# Algebraic Coding Theory

Zitong Gu
McMaster University

## Motivation/Introduction

The discipline of coding theory was initiated in the paper "A Mathematical theory of communication" written by Claude Shannon, in which he stated that **the fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point.** The main goal of coding theory is finding codes with reasonable information content and reasonable error-controlling ability.

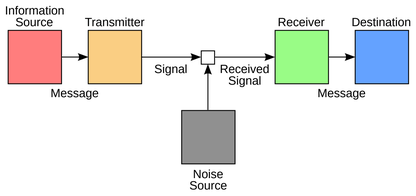


Figure 1. Shannon's model of communication

## Repetition Code

Suppose we would like to send a 1 to signify "yes" and a 0 to signify "no". If we simply send one bit, then there is a chance that noise, due to physical interference, will corrupt that bit and an unintended message will be received. A simple solution is to use a *repetition code*: instead of sending a single bit, we send 11111 to represent 1 and 00000 to represent 0, where 11111 and 00000 are called codewords and 1 and 0 are called message words. We will assume *nearest neighbour decoding* in this example, that is, the receiver decodes a received binary 5-tuple as the closet codeword. For instance, suppose that 11111 is transmitted and 10101 (two errors occurred), then 10101 is decoded as 11111 instead of 00000 since there are more 1s than 0s.

## Basics

- A code is an $(n, m)$-**block code** if the information that is to be coded can be divided into blocks of $m$ binary digits, each of which can be encoded into $n$ binary digits.
- An $(n, m)$-block code consists of an **encoding function** $E : \mathbb{Z}_2^m \to \mathbb{Z}_2^n$ and a **decoding function** $D : \mathbb{Z}_2^n \to \mathbb{Z}_2^m$. A **codeword** is any element in the image of $E$.
- The **Hamming weight** of a codeword $w(\mathsf{c})$ is the number of nonzero components in the codeword.
- The **Hamming distance** of two codewords $\mathsf{x}$ and $\mathsf{y}$, denoted by $d(\mathsf{x}, \mathsf{y})$ is the Hamming weight of of the vector difference $\mathsf{x} - \mathsf{y}$, i.e., $d(\mathsf{x}, \mathsf{y}) = w(\mathsf{x} - \mathsf{y})$.
- The **minimum distance** for a code is the minimum of all distances $d(\mathsf{x}, \mathsf{y})$ of two distinct codewords.
- The Hamming distance is a *metric* on the space of all binary $n$-tuples, i.e., for any $\mathsf{x}, \mathsf{y}, \mathsf{z} \in \mathbb{Z}_2^n$, the following properties are satisfied:
  1. $d(\mathsf{x}, \mathsf{x}) = 0$,
  2. $d(\mathsf{x}, \mathsf{y}) = d(\mathsf{y}, \mathsf{x})$,
  3. $d(\mathsf{x}, \mathsf{z}) \leq d(\mathsf{x}, \mathsf{y}) + d(\mathsf{y}, \mathsf{z})$.

The following notation is popular in coding theory: A $(n, M, d)$ code that has minimum distance $d$ and consist of $M$ codewords, all of length $n$.

## Theorem

- A code $C$ with a minimum distance $d = 2n + 1$ can correct up to $n$ errors and detect up to $2n$ errors.
  ***Proof.*** Suppose that a codeword $\mathsf{x}$ is transmitted and the word $\mathsf{y}$ is received with at most $n$ error. Then $d(\mathsf{x}, \mathsf{y}) \leq n$. Let $\mathsf{z} \neq \mathsf{x}$ be a codeword. Then
  $$2n + 1 \leq d(\mathsf{x}, \mathsf{z}) \leq d(\mathsf{x}, \mathsf{y}) + d(\mathsf{y}, \mathsf{z}) \leq n + d(\mathsf{y}, \mathsf{z}).$$
  Hence, $d(\mathsf{y}, \mathsf{z}) \geq n + 1$ and $\mathsf{y}$ should be correctly decoded as $\mathsf{x}$. Now suppose that $\mathsf{x}$ is transmitted and $\mathsf{y}$ is received with at least one error and at most $2n$ error, i.e., $1 \leq d(\mathsf{x}, \mathsf{y}) \leq 2n$. Then $\mathsf{y}$ cannot be a codeword since the minimum distance is $2n + 1$. Hence, the code can detect up to $2n$ errors.

*Adding Group Structure to Codes*

- A **group code** is a code that is also a subgroup of $\mathbb{Z}_2^n$. Let $d_{\min}$ be the minimum distance for a group code $C$. Then $d_{\min}$ is the minimum weight of all the nonzero codewords in $C$. That is,
  $$d_{\min} = \min\{w(\mathsf{x}) : \mathsf{x} \neq \mathsf{0}\}.$$

## Linear Code

- A code is a **linear code** if it is determined by the null space of some matrix $H \in \mathbb{M}_{m\times n}(\mathbb{Z}_2)$. A linear code is also a group code.
- Given a matrix $H$ in $\mathbb{M}_{m\times n}(\mathbb{Z}_2)$, we can find the code generated by computing the null space of $H$.
- The codewords of a linear code generated by Null$(H)$ can be computed efficiently by the *generator matrix* of the code.

## Hamming Code

Let $H$ be an $m \times n$ matrix over $\mathbb{Z}_2$, where the $i$-th column is the number $i$ written in binary with $m$ bits. The null space of such a matrix is called a **Hamming code**. Consider

$$H = \begin{bmatrix} 0&0&0&1&1&1&1 \\ 0&1&1&0&0&1&1 \\ 1&0&1&0&1&0&1 \end{bmatrix} \sim \begin{bmatrix} 0&1&1&1&1&0&0 \\ 1&0&1&1&0&1&0 \\ 1&1&0&1&0&0&1 \end{bmatrix} = [A \mid I_3].$$

The **generator matrix** associated with $H$ is

$$G = \begin{bmatrix} I_4 \\ \hline A \end{bmatrix}.$$

One can easily verify that $H(G\mathsf{x}) = (HG)\mathsf{x} = \mathsf{0}$. So, an easy way to compute Null$(H)$ is with the generator matrix $G$:

| Message Word $\mathsf{x}$ | Codeword $G\mathsf{x}$ |
|---|---|
| 0000 | 0000000 |
| 0001 | 0001111 |
| 0010 | 0010110 |
| 0011 | 0011001 |
| 0100 | 0100101 |
| ... | ... |
| 1111 | 1111111 |

The generated code is a linear $[7, 16, 3]$ code. From the previous theorems, this code can correct 1 error and detect 2 or fewer errors.

## Efficient Decoding

- Suppose that a codeword $\mathsf{x}$ of a code $C$ is transmitted and an $n$-tuple is received, the decoder produces the codeword
  $$\mathsf{y} = D(\mathsf{y}) = \operatorname*{argmin}_{\mathsf{z}\in C} d(\mathsf{z}, \mathsf{y}).$$
- A linear code $C$ is a subgroup of $\mathbb{Z}_2^n$. We can use the cosets of $C$ for decoding.

## Coset Decoding

Let $C$ be the [5, 2] code $\{00000, 10110, 01101, 11011\}$, which is a subgroup of $\mathbb{Z}_2^5$. From the elements of $\mathbb{Z}_2^5$ that are not in $C$, choose one with the smallest weight, say $e_1 = 10000$. Now we obtain a coset $e_1+C$. From the elements of $\mathbb{Z}_2^5$ that are not in $C$ or $e_1 + C$, choose one with the smallest weight, say $e_2 = 01000$. Then list the coset $e_2+C$. Perform this procedure until all the cosets are listed. The result is as follows:

| Coset Representative | Coset |
|---|---|
| $C$ | {00000, 10110, 01101, 11011} |
| $10000 + C$ | {10000, 00110, 11101, 01011} |
| $00100 + C$ | {00100, 10010, 01001, 11111} |
| $00010 + C$ | {00010, 10100, 01111, 11001} |
| $00001 + C$ | {00001, 10111, 01100, 11010} |
| $11000 + C$ | {11000, 01110, 10101, 00011} |
| $10001 + C$ | {10001, 00111, 11100, 01010} |

Table 1. Cosets of $C$

Suppose that $\mathsf{x}$ is transmitted and $\mathsf{y}$ is received. Let $\mathsf{r}$ denoted the transmission error, then $\mathsf{x} = \mathsf{y} + \mathsf{r}$ or $\mathsf{y} = \mathsf{r} + \mathsf{x}$. This is saying that $\mathsf{y}$ is in the coset $\mathsf{r} + C$. The decoding rule is as follows: Decode a received word $\mathsf{y}$ as the codeword at the top of the column in which $\mathsf{y}$ appears. For example, 00111 in the last row is decoded as 10110.

The cosets of $C$ serves as the role of a look-up table. While the look-up table can faster the process of decoding, there is a potential problem with this method: We might have to examine every coset for the received codeword.

## Cyclic Code

- Let $\phi : \mathbb{Z}_2^k \to \mathbb{Z}_2^n$ be a binary [n, k] linear block code. Then $\phi$ is a **cyclic code** if for every codeword $(a_0, a_1, \ldots, a_{n-1})$, the shifted n-tuple $(a_{n-1}, a_0, \ldots, a_{n-2})$ is also a codeword.
- Cyclic codes are attractive because
  - encoding can be implemented easily on a computer using shift register,
  - cyclic codes have many good algebraic properties that provide practical decoding methods.
- A codeword of a cyclic code is associated with a polynomial through the isomorphism:
  $$\rho : \mathbb{Z}_2^n \to \mathbb{Z}_2[x]/\langle x^n - 1\rangle, \quad \rho(a_0, a_1, \ldots, a_{n-1}) = \sum_{i=0}^{n-1} a_i x^i.$$
- $\mathbb{Z}_2[x]/\langle x^n - 1\rangle$ is a ring of polynomials of the form
  $$p(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}$$
  such that $x^n = 1$.

## Cyclic Code (Cont.)

- A cyclic shift of an $n$-tuple can be described by polynomial multiplication:
  $$(a_0, a_1, \ldots, a_{n-1}) \mapsto f(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}$$
  $$(a_{n-1}, a_0, \ldots, a_{n-2}) \mapsto xf(x) = a_{n-1} + a_0 x + a_1 x^2 + \ldots a_{n-2}x^{n-1}$$
- An important characterization of cyclic codes: A linear code $C \subset \mathbb{Z}_2^n$ is cyclic if and only if it is an ideal in $R_n = \mathbb{Z}_2[x]/\langle x^n - 1\rangle$.

## Construction of Cyclic Codes

We can construct a cyclic code by finding its corresponding ideal in $R_n$. It is easy to describe the ideals in $R_n$ since every ideal $I$ in $R_n$ is principal. So, $I = \langle g(x)\rangle$ for some unique monic polynomial $g(x) \in \mathbb{Z}_2[x]$. Also, by the Correspondence Theorem, every ideal $I = \langle g(x)\rangle$ in $R_n$ contains $\langle x^n - 1\rangle$ and so $g(x)$ must divide $x^n - 1$. Therefore, every ideal $C$ in $R_n$ is of the form

$$C = \langle g(x)\rangle = \{p(x)g(x) : p(x) \in R_n \text{ and } g(x) \mid (x^n - 1) \in \mathbb{Z}_2[x]\}.$$

If a cyclic code $C = \langle g(x)\rangle$, then we call $g(x)$ the generator polynomial of $C$. Consider

$$x^7 - 1 = (1 + x)(1 + x + x^3)(1 + x^2 + x^3)$$

with irreducible factors and each factor is a generator polynomial of a cyclic code. The polynomial $p(x) = 1 + x^2 + x^3$ generates a cyclic $[7, 4]$ block code $C$. To calculate a generator matrix for $C$, we only consider how the polynomials $1, x$, and $x^2$ are encoded: $(1 + x^2 + x^3) \cdot 1 = 1+x^2+x^3$, $(1+x^2+x^3)\cdot x = x+x^3+x^4$, and $(1+x^2+x^3)\cdot x^2 = x^2+x^4+x^5$. Then we obtain the generator matrix

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

With the same procedure, one can compute the generator matrices for each polynomial factors (there are seven factors) of $x^7 - 1$ and obtain all the cyclic codes of length 7.

## References

[1] Jonathan I. Hall.
*Notes on Coding Theory*.
https://users.math.msu.edu/users/halljo/classes/codenotes/coding-notes.html, 2010.

[2] Thomas W. Hungerford.
*Abstract Algebra: An Introduction*.
Brooks/Cole, 2013.

[3] Thomas W. Judson.
*Abstract Algebra: Theory and Applications*.
Orthogonal Publishing L3c, 2022.

[4] Claude E. Shannon.
A mathematical theory of communication.
*Bell System Technical Journal*, 27(3):379–423, 1948.

[5] Sarah A. Spence.
*Introduction to Algebraic Coding Theory*.
https://personal.math.ubc.ca/ lam/Math