ADDITIVE CYCLIC CODES OVER RINGS

by

Jonas Saman

A Thesis Presented to the Faculty of the
American University of Sharjah
College of Arts and Sciences
in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science in
Pure Mathematics

Sharjah, United Arab Emirates
January 2015

Approval Signatures

We, the undersigned, approve the Master's Thesis of Jonas Saman.

Thesis Title: Additive cyclic codes over rings

Signature	**Date of Signature** (dd/mm/yyyy)
Dr. Taher Abualrub Professor Thesis Advisor	
Dr. Faruk Uygul Assistant Professor Thesis Committee Member	
Dr. Ayman Badawi Professor Thesis Committee Member	
Dr. Hana Sulieman Head of Mathematics	
Dr. Pia Anderson CAS Graduate Programs Director	
Dr. Mahmoud Anabtawi Dean of CAS	
Dr. Khaled Assaleh Interim Vice Provost for Research and Graduate Studies	

To everyone.

Abstract

This Master's thesis introduces a reader with an average knowledge of mathematics to coding theory. A background in algebra is included with the required details needed in coding theory. The thesis guides the reader from casual examples to linear codes, as well as cyclic codes. The focus lies on $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic codes. At the end new theorems regarding $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic codes with even β in $\mathbb{Z}_2[x]/(x^\alpha - 1) \times \mathbb{Z}_4[x]/(x^\beta - 1)$ are discussed.

Search Terms: algebraic codes, binary codes, linear codes, cyclic codes, additive codes, coding theory

Table of Contents

Abstract **4**

List of Tables **7**

List of Figures **8**

1 An Introduction to Coding Theory **10**
1.1 Error-Correcting Codes 10
1.2 Hamming Distance 14

2 An Introduction to Algebra **18**
2.1 Groups 18
2.2 Rings 19
2.3 Fields 19
2.4 Homomorphisms 21
2.5 Polynomials 22
2.6 The Division Algorithm 22
2.7 Ideals 23
2.8 Factor Rings 23
2.9 Characteristic of a Ring 24
2.10 Irreducible Polynomials 24
2.11 The Ring of Polynomials modulo f(x) 25
2.12 Vector Spaces 26

3 Linear Codes and Cyclic Codes over Finite Fields **29**
3.1 Linear Codes 29
3.1.1 Encoding. 30
3.2 Introduction to Cyclic Codes 31
3.3 Generator Polynomial 32
3.4 Check Polynomial 33
3.5 Hamming Codes as Cyclic Codes 34
3.6 Encoding with Cyclic Codes 34
3.6.1 The non-systematic method. 34
3.6.2 The systematic method. 35

4 Cyclic Codes over $\mathbb{Z}_4$ **36**
4.1 Introduction 36
4.2 The Construction of Some Non-linear Codes 37

5 Additive Cyclic Codes over $\mathbb{Z}_2\mathbb{Z}_4$ **40**
5.1 $\mathbb{Z}_2\mathbb{Z}_4$-additive Codes 40
5.2 $\mathbb{Z}_2\mathbb{Z}_4$-additive Cyclic Codes 42
5.2.1 $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic codes with odd β. 43
5.2.2 $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic codes with even β. 44
5.2.3 Examples. 46
5.3 Conclusions and Future Work 48

References **49**

Vita **50**

List of Tables

1.1 Three-bit Repetition Code 14
1.2 Error-detection and Error-correction Table 17

5.1 $(x^2 - 1)$ in $\mathbb{Z}_4[x]/(x^4 - 1)$ 46
5.2 Codewords in $\mathbb{Z}_2[x]/(x^3 - 1) \times \mathbb{Z}_4[x]/(x^4 - 1)$ 47

List of Figures

1.1 General Digital Communication System 10
1.2 Binary Symmetric Channel 12

3.1 Encoding Procedure 30

1. An Introduction to Coding Theory

1.1. Error-Correcting Codes

In 1948, a classic paper entitled *"A Mathematical Theory of Communication"* by Claude Shannon was published. This laid the foundation for further work in coding theory. Since then, numerous researches have spent hours scratching their heads in confusion as how they can encode digital information for reliable transmission through noisy channels. Error correcting codes are used in forward error correction, a technique in which errors occurring in data transmission over noisy communication channels are controlled. As expected, the amount of control an error correcting code can exert depends on what code is being used, and this amount is definitely limited. Therefore, it is important that researchers continue developing new codes in order to achieve better forward error correcting. Transmission of information through a noisy network incorporates the use of error correcting codes which allows the rectification of possible errors that are detected during the process.

When Shannon published his paper, algebraic codes (also known as block codes) were the only codes used in forward error correction. Other codes used in forward error correction are convolutional codes. In 1993, turbo codes were developed, these codes combine convolutional codes to create a block code [1]. This paper focusses solely on block codes.

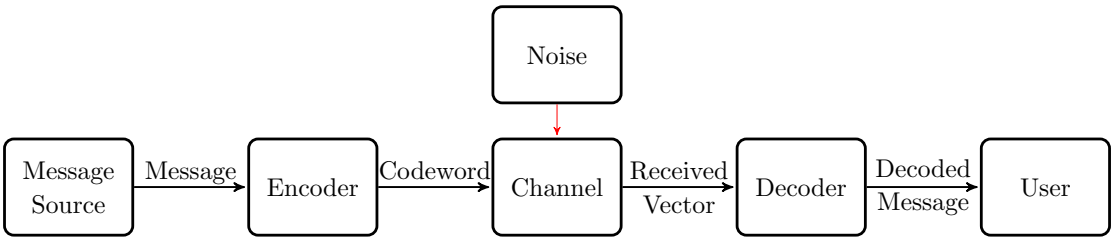

Figure 1.1: General Digital Communication System

Digital communication systems are designed to send messages from a message generating source to its target. Such a communication system can be represented by a diagram as in figure 1.1. The message generated by the source can be of many forms, such as speech, images, or plain text. For sources such as speech or images, a transducer is required in order to convert the source into digital signals,

such as a microphone, or video camera. A transducer is also used at the final destination in order to convert the digital signal into whatever desired form, audio or visual. Once the message is generated, it is transmitted to the encoder.

The encoder, as the name implies, encodes the data. The altered data is then passed on through the channel. Channels include but are not limited to wireless and wired communications, as well as physical data storage devices. The received data then gets decoded, and this decoded message ends up at the target user. However, channels are prone to noise, whether is it human error or malfunctioning equipment, the data may be altered in unpredictable manners. For the sake of simplicity, from hereon we assume that all messages have been converted into digital signals, or data.

The motivation for error correcting codes stems from the requirements of transmitting data over real channels, which are frequently imperfect. One must encode the data that is supposed to be transmitted in a reliable way, such that any noise on a channel does not interfere with the actual data. One has to remember that the received message may not be the same as the original sent message. The required error correction is achieved by introducing redundancy to the transmitted data, this aids in the detection of errors and ensures the possibility of retrieving the original message. This may be done in a simple way of repetition, but this is not the only method of adding redundancy. Usually, complex algorithms are preferred to encode the data with the required redundancy. Shannon's noisy-channel coding theorem states that, under certain conditions, there are certain codes that make the probability of error arbitrarily small [2].

Definition 1.1. *A* binary code *is a set of sequences of elements from the set* $\mathbb{Z}_2 = \{0, 1\}$.

Definition 1.2. *A* q-ary code *is a set of sequences of elements from the set* $F_q = \{\lambda_1, \lambda_2, ..., \lambda_q\}$*, called the* alphabet.

Remark 1.3. *F_q is a finite field with q elements, this will be defined later.*

Example 1.4. *The set of 7-digit phone numbers is a code over the 10-number alphabet* $\{0, 1, 2, ..., 9\}$.

Definition 1.5. *A* codeword *is a single sequence of elements in a code C from the alphabet.*

Example 1.6. *Based on the previous example, a codeword in the set of 7-digit phone numbers is* 5155522.

Let $(F_q)^n$ be the set of all ordered n-tuples $\mathbf{a} = (a_1 a_2 \cdots a_n), a_i \in F_q$. The elements of $(F_q)^n$ are *vectors*. This implies that every codeword and received message is a vector.

When transmitting data through a channel, one has to remember there is always a possibility of interference due to noise. The probability of a bit error in a binary symmetric channel is called the *crossover probability*. The following assumptions are made about the probability of bit errors:

1. Every received bit has the same probability of having an error, this is less than 50%.

2. When an error occurs, the other $q - 1$ elements in the alphabet are equally as likely to occur.

The components present in a communication channel involve a finite alphabet, along with forward channel probabilities $P(a_j \text{ received}|a_i \text{ sent})$ that satisfies the following condition:

$$\sum_{j=1}^{q} P(a_j \text{received}|a_i \text{sent}) = 1$$

When the result of any one transmission is independent of the outcome of prior transmissions, the communication channel is defined to be memoryless.

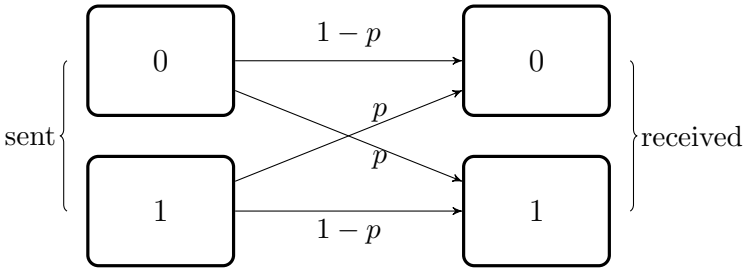

Figure 1.2: Binary Symmetric Channel

In a communications channel model, a common example of such a memoryless channel is the binary symmetric channel which takes a binary input that consists of the code $(0, 1)$ and the following channel probabilities:
$P(1 \text{ received}|0 \text{ sent}) = P(0 \text{ received}|1 \text{ sent}) = p$
$P(0 \text{ received}|0 \text{ sent}) = P(1 \text{ received}|1 \text{ sent}) = 1 - p$

In figure 1.2, p denotes the symbol error probability of the channel. When a codeword of size n is transmitted, the probability of no errors occurring is $(1-p)^n$. The probability for the occurrence of one error in a particular position is $p(1-p)^{n-1}$. In general, the probability of k-errors occurring is $P(k-\text{errors}) = \binom{n}{k}p^k(1-p)^{n-k}$.

Example 1.7. *Let us say that the source messages are '0' and '1' and let* $p = .25$ *be the crossover probability. Without any redundancy, the chance of receiving the correct message is 75%.*

However, using redundancy and introducing the codewords $(000, 111)$ *for '0' and '1' respectively, yields the following probabilities:*

$$\begin{array}{ll} \textit{no-errors} & \binom{3}{0}.25^0(1-.25)^{3-0} = 0.421875 \\ \textit{1-error} & \binom{3}{1}.25^1(1-.25)^{3-1} = 0.421875 \\ \textit{2-errors} & \binom{3}{2}.25^2(1-.25)^{3-2} = 0.140625 \\ \textit{3-errors} & \binom{3}{3}.25^3(1-.25)^{3-3} = 0.015625 \end{array}$$

Transferring one of the codewords of $(000, 111)$ through a noisy channel yields an element of the set $\{000, 001, 010, 100, 111, 110, 101, 011\}$ on the receiving end. A technique called *nearest neighbor decoding* is used to decode the received message. Informally, define this technique to decode the received data to the nearest matching codeword, later this technique will be defined mathematically. The elements of $\{000, 001, 010, 100\}$ will be decoded as the codeword 000, meaning '0', similarly the elements of $\{111, 110, 101, 011\}$ will be decoded as 111, meaning '1', due to the dominating presence of those bits.

As we can see, now one bit of data can be affected by noise without any effect on the interpretation of the message. This implies that the probability of receiving the correct codeword is $P(0-\text{errors}) + P(1-\text{error}) = 0.84375$. Clearly $0.84375 > 0.75$, hence this code increased the chance of receiving the correct message using nearest neighbor decoding.

Source	Codewords Transmitted	Noise	Code Received	Decoded Sequence
1	111		111	1
			011	1
			101	1
			110	1
0	000		000	0
			100	0
			010	0
			001	0

Table 1.1: Three-bit Repetition Code

Albeit simple to implement and use, this is an inefficient method of error correcting due to the fact that the total data transmitted will triple in size. Henceforth, the time required to transmit data is also tripled. Usually, error correcting coding takes into account and permits factors such as quick encoding and decoding of information, easy data transmission and maximum detection or correction capabilities.

1.2. Hamming Distance

In example 1.7 we chose to decode data in a certain way because it was 'closer' to their respective codewords. Formally, we define the following distance function on $(F_q)^n$.

Definition 1.8. *The Hamming Distance, denoted by $d(\boldsymbol{x}, \boldsymbol{y})$, between two vectors of equal length is the amount of positions where they are different.*

Example 1.9. *The Hamming distance between:*

1. mama *and* papa *is 2*
2. *01001010 and 01010011 is 3*
3. *1337 and 2008 is 4*

The Hamming Distance satisfies the following conditions:

1. $0 \leq d(\mathbf{x}, \mathbf{y}) \leq n$
2. $d(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y}$
3. $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x}) \quad \forall \mathbf{x}, \mathbf{y} \in (F_q)^n$
4. $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}) \quad \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in (F_q)^n$

As the Hamming Distance satisfies these conditions it is a *metric* function.

Definition 1.10. *The* minimum distance *of C, denoted by $d(C)$, is defined by*

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

In other words, the smallest amount of places in which two codewords differ is called the minimum distance, denoted by d. The minimum distance is the smallest Hamming Distance between individual codewords.

Example 1.11. *Let $C = \{(1010), (1001), (1101), (1111)\}$ then we have the following Hamming distances:*

1. $d((1010), (1001)) = 2$
2. $d((1010), (1101)) = 3$
3. $d((1010), (1111)) = 2$
4. $d((1001), (1101)) = 1$
5. $d((1001), (1111)) = 2$
6. $d((1101), (1111)) = 1$

Therefore, $d(C) = \min \{1, 2, 3\} = 1$

Definition 1.12. *An (n, M, d)-code is a code of length n, with M codewords, and with minimum distance d.*

Only codewords are transmitted from the encoder to the channel, which are received and corrected if necessary. For instance, if the word received is nanas, which is checked to be a valid codeword, it means the data transmitted is error free. If not, a decoding algorithm must be constructed to find the possible codeword sent through. An example of such a rule is the *maximum-likelihood decoding* rule, which takes into account the forward channel probabilities of the data or codeword received. The maximum-likelihood decoding rule aims to minimize the error probability. In certain cases, such as the binary symmetric channel as in figure 1.2, with crossover probability $p < 1/2$, the maximum-likelihood decoding is equivalent to *nearest-codeword decoding.* In nearest-codeword decoding, a received word is decoded as the closest codeword under the measure of Hamming distance.

Assume codewords are transmitted through a communications channel, and **x'** is received. Then the nearest decoding rule comes into play where **x'** is decoded to **x**, such that $d(\mathbf{x'}, \mathbf{x})$ is minimum. In other words, it will satisfy the following condition
$d(\mathbf{x'}, \mathbf{x}) = min_{\mathbf{x} \in C} d(\mathbf{x'}, \mathbf{x})$.

Theorem 1.13. *A Code C can:*

1. *Detect up to s errors of a codeword if* $d(C) \geq s + 1$.

2. *Correct up to t errors of a codeword if* $d(C) \geq 2t + 1$.

Proof:

1. Let $d(C) \geq s + 1$. Assume a codeword **x** is transmitted and the vector **y** is received with s or fewer errors, then $d(\mathbf{x}, \mathbf{y}) \leq s$. Then the received vector cannot be one of the codewords as $d(C) \geq s + 1$. Therefore, the errors can be detected.

2. Let $d(C) \geq 2t + 1$. Assume a codeword **x** is transmitted and the vector **y** is received with t or fewer errors, then $d(\mathbf{x}, \mathbf{y}) \leq t$.
   Let **x'** be a codeword that is not **x**. $\mathbf{x'} \neq \mathbf{x} \implies d(\mathbf{x'}, \mathbf{y}) \geq t + 1$.
   Otherwise, $d(\mathbf{x'}, \mathbf{y}) \leq t \implies d(\mathbf{x}, \mathbf{x'}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{x'}, \mathbf{y}) \leq 2t$
   But $d(C) \geq 2t + 1$, which is a contradiction.
   So **x** is the nearest codeword to **y** and the nearest neighbor decoding corrects the errors.

□

Corollary 1.14. *When d is the minimum distance of a code C, then the code C can be used for either of the following two:*

1. *Detect up to a total of* $(d - 1)$ *errors.*

2. *Correct up to a total of* $\lfloor \frac{d-1}{2} \rfloor$ *errors in any codeword.*

Proof:

1. $d \geq s + 1 \iff s \leq d - 1$

2. $d \geq 2t + 1 \iff t \leq \frac{d-1}{2}$

□

Example 1.15. *If* $d(C) = 6$ *then C can be used as either a 5-error-detecting code or a 2-error-correcting code.*

The following table shows the relation between the minimum distance and the number of errors detected or corrected by a code C:

$d(C)$	Number of errors detected by C	Number of errors corrected by C
1	0	0
2	1	0
3	2	1
4	3	1
5	4	2
6	5	2
7	6	3
.	.	.
.	.	.
.	.	.

Table 1.2: Error-detection and Error-correction Table

In addition to the size of a code and its length, the minimum distance of a code is an essential attribute of a code. When the length of a code is small, it enables a faster transmission of data; along with a large number of codewords and a large minimum distance, this would be considered an ideal (n, M, d)-code, but these are conflicting aims, and such one value has to be optimized in terms of the other two. The large size, M, contributes to the transmission of diverse data and a large minimum distance, d, allows for more errors to be corrected.

2. An Introduction to Algebra

For the ease of analysing error correction codes, algebraic structures are enforced on code alphabets. With the help of the alphabet, it makes it particularly convenient to perform addition, subtraction, multiplication, and division without restriction. This is done by giving the alphabet, such as $\mathbf{F}_q$, the structure of a field. From prior knowledge, it is known that a field such as $\mathbb{R}$ or a complex field such as $\mathbb{C}$ consists of two operations, specifically addition and multiplication. Hence, $\mathbb{C}$ and $\mathbb{R}$ are fields that contain infinitely many elements, whereas $\mathbf{F}_q$ is a field consisting of only finitely many elements.

2.1. Groups

Definition 2.1. *A group is a nonempty set G, along with a binary operation $*$ where the following properties are satisfied:*

1. $\forall\ a, b, c \in G\ (a * b) * c = a * (b * c)$ *(associativity)*
2. $\exists\ e \in G$ *such that* $e * a = a * e = a$, $\forall\ a \in G$ *(identity)*
3. $\forall\ a \in G$, $\exists\ a^{-1}$ *such that* $a * a^{-1} = a^{-1} * a = e$*(inverse)*

Definition 2.2. *A group G is called an* abelian *group if $a * b = b * a$, $\forall\ a, b \in G$.*

Definition 2.3. *A* subgroup *S of a group G, denoted by $S \leq G$, is a subset of G that is a group under the same operation which is defined on G.*

Example 2.4. *The set $\mathbb{Z}_n = \{0, 1, ..., n-1\}$ is an abelian group under addition modulo n, $a + b = b + a \mod n$. The set $\mathbb{Z}_m$, where $m < n$ is a subgroup of $\mathbb{Z}_n$*

A group G is defined to be finite if it contains a finite number of elements. The *order* of a finite group G, represented by $\mid G \mid$, is the cardinality of this finite group. If $a \in G$, then the smallest positive k for which $a^k = e$ is called the order of the element a, denoted by $|a|$.

Theorem 2.5. *Let G be a group and let $a \in G$. $a^k = e \iff |a| \mid k$.*

Theorem 2.6. *(Lagrange's Theorem) Let G be a finite group, let $H \leq G$ and $g \in G$.*

1. $|H| \mid |G|$.
2. $|g| \mid |G|$.

2.2. Rings

Definition 2.7. *A nonempty set of elements involving two binary operations, namely addition (+) and multiplication ($\cdot$ or juxtaposition) is called a ring if it satisfies the following conditions.*

1. *R is an abelian group under addition.*
2. *$\forall\ a, b, c \in R, (ab)c = a(bc)$ (associativity)*
3. *$\forall\ a, b, c \in R, (a+b)c = ac + bc$ and $c(a+b) = ca + cb$ (distributivity)*

Example 2.8. *The following are rings:*

1. *Rings with infinite elements: $\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}$, i.e., the integers.*
2. *Rings with finite elements: $\mathbb{Z}_4 = \mathbb{Z}/4\mathbb{Z} = \{0, 1, 2, 3\}$*

Definition 2.9.

1. *A ring R is called a ring with identity, or unity, if $\exists\ 1 \in R$ such that $a1 = 1a = a$, $\forall\ a \in R$.*
2. *A ring R is called a commutative ring if $ab = ba\ \forall\ a, b \in R$.*
3. *A ring R with identity is called an integral domain if it is commutative and there are no zero divisors in R. A zero divisor $a \in R$ is a nonzero element for which $ab = 0, b \neq 0, b \in R$.*

Example 2.10. *In $\mathbb{Z}_6 = \{0, 1, 2, 3, 4, 5\}$, 2 and 3 are zero divisors as $2 \cdot 3 = 0$ in $\mathbb{Z}_6$.*

Definition 2.11. *A subring, S, of a ring R is a subset S of R that is a ring under the same operations which are defined on R.*

2.3. Fields

Definition 2.12. *A field $\mathbf{F}$ is a commutative ring with identity in which each nonzero element has an inverse. This means the ring must also satisfy the following condition:*
$\forall\ a \neq 0 \in \mathbf{F}$, $\exists\ a^{-1} \in \mathbf{F}$ (multiplicative inverse) such that $a \cdot a^{-1} = 1$.

In other words, a ring with identity $1 \neq 0$ is a field if $\mathbf{F} \setminus \{0\}$ is an abelian group under multiplication.

Example 2.13. *The following are fields:*

1. *Fields with infinite elements: $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, i.e., the rational numbers, real numbers, and complex numbers.*

2. *Fields with finite elements:* $\mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z} = \{0, 1\}$ *and* $\mathbb{Z}_{11} = \mathbb{Z}/11\mathbb{Z} = \{0, 1, 2, ..., 10\}$.

Lemma 2.14. *Let* $a, b \in \mathbf{F}$*, where* $\mathbf{F}$*is a field. Then*

1. $a0 = 0 \quad \forall a \in \mathbf{F}$

2. $ab = 0 \implies a = 0$ *or* $b = 0$

Proof:

1. $a0 = a(0 + 0) = a0 + a0$. Adding the additive inverse of $a0$ to both sides yields $0 = a0 + (-a0) = a0 + a0 + (-a0) = a0 + 0 = a0$

2. Suppose $ab = 0$. If $a \neq 0$, then a has a multiplicative inverse and so $b = 1b = (a^{-1}a)b = a^{-1}0 = 0$.

□

For any integer $m \geq 2, \mathbb{Z}_m$ is a ring, as the conditions in the definition 2.7 are satisfied. $\mathbb{Z}_m$ is called the ring of integers modulo m. However, $\mathbb{Z}_m$ is not always a field.

Theorem 2.15. *In general,* $\mathbf{F}_p = \mathbb{Z}/p\mathbb{Z}$ *is a field, where* p *is a prime number.*

Proof: When m *is not a prime:* Assuming m is a composite number, let $m = ab$, where $1 < a, b < m$ Thus, $ab \equiv 0(\mod m)$, with $a \neq 0, b \neq 0$. However, $0 = m = a \cdot b \in \mathbb{Z}_m$. The product of the non-zero elements, a and b, is zero. This contradicts lemma 2.14, which states that $ab = 0 \implies a = 0$ or $b = 0$. Hence $\mathbb{Z}_m$ is not a field.

When m *is a prime:* It is sufficient to show that $\mathbb{Z}_m$ has a multiplicative inverse for every nonzero element. Let $a \in \mathbb{Z}_m$, since m is prime, it is relatively prime to every number, including a. $\exists\ c, d \in \mathbb{Z}$ such that $ca + dm = 1 \implies ca \equiv 1(\mod m)$. Thus, $c = a^{-1}$. □

Definition 2.16. *A field with finite elements is called a finite field or a Galois field. This finite number of elements is defined to be the order of said Galois field.*

Theorem 2.17. *A Galois field of order* q *is represented as* $GF(q)$*. A field of order* q *exists if and only if* q *is a prime power. Moreover, only one field of that order exists if* q *is a prime power.*

Definition 2.18. *Given that* $a, b, m \in \mathbb{Z}$ *and* $m > 1$*. Then, two integers* a *and* b *are congruent (modulo* m*) if* $m \mid a - b$*; i.e.,* $a - b$ *is divisible by* m*. We denote congruence by:* $a \equiv b \mod m$

When a and b are not congruent (modulo m), then it is written as $a \not\equiv b \mod m$. If an integer is divided by m, it has a remainder equal to one of the

integers in the set $\mathbb{Z}_m = 0, 1, ..., m-1$. Hence, two integers are known to be congruent if and only if division by m results in the same remainders from $\mathbb{Z}_m$.

Example 2.19.

1. $14 \equiv 4 \mod 10$
2. $a \equiv 0 \mod m$ *means that* $m \mid a$.
3. $a \equiv 0 \mod 2$ *means that* a *is even.*
4. $a \equiv 1 \mod 2$ *means that* a *is odd.*

Theorem 2.20. *Assuming* $a \equiv c \mod m$ *and* $b \equiv d \mod m$. *Then:*

1. $a + b \equiv c + d \mod m$
2. $ab \equiv cd \mod m$

Definition 2.21. *Let* $\mathbf{F}$ *be a given field. The characteristic of* $\mathbf{F}$, *denoted by* $ch(\mathbf{F})$, *is defined as the smallest positive integer* n *such that* $n * 1_{\mathbf{F}} = 0$. *If such an integer does not exist we write* $ch(\mathbf{F}) = 0$.

Theorem 2.22. *Let* $\mathbf{F}$ *be a given field. Then* $ch(\mathbf{F})$ *is either* 0 *or a prime* p.

Example 2.23. *We have the following cases:*

1. $ch(\mathbb{Q}) = 0$, *where* $\mathbb{Q}$ *is the rational number field.*
2. *Let* p *be a prime number. Then* $ch(\mathbf{F}_p) = p$.

2.4. Homomorphisms

A *homomorphism* from one ring, R, into another ring, S, is a mapping $\phi : R \to S$ which preserves addition and multiplication. The formal definition is more specific:

Definition 2.24. *Let* R *and* S *be rings (or fields). A function* $\phi : R \to S$, *where* $\phi(a + b) = \phi(a) + \phi(b)$ *and* $\phi(ab) = \phi(a)\phi(b)$, $\forall\ a, b \in R$, *is a homomorphism.*

An *isomorphism* is a homomorphism that is bijective (injective and surjective). If $\phi : R \to S$ is an isomorphism, we say that the rings, R and S, are isomorphic.

When taking into account the algebraic structure of two rings or fields that are defined to be isomorphic, then their structure is considered to be essentially the same. In other words, isomorphic rings (or fields) can be thought of as different representations of the same ring (or field).

2.5. Polynomials

We will often associate the vector $(a_0 a_1 \cdots a_{n-1})$ with the polynomial $f(x) = a_0 + a_1 x + ... + a_{n-1}x^{n-1}$ under the following isomorphic map:

$$\pi : (\mathbf{F}_q)^n \to \mathbf{F}_q[x]/(x^n - 1)$$
$$(a_0 a_1 ... a_{n-1}) \to a_0 + a_1 x + ... + a_{n-1}x^{n-1} = a(x)$$

If $a_{n-1} \neq 0$, $n-1$ is the *degree* of the polynomial, denoted by $\deg(f(x))$. The *leading coefficient* is a_{n-1}, and if $a_{n-1} = 1$ then the polynomial is said to be *monic*.

Example 2.25. *The polynomials associated with the set* $\{000, 110, 101, 001\}$ *are* 0, $1 + x$, $1 + x^2$, *and* $x + x^2$.

The set of all polynomials over a field $\mathbf{F}$, forms a ring under addition and multiplication of polynomials.
$F[x] = \{f(x) = a_0 + a_1 x + ... + a_{n-1}x^{n-1} | a_i \in \mathbf{F}, \forall i = 0, 1, 2, ..., n-1,\ n \in \mathbb{N}\}$

Remark 2.26. *$F[x]$ is not a field as polynomials of degree larger than zero have no multiplicative inverses.*

2.6. The Division Algorithm

By the division algorithm, for every pair of polynomials $a(x), b(x) \in F[x]$, $b(x) \neq 0$, $\exists\ q(x), r(x)$ such that $a(x) = b(x)q(x) + r(x)$ where $r(x) = 0$ or $\deg(r(x)) < \deg(b(x))$. $q(x)$ and $r(x)$ are called the quotient and remainder, respectively.

This is similar to the quotient and remainder in the division algorithm frequently used with the ring of integers, $\mathbb{Z}$. The polynomials are obtained by division.

Example 2.27. *We can divide* $x^4 + x + 1$ *by* $x^2 + 1$ *in* $\mathbf{F}_2[x]$ *using long division as follows.*

$$\begin{array}{rcl|l} & & x^2 + 1 & \\ x^2 + 1 & & x^4 + x + 1 & \\ & & x^4 + x^2 & \\ \hline & & x^2 + x + 1 & \\ & & x^2 + 1 & \\ \hline & & x & \end{array}$$

From the division algorithm we can see that $x^4 + x + 1 = (x^2 + 1)(x^2 + 1) + x$ *in* $\mathbf{F}_2[x]$.

2.7. Ideals

Definition 2.28. *An ideal is a subset* I *of a ring* R *which satisfies the following conditions:*

1. $a, b \in I \implies a - b \in I$.

2. $a \in R$, $i \in I \implies ai \in I$ *and* $ia \in I$.

When S is a nonempty subset of a ring R, the smallest ideal I of R containing S is called the ideal *generated* by S. Let R is a commutative ring with identity, if $a \in R$, then the ideal generated by a is the set $(a) = \{ra \mid r \in R\}$.

An ideal of this form is called a principal ideal.

Definition 2.29. *If every ideal of a ring,* R*, is principal,* R *is called a* principal ideal domain.

2.8. Factor Rings

Let R be a ring and I be an ideal in R. Then for each $a \in R$, we can form a coset of I, denoted by $a+I = \{a+i \mid i \in I\}$. It is known that $a+I = b+I \iff a - b \in I$, and that any two cosets $a + I$ and $b + I$ are either identical or disjoint. Moreover, the set of all distinct cosets can be interpreted as ring with addition and multiplication defined as follows:

1. $(a + I) + (b + I) = (a + b) + I$

2. $(a + I)(b + I) = ab + I$

The *factor ring* is the ring of all cosets of an ideal, I, denoted by R/I, read as "R mod I."

Example 2.30. *The ring* $\mathbb{Z}_n$ *is isomorphic to the factor ring* $\mathbb{Z}/(n)$*. Addition and multiplication are defined as normal under modulo* n*. The ring* $\mathbb{Z}_n$ *will only be a field if* n *is a prime number.*

Definition 2.31. *A* maximal *ideal* I*, is an ideal* I*,* $I \neq R$*, such that whenever* J *is an ideal and* $I \subseteq J \subseteq R$*, then* $J = I$ *or* $J = R$.

Theorem 2.32. *Suppose* R *is a commutative ring with identity. Then* I *is a maximal ideal if and only if the factor ring* R/I *is a field.*

When $p(x)$ is a polynomial over F_q, then we have the following factor ring: $K = \frac{F_q[x]}{(p(x))} = \{f(x) + I \mid f(x) \in \mathbf{F}_q[x]\}$

We need only consider $r(x) = 0$ or polynomials $f(x)$, where $\deg(f(x)) < \deg p(x)$. Whenever $\deg(f(x)) \geq \deg(p(x))$, then dividing $f(x)$ by $p(x)$ using the division algorithm yields

$$f(x) = q(x)p(x) + r(x)$$

where $\deg(r(x)) < \deg(p(x))$. But as $q(x)p(x) \in I$, we have $q(x)p(x) + I = 0 + I$. Therefore, $f(x) + I = r(x) + I$ and $K = \{r(x) + I \mid \deg(r(x)) < \deg(p(x))\}$.

In other words, the factor ring K may be identified with the set of all polynomials of degree less than $\deg(p(x))$, with addition and multiplication performed modulo $p(x)$.

Remark 2.33. *$r(x) + I$ will be denoted by $\overline{r(x)}$ or simply $r(x)$ when it is clear.*

2.9. Characteristic of a Ring

The characteristic of the ring, denoted by $\text{char}(R)$, is the smallest positive integer, n, for which $na = a + a + ... + a = 0$ for all $a \in R$. If no such integer n exists so that $na = 0$, we have $\text{char}(R) = 0$.

Example 2.34. *Let us look at the characteristics of the the ring $\mathbb{Z}_n$ and the ring of integers $\mathbb{Z}$. We can easily see that $char(R) = n$ and $char(\mathbb{Z}) = 0$.*

Theorem 2.35. *The characteristic of an integral domain is either 0 or a prime number. The characteristic of a finite field is always a prime number.*

Theorem 2.36. *Let the characteristic of a commutative ring, R, be a prime number, p. $q = p^n \implies (a + b)^q = a^q + b^q$ and $(a - b)^q = a^q - b^q$*

2.10. Irreducible Polynomials

Definition 2.37. *An* irreducible *polynomial is a nonconstant polynomial $f(x) \in \mathbf{F}[x]$, such that whenever $f(x) = p(x)q(x)$, then either $p(x)$ or $q(x)$ must be a constant in $\mathbf{F}$. A* reducible *polynomial is a polynomial that can be factored into two polynomials of a lesser degree. Mathematically, $f(x) \in F[x]$, such that $f(x) = a(x)b(x)$ where $\deg(a(x)) < \deg(f(x))$ and $\deg(b(x)) < \deg(f(x))$.*

Lemma 2.38.

1. $f(a) = 0 \iff f(x) = (x - a)g(x)$
2. $x^n - 1 = (x - 1)(x^{n-1} + x^{n-2} + ... + x + 1)$

Theorem 2.39. *Let $f(x)$ be an irreducible polynomial, $f(x) \mid p(x)q(x) \implies f(x) \mid p(x)$ or $f(x) \mid q(x)$.*

Theorem 2.40. *Any nonconstant polynomial in $F[x]$ can be written uniquely (up to permutation) as a product of irreducible polynomials.*

Theorem 2.41. *$p(x)$ is irreducible if and only if the ideal generated by it, $I = (p(x)) \in F[x]$, is maximal.*

2.11. The Ring of Polynomials modulo f(x)

Similarly to how we can consider the ring of integers modulo m, we can also consider $F[x]$ modulo a polynomial $f(x)$, where $f(x)$ is a polynomial in $F[x]$. Two polynomials $g(x)$ and $h(x)$ are said to be congruent modulo $f(x)$ if $f(x) \mid (g(x) - h(x))$. Similarly, we denote this congruence by $g(x) \equiv h(x) \mod f(x)$. By the division algorithm $a(x) = q(x)b(x) + r(x)$ we can see that $a(x) \equiv r(x) \mod b(x)$.

The set of polynomials in $F[x]$ of degree less than $\deg(f(x))$ is denoted by $F[x]/f(x)$. Suppose $a(x) + I, b(x) + I \in F[x]/f(x)$. Addition and multiplication in this set are defined as follows:

1. $a(x) + b(x)$ in $F[x]/f(x)$ is the same as $a(x) + b(x)$ in $F[x]$, as $\deg(a(x) + b(x)) < \deg(f(x))$.

2. $a(x)b(x)$ in $F[x]/f(x)$ is the polynomial in $F[x]/f(x)$ to which $a(x)b(x)$ is congruent modulo $f(x)$.

$F[x]/f(x)$ is a ring, it is called the *ring of polynomials (over F) modulo $f(x)$*. If $f(x) \in F_q[x]$ is of degree n, then the ring $F[x]/f(x)$ consists of polynomials of degree $< n$. The order of the ring of polynomials modulo $f(x)$ is q^n, as every one of the n coefficients of its polynomials belongs to $\mathbf{F}_q$, we may write $\mid \mathbf{F}_q[x]/f(x) \mid = q^n$.

Example 2.42. *In order to find the product of x and $x+1$ in $F_2[x]/(x^2+1)$ we multiply x with $x+1$ in $F[x]$.*

$x(x+1) = x^2 + x \equiv x + 1$ *in* $F_2[x]/(x^2+1)$

In fact, the addition and multiplication tables of $F_2[x]/(x^2+1)$ can easily be computed.

$+$	0	1	x	$x+1$
0	0	1	x	$x+1$
1	1	0	$x+1$	x
x	x	$x+1$	0	1
$x+1$	$x+1$	x	1	0

$\cdot$	0	1	x	$x+1$
0	0	0	0	0
1	0	1	x	$x+1$
x	0	x	1	$x+1$
$x+1$	0	$x+1$	$x+1$	0

Corollary 2.43. *$f(x)$ is irreducible in $F[x] \iff F[x]/f(x)$ is a field.*

Example 2.44. *$x^2 + x + 1$ is irreducible in $F_2[x]$.*

The addition and multiplication tables of $F_2[x]/(x^2+x+1)$ can easily be computed to show it is a field, as each non-zero element has a corresponding multiplicative inverse.

$+$	0	1	x	$x+1$
0	0	1	x	$x+1$
1	1	0	$x+1$	x
x	x	$x+1$	0	1
$x+1$	$x+1$	x	1	0

$\cdot$	0	1	x	$x+1$
0	0	0	0	0
1	0	1	x	$x+1$
x	0	x	$x+1$	1
$x+1$	0	$x+1$	1	0

Remark 2.45. *If $F[x]/f(x)$ is a field, then we have* $\dim(F(x)/f(x)) = \deg(f(x))$.

2.12. Vector Spaces

Definition 2.46. *A set V is a vector space over a finite field $\mathbf{F}$ if it satisfies the axioms of a vector space. $\forall$ $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbf{V}$ and $\forall a, b \in \mathbf{F}$:*

1. $\mathbf{u} + \mathbf{v} \in \mathbf{V}$
2. $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
4. The zero vector $\mathbf{0} = (0, 0, ..., 0) \in \mathbf{V}$ and $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$
5. Given $\mathbf{u} = (u_1, u_2, ..., u_n) \in \mathbf{V}$, the element $\mathbf{u} = (-u_1, -u_2, ..., -u_n) \in \mathbf{V}$ and satisfies $\mathbf{u}+(\text{-}\mathbf{u}) = \mathbf{0}$.
6. $a\mathbf{v} \in \mathbf{V}$ (closure under scalar multiplication)
7. $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$, $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$ (Distributive laws)
8. $(ab)\mathbf{u} = a(b\mathbf{u})$
9. $1\mathbf{u} = \mathbf{u}$, where 1 is the multiplicative identity of $\mathbf{F}$.

Remark 2.47. *When $\mathbf{F}$ is not a field but a ring, and the aforementioned properties of vector spaces are still satisfied, we can say that $\mathbf{V}$ is a module over the ring $\mathbf{F}$, also known as an $\mathbf{F}$-module.*

Definition 2.48. *Let M be an R-module. A nonempty subset N of M is called an R-submodule if the following two conditions are satisfied.*

1. $n_1 + n_2 \in N \quad \forall n_1, n_2 \in N$

2. $rn \in N \quad \forall r \in R, n \in N$

Assuming q is a prime power, a finite field with q elements is represented as $\mathbf{F}_q$, where all of its elements will be known as scalars.

The set $(\mathbf{F}_q)^n$ of all ordered n-tuples over $\mathbf{F}_q$ contains elements called *vectors.*

Within $(\mathbf{F}_q)^n$ we have two operations:

1. Addition of vectors: Suppose $\mathbf{x} = (x_1, x_2, ..., x_n)$ and $\mathbf{y} = (y_1, y,2..., y_n) \in (\mathbf{F}_q)^n$, then: $\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n)$

2. Multiplication of vector by a scalar: Suppose $\mathbf{x} = (x_1, x_2, ..., x_n) \in (\mathbf{F}_q)^n$ and $a \in \mathbf{F}_q$, then: $a\mathbf{x} = (ax_1, ax_2, ..., ax_n)$

Then we can see that $(\mathbf{F}_q)^n$ is a vector space.

Definition 2.49. *A* subspace *of $(\mathbf{F}_q)^n$ is a subset of $(\mathbf{F}_q)^n$, given that it is a vector space under similar scalar multiplication and addition as defined for the set $(\mathbf{F}_q)^n$.*

Remark 2.50. *Furthermore, the space $(\mathbf{F}_q)^n$ and as well as the set $\{\boldsymbol{0}\}$ are said to be subspaces of $(\mathbf{F}_q)^n$. A non-trivial subspace is one that consists of at least one vector that is different than $\boldsymbol{0}$.*

Theorem 2.51. *In $(\mathbf{F}_q)^n$, a non-empty subset C is a subspace if and only if satisfies the following two conditions:*

1. $\boldsymbol{x}, \boldsymbol{y} \in C \implies \boldsymbol{x} + \boldsymbol{y} \in C$

2. $a \in \mathbf{F}_q$ *and* $\boldsymbol{x} \in C, \implies a\boldsymbol{x} \in C$

Definition 2.52. *A* linear combination *of vectors is a vector in the form of $a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + ... + a_r\boldsymbol{v}_r$, where a_i is a scalar in $\mathbf{F}_q$, and $\boldsymbol{v}_i \in (\mathbf{F}_q)^n$*

Remark 2.53. *The set of all linear combinations of a set of vectors of $(\mathbf{F}_q)^n$ is a subspace of $(\mathbf{F}_q)^n$.*

Definition 2.54. *A set of vectors $\{v_1, v_2, ..., v_r\}$ is said to be linearly independent if whenever $a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + ... + a_r\boldsymbol{v}_r = \boldsymbol{0}$ then $a_1 = a_2 = a_r = 0$.*

When a set of vectors is not linearly independent it is said to be *linearly dependent.*

Example 2.55.

1. *Any set, S, that contains $\mathbf{0}$ is linearly dependent.*
2. $\{(0,0,0,1),(0,0,1,0),(0,1,0,0)\}$ *is linearly independent over* $\mathbf{F}_q$.
3. $\{(0,0,0,1),(1,0,0,0),(1,0,0,1)\}$ *is linearly dependent over* $\mathbf{F}_q$.

Assuming that C is a subspace of $(\mathbf{F}_q)^n$, then the subset $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r\}$ of C is said to be a *spanning set* or a *generating set* of C, given that all the vectors in C are a linear combination of $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r$.

Definition 2.56. A basis *of C is a generating set of C that is linearly independent.*

Example 2.57. *A basis of the whole space $(\mathbf{F}_q)^n$ can be expressed by*

$$\{(1,0,0,...,0),(0,1,0,...,0),...,(0,0,...,0,1)\}$$

Theorem 2.58. *If C is a non-trivial subspace of $(\mathbf{F}_q)^n$ then a basis of C must be contained in any generating set of C.*

Theorem 2.59. *Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, ..., \boldsymbol{v}_r\}$ be a basis of a subspace C of $(\mathbf{F}_q)^n$. Then:*

1. *Every vector of C can be uniquely expressed as a linear combination of the vectors in the basis*
2. $|C| = q^r$

Remark 2.60. *Any two bases of C contain the same number of vectors, this is the* dimension *of C, denoted by $dim(C)$, or otherwise referred to as the* rank. *Also note that $dim((\mathbf{F}_q)^n) = n$.*

3. Linear Codes and Cyclic Codes over Finite Fields

3.1. Linear Codes

In this section, we shall assume the alphabet of the codes is a finite field $\mathbf{F}_q$, specifically, the Galois Field $GF(q)$, where q is a prime power. $(\mathbf{F}_q)^n$ is the vector space $V(n,q)$, introduced in the previous section. Vectors, previously written as $\mathbf{x} = (x_1, x_2, ..., x_n)$, will now be represented in a minimalist manner by $\mathbf{x} = (x_1 x_2 \cdots x_n)$.

Definition 3.1. *A* linear code *of length n and rank k, denoted by $[n,k]$, is a k-dimensional linear subspace of the vector space $\mathbf{F}_q^n$ where $\mathbf{F}_q$ is the finite field with q elements.*

Remark 3.2. *The codewords are the vectors in $(\mathbf{F}_q)^n$.*
A linear code, C, of length n, dimension k, and minimum distance d, will be denoted by $[n,k,d]$.

Recall that, in order to fully qualify as a linear code, it means that when C is a subset of $(\mathbf{F}_q)^n$, C must satisfy the following conditions:

1. $\mathbf{0} \in C$
2. $\mathbf{u} + \mathbf{v} \in C, \quad \forall\ \mathbf{u}, \mathbf{v} \in C$
3. $a\mathbf{u} \in C, \quad \forall\ \mathbf{u} \in C, a \in (\mathbf{F}_q)^n$

In other words, it must be nonempty, closed under addition, and scalar multiplication.

Definition 3.3. *The number of non-zero elements of a vector $\boldsymbol{x}$ in $(\mathbf{F}_q)^n$ is defined as the* weight *of $\boldsymbol{x}$, denoted by $w(\boldsymbol{x})$.*

Lemma 3.4. $d(\boldsymbol{x}, \boldsymbol{y}) = w(\boldsymbol{x} - \boldsymbol{y}), \forall\ \boldsymbol{x}, \boldsymbol{y} \in (\mathbf{F}_q)^n$

Proof: The vector $\mathbf{x} - \mathbf{y}$ has non-zero entries where $\mathbf{x}$ and $\mathbf{y}$ are different. □

Theorem 3.5. *Let C be a linear code and let $w(C) = min\{w(x) | x \in C, x \neq 0\}$, then $d(C) = w(C)$.*

Proof: Since C is a linear code, then $\mathbf{x} - \mathbf{y} \in C\ \forall\ \mathbf{x}, \mathbf{y} \in C$. Suppose $d(C) = d(\mathbf{x}, \mathbf{y})$, for some $\mathbf{x}$ and $\mathbf{y} \in C$ such that $d(C) = d(\mathbf{x}, \mathbf{y})$. By lemma 3.4, $d(C) = w(\mathbf{x} - \mathbf{y}) \geq w(C)$. Let $w(x) = w(C)$, $d(C) \leq d(\mathbf{x}, \mathbf{0}) = w(\mathbf{x}) = w(C)$. Clearly, this means $d(C) = w(C)$. □

3.1.1. Encoding.

Definition 3.6. *A* generator matrix *of a linear $[n, k]$-code is a $k \times n$ matrix whose rows form a basis of that code.*

Theorem 3.7. *Two $k \times n$ matrices generate equivalent codes if one can be obtained from the other using operations which preserve linear independence.*

Proof: Operations which preserve linear independence replace the basis by another of the same code. □

Theorem 3.8. *A generator matrix, G, of an $[n, k]$-code can be transformed into* standard form

$$G = [I_k|A], \tag{3.1}$$

using operations which preserve linear independence. In equation 3.1, I_k is the $k \times k$ identity matrix, and A is a $k \times (n - k)$ matrix.

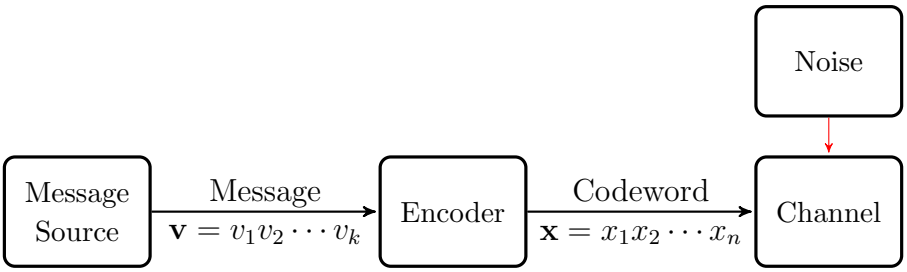

Figure 3.1: Encoding Procedure

As indicated in figure 3.1, the sent message is divided by blocks. Each message block is of length k, which is encoded as a codeword of length $n \geq k$.

Richard Hamming introduced the [7,4] Hamming code in 1950. It encodes four bits into seven bits, applying three bits for redundancy using the following generator matrix:

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Example 3.9. *If we were to encode* $(1,1,0,1)$ *using the above generator matrix we would get:*

$$(1,1,0,1)\begin{bmatrix}1&0&0&0&0&1&1\\0&1&0&0&1&0&1\\0&0&1&0&1&1&0\\0&0&0&1&1&1&1\end{bmatrix} = (1,1,0,1,0,0,1)$$

Definition 3.10. *The* parity check matrix *for a code* C*,* H*, is the* $(n-k)\times n$ *matrix of in the form of:*

$$H = [-A^T|I_{n-k}], \tag{3.2}$$

where A^T *is the transpose of matrix* A*, such that*

$$H\left(x_1x_2\cdots x_n\right) = 0,$$

granted there are no mistakes. The computation is performed in the given field $\mathbf{F}_q$ *and* $\boldsymbol{x}$ *is the received vector.*

Example 3.11. *The parity-check matrix of Hamming (7,4)*

$$H = \begin{bmatrix}0&1&1&1&1&0&0\\1&0&1&1&0&1&0\\1&1&0&1&0&0&1\end{bmatrix}$$

Remark 3.12. *In general, a parity check matrix may not have the form given in (3.2).*

Definition 3.13. *The* dual code *of a linear* $[n,k,d]$*-code* C*, denoted by* $C^\perp$*, is a linear code of dimension* $n-k$*.* $C^\perp = \{\boldsymbol{v}\in(\mathbf{F}_q)^n \mid \boldsymbol{v}\cdot\boldsymbol{u} = 0\ \forall\ \boldsymbol{u}\in C\}$*. Where* $\boldsymbol{v}\cdot\boldsymbol{u}$ *is the dot product of* $\boldsymbol{v}$ *and* $\boldsymbol{u}$*.*

3.2. Introduction to Cyclic Codes

Cyclic codes are important and thoroughly studied for copious reasons. One of these is their ease of encoding and decoding, while another is their mathematical construction.

Definition 3.14. *A linear code* C *is* cyclic *if any cyclic shift of a codeword,* c*, denoted by* $T(c)$ *is also a codeword.*

$$c{=}\ (a_0a_1\cdots a_{n-1})\in C \implies T(c) = (a_{n-1}a_0a_1\cdots a_{n-2})\in C$$

Example 3.15. *The binary linear code* $C = \{000,110,101,011\}$ *is cyclic.*

3.3. Generator Polynomial

In section 2.5 we associated the vector $(a_0 a_1 \cdots a_{n-1})$ with the polynomial $f(x) = a_0 + a_1 x + ... + a_{n-1}x^{n-1}$. Now, we will associate a vector similarly, but the polynomial will be in $F[x]/(x^n - 1)$. Notice that multiplication by x corresponds to a single cyclic shift.

$$\pi(a_0 a_1 ... a_{n-1}) \rightarrow a_0 + a_1 x + ... + a_{n-1}x^{n-1} = a(x)$$

Applying a cyclic shift yields

$$\begin{aligned} \pi(a_{n-1} a_0 ... a_{n-2}) &\rightarrow a_{n-1} + a_0 x + a_1 x + ... + a_{n-2}x^{n-1} \\ &= x(a_0 + a_1 x + ... + a_{n-1}x^{n-1}) \\ &= x \cdot a(x) \in F[x]/(x^n - 1) \end{aligned}$$

Theorem 3.16. *A code C in $F[x]/(x^n - 1)$ is a cyclic code if and only if it is an ideal of $F[x]/(x^n - 1)$.*

Proof: Let C be an ideal of $F[x]/(x^n - 1)$. For any $a(x), b(x) \in C$ and $r(x) \in \mathbf{F}_q[x]/(x^n - 1)$ we have $a(x) + b(x) \in C$ and $r(x)a(x) \in C$. If $r(x)$ is a scalar, then C is a linear code. Let the polynomial

$$a(x) = a_0 + a_1 x + ... + a_{n-2}x^{n-2} + a_{n-1}x^{n-1}$$

be an element of C. Then we can see that

$$\begin{aligned} x \cdot a(x) &= a_0 x + a_1 x^2 + ... + a_{n-2}x^{n-1} + a_{n-1}x^n \\ &= a_{n-1} + a_0 x + a_1 x^2 + ... + a_{n-2}x^{n-1} + a_{n-1}(x^n - 1) \\ &= a_{n-1} + a_0 x + a_1 x^2 + ... + a_{n-2}x^{n-1} \end{aligned}$$

is in C. Thus C is cyclic.

Conversely, suppose C is a cyclic code in $F_q[x]/(x^n - 1)$. Then C is a linear code and therefore $a(x) + b(x) \in C$. We know that for $a(x) \in C$, $x \cdot a(x) \in C$ as it corresponds to a cyclic shift. By induction, assume $x^k \cdot a(x) \in C, k \geq 0$, it is easy to see that $x^{k+1} \cdot a(x)$ corresponds to the next cyclic shift after k. Thus, by the induction principle, C is an ideal of $\mathbf{F}_q[x]/(x^n - 1)$ □

Definition 3.17. *Let $f(x)$ be a polynomial in $F[x]/(x^n - 1)$, then $(f(x))$ denotes the subset of $F[x]/(x^n - 1)$ that contains all the multiples of $f(x)$ modulo $x^n - 1$.* $(f(x)) = \{a(x)f(x) \mid a(x) \in F[x]/(x^n - 1)\}$

Theorem 3.18. *Let C be a cyclic code of length n in $\mathbf{F}_q[x]/(x^n - 1)$. Then*

1. *there is a unique monic polynomial $g(x)$ of smallest degree in C*
2. *$g(x)$ is the* generator polynomial *of C; $C = (g(x))$*

3. $g(x) \mid x^n - 1$

4. *if* $g(x) = g_0 + g_1x + ... + g_rx^r$ *is the generator polynomial, then* $g_0 \neq 0$

5. $\deg(g(x)) = r \implies \dim(C) = n - r$

6. *if* $g(x)$ *is the generator polynomial of degree* r*, then a generator matrix for* C *is given by*

$$G = \begin{bmatrix} g_0 & g_1 & g_2 & \cdots & g_r & 0 & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & g_2 & \cdots & g_r & 0 & \cdots & 0 \\ 0 & 0 & g_0 & g_1 & g_2 & \cdots & g_r & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & g_0 & g_1 & g_2 & \cdots & g_r \end{bmatrix}$$

3.4. Check Polynomial

Let C be a cyclic code, its generator polynomial $g(x)$ must divide $x^n - 1$, and thus $x^n - 1 = g(x)h(x)$ where $h(x)$ is a monic polynomial of degree $n - r$. $h(x)$ is called the *check polynomial* of C.

Theorem 3.19. *Let* C *be a code in* $\mathbf{F}_q[x]/(x^n - 1)$ *and* $h(x) = h_0 + h_1x + ... + h_{n-r}x^{n-r}$ *be its check polynomial.*

1. $C = \{c(x) \in \mathbf{F}_q[x]/(x^n - 1) \mid c(x)h(x) = 0\}$

2. $C^\perp$ *is the cyclic code of dimension* r *generated by the polynomial* $h^\perp(x) = h_0^{-1}(h_{n-r} + h_{n-r-1}x + ... + h_0x^{n-r})$

3. *a parity check matrix for* C *is the following:*

$$H = \begin{bmatrix} h_{n-r} & h_{n-r-1} & h_{n-r-2} & \cdots & h_0 & 0 & 0 & \cdots & 0 \\ 0 & h_{n-r} & h_{n-r-1} & h_{n-r-2} & \cdots & h_0 & 0 & \cdots & 0 \\ 0 & 0 & h_{n-r} & h_{n-r-1} & h_{n-r-2} & \cdots & h_0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & h_{n-r} & h_{n-r-1} & h_{n-r-2} & \cdots & h_0 \end{bmatrix}$$

3.5. Hamming Codes as Cyclic Codes

The $[7, 4]$ Hamming code has the following generator matrix:

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Using operations that preserve the linear independence of the matrix we can get the following equivalent matrix:

$$G' = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

Which can be identified as the generator matrix of a cyclic code, on top of that, it is widely known that the $[7, 4]$ Hamming code is generated by the polynomial $x^3 + x + 1$. The field, $F_2[x]/(x^3 + x + 1) = \{0, 1, x, x^2, x + 1, x^2 + x, x^2 + x + 1, x^2 + 1\}$, is of order 0.

Theorem 3.20. *The binary Hamming code is equivalent to a cyclic code.*

Definition 3.21. *A primitive element of a field, x, is an element such that every element in the field can be written in the form of x^i for some positive integer i.*

Definition 3.22. *$p(x)$ is a primitive polynomial if $p(x)$ is an irreducible polynomial and x is a primitive element of $F_q[x]/f(x)$.*

Theorem 3.23. *Let $p(x)$ be a primitive polynomial over $\mathbb{Z}_2$ and $\deg(p(x)) = r$ then $(p(x))$ generates the binary Hamming code.*

Theorem 3.24. *A q-ary Hamming code of length n, with r redundancy bits, is equivalent to a cyclic code whenever $q - 1$ and r are relatively prime [3].*

3.6. Encoding with Cyclic Codes

There are two simple methods of encoding data using cyclic codes, the non-systematic and systematic method.

3.6.1. The non-systematic method. Let $C = (g(x))$ be a q-ary cyclic code of length n, with $\deg(g(x)) = r$. Then the message to be transmitted must

be of length $n-r$, let this message be in the form of $a_0a_1\cdots a_{n-r-1}$. The *message polynomial* is $a(x) = a_0 + a_1x + ... + a_{n-r-1}x^{n-r-1}$. Then
$C = \{c(x) \mid c(x) = a(x)g(x)\}$

3.6.2. The systematic method. Similarly, let $C = (g(x))$ be a q-ary cyclic code of length n, with $\deg(g(x)) = r$. Then the message to be transmitted must be of length $n-r$, let this message be in the form of $a_0a_1\cdots a_{n-r-1}$. The *message polynomial*, however, is

$$\bar{a}(x) = a_0x^{n-1} + a_1x^{n-2} + ... + a_{n-r-1}x^r$$

Now, let $c(x) = \bar{a}(x) - r(x)$ where $r(x)$ is the remainder of the division of $\bar{a}(x)$ by $g(x)$. If we order the polynomial from highest to lowest degree, we can see that the first $n-r$ bits are the data and the remaining bits are all check symbols.

4. Cyclic Codes over $\mathbb{Z}_4$

4.1. Introduction

In the mid 1960s, the best codes known were the double error-correcting extended Hamming, BCH, and Reed-Muller $(1, 4)$ codes. It is not difficult to show that for a binary code of length 16 and a given minimal distance of 4, 6, or 8, the maximal number of codewords are 2^{11}, 2^8 and 2^5 respectively. While the $[16, 2^{11}, 4]$-code and the $[16, 2^5, 8]$-codes had already been discovered and researched, the existence of a code with $2^8 = 256$ codewords remained unknown, that is, up until J. P. Robinson asked about the possibility of such a code in a high school talk.

A. W. Nordstrom, a fourteen year old student in the audience, felt the need for a challenge, and worked with Robinson to find a binary code of length 16, with 256 codewords, and a minimal distance of 6. The paper describing the now acclaimed Nordstrom-Robinson code was published in 1967. The code had one peculiar aspect to it that made it of certain interest, it is non-linear. In fact, no linear code of length 16 and minimal distance 6 can have more than 128 codewords.

More non-linear codes emerged the following years. In 1968, the Preparata codes were defined. The general Preparata code, $P(m)$, has length 2^m, 2^{2^m-2m} codewords, and a minimal distance of 6 (for even $m \geq 4$). In 1972, the Kerdock codes were classified. They can be described as a sort of dual to the Preparata codes, though not dual as previously defined, as they are not linear. The general Kerdock code, $K(m)$, has length 2^m, 2^{2^m} codewords, and a minimal distance of $2^{m-1} - 2^{(m-2)/2}$, (for even $m \geq 4$). Interestingly, the Nordstrom-Robinson code, the Preparata code $P(4)$, and the Kerdock code $K(4)$ correspond to one another.

Additional non-linear codes and generalizations were discovered by Delsarte, Goethals, and Hergert in the subsequent years. Similar to the Nordstrom-Robinson code, Kerdock codes, and Preparata codes, the Goethals codes and Delsarte-Goethals codes, have more codewords than any comparable linear code. Apart from their error-correcting capabilities these codes are studied for their interesting relations to linear codes. Mathematicians and engineers alike have often wondered whether these codes have a similar underlying algebraic structure. Plenty of researchers have scrutinized these codes and some realized that these codes are mathematically related.

The major development arrived in 1992, with the realization that the Nordstrom-Robinson code is remarkably similar to the *octacode*, a thoroughly studied linear code over $\mathbb{Z}_4$, defined by Sloane in 1992 [4].

Whilst it may seem as a coincidence for some codes, when the way certain codes are defined is altered, as in the paper by Hammons et al. from 1994, it can be seen various families of codes are binary images of a mapping of linear codes in $\mathbb{Z}_4$ [5]. These families of codes are related cyclic codes over $\mathbb{Z}_4$ with more parity check digits.

In 1992, two teams of researchers, Hammons and Kumar, and Calderbank, Sloane, and Solé, started observing the similarities between binary expansions of quaternary codewords and the Kerdock codes. Both teams came to the conclusion that the Kerdock code is the image of a quaternary code. Finally in November 1992, the two teams realized the significant amount of research they had in common, and they started collaborating. The paper they published in 1994 contained a lot of their results, and sparked the interest of others in the field of $\mathbb{Z}_4$ codes.

Working in $\mathbb{Z}_4$, an idea to revisit would be the weight for codes.

Definition 4.1. *The* Lee weight*, wt_L of the elements in $\mathbb{Z}_4$ are as follows*

x	*0*	*1*	*2*	*3*
$wt_L(x)$	*0*	*1*	*2*	*1*

it is extended into $\mathbb{Z}_4^n$ by $wt_L(u) = \sum_{i=1}^{n} wt_L(u_i)$.

4.2. The Construction of Some Non-linear Codes

In order to construct the Nordstrom-Robinson code, Preparata codes, and Kerdock codes, we start with the $[7, 4]$ Hamming code, which has the generator polynomial $g(x) = x^3 + x + 1$, which is a divisor of $x^7 - 1$. We then apply the Hensel lift to $g(x)$ to get the polynomial $G(x) = x^3 + 2x^2 + x - 1$.

Note that in $\mathbb{Z}_2$, $g(x) \equiv G(x)$, but in $\mathbb{Z}_4$, $G(x)$ divides $x^7 - 1 \mod 4$. Notice that $G(x)$ will generate a cyclic code of length 7 over $\mathbb{Z}_4$. We then add a zero-sum check symbol, resulting in a self-dual code of length 8 over $\mathbb{Z}_4$, this is widely known as the *octacode*.

The octacode is a code with 256 codewords and minimal Lee distance 6. In other words, it is an $[8, 256, 6]$-code in $\mathbb{Z}_4$, defined by the generator matrix

$$G = \begin{bmatrix} 3 & 3 & 2 & 3 & 1 & 0 & 0 & 0 \\ 3 & 0 & 3 & 2 & 3 & 1 & 0 & 0 \\ 3 & 0 & 0 & 3 & 2 & 3 & 1 & 0 \\ 3 & 0 & 0 & 0 & 3 & 2 & 3 & 1 \end{bmatrix}$$

Remark 4.2. *As the octacode is self-dual, the generator matrix is also the parity check matrix [6].*

Map this to a binary code under the *Gray map* ϕ defined by:

$\mathbb{Z}_4$	0	1	2	3
$\mathbb{Z}_2^2$	00	01	11	10

and we get the Nordstrom-Robinson code, a nonlinear binary code of length 16 with 256 words and minimal distance 6.

Remark 4.3. *The Gray map preserves the distance between the metric spaces, this is called an* isometry*.*

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

Replacing $g(x)$ by the generator polynomial of a Hamming code of length $2^m - 1$ (m is odd), and repeating the process, yields the Preparata codes. Using the duals of the $\mathbb{Z}_4$-codes we get the Kerdock codes.

Cyclic codes over $\mathbb{Z}_4$ can be generated by $2^l h_1 + h$ elements, where $l = 0$ or 2 and $h_1 \mid h \mid x^n - 1$. Furthermore, incorporating the use of Hensel's Lemma demonstrates that the factorization of $x^n - 1$ for odd n over Z_4 can be found by applying the Hensel lift to the factorization of $x^n - 1$ over Z_2, which can be determined by a methodology called the *Graeffe Method*. Thus, $x^n - 1$ factors into a unique product of monic basic irreducible polynomials over $\mathbb{Z}_4$.

When it comes to the generator polynomial of cyclic codes in $\mathbb{Z}_4[x]/(x^n-1)$, it becomes relatively difficult in the case of even n. In the case of an even n, it is observed that the factorization of $x^n - 1$ over $\mathbb{Z}_4$ does not result in a product of distinct irreducible polynomials. This observation leads to a contradiction of Hensel's Lemma [7].

Theorem 4.4. *Let C be a cyclic code in $\mathbb{Z}_4[x]/(x^n-1)$*

1. *If n is odd, then $\mathbb{Z}_4[x]/(x^n-1)$ is a principal ideal ring and $C = (g(x), 2a(x)) = (g(x) + 2a(x))$ Where $g(x), a(x)$ are polynomials with $a(x) \mid g(x) \mid (x^n - 1) \mod 4$.*

2. *Assume n is even; then either:*

   (a) *C is a free module of generator $C = (g(x)+2p(x))$, Where $g(x) \mid (x^n-1) \mod 2$ and $(g(x) + 2p(x)) \mid (x^n - 1) \mod 4$, or,*

   (b) *$C = (g(x) + 2p(x), 2a(x))$ where $g(x), a(x)$, and $p(x)$ are polynomials with $g(x) \mid (x^n - 1) \mod 4, a(x) \mid g(x) \mod 2, a(x) \mid p(x)\left(\frac{x^n 1}{g(x)}\right) \mod 2$, and $\deg g(x) > \deg a(x) > \deg p(x)$. [8]*

5. Additive Cyclic Codes over $\mathbb{Z}_2\mathbb{Z}_4$

5.1. $\mathbb{Z}_2\mathbb{Z}_4$-additive Codes

Recently, a new family of error-correcting codes has been becoming more popular as it generalizes binary linear codes and quaternary linear codes in one code. This family of codes is called $\mathbb{Z}_2\mathbb{Z}_4$-additive codes. Additive codes were first defined by Delsarte in 1973 [9]. Delsarte defined additive codes in terms of association schemes, which were first introduced by Bose and Shimamoto in 1952 [10]. In an additive Abelian group, F, of order $q \geq 2$, Delsarte focused on the group $X = F^n$, where $n \geq 1$, $n \in \mathbb{Z}$.

Definition 5.1. *An additive code, Y, of length n over F is a subgroup of $X = F^n$*

When working with binary Hamming schemes, X is of order 2^n, and the only structures for X are in the form of $\mathbb{Z}_2^\alpha \times \mathbb{Z}_4^\beta$, where $\alpha + 2\beta = n$ [11].

Definition 5.2. *A $\mathbb{Z}_2\mathbb{Z}_4$-additive code $\mathcal{C}$, is a subgroup of $\mathbb{Z}_2^\alpha \times \mathbb{Z}_4^\beta$*

As any $\mathbb{Z}_2\mathbb{Z}_4$-additive code, $\mathcal{C}$ must be closed under addition, it must also be closed under multiplication with the elements in $\mathbb{Z}_4$. This means that for any $c \in C$, $c = (a_0 a_1 \cdots a_{\alpha-1}, b_0 b_1 \cdots b_{\beta-1})$ and any $n \in \mathbb{Z}_4$ we must have

$$nc = (na_0 na_1 \cdots na_{\alpha-1}, nb_0 nb_1 \cdots nb_{\beta-1}) \in C,$$

where na_i is the product of n and $a_i \mod 2$, $i = 0, 1, ..., \alpha - 1$ and nb_j is the product of n and $b_j \mod 4$, $j = 0, 1, ..., \beta - 1$. Therefore, any $\mathbb{Z}_2\mathbb{Z}_4$-additive code is a $\mathbb{Z}_4$ module (as defined in section 2.12). A similar mapping as used in section 4.2 will be used here, in this case we define an extension of the Gray map as follows

$$\Phi : \mathbb{Z}_2^\alpha \times \mathbb{Z}_4^\beta \to \mathbb{Z}_2^n, \quad n = \alpha + 2\beta$$

$$\Phi(x, y) = (x, \phi(y_1), ..., \phi(y_\beta)) \quad \forall x \in \mathbb{Z}_2^\alpha, \quad \forall y = (y_1, ..., y_\beta) \in \mathbb{Z}_4^\beta$$

ϕ is the Gray map, used in section 4.2

$$\phi : \mathbb{Z}_4 \to \mathbb{Z}_2^2$$

$\mathbb{Z}_4$	0	1	2	3
$\mathbb{Z}_2^2$	00	01	11	10

in other words $\phi(0) = (0,0)$, $\phi(1) = (0,1)$, $\phi(2) = (1,1)$, and $\phi(3) = (1,0)$ This extended Gray map is an isometry, so once again the distance is preserved between these metric spaces, however, in $\mathbb{Z}_2^\alpha \times \mathbb{Z}_4^\beta$ we need to redefine the distance.

Definition 5.3. *For a vector $\boldsymbol{v} = (v_1, v_2) \in \mathbb{Z}_2^\alpha \times \mathbb{Z}_4^\beta$, the weight of $\boldsymbol{v}$, denoted by $wt(\boldsymbol{v})$, is defined by $wt(\boldsymbol{v}) = wt_H(v_1) + wt_L(v_2)$, where $wt_H(v_1)$ is the Hamming weight of v_1 and $wt_L(v_2)$ is the Lee weight of v_2.*

As $\mathcal{C}$ is a subgroup of $\mathbb{Z}_2^\alpha \times \mathbb{Z}_4^\beta$ it must be isomorphic to $\mathbb{Z}_2^\gamma \times \mathbb{Z}_4^\delta$, another Abelian structure. The code, $\mathcal{C}$, has $2^{\gamma+2\delta}$ codewords, as $|\mathcal{C}| = 2^{\gamma+2\delta}$. There are $2^{\gamma+\delta}$ codewords of order two.

Let X be the set of the $\mathbb{Z}_2$ coordinate positions, and let Y be the set of the $\mathbb{Z}_4$ coordinate positions in the code $\mathcal{C}$. Then $|X| = \alpha$ and $|Y| = \beta$. Respecting the order of the vectors in the $\mathbb{Z}_2\mathbb{Z}_4$-additive code, $\mathcal{C}$, we can see that the set X is related to the first α coordinates of $\mathcal{C}$. Likewise, the set Y is related to the last β coordinates of $\mathcal{C}$. Let $\mathcal{C}_X$ be the punctured code of $\mathcal{C}$ by removing the coordinates outside of X. Similarly, let $\mathcal{C}_Y$ be the punctured code of $\mathcal{C}$ by removing the coordinates outside of Y.

Definition 5.4. *Let κ be the dimension of $(\mathcal{C}_b)_X$, where $\mathcal{C}_b$ is the subset of $\mathcal{C}$ which contains all the codewords of order two. When $\alpha = 0$ we will define $\kappa = 0$.*

Remark 5.5. *$(\mathcal{C}_b)_X$ is a binary linear code.*

Taking all the above into consideration, we can say that a code $\mathcal{C}$, or equivalently $C = \Phi(\mathcal{C})$ is of type $(\alpha, \beta, \gamma, \delta, \kappa)$.

Definition 5.6. *Let $\mathcal{C}$ be a $\mathbb{Z}_2\mathbb{Z}_4$-additive code, a subgroup of $\mathbb{Z}_2^\alpha \times \mathbb{Z}_4^\beta$. The binary image $C = \Phi(\mathcal{C})$ is a* $\mathbb{Z}_2\mathbb{Z}_4$-linear code *of type $(\alpha, \beta, \gamma, \delta, \kappa)$, with length $n = \alpha + 2\beta$.*

Remark 5.7. *$\mathbb{Z}_2\mathbb{Z}_4$-linear codes are a generalization of binary linear codes and quaternary linear codes. When $\beta = 0$, C is a binary linear code. Similarly, when $\alpha = 0$, C is a quaternary linear code.*

Definition 5.8. *A vector in C is of the form $(a_0 a_1 \cdots a_{\alpha-1}, b_0 b_1 \cdots b_{\beta-1})$, where $a_i \in \mathbb{Z}_2$ and $b_j \in \mathbb{Z}_4$*

Definition 5.9. *The* inner product *of two vectors $\boldsymbol{u} = (a_0 a_1 \cdots a_{\alpha-1}, b_0 b_1 \cdots b_{\beta-1})$, $\boldsymbol{v} = (d_0 d_1 \cdots d_{\alpha-1}, e_0 e_1 \cdots e_{\beta-1}) \in \mathbb{Z}_2^\alpha \times \mathbb{Z}_4^\beta$, is defined as follows*

$$\begin{aligned} u \cdot v &= (2a_0 d_0 + 2a_1 d_1 + ... + 2a_{\alpha-1} d_{\alpha-1} \\ &\quad + b_0 e_0 + b_1 e_1 + ... + b_{\beta-1} e_{\beta-1}) \mod 4 \\ &= [2 \sum_{i=0}^{\alpha-1} a_i d_i + \sum_{j=0}^{\beta-1} b_j e_j] \mod 4 \end{aligned}$$

Definition 5.10. *The dual code of a $\mathbb{Z}_2\mathbb{Z}_4$-additive code, C, denoted by $C^\perp$, is defined by*

$$C^\perp = \{\boldsymbol{v} \in \mathbb{Z}_2^\alpha \times \mathbb{Z}_4^\beta \mid \boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{0} \quad \forall \boldsymbol{u} \in C\}$$

Theorem 5.11. *Assume C is a $\mathbb{Z}_2\mathbb{Z}_4$-additive code of type $(\alpha, \beta, \gamma, \delta, \kappa)$. Its dual, $C^\perp$, is of type $(\alpha, \beta, \bar{\gamma}, \bar{\delta}, \bar{\kappa})$, where*

$$\bar{\gamma} = \alpha + \gamma - 2\kappa,$$

$$\bar{\delta} = \beta - \gamma - \delta + \kappa,$$

$$\bar{\kappa} = \alpha - \kappa$$

[12]

5.2. $\mathbb{Z}_2\mathbb{Z}_4$-additive Cyclic Codes

Definition 5.12. *A subset C of $\mathbb{Z}_2^\alpha \times \mathbb{Z}_4^\beta$ is called a $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic code if the following two conditions are satisfied*

1. *C is an additive code*

2. *For any codeword $c \in C$,*

$$c = (a_0 a_1 \cdots a_{\alpha-1}, b_0 b_1 \cdots b_{\beta-1}) \in C$$

$$\Longrightarrow$$

$$T(c) = (a_{\alpha-1} a_0 \cdots a_{\alpha-2}, b_{\beta-1} b_0 \cdots b_{\beta-2}) \in C$$

Lemma 5.13. *For any $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic code, C, its dual, $C^\perp$, is also cyclic. [13]*

For an element $c = (a_0 a_1 \cdots a_{\alpha-1}, b_0 b_1 \cdots b_{\beta-1}) \in \mathbb{Z}_2^\alpha \times \mathbb{Z}_4^\beta$ there is a one-to-one correspondence to a module element consisting of two polynomials as follows

$$\begin{aligned} c(x) &= (a_0 + a_1 x + ... + a_{\alpha-1} x^{\alpha-1}, b_0 + b_1 x + ... + b_{\beta-1} x^{\beta-1}) \\ &= (a(x), b(x)) \in R_{\alpha,\beta}, \end{aligned}$$

where $R_{\alpha,\beta} = \mathbb{Z}_2[x]/(x^\alpha - 1) \times \mathbb{Z}_4[x]/(x^\beta - 1)$

Recall that the $\mathbb{Z}_2\mathbb{Z}_4$-additive code is closed under addition and multiplication with elements in $\mathbb{Z}_4$ implies that any $\mathbb{Z}_2\mathbb{Z}_4$-additive code is a $\mathbb{Z}_4$ module. Comparably, once the multiplication is well-defined on $R_{\alpha,\beta}$, we get a similar result. Let $f(x) \in \mathbb{Z}_4[x]$ and $((g(x), h(x)) \in R_{\alpha,\beta}$ and define the multiplication by

$$f(x) * (g(x), h(x)) = (\overline{f(x)g(x)}, f(x)h(x))$$

where $\overline{f(x)g(x)} \equiv f(x)g(x) \mod 2$ and the product is the polynomial product in $\mathbb{Z}_2[x]/(x^\alpha - 1)$ and $\mathbb{Z}_4[x]/(x^\beta - 1)$.

Lemma 5.14. *$R_{\alpha,\beta}$ is a $\mathbb{Z}_4[x]$-module with respect to the above, well-defined multiplication.*

Definition 5.15. *Let C be a subgroup of $R_{\alpha,\beta}$. C is a $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic code if for all*

$$c(x) = (a_0 + a_1x + ... + a_{\alpha-1}x^{\alpha-1}, b_0 + b_1x + ... + b_{\beta-1}x^{\beta-1}) \in C$$

$$\Longrightarrow$$

$$x * c(x) = (a_{\alpha-1} + a_0x + ... + a_{\alpha-2}x^{\alpha-1}, b_{\beta-1} + b_0x + ... + b_{\beta-2}x^{\beta-1}) \in C$$

Theorem 5.16. *A code C is a $\mathbb{Z}_2\mathbb{Z}_4$-cyclic code if and only if C is a $\mathbb{Z}_4[x]$-submodule of $\mathbb{Z}_2[x]/(x^\alpha - 1) \times \mathbb{Z}_4[x]/(x^\beta - 1)$.*

5.2.1. $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic codes with odd β. The majority of the work done on $\mathbb{Z}_4$ cyclic codes and $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic codes is under certain parameters. Many results depend on β being an odd integer [14, 13].

As a $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic code and $\mathbb{Z}_4[x]/(x^\beta - 1)$ are $\mathbb{Z}_4[x]$-submodules of $\mathbb{Z}_2[x]/(x^\alpha - 1) \times \mathbb{Z}_4[x]/(x^\beta - 1)$, we define the following mapping:

$$\Psi : C \to \mathbb{Z}_4[x]/(x^\beta - 1)$$

$$\text{by } \Psi(f_1(x), f_2(x)) = f_2(x)$$

As β is odd, and $\mathrm{Im}(\Psi)$ is an ideal in $\mathbb{Z}_4[x]/(x^\beta - 1)$, $\mathrm{Im}(\Psi) = (g(x) + 2a(x)) = (g(x), 2a(x))$, where $a(x)|g(x)|(x^\beta - 1) \mod 4$ [8]. Note that $\ker(\Psi) = \{(f(x), 0) \in C | f(x) \in \mathbb{Z}_2[x]/(x^\alpha - 1)\}$. Let $I = \{f(x) \in \mathbb{Z}_2[x]/(x^\alpha - 1) | (f(x), 0) \in \ker(\Psi)\}$. Then I is an ideal in $\mathbb{Z}_2[x]/(x^\alpha - 1)$. Hence $I = (f(x))$, where $f(x)|(x^\alpha - 1) \mod 2$.

Theorem 5.17. *Let C be a $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic code in $\mathbb{Z}_2[x]/(x^\alpha - 1) \times \mathbb{Z}_4[x]/(x^\beta - 1)$. Then one of the following statements is true.*

1. *$C = (f(x), 0)$ where $f(x)|(x^\alpha - 1) \mod 2$*
2. *$C = (l(x), g(x) + 2a(x))$ where $(x^\alpha - 1)| \left(l(x) \left(\frac{x^\beta - 1}{a(x)}\right)\right) \mod 2$ and $a(x)|g(x)|(x^\beta - 1) \mod 4$*
3. *$C = ((f(x), 0), (l(x), g(x) + 2a(x)))$ where $f(x)|(x^\alpha - 1) \mod 2$, $a(x)|g(x)|(x^\beta - 1) \mod 4$, $l(x) \in \mathbb{Z}_2[x]$, $\deg(l(x)) < \deg(f(x))$, and $f(x)| \left(l(x) \left(\frac{x^\beta - 1}{a(x)}\right)\right) \mod 2$. [13]*

Theorem 5.18. *Suppose C is a $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic code. Then the generators are unique.*

Proof: Firstly, note that 3 in Theorem 5.17 is a generalization of 1 and 2. Hence, a proof for the theorem using the generators of 3 in Theorem 5.17 is sufficient.
Let $C = ((f(x), 0), (l(x), g(x) + 2a(x)))$. $f(x)$, $g(x)$, and $a(x)$ are unique because $(f(x))$ is a binary cyclic code and $(g(x) + 2a(x))$ is a quaternary cyclic code.
Suppose $((f(x), 0), (l_1(x), g(x) + 2a(x))) = ((f(x), 0), (l_2(x), g(x) + 2a(x)))$.
Then $(l_1(x) - l_2(x), 0) \in C$ which means that $(l_1(x) - l_2(x), 0) \in \ker(\Psi)$. This implies that $l_1(x) - l_2(x) \in (f(x))$.
However, as $\deg(l(x)) < \deg(f(x))$, $l_1(x) - l_2(x) = 0 \implies l_1(x) = l_2(x)$. □

5.2.2. $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic codes with even β. This section brings forth an advance in the field of $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic codes by constructing generators for an even β. The section follows a similar approach to how the isomorphism was constructed and the parameters on generators were found in section 5.2.1 for odd β.

Lemma 5.19. *$\mathbb{Z}_2[x]/(x^\alpha - 1)$ is a $\mathbb{Z}_4[x]$-module using $f(x)g(x) = f(x)g(x) \mod 2$ in $\mathbb{Z}_2[x]/(x^\alpha - 1)$. Where $f(x) \in \mathbb{Z}_4[x]$ and $g(x) \in \mathbb{Z}_2[x]/(x^\alpha - 1)$.*

For the remainder of this section, let C be a $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic codes with even β. Define

$$\Xi : C \to \mathbb{Z}_2[x]/(x^\alpha - 1)$$

$$\text{by } \Xi(f_1(x), f_2(x)) = f_1(x)$$

Ξ is a module homomorphism.

Lemma 5.20. *$\mathrm{Im}(\Xi)$ is a $\mathbb{Z}_2[x]$-submodule of $\mathbb{Z}_2[x]/(x^\alpha - 1)$ [15].*

Since $\mathrm{Im}(\Xi)$ is a $\mathbb{Z}_2[x]$-submodule and hence an ideal in $\mathbb{Z}_2[x]/(x^\alpha - 1)$, then $\mathrm{Im}(\Xi) = (f(x))$, where $f(x)|(x^\alpha - 1) \mod 2$. Note that for Ξ, we have $\ker(\Xi) = \{(0, f_2(x))|(0, f_2(x)) \in C\}$. Let $J = \{b(x)|(0, b(x)) \in \ker(\Xi)\}$.

Lemma 5.21. *J is a $\mathbb{Z}_4[x]$-submodule of $\mathbb{Z}_4[x]/(x^\beta - 1)$*

Proof: Let $b_1(x), b_2(x) \in J$. Then $(0, b_1(x)), (0, b_2(x)) \in \ker(\Xi)$, which means that $(0, b_1(x) + b_2(x)) \in \ker(\Xi)$.
Therefore $b_1(x) + b_2(x) \in J$. Let $r(x) \in \mathbb{Z}_4[x]$.
Then $r(x)b_1(x) \in J$ as $\Xi(0, r(x)b_1(x)) = 0$. □

Since J is a $\mathbb{Z}_4[x]$-submodule and hence an ideal in $\mathbb{Z}_4[x]/(x^\beta - 1)$, then $J = (g(x) + 2p(x)) \oplus (2a(x))$, where $g(x)$, $p(x)$, and $a(x)$ are polynomials such

that $g(x)|(x^\beta - 1) \mod 4$, $a(x)|g(x) \mod 2$, $a(x)|\left(p(x)\left(\frac{x^\beta-1}{g(x)}\right)\right) \mod 2$, and $\deg(a(x)) > \deg(p(x))$. Let $(f(x), l(x)) \in C$ be such that $\Xi(f(x), l(x)) = f(x)$. Then by the first isomorphism theorem we get the following theorem.

Theorem 5.22.

$$C \cong ((f(x), l(x)), (0, g(x) + 2p(x)), (0, 2a(x)))$$

We will write $C = ((f(x), l(x)), (0, g(x) + 2p(x)), (0, 2a(x)))$.

Lemma 5.23. $\left(\frac{x^\alpha-1}{f(x)}\right) l(x) \in (g(x) + 2p(x)) \oplus (2a(x))$

Proof: $\Xi\left(\frac{x^\alpha-1}{f(x)}(f(x), l(x))\right) = \Xi\left(\left(0, \frac{x^\alpha-1}{f(x)}l(x)\right)\right) = 0$
$\implies \left(\frac{x^\alpha-1}{f(x)}(l(x))\right) \in \ker(\Xi)$
$\implies \left(\frac{x^\alpha-1}{f(x)}(l(x))\right) \in (g(x) + 2p(x)) \oplus (2a(x))$
$\implies \left(\frac{x^\alpha-1}{f(x)}(l(x))\right) = \alpha_1(g(x) + 2p(x)) + \alpha_2(2a(x))$ □

Lemma 5.24. *Suppose the leading coefficient of* $l(x)$ *is* 2. *Then*

1. $\deg(l(x)) < \deg(a(x))$
2. $l(x) = 2l_1(x)$

Proof:

1. Suppose $\deg(l(x)) > \deg(a(x))$, with $\deg(l(x)) - \deg(a(x)) = i$. Then
   $$\begin{aligned} D &= ((f(x), l(x)) + x^i(0, 2a(x)), (0, g(x) + 2p(x)), (0, 2a(x))) \\ &= ((f(x), l(x) + x^i 2a(x)), (0, g(x) + 2p(x)), (0, 2a(x))) \subseteq C \end{aligned}$$
   But $(f(x), l(x)) = (f(x), l(x) + 2x^i a(x)) + 2x^i(0, 2a(x)) \implies D \subseteq C$ and $C = D$ with $\deg(l(x)) < \deg(a(x))$.
   If necessary, repeat the process in order to obtain $\deg(l(x)) < \deg(a(x))$.

2. Since $f(x), l(x) \in C$, then $2(f(x), l(x)) = (0, 2l(x)) \in C$ and $\Xi((0, 2l(x))) = 0 \implies 2l(x) \in (g(x) + 2p(x)) \oplus (2a(x)) \implies 2l(x) \in 2a(x) \implies l(x) \in (a(x))$ where $(a(x))$ is the binary code generated by $a(x)$. Since $\deg(l(x)) < \deg(a(x))$, we have that $l(x) = 0$ in $\mathbb{Z}_4[x]$ and thus $l(x) = 2l_1(x)$

□

Lemma 5.25. *If the leading coefficient of* $l(x)$ *is* 1 *or* 3 *then* $\deg(l(x)) < \deg(g(x))$.

Proof: Suppose $deg(g(x)) < \deg(l(x))$ then let $\deg(l(x)) - \deg(g(x)) = i$. Let
$$\begin{aligned} D &= ((f(x), l(x)) + x^i(0, g(x) + 2p(x)), (0, g(x) + 2p(x)), (0, 2a(x))) \\ &= ((f(x), l(x) + x^i(g(x) + 2p(x))), (0, g(x) + 2p(x)), (0, 2a(x))) \subseteq C \end{aligned}$$

But $(f(x), l(x)) = (f(x), l(x) + x^i(g(x) + 2p(x))) + x^i(0, g(x) + 2p(x)) \implies D \subseteq C$ and $C = D$ with $\deg(l(x)) < \deg(g(x))$ If necessary, repeat the process in order to obtain $\deg(l(x)) < \deg(g(x))$. □

5.2.3. Examples. We will now include a simple example. Let $\alpha = 3$ and $\beta = 4$. In other words, $x^\alpha - 1 = x^3 - 1$ in $\mathbb{Z}_2[x]$ and $x^\beta - 1 = x^4 - 1 = (x-1)^4$ in $\mathbb{Z}_4[x]$.
Using these parameters we can have $C = (((x^2 - 1), 0), (0, (x^2 - 1)))$. The first generator yields 0 and $x^2 - 1$. The second generator yields the following:

Codewords in $C_1 = (x^2 - 1)$ in $Z_4[x]/(x^4 - 1)$
0
$x^2 + 3$
$2x^2 + 2$
$3x^2 + 1$
$x^3 + 3x$
$2x^3 + 2x$
$3x^3 + x$
$x^3 + x^2 + 3x + 3$
$2x^3 + 2x^2 + 2x + 2$
$3x^3 + 3x^2 + x + 1$
$x^3 + 2x^2 + 3x + 2$
$3x^3 + 2x^2 + x + 2$
$x^3 + 3x^2 + 3x + 1$
$3x^3 + x^2 + x + 3$
$2x^3 + x^2 + 2x + 3$
$2x^3 + 3x^2 + 2x + 1$

Table 5.1: $(x^2 - 1)$ in $\mathbb{Z}_4[x]/(x^4 - 1)$

For $(x^2-1) = ((x^2+1)+2)$, this means that $g(x) = (x^2+1)$ and $a(x) = 1$, which matches the generators in Theorem 5.22. We combine codewords from (x^2-1) in $\mathbb{Z}_4[x]/(x^4-1)$ with 0 and x^2-1 in the following table to show the codewords.

Codewords in $C = (((x^2-1),0),(0,(x^2-1)))$
$(0,0)$
$(0,x^2+3)$
$(0,2x^2+2)$
$(0,3x^2+1)$
$(0,x^3+3x)$
$(0,2x^3+2x)$
$(0,3x^3+x)$
$(0,x^3+x^2+3x+3)$
$(0,2x^3+2x^2+2x+2)$
$(0,3x^3+3x^2+x+1)$
$(0,x^3+2x^2+3x+2)$
$(0,3x^3+2x^2+x+2)$
$(0,x^3+3x^2+3x+1)$
$(0,3x^3+x^2+x+3)$
$(0,2x^3+x^2+2x+3)$
$(0,2x^3+3x^2+2x+1)$
$(x^2-1,0)$
(x^2-1,x^2+3)
$(x^2-1,2x^2+2)$
$(x^2-1,3x^2+1)$
(x^2-1,x^3+3x)
$(x^2-1,2x^3+2x)$
$(x^2-1,3x^3+x)$
(x^2-1,x^3+x^2+3x+3)
$(x^2-1,2x^3+2x^2+2x+2)$
$(x^2-1,3x^3+3x^2+x+1)$
(x^2-1,x^3+2x^2+3x+2)
$(x^2-1,3x^3+2x^2+x+2)$
(x^2-1,x^3+3x^2+3x+1)
$(x^2-1,3x^3+x^2+x+3)$
$(x^2-1,2x^3+x^2+2x+3)$
$(x^2-1,2x^3+3x^2+2x+1)$

Table 5.2: Codewords in $\mathbb{Z}_2[x]/(x^3-1) \times \mathbb{Z}_4[x]/(x^4-1)$

5.3. Conclusions and Future Work

In this thesis we included some famous codes in coding theory, eventually leading up to cyclic codes and $\mathbb{Z}_2\mathbb{Z}_4$-additive codes. $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic codes have been studied before in coding theory and have been identified as $\mathbb{Z}_4[x]$-submodules of the ring $\mathbb{Z}_2[x]/(x^\alpha - 1) \times \mathbb{Z}_4[x]/(x^\beta - 1)$. The generators for codes with odd β have been described in 2014. This thesis focussed on the case with even β. Now that the $\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic codes have now been classified with odd β as well as even β, the next steps include finding better codes with these parameters, as well as adding more conditions.

References

[1] D. Johnson, D. Sklar, and C. Wang, "Forward error-correction coding," *Crosslink*, vol. 3(1), pp. 26–29, Winter 2001/2002.

[2] C. E. Shannon, "A math. theory of communication," *Bell System Technical J.*, vol. 27(3), pp. 379–423 and 623–656, July 1948.

[3] R. Blahut, *Theory and Practice of Error Control Codes*. Massachusetts, USA: Addison-Wesley, 1983.

[4] G. Forney, M. Trott, and N. Sloane, "The nordstrom-robinson code is the binary image of the octacode," *Coding and Quantization: DIMACS / IEEE Workshop October*, pp. 19–21, 1992.

[5] J. A.R. Hammons, P. Kumar, A. Calderbank, N. Sloane, and P. Solé, "The $\mathbb{Z}_4$-linearity of Kerdock, Preparata, Goethals and related codes," *IEEE Trans. on Inform. Theory*, vol. 40, pp. 301–319, 1994.

[6] N. Sloane, "Algebraic coding theory: Recent developments related to the integers mod 4," *Kyoto University Research Institute for Math. Sci discussion proc.*, vol. 896, pp. 38–52, 1995.

[7] T. Abualrub and R. Oehmke, "Cyclic codes of length 2^e over $\mathbb{Z}_4$," *Discrete Appl. Math.*, vol. 128, pp. 3–9, 2003.

[8] T. Abualrub and I. Siap, "Reversible cyclic codes over $\mathbb{Z}_4$," *Australasian J. of Combinatorics*, vol. 38, pp. 195–205, 2007.

[9] P. Delsarte, *An algebraic approach to the association schemes of coding theory*. Philips Research Laboratories, 1973. Philips Research Rep. Suppl., No. 10.

[10] R. Bose and T. Shimamoto, "Classification and anal. of partially balanced incomplete block designs with two associate classes," *J. of the Amer. Statistical Assoc.*, vol. 47, pp. 151–184, 1952.

[11] P. Delsarte and V. Levenshtein, "Association schemes and coding theory," *IEEE Trans. on Inform. Theory*, vol. 44 (6), pp. 2477–2504, Oct 1998.

[12] J. Borges, C. Fernandez, J. Pujol, J. Rifà, and M. Villanueva, "$\mathbb{Z}_2\mathbb{Z}_4$-linear codes: generator matrices and duality," *Designs, Codes and Cryptography*, vol. 54 (2), pp. 167–179, 2010.

[13] T. Abualrub, I. Siap, and N. Aydin, "$\mathbb{Z}_2\mathbb{Z}_4$-additive cyclic codes," *IEEE Trans. on Inform. Theory*, vol. 60(3), pp. 1508–1514, March 2014.

[14] V. Pless and Z. Qian, "Cyclic codes and quadratic residue codes over $\mathbb{Z}_4$," *IEEE Trans. on Inform. Theory*, vol. 42(5), pp. 1594–1600, 1996.

[15] T. Hungerford, *Algebra, Graduate Texts in Math.* New York, USA: Springer-Verlag, 1974.

Vita

Jonas Saman, born in Belgium, moved to the United Arab Emirates in the summer of 2000. He graduated from the American International School in Abu Dhabi in 2009. He received a merit Scholarship to the American University of Sharjah, where he attained his Bachelor's of Science in Mathematics while being a teacher's assistant in the department of Mathematics and Statistics. He was the president of the Math Club for two years.

Immediately after graduating, Mr. Saman began his Master's of Science in Pure Mathematics at the American University of Sharjah. He received a graduate teaching assistantship at the American University of Sharjah, where he taught undergraduates Calculus and Statistics.