposition 11.10, we deduce

Proposition 11.13. *If $p \equiv -1 \pmod 8$, $\bar{Q}_4$ and $\bar{N}_4$ are self-dual codes. If $p \equiv 1 \pmod 8$, $\bar{Q}_4^{\perp} = \tilde{Q}_4$ and $\bar{N}_4^{\perp} = \tilde{N}_4$.*

Proof. By Lemma 11.12, $\bar{Q}_4$ has generator matrix (11.5), where G is a generator matrix of Q'_4. If $p \equiv -1 \pmod 8$, by Proposition 11.10 every row of G is orthogonal to every row of (11.4). It follows that any two rows of G are orthogonal. Thus $\bar{Q}_4^{\perp} \subseteq \bar{Q}_4$. But $|\bar{Q}_4| = |Q_4| = 4^{(p+1)/2}$. By Proposition 1.2, $|\bar{Q}_4^{\perp}| = 4^{p+1-(p+1)/2} = 4^{(p+1)/2}$. Therefore $\bar{Q}_4^{\perp} = \bar{Q}_4$, i.e., $\bar{Q}_4$ is self-dual. Similarly, if $p \equiv -1 \pmod 8$, $\bar{N}_4$ is also self-dual.

If $p \equiv 1 \pmod 8$, $\bar{Q}_4$ (or $\bar{N}_4$) has generator matrix (11.6),

$$\begin{pmatrix} 3 & 1 & 1 & \cdots & 1 \\ 0 & & & & \\ \vdots & & G & & \\ 0 & & & & \end{pmatrix}$$

where G is a generator matrix of Q'_4 (or N'_4, respectively). It also follows from Proposition 11.10 that $\bar{Q}_4^{\perp} = \tilde{Q}_4$ and $\bar{N}_4^{\perp} = \tilde{N}_4$. □

A $\mathbb{Z}_4$-linear code C is called *isodual* if C is equivalent to its dual $C^{\perp}$.

Clearly, when $p \equiv 1 \pmod 8$, $\tilde{Q}_4$ (or $\tilde{N}_4$) is equivalent to $\bar{Q}_4$ (or $\bar{N}_4$, respectively). Therefore the second assertion of Proposition 11.13 implies

Corollary 11.14. *If $p \equiv 1 \pmod 8$, both $\bar{Q}_4$ and $\bar{N}_4$ are isodual.* □

By Proposition 11.9 Q_4 and N_4 are permutation-equivalent. Therefore $\bar{Q}_4$ and $\bar{N}_4$ are permutation-equivalent and when $p \equiv 1 \pmod 8$, $\tilde{Q}_4$ and $\tilde{N}_4$ are permutation-equivalent. Therefore it is sufficient to study Q_4, $\bar{Q}_4$ and $\tilde{Q}_4$.

Parallel to Proposition 11.16 we have

Proposition 11.15. $\mathrm{PSL}_2(\mathbb{F}_p) \subseteq \mathrm{Aut}\,\bar{Q}_4$. *If $p \equiv 1 \pmod 8$, we also have* $\mathrm{PSL}_2(\mathbb{F}_p) \subseteq \mathrm{Aut}\,\tilde{Q}_4$. □

The proof of Proposition 11.15 is almost the same as that of Proposition 11.6 and is omitted.

Proposition 11.16. *Let d be the minimum Lee weight of the quaternary quadratic residue code $\bar{Q}_4$ of length p, where p is a prime $\equiv \pm 1 \pmod 8$. Then*

(i) *if there is a minimum Lee weight codeword in $\bar{Q}_4$, which has a component equal to 2, then*

$$(d-1)^2 - (d-1) + 1 - 4n_2\,(n_2 - 1) \geq 2q + 1\,,$$

where n_2 is the number of components equal to 2 of that codeword.

(ii) *If all minimum Lee weight codewords in $\bar{Q}_4$ have no component equal to 2, then*

$$d^2 \geq 3q/2\,.$$

Proof. By Proposition 11.9, Q_4 and N_4 are permutation-equivalent, so they have the same minimum Lee weight d. We know that

$$Q_4 = (g(X)), \qquad N_4(X) = (h(X))\,,$$

and

$$g(X)\,h(X) = \frac{X^p - 1}{X - 1}\,.$$

Therefore the intersection of the $\mathbb{Z}_4$-cyclic codes Q_4 and N_4 is the code

$$(1 + X + \cdots + X^{p-1}) = \{\varepsilon 1^p \mid \varepsilon \in \mathbb{Z}_4\}\,.$$

Let $\mathbf{f} = (f_\infty, f_0, f_1, \ldots, f_{p-1})$ be a codeword of minimum Lee weight d in $\bar{Q}_4$ and let

$$n_2 = |\{i \in \{\infty, 0, 1, \ldots, p-1\} \mid f_i = 2\}|,$$

i.e., n_2 is the number of components of $\mathbf{f}$ which are equal to 2. Then the number of components of $\mathbf{f}$ which are equal to 1 or 3 is $d - 2n_2$.

Consider first the case $n_2 \neq 0$. By Proposition 11.15, $\mathrm{PSL}_2(\mathbb{F}_p) \subseteq \mathrm{Aut}\,\bar{Q}_4$. Since $\mathrm{PSL}_2(\mathbb{F}_2)$ acts transitively on $\mathrm{PG}(1, \mathbb{F}_p) = \{\infty, 0, 1, \ldots, p-1\}$, by applying some automorphism in $\mathrm{PSL}_2(\mathbb{F}_p)$, we can assume that $f_\infty = 2$. We assert that not all nonzero f_i $(0 \leq i \leq p-1)$ are equal to 2; otherwise, $(\beta(f_0), \beta(f_1), \ldots, \beta(f_{p-1}))$ would be a codeword in Q_4 and then $(-\sum_{i=0}^{p-1} \beta(f_i), \beta(f_0), \beta(f_1), \ldots, \beta(f_{p-1}))$ would be a codeword of Lee weight less than d in $\bar{Q}_4$, which is a contradiction.

Let j be an integer with $0 < j < p$ and $j \in N$. Then $\pi_j(\bar{Q}_4) = \bar{N}_4$ and $\pi_j(\mathbf{f}) \in \bar{N}_4$. Let $\pi_j(\mathbf{f}) = \mathbf{k} = (k_\infty, k_0, k_1, \ldots, k_{p-1})$. Applying some automorphism in $\mathrm{PSL}_2(\mathbb{F}_p)$ we can assume that $k_\infty = \pm 1$.

Let

$$f(X) = \sum_{i=0}^{p-1} f_i X^i \qquad \text{and} \qquad k(X) = \sum_{i=0}^{p-1} k_i X^i,$$

then $f(X) \in Q_4$ and $k(X) \in N_4$. Since $Q_4 \cap N_4 = (1 + X + \cdots + X^{p-1})$, we have

$$f(X)\, k(X) = \alpha(1 + X + \cdots + X^{p-1}), \qquad \text{where} \qquad \alpha \in \mathbb{Z}_4.$$

Substituting $X = 1$ into the above equation, we obtain

$$f_\infty\, k_\infty = \alpha p \,(\mathrm{mod}\, 4).$$

Since $f_\infty = 2$, $k_\infty = \pm 1$, and $p \equiv \pm 1 \pmod 4$, we must have $\alpha = 2$. It follows that

$$f(X)\, k(X) = 2(1 + X + \cdots + X^{p-1}),$$

all the p coefficients of which are equal to 2. Now let us examine how products $f_i k_j$ might combine to give a coefficient 2 in $1 + X + \cdots + X^{p-1}$. Clearly

$$w_2(f(X)) = n_2 - 1, \qquad w_1(f(X)) + w_3(f(X)) = d - 2n_2,$$

$$w_2(k(X)) = n_2, \qquad w_1(k(X)) + w_3(k(X)) = d - 2n_2 - 1.$$

Therefore

$$(n_2 - 1)(d - 2n_2 - 1) + (d - 2n_2) n_2 + (d - 2n_2)(d - 2n_2 - 1)/2 \geq p.$$

Simplifying, we get

$$(d-1)^2 - (d-1) + 1 - 4n_2(n_2-1) \geq 2p + 1\,.$$

Thus (i) is proved.

(ii) can be proved in a similar way. □

Example 11.4. Let us study the quaternary quadratic residue code $Q_4(7)$. We have

$$X^7 - 1 = (X-1)\,g_2(X)\,h_2(X) \text{ over } \mathbb{Z}_2\,,$$

where

$$g_2(X) = X^3 + X + 1\,,$$
$$h_2(X) = X^3 + X^2 + 1\,.$$

By Proposition 5.15, the Hensel lifts of $g_2(X)$ and $h_2(X)$ can be computed and are

$$g(X) = X^3 + 2X^2 + X - 1\,,$$

and

$$h(X) = X^3 - X^2 - 2X - 1\,,$$

respectively. We have

$$X^7 - 1 = (X-1)\,g(X)\,h(X) \text{ over } \mathbb{Z}_4\,.$$

The quaternary quadratic residue code $Q_4(7)$ is the $\mathbb{Z}_4$-cyclic code of length 7 with generating polynomial $g(X)$ and the extended quaternary quadratic residue code $\overline{Q_4(7)}$ is the $\mathbb{Z}_4$-linear code with generator matrix

$$\begin{pmatrix} 1 & 1 & 2 & 1 & -1 & & & \\ 1 & & 1 & 2 & 1 & -1 & & \\ 1 & & & 1 & 2 & 1 & -1 & \\ 1 & & & & 1 & 2 & 1 & -1 \end{pmatrix}$$

It is easy to prove that the above matrix is equivalent to (1.6). Hence $\overline{Q_4(7)}$ is equivalent to the octacode. By Proposition 11.13, $\overline{Q_4(7)}$ is self-dual, but we already knew that the octacode is self-dual in Example 1.3. □

Example 11.5. Let us study the quaternary quadratic residue code $Q_4(23)$. We have

$$X^{23} - 1 = (X-1)\,g_2(X)\,h_2(X) \text{ over } \mathbb{Z}_2\,,$$

where

$$g_2(X) = X^{11} + X^9 + X^7 + X^6 + X^5 + X + 1\,,$$

and

$$h_2(X) = X^{11} + X^{10} + X^6 + X^5 + X^4 + X^2 + 1\,.$$

By Proposition 5.15, the Hensel lifts of $g_2(X)$ and $h_2(X)$ can be computed, and they are

$$g(X) = X^{11} + 2X^{10} + 3X^9 + 3X^7 + 3X^6 + 3X^5 + 2X^4 + X + 3$$

and

$$h(X) = X^{11} + 3X^{10} + 2X^7 + X^6 + X^5 + X^4 + X^2 + 2X + 3$$

respectively. We have

$$X^{23} - 1 = (X - 1)\, g(X)\, h(X) \text{ over } \mathbb{Z}_4\,.$$

The quaternary quadratic residue code $Q_4(23)$ is the $\mathbb{Z}_4$-cyclic code of length 23 with generating polynomial $g(X)$ and the extended quaternary quadratic residue code $\overline{Q_4(23)}$ is a $\mathbb{Z}_4$-linear code, the generator matrix of which can be easily written down. $Q_4(23)$ and $\overline{Q_4(23)}$ are also called the *quaternary Golay code* and the *extended quaternary Golay code*, respectively. By Proposition 11.13, $\overline{Q_4(23)}$ is self-dual. Moreover, $\overline{Q_4(23)}$ has 4^{12} codewords and minimum Lee weight 12. □

Example 11.6. Let us study the quaternary quadratic residue codes $Q_4(17)$. We have

$$X^{17} - 1 = (X - 1)\, g_2(X)\, h_2(X) \text{ over } \mathbb{Z}_2\,,$$

where

$$g_2(X) = X^8 + X^5 + X^4 + X^3 + 1$$

and

$$h_2(X) = X^8 + X^7 + X^6 + X^4 + X^2 + X + 1\,.$$

By Proposition 5.15 we can compute the Hensel lifts of $g_2(X)$ and $h_2(X)$, and they are

$$g(X) = X^8 + 2X^6 + 3X^5 + X^4 + 3X^3 + 2X^2 + 1$$

and

$$h(X) = X^8 + X^7 + 3X \quad + 3X^4 + 3X^2 + X + 1$$

respectively. We have

$$X^{17} - 1 = (X-1)\,g(X)\,h(X) \text{ over } \mathbb{Z}_4 .$$

The quaternary quadratic residue code $Q_4(17)$ is the $\mathbb{Z}_4$-cyclic code of length 17 with generator polynomial $g(X)$. By Corollary 11.14, the extended quaternary quadratic residue code $\overline{Q_4(17)}$ is isodual, or more precisely, $\overline{Q_4(17)}^{\perp} = \widetilde{Q_4(17)}$, where $\widetilde{Q_4(17)}$ is the $\mathbb{Z}_4$-linear code with generator matrix (11.7), where G is a generator matrix of the $\mathbb{Z}_4$-cyclic code $Q_4'(17)$ with generator polynomial $(X-1)g(X)$. $\overline{Q_4(17)}$ has 4^9 codewords and minimum Lee weight 8. □

For later applications let us introduce the *Euclidean weights* of vectors in $\mathbb{Z}_4^n$. First, the Euclidean weights of $0,1,2,3$ of $\mathbb{Z}_4$ are defined to be $0,1,4,1$, respectively. Then the Euclidean weight of an n-tuple in $\mathbb{Z}_4^n$ is defined to be the integral sum of the Euclidean weights of its components.

Proposition 11.17. *Let p be an odd prime and assume that $p \equiv -1 \pmod 8$. Then all Euclidean weights of the codewords in the extended quaternary quadratic residue codes $\overline{Q}_4$ and $\overline{N}_4$ are divisible by* 8.

Proof. By Proposition 11.9, Q_4 and N_4 are equivalent. So we consider only Q_4. Write $p+1 = 8r$, where r is a positive integer. Let $e(X)$ be the generating idempotent of Q_4 and C the $p \times p$ circulant matrix with the coefficients of $X^0, X^1, \ldots, X^{p-1}$ of $e(X)$ as its first row. Then the rows of C span Q_4.

If r is odd, then $e(X) = \theta_Q(X) + 2\theta_N(X)$ and there are $\frac{p-1}{2} = 4r-1$ coefficients of $e(X)$ equal to 1, $4r-1$ coefficients equal to 2, and all other coefficients equal to 0. So the zero-sum check symbol of $e(X)$ is -1. Thus the rows of the matrix

$$\begin{pmatrix} -1 & \\ \vdots & C \\ -1 & \end{pmatrix}$$

span $\bar{Q}_4$ over $\mathbb{Z}_4$. The Euclidean weight of each row of the above matrix is equal to

$$1 + (4r-1) + (4r-1)\cdot 4 = 20r - 4 \equiv 0 \pmod 8 .$$

If r is even, then $e(X) = 3\theta_Q(X)$, there are $4r-1$ coefficients of $e(X)$ equal to 3 and all other coefficients equal to 0. So the zero-sum check symbol of $e(X)$ is also -1. Thus the rows of the matrix

$$\begin{pmatrix} -1 & \\ \vdots & C \\ -1 & \end{pmatrix}$$

span $\bar{Q}_4$ over $\mathbb{Z}_4$. The Euclidean weight of each row of the above matrix is equal to $1 + 4r - 1 = 4r \equiv 0 \pmod 8$.

Then we can use induction to prove that the Euclidean weight of every codeword in $\bar{Q}_4$ is divisible by 8. This follows from the identity

$$\| \mathbf{x} + \mathbf{y} \|^2 \equiv \| \mathbf{x} \|^2 + \| \mathbf{y} \|^2 + 2\mathbf{x} \cdot \mathbf{y} \pmod 8,$$

where $\| \mathbf{x} \|^2$ denote the Euclidean weight of $\mathbf{x} \in \mathbb{Z}_4^n$, and the fact that $\bar{Q}_4$ is self-dual. $\square$

For the extended quaternary Golay code we have

Proposition 11.18. *The extended quaternary Golay code has minimum Lee weight* 12, *minimum Euclidean weight* 16, *and minimum Hamming weight* 8. $\square$

For the proof of Proposition 11.18, see Bonnecaze *et al.* (1995).

Changing from a binary alphabet to a quaternary alphabet provides extra flexibility in constructing self-dual and isodual codes.

Definition 11.4. The supplementary quaternary quadratic residue codes $S_Q(p)$ and $S_N(p)$ are defined to be the $\mathbb{Z}_4$-linear codes obtained by supplementing the codes $Q_4'(p)$ and $N_4'(p)$, respectively, with the all 2 p-tuple $2(1^p)$ $(2, 2, \ldots, 2)$. That is $S_Q(p) = \langle Q_4'(p), 2(1^p)\rangle$ and $S_N(p) = \langle N_4'(p), 2(1^p)\rangle$. $\square$

We write S_Q and S_N for $S_Q(p)$ and $S_N(p)$, respectively, if no ambiguity arises.

From Proposition 11.10 we deduce also

Proposition 11.19. *If* $p \equiv -1 \pmod 8$, *then* S_Q *and* S_N *are self-dual. If* $p \equiv 1 \pmod 8$, *then* $S_Q^{\perp} = S_N$ *and* S_Q *and* S_N *are isodual.*

Proof. We consider only S_Q for S_N can be treated in a similar way.

First we prove that the word $2(1^p) \notin Q'_4$, where Q'_4 is the $\mathbb{Z}_4$-cyclic code of length p with generator polynomial $(X-1)\,g(X)$. $2(1^p)$ can be expressed as the polynomial

$$2 + 2X + 2X^2 + \cdots + 2X^{p-1}$$

Substituting $X = 1$ into this polynomial, we obtain $2p \not\equiv 0 \pmod 4$. Therefore $2 + 2X + 2X^2 + \cdots + 2X^{p-1}$ is not a multiple of $X - 1$. Hence $2(1^p) \notin Q'_4$. It follows that S_Q has generator matrix

$$\begin{pmatrix} G \\ 2\,2\cdots 2 \end{pmatrix}, \tag{11.8}$$

where G is a generator matrix of Q'_4.

Consider the case $p \equiv -1 \pmod 8$. By Lemma 11.12, (11.4)

$$\begin{pmatrix} 1\,1\cdots 1 \\ G \end{pmatrix}$$

is a generator matrix of Q_4. By Proposition 11.10, every row of G is orthogonal to every row of (11.4). Clearly, the last row of (11.8) is orthogonal to itself. Therefore any two rows of (11.8) are orthogonal. It follows that $S_Q^\perp \subset S_Q$. But $|S_Q| = |Q'_4| \cdot 2 = 4^{(p-1)/2}\,2$. By Proposition 1.2, we also have $|S_Q^\perp| = 4^{(p-1)/2} \cdot 2$. Therefore $S_Q^\perp = S_Q$, i.e., S_Q is self-dual.

Then consider the case $p \equiv 1 \pmod 8$. By Proposition 11.10, $Q_4 = N_4'^{\perp}$ and $N_4 = Q_4'^{\perp}$. But $S_Q \subset Q_4$, so $S_Q \subset N_4'^{\perp}$. From $1^p \in N_4$ we deduce that 1^p is orthogonal to every row of G. Therefore $2(1^p)$ is also orthogonal to every row of G and that $S_Q \subset \langle 2(1^p)\rangle^\perp$. It follows that $S_Q \subset \langle N'_4, 2(1^p)\rangle^\perp = S_N^\perp$. But $|S_Q| = |S_N| = 4^{(p-1)/2}\ \ 2$, therefore $S_Q = S_N^\perp$ and $S_Q^\perp = S_N$. By Proposition 11.9, Q_4 and N_4 are permutation-equivalent, so are S_Q and S_N. Hence S_Q is isodual. □

Example 11.7. $S_Q(7)$ is a self-dual $\mathbb{Z}_4$-code of length 7 and its binary image $\phi(S_Q(7))$ is a formally self-dual binary code of length 14. Clearly $|S_Q(7)| = 4^3 \cdot 2 = 2^7$ and the minimum Lee weight of $S_Q(7)$ is 4. Therefore $|\phi(S_Q(7))| = 2^7$ and the minimum Hamming distance of $\phi(S_Q(7))$ is also 4.

$S_Q(17)$ is an isodual code of length 17, it has 2^{17} codewords and its minimum Lee weight is 6.

$S_Q(23)$ is a self-dual $\mathbb{Z}_4$-code of length 23, it has 2^{23} codewords. It can be proved that the minimum Euclidean weight of $S_Q(23)$ is 12. □

Most of this section are from Bonnecaze and Solé (1994) and Bonnecaze *et al.* (1995).

CHAPTER 12

QUATERNARY CODES AND LATTICES

12.1. Lattices

We state some definitions and facts on lattices below; for details, see Conway and Sloane (1993).

Let $\mathbb{R}^n = \{(x_1, \ldots, x_n) \mid x_i \in \mathbb{R}\}$ be the n-dimensional row vector space over $\mathbb{R}$. For $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ define the *inner product* of $\mathbf{x}$ and $\mathbf{y}$ by

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^t \mathbf{y} = x_1 y_1 + \cdots + x_n y_n ,$$

then $\mathbb{R}^n$ together with the inner product is called the *n-dimensional Euclidean space*, which is also denoted by $\mathbb{R}^n$. For any $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}^2 = \mathbf{x} \cdot \mathbf{x}$ is called the norm of $\mathbf{x}$.

A *lattice* L in the n-dimensional Euclidean space $\mathbb{R}^n$ is a free abelian subgroup of rank n of the additive group of $\mathbb{R}^n$, i.e., there exists a basis $\{\varepsilon_1, \ldots, \varepsilon_n\}$ of $\mathbb{R}^n$ such that $L = \mathbb{Z}\varepsilon_1 + \cdots + \mathbb{Z}\varepsilon_n$. L is also called an *n-dimensional lattice.*

For example, $\mathbb{Z}^n = \{(z_1, \ldots, z_n) | z_i \in \mathbb{Z}\}$ is a lattice in $\mathbb{R}^n$. There exists a basis

$$\{\mathbf{e}_1 = (1, 0, \ldots, 0), \ldots, \mathbf{e}_n = (0, \ldots, 0, 1)\}$$

of $\mathbb{R}^n$ such that $\mathbb{Z}^n = \mathbb{Z}\mathbf{e}_1 + \cdots + \mathbb{Z}\mathbf{e}_n$. $\mathbb{Z}^n$ is called the *standard lattice* of $\mathbb{R}^n$ and $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a basis of $\mathbb{Z}^n$.

Let $\varepsilon_1 = (1, 0)$ and $\varepsilon_2 = (\sqrt{2}, \sqrt{2})$, then $L = \mathbb{Z}\varepsilon_1 + \mathbb{Z}\varepsilon_2$ is a lattice in $\mathbb{R}^2$, (see Fig. 12.1).

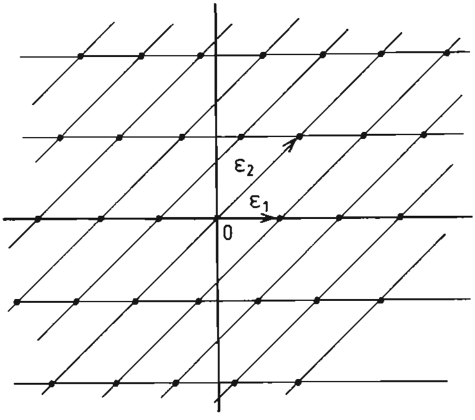

Fig. 12.1.

Let L be a lattice in $\mathbb{R}^n$ A basis $\{\varepsilon_1, \ldots, \varepsilon_n\}$ of $\mathbb{R}^n$ such that $L = \mathbb{Z}\varepsilon_1 + \cdots + \mathbb{Z}\varepsilon_n$ is called a *basis* of L. Let $\{\varepsilon_1, \ldots, \varepsilon_n\}$ be a basis of L, $q_{ij} \in \mathbb{Z}$ $(1 \leq i, j \leq n)$, and

$$\eta_i = \sum_{j=1}^{n} q_{ij}\varepsilon_j, \qquad i = 1, \ldots, n,$$

then $\{\eta_1, \ldots, \eta_n\}$ is also a basis of L if and only if the matrix

$$Q = (q_{ij})_{1 \leq i, j \leq n}$$

is *unimodular*, i.e., det $Q = \pm 1$.

A *fundamental region* of a lattice L in $\mathbb{R}^n$ is a set of vectors in $\mathbb{R}^n$ that contains one and only one vector from each coset of $\mathbb{R}^n$ relative to L. Let $\varepsilon_1, \ldots, \varepsilon_n$ be a basis of L, then the parallelogram

$$P_L = \{x_1\varepsilon_1 + \cdots + x_n\varepsilon_n \,|\, 0 \leq x_i < 1\}$$

is an example of a fundamental region of L, called a *fundamental parallelogram*. The *volume* of P_L is

$$\operatorname{vol} P_L = \left| \det \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} \right| .$$

Clearly,

$$(\operatorname{vol} P_L)^2 = \det(\varepsilon_i \cdot \varepsilon_j)_{1 \leq i, j \leq n}$$

and $(\text{vol } P_L)^2$ is independent of the particular choice of the basis of L. Define

$$\text{disc } L = \det\,(\varepsilon_i \cdot \varepsilon_j)_{1 \leq i, j \leq n}\,,$$

which is called the *discriminant* of L. Clearly

$$\text{disc } L = (\text{vol } P_L)^2\,.$$

If L and L' are lattices in $\mathbb{R}^n$ and $L' \subseteq L$, then we have

$$\text{disc } L' = \text{disc } L\,|\,L/L'|^2\,, \tag{12.1}$$

where $|L/L'|$ is the index of L' in L.

For the standard lattice Z^n in $\mathbb{R}^n$ we have disc $\mathbb{Z}^n = 1$. For the lattice $L = \mathbb{Z}\varepsilon_1 + \mathbb{Z}\varepsilon_2$ where $\varepsilon_1 = (1,\,0)$ and $\varepsilon_2 = (\sqrt{2},\,\sqrt{2})$, we have disc $L = 2$.

Let L be a lattice in $\mathbb{R}^n$. The *dual* of L, denoted by L^* is defined by

$$L^* = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{x} \cdot \mathbf{y} \in \mathbb{Z} \text{ for all } \mathbf{y} \in L\}\,.$$

For example, $(\mathbb{Z}^n)^* = \mathbb{Z}^n$ and for $L = \mathbb{Z}\varepsilon_1 + \mathbb{Z}\varepsilon_2$, where $\varepsilon_1 = (1,\,0)$ and $\varepsilon_2 = (\sqrt{2},\,\sqrt{2})$, we have $L^* = \mathbb{Z}(1,\,-1) + \mathbb{Z}(0,\,\frac{1}{\sqrt{2}})$.

It is easy to prove that if L is an n-dimensional lattice with $\{\varepsilon_1,\,\ldots,\,\varepsilon_n\}$ as a basis, then L^* is also an n-dimensional lattice with $\{\varepsilon_1^*,\,\ldots,\,\varepsilon_n^*\}$ as a basis, where $\varepsilon_1^*,\,\ldots,\,\varepsilon_n^*$ are defined by

$$\varepsilon_i \cdot \varepsilon_j^* = \delta_{ij}\,, \qquad 1 \leq i,\, j \leq n\,.$$

A lattice L in $\mathbb{R}^n$ is said to be *integral*, if $L \subseteq L^*$, in other words, if $\mathbf{x} \cdot \mathbf{y} \in \mathbb{Z}$ for all $\mathbf{x}$, $\mathbf{y} \in L$. L is said to be *even*, if $\mathbf{x}^2 \in 2\mathbb{Z}$ for all $\mathbf{x} \in L$. L is said to be *unimodular*, if $L^* = L$.

It is clear that L is unimodular if and only if L is integral and disc $L = 1$. It is also clear that if L is even, then it is also integral.

The *theta series* $\theta_L(q)$ of the integral lattice L is the formal power series

$$\theta_L(q) = \sum_{\mathbf{x} \in L} q^{\mathbf{x}^2} = \sum_{m=0}^{\infty} N_m q^m\,,$$

where N_m is the number of vectors $\mathbf{x} \in L$ with norm m.

Let L be an n-dimensional lattice, L_1 and L_2 be lattices contained in L and of dimensions n_1 and n_2, respectively, and $n = n_1 + n_2$. Assume that every vector of L can be expressed uniquely as a sum of a vector of L_1 and a vector of L_2 and that $\mathbf{x} \cdot \mathbf{y} = 0$ for any $\mathbf{x} \in L_1$ and $\mathbf{y} \in L_2$. Then we say that L is the *orthogonal direct sum* of L_1 and L_2 and write $L = L_1 \perp L_2$.

Two lattices L_1 and L_2 in $\mathbb{R}^n$ are said to be *isomorphic*, if there exist an orthogonal transformation σ, i.e., an element $\sigma \in O_n(\mathbb{R}^n)$ such that $\sigma(L_1) = L_2$.

12.2. A Construction of Lattices from Quaternary Linear Codes

Let ρ be the natural homomorphism

$$\begin{aligned} \rho : \mathbb{Z} &\longrightarrow \mathbb{Z}_4 \\ n &\mapsto n + (4) \end{aligned}$$

from the ring of integers to the residue class ring of $\mathbb{Z}$ modulo the ideal (4). As before, the elements of $\mathbb{Z}_4$ are denoted by $0, 1, 2$ and 3, that is, they represent the residue class (4), $1 + (4)$, $2 + (4)$, and $3 + (4)$, respectively.

The map ρ can be extended to a map from the standard lattice $\mathbb{Z}^n$ in $\mathbb{R}^n$ to $\mathbb{Z}_4^n$, the additive group of n-tuples over $\mathbb{Z}_4$, which is denoted also by ρ, as follows:

$$\begin{aligned} \rho : \mathbb{Z}^n &\longrightarrow \mathbb{Z}_4^n \\ (x_1, \ldots, x_n) &\mapsto (\rho(x_1), \ldots, \rho(x_n)) . \end{aligned} \tag{12.2}$$

Clearly, this is a group homomorphism.

Let C be a quaternary linear code of length n and type $4^{k_1}2^{k_2}$. Denote the complete inverse image of C under ρ by $\rho^{-1}(C)$. Then we have

Proposition 12.1. *Let $\rho : \mathbb{Z}^n \longrightarrow \mathbb{Z}_4^n$ be the map* (12.2) *and C be a quaternary linear code of length n and type $4^{k_1}2^{k_2}$. Then $\rho^{-1}(C)$ is a lattice in $\mathbb{R}^n$ and* disc $\rho^{-1}(C) = 4^{2n-2k_1-k_2}$.

Proof. Since $|C| = 4^{k_1}2^{k_2}$, $|\mathbb{Z}_4^n/C| = 4^{n-k_1-k_2}2^{k_2}$. By the first isomorphism theorem,

$$\mathbb{Z}^n/\rho^{-1}(C) \simeq \mathbb{Z}_4^n/C .$$

Consequently, $|\mathbb{Z}^n/\rho^{-1}(C)| = 4^{n-k_1-k_2}2^{k_2}$. Therefore $\rho^{-1}(C)$ is a free abelian subgroup of $\mathbb{Z}^n$ of rank n and hence, is a lattice in $\mathbb{R}^n$ Moreover, by (12.1)

$$\begin{aligned} \operatorname{disc} \rho^{-1}(C) &= \operatorname{disc} \mathbb{Z}^n \, |\mathbb{Z}^n/\rho^{-1}(C)|^2 \\ &= 4^{2n-2k_1-k_2} . \end{aligned}$$

□

Definition 12.1. Let C be a quaternary linear code of length n. The lattice

$$L_C = \frac{1}{2}\,\rho^{-1}(C)$$

is called the *lattice in* $\mathbb{R}^n$ *associated with* C. □

Clearly,

$$L_C = \left\{ \frac{1}{2}\,(\mathbf{c} + 4\mathbf{z}) \,|\, \mathbf{c} \in C,\ \mathbf{z} \in \mathbb{Z}^n \right\},$$

where $\mathbf{c}$ is regarded as n-tuples with integers 0, 1, 2, 3 as components.

Proposition 12.2. *Let C be a quaternary linear code of length n and $L_C = \frac{1}{2}\,\rho^{-1}(C)$. Then*

(i) *L_C is integral if and only if C is self-orthogonal.*
(ii) *L_C is unimodular if and only if C is self-dual.*
(iii) *L_C is even if and only if the Euclidean weights of all codewords of C are divisible by 8.*

Proof. Let

$$\mathbf{x}_1 = \frac{1}{2}\,(\mathbf{c}_1 + 4\mathbf{z}_1) \qquad \text{and} \qquad \mathbf{x}_2 = \frac{1}{2}\,(\mathbf{c}_2 + 4\mathbf{z}_2)$$

be any two vectors of L_C, i.e., $\mathbf{c}_1, \mathbf{c}_2 \in C$ and $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{Z}^n$. We manipulate in $\mathbb{R}$,

$$\mathbf{x}_1 \cdot \mathbf{x}_2 = \frac{1}{4}\,\mathbf{c}_1 \cdot \mathbf{c}_2 + \mathbf{c}_1 \cdot \mathbf{z}_2 + \mathbf{z}_1 \cdot \mathbf{c}_2 + 4\mathbf{z}_1 \cdot \mathbf{z}_2\,. \tag{12.3}$$

(i) If C is self-orthogonal, $\mathbf{c}_1 \cdot \mathbf{c}_2 = 0$ when it is manipulated in $\mathbb{Z}_4$. It follows that if $\mathbf{c}_1 \cdot \mathbf{c}_2$ is manipulated in $\mathbb{Z}$ we have $\mathbf{c}_1 \cdot \mathbf{c}_2 \in 4\mathbb{Z}$. Therefore $\mathbf{x}_1 \cdot \mathbf{x}_2 \in \mathbb{Z}$. Hence L_C is integral.

Conversely, if L_C is integral, i.e., $\mathbf{x}_1 \cdot \mathbf{x}_2 \in \mathbb{Z}$ for all $\mathbf{x}_1, \mathbf{x}_2 \in L_C$. It follows from (12.3) that $\mathbf{c}_1 \cdot \mathbf{c}_2 \in 4\mathbb{Z}$. Therefore $\mathbf{c}_1 \cdot \mathbf{c}_2 = 0$ in $\mathbb{Z}_4$. Hence C is self-orthogonal.

(ii) Let C be of type $4^{k_1}2^{k_2}$ By Proposition 12.1, disc $L_C = 4^{-n}\,4^{2n-2k_1-k_2} = 4^{n-2k_1-k_2}$ We deduce disc $L_C = 1$ if and only if $n = 2k_1 + k_2$. Therefore disc $L_C = 1$ if and only if $|C| = 2^n$. By (i), L_C is integral if and only if $C \subseteq C^\perp$. Hence

$$\begin{aligned} L_C \text{ is unimodular} &\Leftrightarrow L_C \text{ is integral and disc } L_C = 1 \\ &\Leftrightarrow C \subseteq C^\perp \quad \text{and} \quad |C| = 2^n \\ &\Leftrightarrow C = C^\perp \end{aligned}$$

(iii) Let $\mathbf{x} = \frac{1}{2}(\mathbf{c} + 4\mathbf{z})$ be any vector of L_C, i.e., $\mathbf{c} \in \mathcal{C}$ and $\mathbf{z} \in \mathbb{Z}^n$. Manipulating in $\mathbb{R}$, we have

$$\mathbf{x} \cdot \mathbf{x} = \frac{1}{4}\mathbf{c} \cdot \mathbf{c} + 2\mathbf{c} \cdot \mathbf{z} + 4\mathbf{z} \cdot \mathbf{z}\,.$$

Therefore L_C is even if and only if $\mathbf{c} \cdot \mathbf{c} \in 8\mathbb{Z}$. But $\mathbf{c} \cdot \mathbf{c}$ is the Euclidean weight of $\mathbf{c}$. □

Corollary 12.3. *Let C be a self-dual quaternary linear code such that the Euclidean weights of all codewords of C are divisible by* 8. *Then L_C is an even unimodular lattice.* □

Corollary 12.4. *Let $\overline{Q_4(p)}$ be the extended quaternary quadratic residue code of length $p+1$, where p is a prime $\equiv -1 \pmod 8$. Then $L_{\overline{Q_4(p)}}$ is an even unimodular lattice of dimension $p+1$.*

Proof. By Proposition 11.13, when $p \equiv -1 \pmod 8$, $\overline{Q_4(p)}$ is self-dual and by Proposition 11.17, the Euclidean weights of all codewords of $\overline{Q_4(p)}$ are divisible by 8. □

Example 12.1. Let $\mathcal{O}_8$ be the octacode. By Example 11.3 $\mathcal{O}_8$ is equivalent to $\overline{Q_4(7)}$. Therefore the lattice $L_{\mathcal{O}_8}$ is an even unimodular lattice. But $L_{\mathcal{O}_8}$ is an eight-dimensional lattice and the Gosset lattice E_8 is the unique even unimodular lattice of dimension 8 to within isomorphism, (see Conway and Sloane (1993)). Therefore $L_{\mathcal{O}_8}$ is isomorphic to E_8. It is known that the first three terms of the theta series $\theta_{E_8}(q)$ are as follows:

$$\theta_{E_8}(q) = 1 + 240\,q^2 + \text{higher terms}\,,$$

where 240 is the number of vectors of norm 2 in the lattice E_8, (see Conway and Sloane (1993), p. 122). □

Example 12.2. Let $\overline{Q_4(23)}$ be the extended quaternary Golay code of length 24. By Corollary 12.4, $L_{\overline{Q_4(23)}}$ is an even unimodular lattice of dimension 24. By Proposition 11.18, $\overline{Q_4(23)}$ has minimum Euclidean weight 16. It follows that the minimum norm of $L_{\overline{Q_4(23)}}$ is 4. But the Leech lattice Λ_{24} is the unique even unimodular lattice without vectors of norm 2 in $\mathbb{R}^{24}$ to within isomorphism, (see Conway and Sloane (1993), Chap. 12 or Wan (1997)). Therefore

$L_{\overline{Q_4(23)}}$ is isomorphic to Λ_{24}. It is known that the first five terms of the theta series $\theta_{\Lambda_{24}}(q)$ are as follows:

$$\theta_{\Lambda_{24}}(q) = 1 + 196,560q^4 + \text{higher terms},$$

where 196,560 is the number of vectors of minimum norm 4 in the lattice Λ_{24}, (see Conway and Sloane (1993), p. 131). □

Corollary 12.5. *Let $S_Q(p)$ be the supplemented quaternary quadratic residue code of length p, where p is a prime and $p \equiv -1 \pmod 8$. Then $L_{S_Q(p)}$ is a unimodular lattice.*

Proof. By Proposition 11.19, when $p \equiv -1 \pmod 8$, $S_Q(p)$ is self-dual. Therefore by Proposition 12.2 (ii), $L_{S_Q(p)}$ is unimodular. □

Example 12.3. Let $S_Q(7)$ be supplemented quaternary quadratic residue code of length 7. By Corollary 12.5, $L_{S_Q(7)}$ is a unimodular lattice of dimension 7. It is known that $\mathbb{Z}^7$ is the unique unimodular lattice of dimension 7 to within isomorphism, (see Conway and Sloane (1993), Chap. 2, §2.4). Therefore $L_{S_Q(7)}$ is isomorphic to $\mathbb{Z}^7$, i.e., there is an orthogonal transformation $\sigma \in O_7(\mathbb{R})$ such that $\sigma(L_{S_Q(7)}) = \mathbb{Z}^7$. □

Example 12.4. Let $S_Q(23)$ be the supplemented quaternary quadratic residue code of length 23. By Corollary 12.5, $L_{S_Q(23)}$ is a unimodular lattice of dimension 23. But $S_Q(23)$ has minimum Euclidean weight 12 (Example 11.7). It follows that $L_{S_Q(23)}$ has minimum norm 3. But there is a unique unimodular lattice of minimum norm 3 to within isomorphism, which is denoted by O_{23}, (see Conway and Sloane (1993), Chaps. 16 and 19). Therefore $L_{S_Q(23)}$ is isomorphic to O_{23}. □

We end this section with more examples.

Example 12.5. Let $\mathcal{K}_4$ be the $\mathbb{Z}_4$-linear code with generator matrix (1.3)

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 2 & 2 \end{pmatrix},$$

introduced in Example 1.1. It is known that $\mathcal{K}_4$ is a self-dual code. By Proposition 12.2 (ii) $L_{\mathcal{K}_4}$ is a unimodular lattice. It is known that $\mathbb{Z}^4$ is the unique unimodular lattice of dimension 4 to within isomorphism, (see Conway and Sloane (1993), Chap. 2, §2.4). Therefore $L_{\mathcal{K}_4}$ is isomorphic to $\mathbb{Z}^4$ □

Example 12.6. Let $\mathcal{K}_8$ be the $\mathbb{Z}_4$-linear code with generator matrix (1.7)

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 \end{pmatrix},$$

introduced in Example 1.4. It is known that $\mathcal{K}_8$ is a self-dual code. It is also clear that every row of the matrix (1.7) is of Euclidean weight 8. As in the proof of Proposition 11.17 we can show that the Euclidean weights of all codewords of $\mathcal{K}_8$ are divisible by 8. Therefore by Corollary 12.3, $L_{\mathcal{K}_8}$ is an even unimodular lattice of dimension 8. But the Gosset lattice E_8 is the unique even unimodular lattice in $\mathbb{R}^8$ to within isomorphism. Therefore $L_{\mathcal{K}_8}$ is isomorphic to E_8. Together with $L_{\mathcal{O}_8}$ of Example 12.1 we already have two constructions of E_8 from quaternary codes. □

Example 12.7. Besides $\mathcal{O}_8$ and $\mathcal{K}_8$, there are two more self-dual $\mathbb{Z}_4$-codes of length 8, denoted by $\mathcal{K}_8'$ and $\mathcal{Q}_8$, respectively, which were introduced by Conway and Sloane (1993a). $\mathcal{K}_8'$ has generator matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 & 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 \end{pmatrix} \tag{12.4}$$

and $\mathcal{Q}_8$ has generator matrix

$$\begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 2 & 1 & 3 \\ 0 & 0 & 0 & 2 & 1 & 3 & 1 & 1 \\ 1 & 1 & 0 & 2 & 0 & 0 & 1 & 3 \\ 0 & 2 & 0 & 2 & 0 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 \end{pmatrix} \tag{12.5}$$

Clearly the Euclidean weight of every row of (12.4) and (12.5) is 8. As in the proof of Proposition 11.17 the Euclidean weights of all codewords of $\mathcal{K}_8'$ and $\mathcal{Q}_8$ are divisible by 8. By Corollary 12.3, both $L_{\mathcal{K}_8'}$ and $L_{\mathcal{Q}_8}$ are even unimodular lattice in $\mathbb{R}^8$.

Thus we have altogether four constructions of the Gosset lattice E_8 from quaternary codes; they are $L_{\mathcal{O}_8}$, $L_{\mathcal{K}_8}$, $L_{\mathcal{K}_8'}$, and $L_{\mathcal{Q}_8}$. □

Example 12.8. Let RM(r, m) be the rth-order Reed–Muller code of length 2^m, where $0 \le r \le m$. Consider the $\mathbb{Z}_4$-code

$$\mathrm{RM}(1, m) + 2\mathrm{RM}(m-2, m).$$

It is known that RM$(1, m)$ is doubly even, that $\mathrm{RM}(1, m)^\perp = \mathrm{RM}(m-2, m)$, and that for any $\mathbf{a}, \mathbf{a}' \in \mathrm{RM}(1, m)$ $\mathbf{a} * \mathbf{a}' \in \mathrm{RM}(m-2, m)$. By Proposition 3.20 the code $\mathrm{RM}(1, m) + 2\mathrm{RM}(m-2, m)$ is a self-dual $\mathbb{Z}_4$-linear code. It is easy to verify that the Euclidean weights of all its codewords are divisible by 8. By Corollary 12.3 the lattice

$$L_{\mathrm{RM}(1,m)+2\mathrm{RM}(m-2,m)} = \frac{1}{2}\left(\mathrm{RM}(1, m) + 2\mathrm{RM}(m-2, m) + 4\mathbb{Z}^{2^m}\right)$$

associated with the $\mathbb{Z}_4$-linear code $\mathrm{RM}(1, m) + 2\mathrm{RM}(m-2, m)$ is an even unimodular lattice of dimension 2^m. Denote it by L_m.

When $m = 4$, it is easy to check that the minimum Euclidean weight of $\mathrm{RM}(1, 4) + 2\mathrm{RM}(2, 4)$ is 8 and the minimum norm of L_4 is 2. It is not difficult to prove that L_4 is isomorphic to $E_8 \perp E_8$.

When $m = 5$, it is easy to check that the minimum Euclidean weight of $\mathrm{RM}(1, 5) + 2\mathrm{RM}(3, 5)$ is 16 and the minimum norm of L_5 is 4. L_5 is the Barnes–Wall lattice of dimension 32 and is usually denoted by BW_{32}. □

Most of this section are from Bonnecaze and Solé (1994) and Bonnecaze *et al.* (1995).

CHAPTER 13

SOME INVARIANT THEORY

In this chapter we review some classical invariant theory, which will be needed in the study of weight enumerators of self-dual quaternary codes in Chap. 14.

13.1. The Poincaré Series

Let $\mathbb{C}$ denote the complex field, n be an integer ≥ 1, $X_1, \ldots, X_n$ be n independent indeterminates, and $\mathbb{C}[X_1, \ldots, X_n]$ be the polynomial algebra in $X_1, \ldots, X_n$ with coefficients in $\mathbb{C}$. We write $A = \mathbb{C}[X_1, \ldots, X_n]$ for simplicity. Denote by A_m the set of homogeneous polynomials of degree m, where $m \geq 0$. Then A_m $(m = 0, 1, 2, \ldots)$ are subspaces of A, $A_0 = \mathbb{C}$, $A_1 = \mathbb{C}X_1 \dot{+} \cdots \dot{+} \mathbb{C}X_n$,

$$A = \dot{\sum_{m=0}^{\infty}} A_m \,,$$

and

$$A_i\, A_j \subset A_{i+j} \qquad \text{for all} \quad i, j \geq 0\,.$$

Here and after we use $\dot{+}$ to denote the direct sum of subspaces.

Define the *Poincaré series* of $A = \mathbb{C}[X_1, \ldots, X_n]$ as the formal power series

$$\Phi(A, \lambda) = \sum_{m=0}^{\infty} (\dim A_m)\, \lambda^m \,,$$

where λ is an indeterminate. It is well known that

$$\{X_1^{m_1} \cdots X_n^{m_n} \mid m_i \geq 0 \text{ and } m_1 + \cdots + m_n = m\}$$

is a basis of A_m, therefore dim A_m is the number of partitions of m into n non-negative integers $m_1, \ldots, m_n$ and is known to be equal to

$$\binom{n+m-1}{m}.$$

Clearly, we have

$$\Phi(A, \lambda) = (1-\lambda)^{-n}.$$

More generally, let S be a subspace of $A = \mathbb{C}[X_1, \ldots, X_n]$ and assume that S is *homogeneous*, which means that for any $f \in S$ if we express f as a sum of homogeneous polynomials of different degrees $f = f_1 + \cdots + f_s$ (say), where f_i's are homogeneous and $\deg f_i \neq \deg f_j$ for $i \neq j$, then $f_i \in S$ for all $i = 1, 2, \ldots, s$. Let S_m be the set of homogeneous polynomials of degree m in S, then S_m $(m = 0, 1, 2, \ldots)$ are subspaces of S and

$$S = \dot{+}_{m=0}^{\infty} S_m .$$

Define the *Poincaré series* of S as the formal power series

$$\Phi(S, \lambda) = \sum_{m=0}^{\infty} (\dim S_m)\, \lambda^m$$

Proposition 13.1. *Let $g_1, \ldots, g_r$ be r algebraically independent homogeneous polynomials in $\mathbb{C}[X_1, \ldots, X_n]$, where $r \leq n$, and let* $\deg g_i = d_i$ $(i = 1, 2, \ldots, r)$. *Denote $B = \mathbb{C}[g_1, \ldots, g_r]$. Then B is homogeneous and the Poincaré series of B can be expressed as*

$$\Phi(B, \lambda) = \prod_{i=1}^{r} (1-\lambda^{d_i})^{-1}$$

Proof. Clearly,

$$\{g_1^{m_1} \cdots g_r^{m_r} \mid m_i \geq 0 \text{ and } m_1 d_1 + \cdots + m_r d_r = m\}$$

is a basis of B_m, thus $\dim B_m$ is the number of partitions of m into a sum of some d_1, some d_2, $\ldots$, and some d_r. But expanding

$$\prod_{i=1}^{r} (1-\lambda^{d_i})^{-1}$$

into a formal power series in λ, the coefficient of λ^m is also equal to this number. Therefore

$$\Phi(B, \lambda) = \prod_{i=1}^{r} (1-\lambda^{d_i})^{-1} . \qquad \square$$

Corollary 13.2. *Let $g_1, \ldots, g_r$ be r algebraically independent homogeneous polynomials in $\mathbb{C}[X_1, \ldots, X_n]$, where $r \le n$, and g_{r+1} be another homogeneous polynomial in $\mathbb{C}[X_1, \ldots, X_n]$. Assume that the subspace*

$$D = \mathbb{C}[g_1, \ldots, g_r] \dotplus g_{r+1}\, \mathbb{C}[g_1, \ldots, g_r]$$

is a direct sum. Let $\deg g_i = d_i$ $(i = 1, \ldots, r+1)$. *Then*

$$\Phi(D, \lambda) = (1 - \lambda^{d_{r+1}}) \prod_{i=1}^{r} (1 - \lambda^{d_i})^{-1} \qquad \square$$

13.2. Molien's Theorem

It is well known that the set of $n \times n$ nonsingular matrices over $\mathbb{C}$ form a group with respect to the matrix multiplication. This group is called the *general linear group* of degree n over $\mathbb{C}$ and denoted by $\mathrm{GL}_n(\mathbb{C})$. Denote the polynomial ring $\mathbb{C}[X_1, \ldots, X_n]$ in n independent indeterminates $X_1, \ldots, X_n$ over $\mathbb{C}$ again by A. For any $\sigma = (\sigma_{ij})_{1 \le i, j \le n} \in \mathrm{GL}_n(\mathbb{C})$ and $f \in A$ define

$$(\sigma \cdot f)(X_1, \ldots, X_n) = f((X_1, \ldots, X_n)\, {}^t\sigma).$$

Then for any $\sigma, \tau \in \mathrm{GL}_n(\mathbb{C})$,

$$\begin{aligned}
\tau \cdot (\sigma \cdot f)(X_1, \ldots, X_n) &= \sigma \cdot f((X_1, \ldots, X_n)\, {}^t\tau) \\
&= f((X_1, \ldots, X_n)\, {}^t\sigma\, {}^t\tau) \\
&= f((X_1, \ldots, X_n)\, {}^t(\tau\sigma)) \\
&= ((\tau\sigma)\ f)(X_1, \ldots, X_n).
\end{aligned}$$

Therefore the map

$$\begin{aligned}
\mathrm{GL}_n(\mathbb{C}) \times A &\to A \\
(\sigma, f) &\mapsto \sigma \cdot f
\end{aligned}$$

defines an *action* of $\mathrm{GL}_n(\mathbb{C})$ on A. Clearly, $\sigma\ A_m = A_m$, where

$$\sigma\ A_m = \{\sigma \cdot f \mid f \in A_m\}$$

Let G be a subgroup of $\mathrm{GL}_n(\mathbb{C})$. For $f \in A$, if $\sigma \cdot f = f$ for all $\sigma \in G$, then f is called a *G-invariant polynomial.* Let

$$A^G = \{f \in A \mid \sigma \cdot f = f\ \forall \sigma \in G\},$$

then A^G is a subalgebra of A, called the *algebra of G-invariant polynomials.* Clearly, $\square$

Proposition 13.3. *A^G is homogeneous. More precisely, let*

$$A_m^G = \{f \in A_m \mid \sigma \quad f = f \quad \forall \sigma \in G\},$$

then

$$A_m^G = A_m \cap A^G$$

and

$$A^G = \dot{+}_{m=0}^{\infty} (A_m \cap A^G) = \dot{+}_{m=0}^{\infty} A_m^G \qquad \square$$

Let $d_m(G) = \dim A_m^G$, which is the number of linearly independent G-invariant homogeneous polynomials of degree m. The Poincaré series of A^G, $\Phi(A^G, \lambda)$, will be abbreviated as $\Phi_G(\lambda)$, i.e.,

$$\Phi_G(\lambda) = \sum_{m=0}^{\infty} d_m(G)\, \lambda^m ,$$

which is also called the *Molien series* of A^G.

We mentioned before that for any $\sigma \in G$, $\sigma \quad A_m = A_m$. Denote the restriction of σ to A_m by $\sigma|_{A_m}$ and the trace of the linear operator $\sigma|_{A_m}$ on A_m by $\mathrm{Tr}\,(\sigma|_{A_m})$. Then we have

Lemma 13.4. *Let G be a finite subgroup of $\mathrm{GL}_n(\mathbb{C})$ and $\sigma \in G$. Then*

$$\sum_{m=0}^{\infty} \mathrm{Tr}\,(\sigma|_{A_m})\, \lambda^m = \frac{1}{\det\,(I - \lambda\sigma)} .$$

Proof. We can assume that

$$\sigma = P^{-1} \operatorname{diag}\{\lambda_1, \ldots, \lambda_n\}\, P ,$$

where $\operatorname{diag}\{\lambda_1, \ldots, \lambda_n\}$ is a diagonal matrix whose diagonal entries are $\lambda_1, \ldots, \lambda_n$ in succession. Let

$$(Y_1, \ldots, Y_n) = (X_1, \ldots, X_n)\, {}^tP\,,$$

then

$$\{Y_1^{i_1} \cdots Y_n^{i_n} \mid i_1, \ldots, i_n \geq 0 \text{ and } i_1 + \cdots + i_n = m\}$$

is a basis of A_m. For any $f \in A$, we have

$$\begin{aligned}
\sigma \cdot f(Y_1 \ldots, Y_n) &= \sigma \cdot f\left((X_1, \ldots, X_n)\, {}^tP\right) \\
&= f\,(X_1, \ldots, X_n)\, {}^t\sigma\, {}^tP\,, \\
&= f\,(Y_1, \ldots, Y_n)\, {}^tP^{-1}\, {}^t\sigma\, {}^tP \\
&= f\,(Y_1, \ldots, Y_n) \operatorname{diag}\{\lambda_1, \ldots, \lambda_n\}) \\
&= f\,(\lambda_1 Y_1, \ldots, \lambda_n Y_n)\,.
\end{aligned}$$

In particular,

$$\sigma\,(Y_1^{i_1} \cdots Y_n^{i_n}) = \lambda_1^{i_1} \cdots \lambda_n^{i_n}\, Y_1^{i_1} \cdots Y_n^{i_n}\,.$$

Therefore

$$\operatorname{Tr}(\sigma\,|_{A_m}) = \sum_{\substack{i_1, \ldots, i_n \geq 0 \\ i_1 + \cdots + i_n = m}} \lambda_1^{i_1} \cdots \lambda_n^{i_n}\,.$$

On the other hand,

$$\begin{aligned}
\frac{1}{\det(I - \lambda\sigma)} &= \frac{1}{(1 - \lambda_1\lambda) \cdots (1 - \lambda_n\lambda)} \\
&= \prod_{i=1}^{n} (1 + \lambda_i\lambda + \lambda_i^2\lambda^2 + \cdots) \\
&= \sum_{m=0}^{\infty} \left(\sum_{\substack{i_1, \ldots, i_n \geq 0 \\ i_1 + \cdots + i_n = m}} \lambda_1^{i_1} \cdots \lambda_n^{i_n} \right) \lambda^m \\
&= \sum_{m=0}^{\infty} \operatorname{Tr}(\sigma\,|_{A_m})\, \lambda^m\,.
\end{aligned}$$

□

Let G be a finite subgroup of $\mathrm{GL}_n(\mathbb{C})$. The element

$$z = \frac{1}{|G|} \sum_{\sigma \in G} \sigma$$

can be regarded as an operator acting on A, which is defined by

$$z \cdot f = \frac{1}{|G|} \sum_{\sigma \in G} \sigma \ f \qquad \text{for all} \quad f \in A\,,$$

and is called the *averaging operator* of G. Denote $z \cdot f$ simply by $\tilde{f}$. Clearly we have

(i) $\widetilde{(f_1 + f_2)} = \tilde{f}_1 + \tilde{f}_2$ for $f_1,\ f_2 \in A$,
(ii) $\widetilde{cf} = c\tilde{f}$ for $c \in \mathbb{C}$ and $f \in A$,
(iii) $\widetilde{fh} = \tilde{f}h$ for $f \in A$ and $h \in A^G$

Lemma 13.5. *Let G be a finite subgroup of* $\mathrm{GL}_n(\mathbb{C})$ *and* $z = \frac{1}{|G|} \sum_{\sigma \in G} \sigma$. *Then* $d_m(G) = \mathrm{Tr}\,(z\,|_{A_m})$.

Proof. It is easy to verify that $z^2 = z$, so the eigenvalues of $z\,|_{A_m}$ are only 0 or 1. Let $A_m = A_m^{(0)} \dotplus A_m^{(1)}$, where $A_m^{(i)}$ is the eigenspace corresponding to eigenvalue i ($i = 0,\ 1$). Then

$$\mathrm{Tr}\,(z\,|_{A_m}) = \mathrm{Tr}\,(z\,|_{A_m^{(0)}}) + \mathrm{Tr}\,(z\,|_{A_m^{(1)}}) = \dim\ A_m^{(1)}\,.$$

Clearly, $A_m^G \subseteq A_m^{(1)}$ Conversely, assume that $v \in A_m^{(1)}$, then $z \cdot v = v$ and $\sigma \cdot v = \sigma \cdot (z \cdot v) = (\sigma z) \cdot v = z \cdot v = v$ for all $\sigma \in G$, therefore $v \in A_m^G$. Hence $A_m^G = A_m^{(1)}$ and $\mathrm{Tr}(z\,|_{A_m}) = \dim\ A_m^{(1)} = \dim\ A_m^G = d_m(G)$. □

Theorem 13.6. (Molien) *Let G be a finite subgroup of* $\mathrm{GL}_n(\mathbb{C})$. *Then*

$$\Phi_G(\lambda) = \frac{1}{|G|} \sum_{\sigma \in G} \frac{1}{\det\,(1 - \sigma\lambda)}$$

Proof. By Lemmas 13.4 and 13.5

$$\begin{aligned}\frac{1}{|G|} \sum_{\sigma \in G} \frac{1}{\det\,(1 - \sigma\lambda)} &= \frac{1}{|G|} \sum_{\sigma \in G} \sum_{m=0}^{\infty} \mathrm{Tr}\,(\sigma\,|_{A_m})\,\lambda^m \\ &= \sum_{m=0}^{\infty} \left(\frac{1}{|G|} \sum_{\sigma \in G} \mathrm{Tr}\,(\sigma\,|_{A_m}) \right) \lambda^m \\ &= \sum_{m=0}^{\infty} \mathrm{Tr}\,(z\,|_{A_m})\,\lambda^m\end{aligned}$$

$$= \sum_{m=0}^{\infty} d_m(G)\,\lambda^m$$
$$= \Phi_G(\lambda)\,. \qquad \square$$

Molien's theorem helps us to compute the Molien series of A^G for any finite subgroup G of $\mathrm{GL}_n(\mathbb{C})$ and the latter can be used to determine whether a set of G-invariant polynomials generates $\mathbb{C}[X_1, \ldots, X_n]^G$, as the following examples show.

Example 13.1. Let $G \subset \mathrm{GL}_2(\mathbb{C})$ be the group consisting of the following four elements:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Denote them by 1, -1, σ, $-\sigma$, respectively. We compute

$$\det(I - 1t) = (1-t)^2\,,$$
$$\det(I - (-1)t) = (1+t)^2\,,$$
$$\det(I - \sigma t) = \det(I - (-\sigma)\,t) = (1-t)\,(1+t)\,.$$

By Molien's theorem, the Molien's series of $\mathbb{C}[X_1, X_2]^G$ is

$$\Phi_G(t) = \frac{1}{4}\left(\frac{1}{(1-t)^2} + \frac{1}{(1+t)^2} + \frac{2}{(1-t)\,(1+t)}\right)$$
$$= \frac{1}{(1-t^2)^2}$$
$$= \sum_{k=0}^{\infty} (k+1)\,t^{2k}\,.$$

It follows that $\dim \mathbb{C}[X_1, X_2]^G_{2k} = k+1$ and $\dim \mathbb{C}[X_1, X_2]^G_{2k+1} = 0$ for any non-negative integer k. Clearly, $f_1 = X_1^2 + X_2^2$ and $f_2 = X_1^2 - X_2^2$ are G-invariant homogeneous polynomials of degree 2 and they are linearly independent. Hence f_1 and f_2 form a basis of $\mathbb{C}[X_1, X_2]^G_2$. It is easy to verify that $f_1^k, f_1^{k-1} f_2, \ldots, f_1 f_2^{k-1}, f_2^k$ form a basis of $\mathbb{C}[X_1, X_2]^G_{2k}$. Therefore $\mathbb{C}[X_1, X_2]^G = \mathbb{C}[f_1, f_2]$. $\square$

Example 13.2. Let

$$\sigma = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

which is an element of order 4, and G be the cyclic group generated by σ, i.e.,

$$G = \{1, \sigma, \sigma^2, \sigma^3\}.$$

We have

$$\begin{aligned} \det(I - 1t) &= (1-t)^2, \\ \det(I - \sigma t) &= \det(I - \sigma^3 t) = 1 + t^2, \\ \det(I - \sigma^2 t) &= (1+t)^2. \end{aligned}$$

By Molien's theorem, the Molien series of $\mathbb{C}[X_1, X_2]^G$ is

$$\begin{aligned} \Phi_G(t) &= \frac{1}{4}\left(\frac{1}{(1-t)^2} + \frac{2}{1+t^2} + \frac{1}{(1+t)^2}\right) \\ &= \frac{1+t^4}{(1-t^2)(1-t^4)} \\ &= \sum_{k=0}^{\infty} (2k+1)(t^{4k} + t^{4k+2}). \end{aligned}$$

Thus

$$\begin{aligned} \dim \mathbb{C}[X_1, X_2]^G_{4k} &= \dim \mathbb{C}[X_1, X_2]^G_{4k+2} = 2k+1 \\ \dim \mathbb{C}[X_1, X_2]^G_{4k+1} &= \dim \mathbb{C}[X_1, X_2]^G_{4k+3} = 0 \end{aligned} \tag{13.1}$$

for any non-negative integer k.

Clearly, $f_1 = X_1^2 + X_2^2$, $f_2 = X_1^2 X_2^2$ and $f_3 = X_1^3 X_2 - X_1 X_2^3$ are G-invariant homogeneous polynomials of degrees 2, 2 and 4 respectively. For any $p(Y_1, Y_2)$, $q(Y_1, Y_2) \in \mathbb{C}[Y_1, Y_2]$, $p(f_1, f_2) + f_3 q(f_1, f_2)$ is also G-invariant. Furthermore, using (13.1) we can show that any G-invariant polynomial can be expressed uniquely in this form. Therefore $\mathbb{C}[X_1, X_2]^G = \mathbb{C}[f_1, f_2, f_3]$. But f_1, f_2 and f_3 are not algebraically independent over $\mathbb{C}$. In fact, we have $f_1^2 f_2 - 4f_2^2 - f_3^2 = 0$, which is called a *syzygy* relating f_1, f_2 and f_3. □

13.3. Hilbert's Finite Generation Theorem

An algebra D over $\mathbb{C}$ is said to be *finitely generated* if there are finitely many elements $f_1, \ldots, f_m$ in D such that $D = \mathbb{C}[f_1, \ldots, f_m]$. For example, the algebras of G-invariant polynomials $\mathbb{C}[X_1, X_2]^G$ for the finite subgroups G of $\mathrm{GL}_2(\mathbb{C})$ considered in Examples 13.1 and 13.2 are finitely generated. In fact, for the subgroup G considered in Example 13.1 we have $\mathbb{C}[X_1, X_2]^G$

$= \mathbb{C}[f_1, f_2]$, where $f_1 = X_1^2 + X_2^2$ and $f_2 = X_1^2 - X_2^2$, and for the subgroup G considered in Example 13.2 we have $\mathbb{C}[X_1, X_2]^G = \mathbb{C}[f_1, f_2, f_3]$, where $f_1 = X_1^2 + X_2^2$, $f_2 = X_1^2 X_2^2$, and $f_3 = X_1^3 X_2 - X_1 X_2^3$.

More generally, Hilbert proved the following famous finite generation theorem of the algebra of G-invariant polynomials for any finite subgroup G of $\mathrm{GL}_n(\mathbb{C})$.

Theorem 13.7. (Hilbert) *Let G be any finite subgroup of $\mathrm{GL}_n(\mathbb{C})$. Then these are finitely many G-invariant polynomials $f_1, \ldots, f_m$, say, such that $\mathbb{C}[X_1, \ldots, X_n]^G = \mathbb{C}[f_1, \ldots, f_m]$.*

To prove Theorem 13.7 we need the following Hilbert's basis theorem.

Theorem 13.8. (Hilbert) *Let $X_1, \ldots, X_n$ be n indeterminates over $\mathbb{C}$. Then every ideal I of $\mathbb{C}[X_1, \ldots, X_n]$ has a finite basis, i.e., there are finitely many polynomials $f_1, \ldots, f_m$ in I such that*

$$I = (f_1, \ldots, f_m) = \{g_1 f_1 + \cdots + g_m f_m \mid g_i \in \mathbb{C}[X_1, \ldots, X_n]\}.$$

Proof. Apply induction on n. For $n = 1$, it is well known that $\mathbb{C}[X_1]$ is a principal ideal domain, i.e., for any ideal I of $\mathbb{C}[X_1]$ there is a polynomial f in I such that $I = (f)$. Therefore our theorem is true for $n = 1$.

Assume that our theorem is true for $n-1$. That is, every ideal of $\mathbb{C}[X_1, \ldots, X_{n-1}]$ has a finite basis. Write $B = \mathbb{C}[X_1, \ldots, X_{n-1}]$, then $\mathbb{C}[X_1, \ldots, X_n] = B[X_n]$. Let I be any ideal of $B[X_n]$. For any $f \in I$ we can write $f = a_0 + a_1 X_n + \cdots + a_l X_n^l$, where $a_0, \ldots, a_l \in B$ and $a_l \neq 0$. We call a_l the leading coefficient of f. Denote by I_0 the set of the leading coefficients of polynomials in I. It is clear that I_0 is an ideal of B. By induction hypothesis, there are elements $a_1, \ldots, a_m \in I$ such that $I_0 = (a_1, \ldots, a_m)$. Let $f_1, \ldots, f_m$ be polynomials in I whose leading coefficients are $a_1, \ldots, a_m$, respectively. Let $d = \max\{\deg f_1, \ldots, \deg f_n\}$. For any $f \in I$ and $\deg f \geq d$, since the leading coefficient of f belongs to I_0, subtracting a linear combination of $f_1, \ldots, f_m$ with coefficients in B from f we obtain a polynomial in I whose degree is lower than f. Continuing in this way we obtain finally a polynomial in I whose degree is $< d$.

For any i $(0 \leq i \leq d-1)$ denote by I_i the set of leading coefficients of polynomials of degree i in I. It is clear that all I_i are ideals of B. By induction hypothesis there are elements $a_{i1}, \ldots, a_{im} \in I_i$ such that $I_i = (a_{i1}, \ldots, a_{im_i})$. Let

$f_{i1}, \ldots, f_{im_i}$ be polynomials in I whose leading coefficients are $a_{i1}, \ldots, a_{im_i}$, respectively. Clearly we have

$$I = (f_1, \ldots, f_m, f_{d-1,1}, \ldots, f_{d-1,m_{d-1}}, \ldots, f_{01}, \ldots, f_{0m_0}). \qquad \square$$

Proof of Theorem 13.7. By Proposition 13.3, $\mathbb{C}[X_1, \ldots, X_n]^G$ is homogeneous, i.e.,

$$\begin{aligned}\mathbb{C}[X_1, \ldots, X_n]^G = \mathbb{C}[X_1, \ldots, X_n]_0^G \dotplus [X_1, \ldots, X_n]_1^G \\ \dotplus \mathbb{C}[X_1, \ldots, X_n]_2^G \dotplus \cdots .\end{aligned}$$

Let

$$\mathbb{C}[X_1, \ldots, X_n]_+^G = \mathbb{C}[X_1, \ldots, X_n]_1^G \dotplus \mathbb{C}[X_1, \ldots, X_n]_2^G \dotplus \cdots$$

Denote by I the ideal of $\mathbb{C}[X_1, \ldots, X_n]$ generated by $\mathbb{C}[X_1, \ldots, X_n]_+^G$. By Theorem 13.8, I has a finite basis, i.e., there are polynomials $f_1, \ldots, f_m$ in I such that $I = (f_1, \ldots, f_m)$. Without loss of generality, we can assume that all $f_1, \ldots, f_m$ are G-invariant homogeneous polynomials of degrees ≥ 1.

Let $f \in \mathbb{C}[X_1, \ldots, X_n]_d^G$. We apply induction on d to show that f can be expressed as polynomials in $f_1, \ldots, f_m$. When $d = 0$, this is trivial. Now assume that $d > 0$. We can express f as

$$f = h_1 f_1 + \cdots + h_m f_m, \qquad h_1, \ldots, h_m \in \mathbb{C}[X_1, \ldots, X_n]. \tag{13.2}$$

Canceling all the terms in $h_1 f_1, \ldots, h_m f_m$ which are of degrees $\neq d$, we can assume that all $h_1, \ldots, h_m$ are homogeneous. Applying the averaging operator

$$z = \frac{1}{|G|} \sum_{\sigma \in G} \sigma$$

to both sides of (13.2), we obtain

$$f = \tilde{h}_1 f_1 + \cdots + \tilde{h}_m f_m,$$

where $\tilde{h}_i = z \cdot h_i$ $(i = 1, \ldots, m)$ are G-invariant homogeneous polynomials with $\deg \tilde{h}_i = \deg f - \deg f_i < d$. By induction hypothesis, $\tilde{h}_1, \ldots, \tilde{h}_m$ can be expressed as polynomials in $f_1, \ldots, f_m$, so is f. $\square$

Moreover, we have

Proposition 13.9. *Let G be a finite subgroup of $\mathrm{GL}_n(\mathbb{C})$ and $f_1, \ldots, f_m$ be G-invariant polynomials such that $\mathbb{C}[X_1, \ldots, X_n]^G = \mathbb{C}[f_1, \ldots, f_m]$. Then $m \geq$*

n. In particular, if $m = n$ then $f_1, \ldots, f_n$ are algebraically independent over $\mathbb{C}$ and if $m > n$ then there are some polynomial relations among $f_1, \ldots, f_m$, (which are called syzygies relating $f_1, \ldots, f_m$).

Proof. For any $h \in \mathbb{C}(X_1, \ldots, X_n)$, let $h = \frac{f}{g}$ where $f, g \in \mathbb{C}[X_1, \ldots, X_n]$. Define

$$\sigma \cdot h = \frac{\sigma \cdot f}{\sigma \cdot h}.$$

It is easy to verify that this definition is well-defined, i.e., independent of the representation of h as a quotient of two polynomials. It is also easy to verify that σ is an automorphism of the field $\mathbb{C}(X_1, \ldots, X_n)$.

Let

$$\mathbb{C}(X_1, \ldots, X_n)^G = \{h \in \mathbb{C}(X_1, \ldots, X_n) \mid \sigma \cdot h = h\}.$$

Then $\mathbb{C}(X_1, \ldots, X_n)^G$ is a subfield of $\mathbb{C}(X_1, \ldots, X_n)$, called the *fixed field* of G. Clearly, $\mathbb{C}(f_1, \ldots, f_m) \subset \mathbb{C}(X_1, \ldots, X_n)^G$. Conversely, for any $h \in \mathbb{C}(X_1, \ldots, X_n)^G$, let $h = \frac{f}{g}$, where $f, g \in \mathbb{C}[X_1, \ldots, X_n]$, then

$$\frac{f}{g} = \left(f \prod_{\substack{\sigma \in G \\ \sigma \neq 1}} \sigma \cdot g \right) \Bigg/ \left(\prod_{\sigma \in G} \sigma \cdot g \right).$$

Since the left-hand side of the above equation belongs to $\mathbb{C}(X_1, \ldots, X_n)^G$ and the denominator of the right-hand side belongs to $\mathbb{C}[X_1, \ldots, X_n]^G$, the numerator of the right-hand side also belongs to $\mathbb{C}[X_1, \ldots, X_n]^G$. We assumed $\mathbb{C}[X_1, \ldots, X_n]^G = \mathbb{C}[f_1, \ldots, f_m]$. Therefore $h \in \mathbb{C}(f_1, \ldots, f_m)$. Hence $\mathbb{C}(X_1, \ldots, X_n)^G = \mathbb{C}(f_1, \ldots, f_m)$.

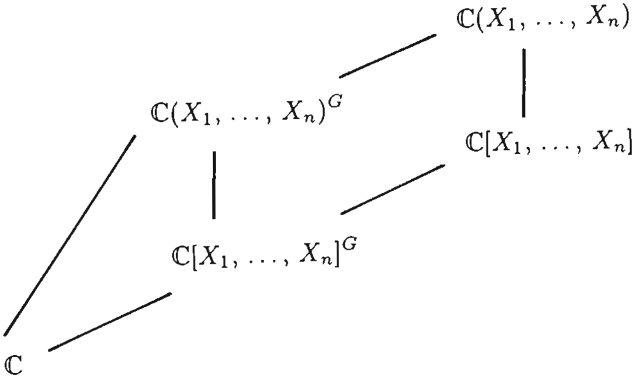

Fig. 13.1

For any $h \in \mathbb{C}(X_1, \ldots, X_n)$, h satisfies the polynomial

$$\prod_{\sigma \in G} (Y - \sigma \cdot h)$$

with coefficients in $\mathbb{C}(X_1, \ldots, X_n)^G$ Therefore $\mathbb{C}(X_1, \ldots, X_n)$ is algebraic over $\mathbb{C}(X_1, \ldots, X_n)^G$. The transcendental degree of $\mathbb{C}(X_1, \ldots, X_n)$ over $\mathbb{C}$ is n, so is that of $\mathbb{C}(X_1, \ldots, X_n)^G$ over $\mathbb{C}$. But $\mathbb{C}(X_1, \ldots, X_n)^G = \mathbb{C}(f_1, \ldots, f_m)$. Hence $m \geq n$. $\square$

For example, in Example 13.1 we have $m = n = 2$ and in Example 13.2 we have $m = 3 > n = 2$ and $f_1^2 f_2 - 4f_2^2 - f_3^2 = 0$ is a *syzygy*.

CHAPTER 14

SELF-DUAL QUATERNARY CODES AND THEIR WEIGHT ENUMERATORS

14.1. Examples of Self-dual Quaternary Codes

Recall that a $\mathbb{Z}_4$-linear code $\mathcal{C}$ is called self-dual if $\mathcal{C}^\perp = \mathcal{C}$.

Example 14.1. Among the three $\mathbb{Z}_4$-linear codes of length 1 listed in Sec. 1.1 only the code $\{(0),(2)\}$ is self-dual. Denote

$$\mathcal{A}_1 = \{(0),(2)\}.$$

The complete weight enumerator and the symmetrized weight enumerator of $\mathcal{A}_1$ are the same.

$$\begin{aligned}\mathrm{cwe}_{\mathcal{A}_1}(X_0, X_1, X_2, X_3) &= \mathrm{swe}_{\mathcal{A}_1}(X_0, X_1, X_2)\\ &= X_0 + X_2.\end{aligned}$$ □

It is easy to prove that there is no self-dual code of length 2 and 3.

Example 14.2. Consider the $\mathbb{Z}_4$-linear code $\mathcal{K}_4$ with generator matrix (1.3)

$$\begin{pmatrix} 1 & 1 & 1 & 1\\ 0 & 2 & 0 & 2\\ 0 & 0 & 2 & 2\end{pmatrix}$$

introduced in Example 1.1. We know that $\mathcal{K}_4$ is of type $4^1 2^2$ and is self-dual. From Examples 2.2 and 2.5, we have

$$\mathrm{cwe}_{\mathcal{K}_4}(X_0, X_1, X_2, X_3) = X_0^4 + X_1^4 + X_2^4 + X_3^4 + 6X_0^2X_2^2 + 6X_1^2X_3^2$$

and

$$\mathrm{swe}_{\mathcal{K}_4}(X_0, X_1, X_2) = X_0^4 + 8X_1^4 + X_2^4 + 6X_0^2X_2^2.$$

Denote the $\mathbb{Z}_4$-linear code with generator matrix

$$\begin{pmatrix} 1 & 3 & 3 & 3 \\ 0 & 2 & 0 & 2 \\ 0 & 0 & 2 & 2 \end{pmatrix}$$

by $\mathcal{K}_4'$. $\mathcal{K}_4'$ is also of type $4^1 2^2$ and is self-dual. We have

$$\mathrm{cwe}_{\mathcal{K}_4'}(X_0, X_1, X_2, X_3) = X_0^4 + X_2^4 + 6X_0^2X_2^2 + 4(X_1^3X_3 + X_1X_3^3)$$

and

$$\mathrm{swe}_{\mathcal{K}_4'}(X_0, X_1, X_2) = X_0^4 + 8X_1^4 + X_2^2 + 6X_0^2X_2^2.$$

$\mathcal{K}_4$ and $\mathcal{K}_4'$ have the same symmetrized weight enumerator but different complete weight enumerators. Actually they are equivalent but not permutation-equivalent. □

Example 14.3. We have already met four self-dual codes of length 8; they are $\mathcal{O}_8$ (see Example 1.3), $\mathcal{K}_8$ (see Example 1.4), $\mathcal{K}_8'$ and $\mathcal{Q}_8$ (see Example 12.7). $\mathcal{O}_8$ and $\mathcal{K}_8$ have generator matrices

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 3 & 1 & 2 & 1 \\ 0 & 1 & 0 & 0 & 1 & 2 & 3 & 1 \\ 0 & 0 & 1 & 0 & 3 & 3 & 3 & 2 \\ 0 & 0 & 0 & 1 & 2 & 3 & 1 & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 \end{pmatrix}$$

respectively. They are of type 4^4 and $4^1 2^6$, respectively. From Example 2.6 we have (2.11):

$$\begin{aligned}\mathrm{cwe}_{\mathcal{O}_8}(X_0, X_1, X_2, X_3) = {} & X_0^8 + X_1^8 + X_2^8 + X_3^8 + 14(X_0^4 X_2^4 + X_1^4 X_3^4) \\ & + 56(X_0^3 X_1^3 X_2 X_3 + X_0^3 X_1 X_2 X_3^3 \\ & + X_0 X_1^3 X_2^3 X_3 + X_0 X_1 X_2^3 X_3^3)\end{aligned}$$

and (2.12)

$$\mathrm{swe}_{\mathcal{O}_8}(X_0, X_1, X_2) = X_0^8 + 16X_1^8 + X_2^8 + 14X_0^4 X_2^4 + 112X_0 X_1^4 X_2(X_0^2 + X_2^2).$$

We can compute

$$\mathrm{cwe}_{\mathcal{K}_8}(X_0, X_1, X_2, X_3) = \frac{1}{2}((X_0+X_2)^8+(X_1+X_3)^8+(X_0-X_2)^8+(X_1-X_3)^8),$$

and

$$\mathrm{swe}_{\mathcal{K}_8}(X_0, X_1, X_2) = \frac{1}{2}((X_0 + X_2)^8 + (X_0 - X_2)^8 + (2X_1)^8).$$

$\mathcal{K}_8'$ and $\mathcal{Q}_8$ have generator matrices

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 & 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 \end{pmatrix}$$

and

$$\begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 2 & 1 & 3 \\ 0 & 0 & 0 & 2 & 1 & 3 & 1 & 1 \\ 1 & 1 & 0 & 2 & 0 & 0 & 1 & 3 \\ 0 & 2 & 0 & 2 & 0 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 \end{pmatrix}$$

respectively. $\mathcal{K}_8'$ is of type $4^2 2^4$ and $\mathcal{Q}_8$ is of type $4^3 2^2$. We have

$$\begin{aligned}\mathrm{swe}_{\mathcal{K}_8'}(X_0, X_1, X_2) = {} & X_0^8 + 64X_1^8 + X_2^8 + 12X_0^2 X_2^2(X_0^4 + X_2^4) + 38X_0^4 X_2^4 \\ & + 64X_0 X_1^4 X_2(X_0^2 + X_2^2)\end{aligned}$$

and

$$\mathrm{swe}_{Q_8}(X_0, X_1, X_2) = X_0^8 + 32X_1^8 + X_2^8 + 4X_0^2X_2^2(X_0^4 + X_2^4) + 22X_0^4X_2^4 + 96X_0X_1^4X_2(X_0^2 + X_2^2). \qquad \square$$

Example 14.4. The self-dual $\mathbb{Z}_4$-codes $\mathcal{K}_4$ and $\mathcal{K}_8$ can be generalized to $\mathcal{K}_{4m}(m \geq 1)$. It has generator matrix

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 0 & 2 & 0 & \cdots & 0 & 2 \\ 0 & 0 & 2 & \cdots & 0 & 2 \\ . & . & . & \cdots & . & . \\ 0 & 0 & 0 & \cdots & 2 & 2 \end{pmatrix},$$

which is a $(4m-1) \times 4m$ matrix. $\mathcal{K}_{4m}$ was introduced by Klemm (1989). We have

$$\mathrm{cwe}_{\mathcal{K}_{4m}}(X_0, X_1, X_2, X_3) = \frac{1}{2}((X_0 + X_2)^{4m} + (X_1 + X_3)^{4m} + (X_0 - X_2)^{4m} + (X_1 - X_3)^{4m})$$

and

$$\mathrm{swe}_{\mathcal{K}_{4m}}(X_0, X_1, X_2) = \frac{1}{2}((X_0 + X_2)^{4m} + (X_0 - X_2)^{4m}) + 2^{4m-1}X_1^{4m}. \qquad \square$$

Example 14.5. $\mathcal{C}_{10}$ is the self-dual $\mathbb{Z}_4$-code of length 10 with generator matrix

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 3 & 3 & 1 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 2 \end{pmatrix}$$

and type $4^2 2^6$, see Conway and Sloane (1993a). $\square$

Example 14.6. We have the self-dual code $\mathcal{C}_{16}$ with generator matrix

$$\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 3 & 3 & 1 & 0 & 3 & 2 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 3 & 3 & 1 & 1 & 2 & 3 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 2 & 2 & 0 & 3 & 3 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 3 & 0 & 2 & 3 & 3 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 3 & 2 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 2 & 0 & 2 & 0 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 2 & 2 & 2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 2 & 0 & 0
\end{pmatrix}.$$

$\mathcal{C}_{16}$ was introduced by Conway and Sloane (1993a). □

14.2. Complete Weight Enumerators of Self-dual $\mathbb{Z}_4$-Codes

Let $\mathcal{C}$ be a self-dual $\mathbb{Z}_4$-code of length n. By Corollary 1.3, $|\mathcal{C}| = 2^n$ and the MacWilliams identity in Theorem 2.2. becomes

$$\begin{aligned}
&\mathrm{cwe}_{\mathcal{C}}(X_0, X_1 X_2, X_3) \\
&= \frac{1}{2^n}\,\mathrm{cwe}_{\mathcal{C}}(X_0 + X_1 + X_2 + X_3,\ X_0 + iX_1 - X_2 - iX_3, \\
&\quad X_0 - X_1 + X_2 - X_3,\ X_0 - iX_1 - X_2 + iX_3). \\
&= \mathrm{cwe}_{\mathcal{C}}\left(\frac{1}{2}(X_0 + X_1 + X_2 + X_3),\ \frac{1}{2}(X_0 + iX_1 - X_2 - iX_3),\right. \\
&\quad \left.\frac{1}{2}(X_0 - X_1 + X_2 - X_3),\ \frac{1}{2}(X_0 - iX_1 - X_2 + iX_3)\right).
\end{aligned}$$

This means that $\mathrm{cwe}_{\mathcal{C}}$ is invariant under the linear transformation

$$\mu : (X_0, X_1, X_2, X_3) \to (X_0, X_1, X_2, X_3)\,{}^t\mu,$$

where

$$\mu = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{pmatrix}, \tag{14.1}$$

i.e., $\mu \cdot \mathrm{cwe}_{\mathcal{C}} = \mathrm{cwe}_{\mathcal{C}}$.

For all $\mathbf{c} \in C$, $\mathbf{c} \cdot \mathbf{c} = 0$, from which follows:

$$w_1(\mathbf{c}) + w_3(\mathbf{c}) \equiv 0 \pmod 4. \tag{14.2}$$

Therefore cwe_C is also invariant under

$$\delta = \mathrm{diag}\{1, i, 1, i\}. \tag{14.3}$$

We also have

$$w_a(-\mathbf{c}) = w_{-a}(\mathbf{c}) \quad \text{for all} \quad \mathbf{c} \in C \quad \text{and} \quad a \in \mathbb{Z}_4.$$

If $-\mathbf{c} = \mathbf{c}$ for some $\mathbf{c} \in C$, then $w_a(\mathbf{c}) = w_{-a}(\mathbf{c})$ and, hence,

$$X_0^{w_0(\mathbf{c})} X_1^{w_1(\mathbf{c})} X_2^{w_2(\mathbf{c})} X_3^{w_3(\mathbf{c})}$$

is invariant under

$$\pi = \begin{pmatrix} 1 & & & \\ & 0 & & 1 \\ & & 1 & \\ & 1 & & 0 \end{pmatrix} \tag{14.4}$$

If $-\mathbf{c} \neq \mathbf{c}$ for some $\mathbf{c} \in C$, then $-\mathbf{c} \in C$ and

$$X_0^{w_0(\mathbf{c})} X_1^{w_1(\mathbf{c})} X_2^{w_2(\mathbf{c})} X_3^{w_3(\mathbf{c})} + X_0^{w_0(-\mathbf{c})} X_1^{w_1(-\mathbf{c})} X_2^{w_2(-\mathbf{c})} X_3^{w_3(-\mathbf{c})}$$

is also invariant under π.

Let

$$G = \langle \mu, \delta, \pi \rangle. \tag{14.5}$$

We have proved

Proposition 14.1. *For any self-dual $\mathbb{Z}_4$-code C of length n, let $\mathrm{cwe}_C(X_0, X_1, X_2, X_3)$ be its complete weight enumerator. Then*

$$\mathrm{cwe}_C(X_0, X_1, X_2, X_3) \in \mathbb{C}[X_0, X_1, X_2, X_3]_n^G,$$

where G is defined by (14.5). □

It will be helpful to determine $\mathbb{C}[X_0, X_1, X_2, X_3]^G$.

For any polynomial $f(X_0, X_1, X_2, X_3) \in \mathbb{C}[X_0, X_1, X_2, X_3]$, define

$$f^*(X_0, X_1, X_2, X_3) = f\left(\frac{X_0 + X_2}{\sqrt{2}}, \frac{X_1 + X_3}{\sqrt{2}}, \frac{X_0 - X_2}{\sqrt{2}}, \frac{X_1 - X_3}{\sqrt{2}}\right).$$

Let

$$Z_0 = \frac{X_0 + X_2}{\sqrt{2}}\,,\; Z_1 = \frac{X_1 + X_3}{\sqrt{2}}\,,\; Z_2 = \frac{X_0 - X_2}{\sqrt{2}}\,,\; Z_3 = \frac{X_1 - X_3}{\sqrt{2}}\,,$$

then

$$f^*(X_0, X_1, X_2, X_3) = f(Z_0, Z_1, Z_2, Z_3).$$

Let $\mathrm{cwe}_{\mathcal{C}}(X_0, X_1, X_2, X_3)$ be the complete weight enumerator of a $\mathbb{Z}_4$-code $\mathcal{C}$, both $\mathrm{cwe}^*_{\mathcal{C}}(X_0, X_1, X_2, X_3)$ and $\mathrm{cwe}_{\mathcal{C}}(Z_0, Z_1, Z_2, Z_3)$ are called the *transformed complete weight enumerator* of $\mathcal{C}$.

Let

$$\rho = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & & 1 & \\ & 1 & & 1 \\ 1 & & -1 & \\ & 1 & & -1 \end{pmatrix}. \tag{14.6}$$

For any finite subgroup H of $GL_n(\mathbb{C})$, define

$$H^* = \rho H \rho^{-1}.$$

Lemma 14.2. *Let $f(X_0, X_1, X_2, X_3)$ be any polynomial in $\mathbb{C}[X_0, X_1, X_2, X_3]$ and H be any finite subgroup of $GL_n(\mathbb{C})$. Then*

(i) $\rho \cdot f = f^*$.

(ii) $(f^*)^* = f$.

(iii) $f \in \mathbb{C}[X_0, X_1, X_2, X_3]^H \Leftrightarrow f^* \in \mathbb{C}[X_0, X_1, X_2, X_3]^{H^*}$
$\qquad\Leftrightarrow f(Z_0, Z_1, Z_2, Z_3) \in \mathbb{C}[Z_0, Z_1, Z_2, Z_3]^{H^*}$.

(iv) $\Phi_H(\lambda) = \Phi_{H^*}(\lambda)$.

Proof. (i) is clear from the definition of f^*.

(ii) follows from $\rho^2 = 1$.

For (iii), we have

$$\begin{aligned} f \in \mathbb{C}[X_0, X_1, X_2, X_3]^H &\Leftrightarrow \sigma \cdot f = f \quad \text{for all} \quad \sigma \in H \\ &\Leftrightarrow \rho\sigma\rho^{-1} \cdot (\rho \cdot f) = \rho \cdot f \quad \text{for all} \quad \sigma \in H \\ &\Leftrightarrow \rho\sigma\rho^{-1} \cdot f^* = f^* \quad \text{for all} \quad \sigma \in H \\ &\Leftrightarrow f^* \in \mathbb{C}[X_0, X_1, X_2, X_3]^{H^*}. \end{aligned}$$

(iv) is a consequence of (iii). $\square$

It follows from Lemma 14.2 that to determine $\mathbb{C}[X_1, X_2, X_3, X_4]^H$ is equivalent to determine $\mathbb{C}[Z_0, Z_1, Z_2, Z_3]^{H^*}$.

Now let us return to the group G defined by (14.5). Let $G^* = \rho G \rho^{-1}$. Let $\mu^* = \rho\mu\rho^{-1}$, $\delta^* = \rho\delta\rho^{-1}$, and $\pi^* = \rho\pi\rho^{-1}$, then

$$G^* = \langle \mu^*, \delta^*, \pi^* \rangle.$$

We have

$$\mu^* = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & i \end{pmatrix},$$

$$\delta^* = \operatorname{diag}\{1, i, 1, i\} = \delta,$$

$$\pi^* = \operatorname{diag}\{1, 1, 1, -1\}$$

and

$$\mu^* \delta^* {\mu^*}^{-1} = \operatorname{diag}\{1, 1, i, i\}.$$

Let

$$N = \langle \operatorname{diag}\{1, i, 1, i\},\ \operatorname{diag}\{1, 1, i, i\},\ \operatorname{diag}\{1, 1, 1, -1\} \rangle.$$

Then N is an abelian group of order 32,

$$G^* = \langle N, \mu^* \rangle,\ N \triangleleft G^*,\ [G^* : N] = 2 \quad \text{and} \quad |G^*| = 64.$$

Theorem 14.3. *Let C be a self-dual $\mathbb{Z}_4$-code and $cwe_C(Z_0, Z_1, Z_2, Z_3)$ be its transformed complete weight enumerator. Then*

(i) $\mathrm{cwe}_C(Z_0, Z_1, Z_2, Z_3) \in \mathbb{C}[Z_0, Z_1, Z_2, Z_3]_n^{G^*}$.
(ii) *The Molien series of* $\mathbb{C}[Z_0, Z_1, Z_2, Z_3]^{G^*}$ *is*

$$\Phi_{G^*}(\lambda) = (1 + \lambda^{10})/(1-\lambda)(1-\lambda^4)^2(1-\lambda^8).$$

(iii) $\mathbb{C}[Z_0, Z_1, Z_2, Z_3]^{G^*} = \mathbb{C}[\theta_1, \theta_{4a}, \theta_{4b}, \theta_8] + \theta_{10}\mathbb{C}[\theta_1, \theta_{4a}, \theta_{4b}, \theta_8]$, *where*

$$\theta_1 = Z_0,\ \theta_{4a} = Z_1^4 + Z_2^4,\ \theta_{4b} = Z_3^4,\ \theta_8 = Z_1^4 Z_2^4,$$

$$\theta_{10} = Z_1^6 Z_2^2 Z_3^2 - Z_1^2 Z_2^6 Z_3^2.$$

Proof. (i) follows from Proposition 14.1 and Lemma 14.2 (iii).

(ii) First let us determined the explicit structure of $\mathbb{C}[Z_0, Z_1, Z_2, Z_3]^N$ Let $f(Z_0, Z_1, Z_2, Z_3) \in \mathbb{C}[Z_0, Z_1, Z_2, Z_3]^N$. Assume that $Z_0^a Z_1^b Z_2^c Z_3^d$ appears in f with a nonzero coefficient. Clearly

$$\begin{aligned} \delta^* \cdot Z_0^a Z_1^b Z_2^c Z_3^d &= i^{b+d} Z_0^a Z_1^b Z_2^c Z_3^d, \\ (\mu^* \delta^* \mu^{*-1}) \cdot Z_0^a Z_1^b Z_2^c Z_3^d &= i^{c+d} Z_0^a Z_1^b Z_2^c Z_3^d, \\ \pi^* \cdot Z_0^a Z_1^b Z_2^c Z_3^d &= (-1)^d Z_0^a Z_1^b Z_2^c Z_3^d. \end{aligned}$$

Since f is N-invariant, we have

$$\begin{aligned} b + d &\equiv 0 \pmod 4, \\ c + d &\equiv 0 \pmod 4, \\ d &\equiv 0 \pmod 2). \end{aligned}$$

There are two possibilities

1° $d \not\equiv 0 \pmod 4$. Then $b \equiv c \equiv d \equiv 0 \pmod 2$, $b \not\equiv 0 \pmod 4$, $c \not\equiv 0 \pmod 4$.

2° $d \equiv 0 \pmod 4$. Then $b \equiv c \equiv 0 \pmod 4$.

Therefore

$$\mathbb{C}[Z_0, Z_1, Z_2, Z_3]^N = \mathbb{C}[Z_0, Z_1^4, Z_2^4, Z_3^4] + Z_1^2 Z_2^2 Z_3^2 \mathbb{C}[Z_0, Z_1^4, Z_2^4, Z_3^4].$$

By Corollary 13.2

$$\Phi_N(\lambda) = (1 + \lambda^6)/(1 - \lambda)(1 - \lambda^4)^3. \tag{14.7}$$

Now let us compute $\Phi_{G^*}(\lambda)$. We have $G^* = N \cup N\mu^*$ By Theorem 13.5,

$$\begin{aligned} \Phi_{G^*}(\lambda) &= \frac{1}{|G^*|} \sum_{\sigma \in G^*} \frac{1}{1 - \sigma\lambda} \\ &= \frac{1}{2|N|} \sum_{\sigma \in N} \frac{1}{1 - \sigma\lambda} + \frac{1}{|G^*|} \sum_{\sigma \in N\mu^*} \frac{1}{1 - \sigma\lambda} \\ &= \frac{1}{2} \Phi_N(\lambda) + \frac{1}{|G^*|} \sum_{\sigma \in N\mu^*} \frac{1}{1 - \sigma\lambda}. \end{aligned} \tag{14.8}$$

By routine computation, we have

$$\frac{1}{|G^*|} \sum_{\sigma \in N\mu^*} \frac{1}{1 - \sigma\lambda} = \frac{1}{2} \left(\frac{1 + \lambda^2 + \lambda^4}{(1 - \lambda)(1 + \lambda^2)(1 - \lambda^8)} \right). \tag{14.9}$$

Substituting (14.7) and (14.9) into (14.8) and simplifying, we obtain

$$\Phi_{G^*}(\lambda) = \frac{1+\lambda^{10}}{(1-\lambda)(1-\lambda^4)^2(1-\lambda^8)}.$$

(iii) It is easy to verify that $\theta_1 = Z_0$, $\theta_{4a} = Z_1^4 + Z_2^4$, $\theta_{4b} = Z_3^4$, $\theta_8 = Z_1^4 Z_2^4$, and $\theta_{10} = Z_1^6 Z_2^2 Z_3^2 - Z_1^2 Z_2^6 Z_3^2$ are all invariant under G^*. Clearly, $\theta_1, \theta_{4a}, \theta_{4b}$ and θ_8 are algebraically independent over $\mathbb{C}$, and $\theta_{10}^2 \in \mathbb{C}[\theta_1, \theta_{4a}, \theta_{4b}, \theta_8]$. Therefore,

$$\mathbb{C}[Z_0, Z_1, Z_2, Z_3]^{G^*} = \mathbb{C}[\theta_1, \theta_{4a}, \theta_{4b}, \theta_8] \dotplus \theta_{10}\mathbb{C}[\theta_1, \theta_{4a}, \theta_{4b}, \theta_8]. \qquad \square$$

Now let C be a self-dual $\mathbb{Z}_4$-code of length n and contain the all 1 codeword 1^n. We already know that $W_C(X_0, X_1, X_2, X_3)$ is invariant under

$$G = \langle \mu, \delta, \pi \rangle,$$

where μ, δ, π are defined by (14.1), (14.3), (14.4), respectively. Since $1^n \in C$ and C is self-dual, $1^n \cdot 1^n = 0$ and $1^n \cdot \mathbf{c} = \mathbf{c} \cdot \mathbf{c} = 0$ for all $\mathbf{c} \in C$. It follows that

$$w_0(\mathbf{c}) + w_1(\mathbf{c}) + w_2(\mathbf{c}) + w_3(c) \equiv 0 \pmod 4, \tag{14.10}$$

$$w_1(\mathbf{c}) + 2w_2(\mathbf{c}) - w_3(\mathbf{c}) \equiv 0 \pmod 4 \tag{14.11}$$

and (14.2)

$$w_1(\mathbf{c}) + w_3(\mathbf{c}) \equiv 0 \pmod 4.$$

From (14.2) and (14.10) we deduce

$$w_0(\mathbf{c}) + w_2(\mathbf{c}) \equiv 0 \pmod 4.$$

Therefore cwe_C is invariant under

$$\delta_1 = \mathrm{diag}\,\{i, 1, i, 1\}. \tag{14.12}$$

Adding (14.2) and (14.11) together, we obtain

$$w_1(\mathbf{c}) + w_2(\mathbf{c}) \equiv 0 \pmod 2.$$

Therefore cwe_C is invariant under

$$\delta_2 = \mathrm{diag}\,\{1, -1, -1, 1\}. \tag{14.13}$$

Since $\mathbf{c} + 1^n \in C$ for all $\mathbf{c} \in C$, we have

$$w_r(\mathbf{c} + 1^n) = w_{r-1}(\mathbf{c}) \quad \text{for all} \quad r \in \mathbb{Z}_4.$$

For all $\mathbf{c} \in \mathcal{C}$, $\mathbf{c} + 1^n, \mathbf{c} + 2(1^n)$, $\mathbf{c} + 3(1^n) \in \mathcal{C}$, thus all

$$X_0^{w_0(\mathbf{c})} X_1^{w_1(\mathbf{c})} X_2^{w_2(\mathbf{c})} X_3^{w_3(\mathbf{c})},\ X_0^{w_3(\mathbf{c})} X_1^{w_0(\mathbf{c})} X_2^{w_1(\mathbf{c})}, X_3^{w_2(\mathbf{c})},$$

$$X_0^{w_2(\mathbf{c})} X_1^{w_3(\mathbf{c})} X_2^{w_0(\mathbf{c})} X_3^{w_1(\mathbf{c})} \text{ and } X_0^{w_1(\mathbf{c})} X_1^{w_2(\mathbf{c})} X_2^{w_3(\mathbf{c})} X_3^{w_0(\mathbf{c})}$$

appear in $\mathrm{cwe}_{\mathcal{C}}$. Hence $\mathrm{cwe}_{\mathcal{C}}$ is invariant under

$$\xi = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

Let

$$G_1 = \langle G, \delta_1, \delta_2, \xi \rangle. \tag{14.14}$$

Therefore we have proved

Proposition 14.4. *For any self-dual $\mathbb{Z}_4$-code C of length n and containing 1^n, let $cwe_C(X_0, X_1, X_2, X_3)$ be its complete weight enumerator. Then*

$$\mathrm{cwe}_{\mathcal{C}}(X_0, X_1, X_2, X_3) \in \mathbb{C}[X_0, X_1, X_2, X_3]_n^{G_1},$$

where G_1 is defined by (14.14). □

Let $G_1^* = \rho G_1 \rho^{-1}$, where ρ is defined by (14.6). Then

$$G_1^* = \langle G^*, \delta_1^*, \delta_2^*, \xi^* \rangle,$$

where $\delta_1^* = \rho\delta_1\rho^{-1}$, $\delta_2^* = \rho\delta_2\rho^{-1}$, $\xi^* = \rho\xi\rho^{-1}$. We have

$$\delta_1^* = \operatorname{diag}\{i, 1, i, 1\} = \delta_1,$$

$$\delta_2^* = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix},$$

$$\xi^* = \begin{pmatrix} 0 & 1 & & \\ 1 & 0 & & \\ & & 0 & 1 \\ & & -1 & 0 \end{pmatrix}.$$

Let

$$N_1 = \langle N, \delta_1^* = \operatorname{diag}\{i, 1, i, 1\}\rangle.$$

Clearly, N_1 is an abelian group of order 128. Let

$$\xi_1^* = \pi^* \xi^* = \begin{pmatrix} 0 & 1 & & \\ 1 & 0 & & \\ & & 0 & 1 \\ & & 1 & 0 \end{pmatrix}$$

$$\xi_2^* = \pi^* \delta_2^* \pi^* = \begin{pmatrix} & & 1 & 0 \\ & & 0 & 1 \\ 1 & 0 & & \\ 0 & 1 & & \end{pmatrix},$$

then

$$G_1^* = \langle N, \mu^*, \xi_1^*, \xi_2^* \rangle,$$

$$N_1 \lhd G_1^*,$$

G_1^*/N_1 is isomorphic to the Dieder group of order 8 generated by permutations (01)(23), (02)(13) and (12), and $|G_1^*| = 1024$.

Parallel to Theorem 14.3 we have

Theorem 14.5. *Let C be a self-dual $\mathbb{Z}_4$-code of length n containing the all 1 codeword 1^n and $\mathrm{cwe}_C(Z_0, Z_1, Z_2, Z_3)$ be its transformed complete weight enumerator. Then*

(i) $\mathrm{cwe}_C(Z_0, Z_1, Z_2, Z_3) \in \mathbb{C}[Z_0, Z_1, Z_2, Z_3]_n^{G_1^*}$.

(ii) *The Molien series of $\mathbb{C}[Z_0, Z_1, Z_2, Z_3]^{G_1^*}$ is*

$$\Phi_{G_1^*}(\lambda) = (1 + \lambda^8 + \lambda^{16})(1 + \lambda^{16})/(1 - \lambda^4)(1 - \lambda^8)(1 - \lambda^{12})(1 - \lambda^{16}).$$

(iii) $\mathbb{C}[Z_0, Z_1, Z_2, Z_3]^{G_1^*} = R \dotplus \sigma_8 R \dotplus \sigma_8^2 R \dotplus \sigma_{16} R \dotplus \sigma_8 \sigma_{16} R \dotplus \sigma_8^2 \sigma_{16} R$, *where R is the $\mathbb{C}$-algebra of symmetric functions of Z_0^4, Z_1^4, Z_2^4 and Z_3^4 and*

$$\sigma_8 = Z_0^4 Z_3^4 + Z_1^4 Z_3^4,$$

$$\sigma_{16} = (Z_0 Z_1 Z_2 Z_3)^2 (Z_0^4 Z_1^4 + Z_2^4 Z_3^4 - Z_0^4 Z_2^4 - Z_1^4 Z_3^4).$$

Proof. (i) follows from Proposition 14.1 and Lemma 14.2 (iii).

(ii) As in the proof of Theorem 14.3 (ii) first we prove

$$\mathbb{C}[Z_0, Z_1, Z_2, Z_3]^{N_1} = \mathbb{C}[Z_0^4, Z_1^4, Z_2^4, Z_3^4] \dotplus Z_0^2 Z_1^2 Z_2^2 Z_3^2 \mathbb{C}[Z_0^4, Z_1^4, Z_2^4, Z_3^4],$$

from which we deduce

$$\Phi_{N_1}(\lambda) = (1+\lambda^8)/(1-\lambda^4)^4,$$

and then we compute $\Phi_{G_1^\bullet}(\lambda)$. The details of the proof will be omitted.

(iii) Clearly, $R \dotplus \sigma_8 R \dotplus \sigma_8^2 R \dotplus \sigma_{16}(R \dotplus \sigma_8 R \dotplus \sigma_8^2 R) \subseteq \mathbb{C}[Z_0, Z_1, Z_2, Z_3]^{G_1^\bullet}$, and $\sigma_8 \notin R$, $\sigma_8^2 \notin R \dotplus \sigma_8 R$, and $\sigma_{16} \notin R \dotplus \sigma_8 R \dotplus \sigma_8^2 R$. The symmetric group S_4 on four letters has the Molein series

$$\frac{1}{(1-\lambda)(1-\lambda^2)(1-\lambda^3)(1-\lambda^4)}.$$

It follows from (ii) that

$$\mathbb{C}[Z_0, Z_1, Z_2, Z_3]^{G_1^\bullet} = R \dotplus \sigma_8 R \dotplus \sigma_8^2 R \dotplus \sigma_{16}(R \dotplus \sigma_8 R \dotplus \sigma_8^2 R). \qquad \square$$

Gleason proved that the weight enumerator of any self-dual doubly even binary code can be expressed as a polynomial in the weight enumerators of the binary Hamming code H_8 and the Golay code G_{24}. For self-dual $\mathbb{Z}_4$-codes we have the following analogs of Gleason's theorem which can be derived from Theorems 14.3 and 14.5 by routine verification.

Theorem 14.6. *The complete weight enumerator of any self-dual $\mathbb{Z}_4$-code can be expressed as a polynomial of the complete weight enumerators of the self-dual $\mathbb{Z}_4$-codes $\mathcal{A}_1$, $\mathcal{K}_4$, $\mathcal{K}_4'$, $\mathcal{O}_8$ and $\mathcal{C}_{10}$.* $\square$

Theorem 14.7. *The complete weight enumerator of any self-dual $\mathbb{Z}_4$-code containing the all 1 codeword can be expressed as a polynomial of the complete weight enumerators of self-dual $\mathbb{Z}_4$-codes $\mathcal{K}_4$, $\mathcal{K}_8$, $\mathcal{K}_{12}$, $\mathcal{K}_{16}$, $\mathcal{O}_8$ and $\mathcal{C}_{16}$, all containing the all 1 codewords.* $\square$

14.3. Symmetrized Weight Enumerators of Self-dual $\mathbb{Z}_4$-Codes

The analogous theorems to Theorems 14.3 and 14.5 for symmetrized weight enumerators of self-dual $\mathbb{Z}_4$-codes follow easily from Theorems 14.3 and 14.5.

Theorem 14.8. *The symmetrized weight enumerator of any self-dual $\mathbb{Z}_4$-code belongs to the ring*

$$\mathbb{C}[\tilde{Z}_0,\ \tilde{Z}_1^4 + \tilde{Z}_2^4,\ \tilde{Z}_1^4\tilde{Z}_2^4]\,,$$

where

$$\tilde{Z}_0 = \frac{X_0 + X_2}{\sqrt{2}},\ \tilde{Z}_1 = \sqrt{2}X_1,\ \tilde{Z}_2 = \frac{X_0 - X_2}{\sqrt{2}}\,.$$

The ring has the Molien series

$$\frac{1}{(1-\lambda)(1-\lambda^4)(1-\lambda^8)}$$

An alternate basis is given by the polynomials

$$\begin{aligned}\phi_1 &= X_0 + X_2,\\ \phi_4 &= 2X_1^4 - X_0X_2(X_0^2 + X_2^2),\\ \phi_8 &= X_1^4(X_0 - X_2)^4.\end{aligned}$$

□

Theorem 14.9. *The symmetrized weight enumerator of any self-dual $\mathbb{Z}_4$-code containing the all 1 codeword belongs to the ring*

$$S \dotplus \tilde{Z}_1^4\tilde{Z}_2^4 S \dotplus \tilde{Z}_1^8\tilde{Z}_2^8 S,$$

where S is the ring of symmetric functions of $\tilde{Z}_0^4$, $\tilde{Z}_1^4$, $\tilde{Z}_2^4$. This ring has the Molien series

$$\frac{1+\lambda^8+\lambda^{16}}{(1-\lambda^4)(1-\lambda^8)(1-\lambda^{12})}$$

An explicit basis for S is given by the polynomials

$$\begin{aligned}\Phi_4 &= X_0^4 + 6X_0^2X_2^2 + 8X_1^4 + X_3^4,\\ \Phi_8 &= (X_0^2X_2^2 - X_1^4)((X_0^2 + X_2^2)^2 - 4X_1^4),\\ \Phi_{12} &= X_1^4(X_0^2 - X_2^2)^4,\end{aligned}$$

and then the ring is $S \dotplus \Psi_8 S \dotplus \Psi_8^2 S$, where

$$\Psi_8 = X_1^4(X_0 - X_2)^4$$

with

$$\Psi_8^3 = \Psi_8^2\left(\frac{1}{16}\Phi_4^2 - \Phi_8\right) - \frac{1}{8}\Psi_8\Phi_4\Phi_{12} + \frac{1}{16}\Phi_{12}^2.$$

□

We also have the following analogs of Gleason's theorem

Theorem 14.10. *The symmetrized weight enumerator of any self-dual $\mathbb{Z}_4$-code can be expressed as a polynomial of the symmetrized weight enumerators of the self-dual $\mathbb{Z}_4$-codes $\mathcal{A}_1$, $\mathcal{K}_4$, $\mathcal{O}_8$.* □

Theorem 14.11. *The symmetrized weight enumerator of any self-dual $\mathbb{Z}_4$-code containing the all 1 codeword can be expressed as a polynomial of the symmetrized weight enumerators of the self-dual $\mathbb{Z}_4$-codes $\mathcal{K}_4, \mathcal{K}_8, \mathcal{K}_{12}$ and $\mathcal{O}_8$, all containing the all 1 codewords.* □

Theorems 14.3 and 14.5 are due to the Klemm (1987, 1989). Theorems 14.6–14.11 are due to Conway and Sloane (1993a). More results on self-dual quaternary codes can be found in Conway and Sloane (1993a), Bonnecaze *et al.* (1997), Calderbank and Sloane (1997), and Pless *et al.* (1997).

BIBLIOGRAPHY

Adoul, P. (1987). Fast ML decoding algorithm for the Nordstrom–Robinson code, *IEEE Trans. Inform. Theory* **33**, 931–933.

Baker, R. D., van Lint, J. H. and Wilson, R. M. (1983). On the Preparata and Goethals codes, *IEEE Trans. Inform. Theory* **29**, 342–345.

Berlekamp, E. R. (1971). Coding theory and the Mathieu groups, *Inform. Cont.* **18**, 40–64.

Best, M. R. (1980). Binary codes with a minimum distance of four, *IEEE Trans. Inform. Theory* **26**, 738–742.

Blake, I. F. (1972). Codes over certain rings, *Inform. Cont.* **20**, 396–404.

Blake, I. F. (1975). Codes over integer residue rings, *Inform. Cont.* **29**, 295–300.

Bonnecaze, A. and Duursma, I. M. (1997). Translates of Linear codes over $\mathbb{Z}_4$, *IEEE Trans. Inform. Theory* **43**, 1218–1230.

Bonnecaze, A. and Solé, P. (1994). Quaternary constructions of formally self-dual binary codes and unimodular lattices, in *Proc. 1st Franco-Israeli Workshop on Coding Theory* (Paris, July 1993), *Lecture Notes in Computer Science* **781** (Springer-Verlag), 194–206.

Bonnecaze, A., Solé, P. and Calderbank, A. R. (1995). Quaternary quadratic residue codes and unimodular lattices, *IEEE Trans. Inform. Theory* **41**, 366–377.

Bonnecaze, A., Solé, P., Bachoc, C. and Mourrain, B. (1997). Type II codes over $\mathbb{Z}_4$, *IEEE Trans. Inform. Theory* **43**, 969–976.

Boztaş, S., Hammons, A. R. Jr. and Kumar, P. V. (1992). Four-phase sequences with near-optimum correlation properties, *IEEE Trans. Inform. Theory* **38**, 1101–1113.

Boztaş, S. and Kumar, P. V. (1994). Binary sequences with Gold-like correlation properties but larger linear span, *IEEE Trans. Inform. Theory* **40**, 532–537.

Calderbank, A. R., Cameron, P. J., Kantor, W. M. and Seidel, J. J. (1997). $\mathbb{Z}_4$-Kerdock codes, orthogonal spreads, and extremal Euclidean line-sets, to appear in *Proc. London Math. Soc.*

Calderbank, A. R., Li, W.-C. W and Poonen, B. (1997). A 2-adic approach to the analysis of cyclic codes, *IEEE Trans. Inform. Theory* **43**, 977–986.

Calderbank, A. R. and McGuire, G. (1995). $\mathbb{Z}_4$-linear codes obtained as projections of Kerdock and Delsarte–Goethals codes, *Linear Algebra Appl.* **226–228**, 647–665.

Calderbank, A. R., McGuire, G., Kummer, P. V. and Helleseth, T. (1996). Cyclic codes over $\mathbb{Z}_4$, locator polynomials, and Newton's identities, *IEEE Trans. Inform. Theory* **42**, 217–226.

Calderbank, A. R. and Sloane, N. J. A. (1995). Modular and p-adic cyclic codes, *Designs, Codes and Cryptography* **6**, 21–35.

Calderebank, A. R. and Sloane, N. J. A. (1997). Double circulant codes and even unimodular lattices, *J. Algebraic Combinatorics* **6**, 119–131.

Carlet, C. (1989). A simple description of Kerdock codes, in *Lecture Notes in Computer Science* **388** (Springer-Verlag), 202–208.

Carlet, C. (1991). The automorphism groups of the Kerdock codes, *J. Inform. Optimization Sci.* **12**, 387–400.

Carlet, C. (1992). Les groupes d'automorphism des codes de Delsarte–Goethals, *C. R. Acad. Sci. Paris, Série I* **315**, 475–478.

Carlet, C. (1993). The automorphism groups of the Delsarte–Goethals codes, *Designs, Codes and Cryptograph* **3**, 237–249.

Conway, J. H. and Sloane, N. J. A. (1990). Orbit and coset analysis of the Golay and related codes, *IEEE Trans. Inform. Theory* **36**, 1038–1050.

Conway, J. H. and Sloane, N. J. A. (1993). *Sphere Packings, Lattices and Groups* (Springer-Verlag), 2nd edition.

Conway, J. H. and Sloane, N. J. A. (1993a). Self-dual codes over the integers modulo 4, *J. Comb. Theory, Series A* **62**, 30–45.

Conway, J. H. and Sloane, N. J. A. (1994). Quaternary constructions of the binary single-error correcting codes of Julin, Best, and others, *Designs, Codes and Cryptography* **4**, 31–42.

Delsarte, P. (1973). Four fundamental parameters of a code and their combinatorial significance, *Inform. Cont.* **23**, 407–438.

Delsarte, P. (1973a). An algebraic approach to the association schemes of coding theory, *Philips Research Reports Supplements*, No. 10.

Delsarte, P. and Goethals, J.-M. (1975). Alternating bilinear forms over $GF(q)$, *J. Combinatorial Theory, Series A* **19**, 26–50.

Forney, G. D., Jr., Sloane, N. J. A. and Trott, M. D. (1993). The Nordstrom–Robinson code is the binary image of the octacode, in *Coding and Quantization*, eds. Calderbank, R., Forney, G. D., Jr. and Moayeri, N., *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* **14** (AMS), 19–26.

Gaborit, P. (1996). Mass formulas for self-dual codes over $\mathbb{Z}_4$ and $\mathbb{F}_q + u\mathbb{F}_q$ rings, *IEEE Trans. Inform. Theory* **42**, 1222–1228.

Goethals, J. M. (1974). Two dual families of nonlinear binary codes. *Electronics Lett.* **10**, 471–472.

Goethals, J. M. (1976). Nonlinear codes defined by quadratic forms over $GF(2)$. *Inform. Cont.* **31**, 43–74.

Hammons, A. R., Jr., Kumar, P. V., Calderbank, A. R., Sloane, N. J. A. and Solé, P. (1994). The $\mathbb{Z}_4$-linearity of Kerdock, Preparata, Goethals, and related codes, *IEEE Trans. Inform. Theory* **40**, 301–319.

Helleseth, T. (1996). Private communication.

Helleseth, T. and Kumar, P. V. (1995). The algebraic decoding of the $\mathbb{Z}_4$-linear Goethals code, *IEEE Trans. Inform. Theory* **41**, 2040–2048.

Helleseth, T., Kumar, P V. and Shanbhag, A. (1996). Codes with the same weight distributions as the Goethals codes and the Delsarte–Goethals codes, *Designs, Codes and Cryptography* **9**, 257–266.

Helleseth, T., Kumar, P. V. and Shanbhag, A. (1996a). Exponential sums over Galois rings and their applications, in *Finite Fields and Their Applications*, eds. Cohen, S. and Niederreiter, H. (Cambridge University Press), 109–128.

Hergert, F. B. (1990). On the Delsarte–Goethals codes and their formal duals, *Discrete Math.* **83**, 249–263.

Hou, X.-D., Koponen, S. and Lahtonen, J. (1997). On the $\mathbb{Z}_4$-linearity of the Reed–Muller codes, *Abstracts IEEE Inform. Theory Workshop*, Longyearbyen, Norway, July 6–12, 11.

Julin, D. (1965). Two improved block codes, *IEEE Trans. Inform. Theory* **11**, 459.

Kantor, W. (1982). Spreads, translation planes and Kerdock sets, I, II, *SIAM J. Alg. Discrete Math.* **3**, 151–165, 308–318.

Kantor, W. (1982a). An exponential number of generalized Kerdock codes, *Inform. Cont.* **53**, 74–80.

Kantor, W. (1983). On the inequivalence of generalized Preparata codes, *IEEE Trans. Inform. Theory* **29**, 345–348.

Kantor, W. (1995). Codes, quadratic forms and finite geometry, in *Different*

Aspects of Coding Theory, ed. Calderbank, R., *Proceedings of Symposia in Applied Mathematics* **50** (AMS), 153–177.

Kantor, W. (1995a). Quaternionic line-sets and quaternionic Kerdock codes, *Linear Algebra Appl.* **226–228**, 749–779.

Kantor, W. (1996). Orthogonal spreads and translation planes, in *Progress in Algebraic Combinatorics*, eds. Bannai, E. and Munemasa, A., *Advanced Studies in Pure Mathematics* **24** (Math. Soc. of Japan), 227–242.

Kerdock, A. M. (1972). A class of low-rate nonlinear codes, *Inform. Cont.* **20**, 182–187.

Klemm, M. (1987). Über die Identität von MacWilliams für die Gewichtsfunktion von Codes, *Arch. Math.* **49**, 400–406.

Klemm, M. (1989). Selbstduale Codes über dem Ring der ganzen Zahlen modulo 4, *Arch. Math.* **53**, 201–207.

Krawtchouk, M. (1929). Sur une généralisation des polynomes d'Hermite, *Comptes Rendus* **189**, 620–622.

Krawtchouk, M. (1933). Sur les distribution des racines des polynomes orthogonaux, *Comptes Rendus* **196**, 739–741.

Krull, W. (1924). Algebraische theorie der ringe, *Math. Ann.* **92**, 183–213.

Kumar, P. V., Helleseth, T. and Calderbank, A. R. (1995). An upper bound for Weil exponential sums over Galois rings and applications, *IEEE Trans. Inform. Theory* **41**, 456–468.

Kumar, P. V., Helleseth, T., Calderbank, A. R. and Hammons, A. R., Jr. (1996). Large families of quaternary sequences with low correlation, *IEEE Trans. Inform. Theory* **42**, 579–592.

Kuzmin, A. S. and Nechaev, A. A. (1993). Construction of noise stable codes using linear recurring sequences over Galois rings, *Usp. Math. Nauk* **48**, 197–198 (in Russian). English translation: *Russian Math. Surv.*

Kuzmin, A. S. and Nechaev, A. A. (1994). Linearly presented codes and Kerdock codes over arbitrary Galois field of characteristic 2, *Usp. Math. Nauk* **49**, 165–166 (in Russian). English translation: *Russian Math. Surv.*

Lahtonen, J. (1995). Certain exponential sums over Galois rings and related constructions of families of sequences, in *Proc. IEEE Int. Symp. Inform. Theory* (Whister, Canada), 85.

MacDonald, B. R. (1974). *Finite Rings with Identity* (Marcel Dekker).

MacWilliams, F. J. and Sloane, N. J. A. (1977). *The Theory of Error-Correcting Codes* (North-Holland).

Nechaev, A. A. (1989). Kerdock code in a cyclic form, *Diskretnaya Mat. (USSR)* **1**, 123–139 (in Russian). English translation: *Discrete Math.*

Appl. **1** (1991), 365–384.

Nordstrom, A. W. and Robinson, J. P. (1967). An optimum nonlinear code, *Inform. Cont.* **11**, 613–616.

Pless, V. and Qian Z. (1996). Cyclic codes and quadratic residue codes over $\mathbb{Z}_4$, *IEEE Trans. Inform. Theory* **42**, 1594–1600.

Pless, V., Solé, P. and Qian, Z. (1997). Cyclic self-dual $\mathbb{Z}_4$-codes, *Finite Fields and Their Applications* **3**, 48–69.

Preparata, F. P. (1968). A class of optimum nonlinear double-error correcting codes, *Inform. Cont.* **13**, 378–400.

Semankov, N. V. and Zinovév, A. (1969). Balanced codes and tactical configurations, *Problems of Inform. Trans.* **5**, 22–28.

Serre, J. P. (1973). *A Course in Arithmetic* (Springer-Verlag).

Shankar, P. (1979). On BCH codes over arbitrary integer rings, *IEEE Trans. Inform. Theory* **25**, 480–483.

Shanbhag, A. G., Kumar, P. V. and Helleseth, T. (1996). Upper bound for a hybrid sum over Galois rings with applications to aperiodic correlation for some q-ary sequences, *IEEE Trans. Inform. Theory* **42**, 250–254.

Sloane, N. J. A. and Whitehead, D. S. (1970). A new family of single-error correcting codes, *IEEE Trans. Inform. Theory* **16**, 717–719.

Sloane, N. J. A. (1977). Error-correcting codes and invariant theory: New applications of a nineteenth-century technique, *Amer. Math. Monthly* **84**, 82–107.

Solé, P. (1989). A quaternary cyclic code, and a family of quadriphase sequences with low correlation properties, *Lectures Notes in Computer Science* **388** (Springer-Verlag), 193–201.

Solé, P. (1990). An inversion formula for Krawtchouk polynomials with application to coding theory, *J. Inform. Optim. Sci.* **11**, 207–213.

Solé, P. (1993). Generalized theta functions for lattice vector quantization, in *Coding and Quantization*, eds. Calderbank, R., Forney, G. D., Jr. and Moayeri, N., *DIMACS Series in Discrete Mathematics and Theoretical Computer Science,* **14** (AMS), 27–32.

Udaya, P. and Siddiqi, M. U. (1991). Large linear complexity sequences over $\mathbb{Z}_4$ for quadriphase modulated communication systems having good correlation properties, *Proc. IEEE Int. Symp. Inform. Theory* (Budapest, Hungary), 386.

Uspensky, J. V. (1948). *Theory of Equations* (McGraw-Hill).

van Lint, J. H. (1983). Kerdock and Preparata codes, *Cong. Numer.* **39**, 25–41.

Wan, Z.-X. (1992). *Introduction to Abstract and Linear Algebra* (Chatwell-Bratt, Bromley, United Kingdom and Studentlitteratur, Lund, Sweden).

Wan, Z.-X. (1997). On the uniqueness of the Leech lattice, *European J. Combinatorics*, **18**, 455–459.

Wasan, S. K. (1982). On codes over $\mathbb{Z}_m$, *IEEE Trans. Inform. Theory* **28**, 117–120.

Yang, K., Helleseth, T., Kumar, P. V. and Shanbag, A. G. (1996). On the weight hierarchy of Kerdock codes over $\mathbb{Z}_4$, *IEEE Trans. Inform. Theory* **42**, 1587–1593.

Zariski, O. and Samuel, P. (1958). *Commutative Algebra* (Van Norstrand), Vol. 1.

SUBJECT INDEX