CO 331 - Coding Theory

Cameron Roopnarine

Last updated: February 18, 2020

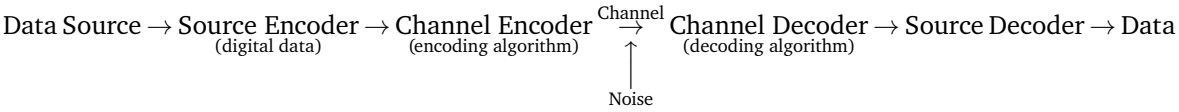
Contents

Chapter 1

Introduction and Fundamentals

2020-01-06

1.1 An Introduction to Coding Theory



EXAMPLE 1.1.1 (Repetition Code)**.**

source message → codeword	# errors/codeword that can be detected	# errors/codeword that can be corrected	rate
$0 \to 0$ $1 \to 1$	0	0	1
$0 \to 00$ $1 \to 11$	1	0	$1/2$
$0 \to 000$ $1 \to 111$	2	1	$1/3$
$0 \to 00000$ $1 \to 11111$	4	2	$1/5$

Goal of Coding Theory

Design codes such that:

- High information rate
- High error-correcting capability
- Efficient encoding and decoding algorithms

Codes $\supset$ Block codes $\supset$ Linear codes $\supset$ Cyclic codes $\supset$ BCH Codes $\supset$ RS Codes

Codes not covered in this course:

- Flamming codes
- Golay codes

- Raptor codes
- LDPC codes
- Turbo codes

Requirements for this course:

- MATH 136
- Not required (but required to take the course): MATH 235
- Familiarity with: Groups, Fields, Ideals, Rings (these will be taught)
- Useful, if you have completed these you might be bored: PMATH 336, PMATH 334 [or the advanced equivalents]

The big picture

In its broadest sense, coding deals with the reliable, efficient, and secure transmissions of data over channels that are subject to inadvertent noise and malicious intrusion.

Data source $\to$ Source encoder (data compression) $\to$ Encryptory authentication (cryptography; co 487) $\to$ Channel encoder (error correction codes) $\xrightarrow{\text{channel}}$

noise (non-malicious) ↑ adversarial intrusion (malicious) ↑

$\to$ Decriptory verification $\to$ Source decoder (data decompression) $\to$ Data

2020-01-08

1.2 Fundamental Concepts

Message source $\to$ Source encoder (binary strings) $\to$ Channel encoder (adds redundancy to message) $\xrightarrow{\text{channel}}$ Channel decoder $\to$ Source Decoder $\to$

noise ↑

this course

Message

DEFINITION 1.2.1. An **alphabet** A is a finite set of $|A| = q \geqslant 2$ symbols.

DEFINITION 1.2.2. A **word** is a finite sequence (**tuples** or **vectors**) of symbols from an alphabet A.

DEFINITION 1.2.3. The **length** of a word is the number of symbols in it.

DEFINITION 1.2.4. A **code** C over A is a finite set of words in A with $|C| \geqslant 2$.

DEFINITION 1.2.5. A **codeword** $\boldsymbol{c}$ is a word in code C.

DEFINITION 1.2.6. A **block code** is a code where all codewords have the same length. A block code C of length n containing M codewords over A is a subset $C \subseteq A^n$, with $|C| = M$. We refer to such a block code as an $[n, M]$-code over A.

EXAMPLE 1.2.7 (Block Code). Let $A = \{0, 1\}$ and $C = \{00000,\ 11100,\ 00111,\ 10101\}$. C is a $[5, 4]$-code over $\{0, 1\}$.

Messages $\to$ Codewords
$00 \to 00000$
$10 \to 11100$
$01 \to 00111$
$11 \to 10101$

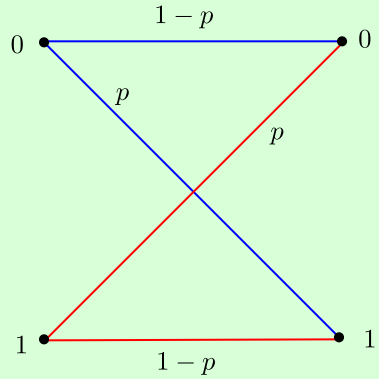
The encoding is a one-to-one map.

The channel encoder transmits only codewords, but what's received by the channel decoder might not be a codeword. For example, suppose the channel decoder receives $\boldsymbol{r} = 11001$. What should it do? In our above example, we can see that $\boldsymbol{r}$ is closest to 11100 and 10101 (only two bits are different), so it's possible that the codeword was one of those two. However, this may not be the case in practice.

1.3 Assumptions About the Communications Channel

1) The channels only transmit symbols from A.
2) No symbols are deleted, added, or transposed.
3) Errors are random

EXAMPLE 1.3.1 (Binary Symmetric Channel, BSC). Let $A = \{0, 1\}$, and p denote the symbol error probability. The encoding map is:



A similar encoding map can be drawn for $A = \{0, 1, 2\}$, with symbol error probability $p/2$.

Suppose that the symbols transmitted are $X_1, X_2, \ldots$, and the symbols received are $Y_1, Y_2, \ldots$. Then for all $i \geqslant 1$, $j \geqslant 1$, $k \leqslant q$, the probability that Y_i is received, given that X_i is transmitted is:

$$P(Y_i = a_j \mid X_i = a_k) = \begin{cases} 1 - p, & \text{if } j = k \\ \frac{p}{q-1}, & \text{if } j \neq k \end{cases}$$

1.4 Notes about BSC

(i) If $p = 0$, the channel is perfect.

(ii) If $p = 1/2$, the channel is useless.

(iii) If $1/2 < p \leqslant 1$, then simply flip all bits that are received.

(iv) WLOG, we can assume $0 < p < 1/2$.

(v) Analogously, for a q-ary channel, we can assume that $0 < p < \frac{q-1}{q}$.

DEFINITION 1.4.1. If $\boldsymbol{x}, \boldsymbol{y} \in A^n$, the **Hamming distance** $d(\boldsymbol{x}, \boldsymbol{y})$ is the number of coordinate positions in which $\boldsymbol{x}$ and $\boldsymbol{y}$ differ.

EXAMPLE 1.4.2 (Hamming Distance)**.** Let $\boldsymbol{x} = 10111$ and $\boldsymbol{y} = 01010$. The Hamming distance of $\boldsymbol{x}$ and $\boldsymbol{y}$ is $d(\boldsymbol{x}, \boldsymbol{y}) = 4$ since $\boldsymbol{x}$ and $\boldsymbol{y}$ differ in the coordinate positions 1, 2, 3, and 5.

DEFINITION 1.4.3. Let C be an $[n, M]$-code. The **Hamming distance** d **of a code** C is

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

THEOREM 1.4.4. *d is a **metric**. For all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in A^n$:*
(1) $d(\boldsymbol{x}, \boldsymbol{y}) \geqslant 0$
(2) $d(\boldsymbol{x}, \boldsymbol{y}) = 0$ if and only if $\boldsymbol{x} = \boldsymbol{y}$
(3) $d(\boldsymbol{x}, \boldsymbol{y}) = d(\boldsymbol{y}, \boldsymbol{x})$
(4) (Triangle inequality): $d(\boldsymbol{x}, \boldsymbol{z}) \leqslant d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{y}, \boldsymbol{z})$

Proof. (1)-(3) are trivially true.

(4) Let $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in A^n$. Suppose that $\boldsymbol{x}$ and $\boldsymbol{z}$ differ in exactly a positions; that is, $d(\boldsymbol{x}, \boldsymbol{z}) = a$. Out of the a positions in which $\boldsymbol{x}$ and $\boldsymbol{z}$ differ, there are b positions in which $\boldsymbol{y}$ is identical to $\boldsymbol{x}$, but not $\boldsymbol{z}$. Out of the a positions, there are $a - b$ positions in which $\boldsymbol{y}$ is identical to $\boldsymbol{z}$, but not $\boldsymbol{x}$. Lastly, in the $n - a$ positions where $\boldsymbol{x}$ is identical to $\boldsymbol{z}$, there are c positions in which $\boldsymbol{y}$ does not match either $\boldsymbol{x}$ or $\boldsymbol{z}$. We can see that $d(\boldsymbol{x}, \boldsymbol{y}) = b + c$ and $d(\boldsymbol{y}, \boldsymbol{z}) = a - b + c$. We get

$$d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{y}, \boldsymbol{z}) = (b + c) + (a - b + c) = a + 2c \geqslant a$$

Therefore d is a metric. □

DEFINITION 1.4.5. The **rate** (or **information rate**) of an $[n, M]$-code C over A, is

$$R = \frac{\log_q(M)}{n}$$

where $q = |A|$.
If the source messages are all k-tuples over A, then $M = q^k$, so we have

$$R = \frac{\log_q(q^k)}{n} = \frac{k}{n}$$

EXAMPLE 1.4.6 (Rate & Distance of Code)**.** Let $A = \{0, 1\}$ and $C = \{00000,\ 11100,\ 00111,\ 10101\}$ which is a $[2, 4]$-code over $\{0, 1\}$.

- Rate of code: $R = 2/5$
- Distance of code: $d(C) = 2$, since the minimum distance are from the pair of codewords 00111 and 10101 which have Hamming distance of 2 as they differ in coordinate positions 1 and 4.

1.5 Decoding Strategy

Suppose we have an $[n, M]$-code C over A of distance d. We need to adopt a strategy for the channel decoder (henceforth called the decoder). When the decoder receives an n-tuple $\boldsymbol{r} \in A^n$ it must make some decision. This decision may be one of

(i) no errors hace occurred; accept $\boldsymbol{r}$ as a codeword.

(ii) errors have occurred; correct $\boldsymbol{r}$ to a codeword $\boldsymbol{c}$; e.g. $0 \to 0000$, $1 \to 1111$, $r = 0001$ corrected to 0000.

(iii) errors have occurred; no correction is possible.

1.6 Nearest Neighbour Decoding

1.6.1 Incomplete Maximum Likelihood Decoding (IMLD)

Correct $\boldsymbol{r}$ to the unique codeword $\boldsymbol{c}$ for which $d(\boldsymbol{r}, \boldsymbol{c})$ is smallest. If $\boldsymbol{c}$ is not unique, reject $\boldsymbol{r}$.

1.6.2 Complete Maximum Likelihood Decoding (CMLD)

Same as IMLD, except ties are broken arbitrarily.

Question: Is IMLD a reasonable strategy?

THEOREM 1.6.1. *IMLD selects the codeword $\boldsymbol{c}$ that maximizes $P(\boldsymbol{r} \mid \boldsymbol{c})$; that is, it maximizes the probability $\boldsymbol{r}$ is received, given $\boldsymbol{c}$ was sent.*

We actually want to maximize $P(\boldsymbol{c} \mid \boldsymbol{r})$, but we will ignore that for now.

Proof. Suppose $\boldsymbol{c}_1, \boldsymbol{c}_2 \in C$ with $d(\boldsymbol{c}_1, \boldsymbol{r}) = d_1$ and $d(\boldsymbol{c}_2, \boldsymbol{r}) = d_2$. Suppose $d_1 > d_2$. Now,

$P(\boldsymbol{r} \mid \boldsymbol{c}_1) = (1-p)^{n-d_1} \left(\frac{p}{q-1}\right)^{d_1}$ and $P(\boldsymbol{r} \mid \boldsymbol{c}_2) = (1-p)^{n-d_2} \left(\frac{p}{q-1}\right)^{d_2}$.

Hence,

$$\begin{aligned} \frac{P(\boldsymbol{r} \mid \boldsymbol{c}_1)}{P(\boldsymbol{r} \mid \boldsymbol{c}_2)} &= (1-p)^{d_2-d_1} \left(\frac{p}{q-1}\right)^{d_1-d_2} \\ &= \left[\frac{p}{(1-p)(q-1)}\right]^{d_1-d_2} \end{aligned}$$

Recall that, for a q-ary channel, we can assume that $p < \frac{q-1}{q}$. Thus,

$$\begin{aligned} &\implies pq < q-1 \\ &\implies 0 < q-1-pq \\ &\implies p < q-1-pq+p \\ &\implies p < (1-p)(q-1) \\ &\implies \frac{p}{(1-p)(q-1)} < 1 \end{aligned}$$

Since $d_1 > d_2$, we get $\frac{P(\boldsymbol{r}|\boldsymbol{c}_1)}{P(\boldsymbol{r}|\boldsymbol{c}_2)} < 1$, and so $P(\boldsymbol{r} \mid \boldsymbol{c}_1) < P(\boldsymbol{r} \mid \boldsymbol{c}_2)$. □

The ideal strategy is to correct $\boldsymbol{r}$ to $\boldsymbol{c} \in C$ such that $P(\boldsymbol{c} \mid \boldsymbol{r})$ is maximized. This is **Minimum Error Decoding (MED)**.

EXAMPLE 1.6.2 (IMLD != MED). Let $C = \{\underbrace{000}_{\boldsymbol{c}_1}, \underbrace{111}_{\boldsymbol{c}_2}\}$, $P(\boldsymbol{c}_1) = 0.1$, $P(\boldsymbol{c}_2) = 0.9$, $p = 1/4$, and $\boldsymbol{r} = 100$.
IMLD $\boldsymbol{r}$ is decoded to to $\boldsymbol{c}_1 = 000$.
MED

$$\begin{aligned} P(\boldsymbol{c}_1 \mid \boldsymbol{r}) &= \frac{P(\boldsymbol{r} \mid \boldsymbol{c}_1)P(\boldsymbol{c}_1)}{P(\boldsymbol{r})} \\ &= \frac{p(1-p)^2(0.1)}{P(\boldsymbol{r})} \\ &= \frac{0.0140625}{P(\boldsymbol{r})} \end{aligned}$$

$$\begin{aligned} P(\boldsymbol{c}_2 \mid \boldsymbol{r}) &= \frac{P(\boldsymbol{r} \mid \boldsymbol{c}_2)P(\boldsymbol{c}_2)}{P(\boldsymbol{r})} \\ &= \frac{p^2(1-p)(0.9)}{P(\boldsymbol{r})} \\ &= \frac{0.0421875}{P(\boldsymbol{r})} \end{aligned}$$

Since $P(\boldsymbol{c}_1 \mid \boldsymbol{r}) < P(\boldsymbol{c}_2 \mid \boldsymbol{r})$, $\boldsymbol{r}$ is decoded to $\boldsymbol{c}_2 = 111$.

Notes:

(i) IMLD selects $\boldsymbol{c}$ such that $P(\boldsymbol{r} \mid \boldsymbol{c})$ is maximum.

(ii) MED selects $\boldsymbol{c}$ such that $P(\boldsymbol{c} \mid \boldsymbol{r})$ is maximum.

(iii) MED has a drawback that it requires knowledge of $P(\boldsymbol{c}_i)$ for each $i \in [1, M]$.

(iv) Suppose source messages are equally likely, so $P(\boldsymbol{c}_i) = \frac{1}{M}$ for each $i \in [1, M]$. Then,

$$P(\boldsymbol{r} \mid \boldsymbol{c}_i) = \frac{P(\boldsymbol{c}_i \mid \boldsymbol{r})P(\boldsymbol{r})}{P(\boldsymbol{c}_i)} = P(\boldsymbol{c}_i \mid \boldsymbol{r}) \underbrace{MP(r)}_{\text{constant}}$$

So, maximizing $P(\boldsymbol{r} \mid \boldsymbol{c}_i)$ is the same as maximizing $P(\boldsymbol{c}_i \mid \boldsymbol{r})$. Thus, IMLD is the same as MED in this case.

In the remainder of the course, we will use IMLD/CMLD.

2020-01-13

1.7 Error Correcting & Detecting Capabilities of a Code

- If C is used for error correction, the strategy is IMLD/CMLD.
- If C is used for error detection only, the strategy is to reject $\boldsymbol{r}$ if $\boldsymbol{r} \notin C$, otherwise accept $\boldsymbol{r}$.

DEFINITION 1.7.1. A code C is called an ***e*-error correcting code** if the decoder always makes the correct decision if at most e errors per codeword are introduced per transmission. We define ***e*-error detecting code** similarly.

EXAMPLE 1.7.2 (Error Detecting and Correcting Codes).

- $C = \{\underbrace{0\cdots0}_{m}, \underbrace{1\cdots1}_{m}\}$ is a $\lfloor\frac{m-1}{2}\rfloor$-error correcting code.
- $C = \{0000, 1111\}$ is a 3-error detecting code.
- $C = \{0000, 1111\}$ is a 1-error correcting code, but not a 2-error correcting code.

THEOREM 1.7.3. *Suppose $d(C) = d$, then C is a $(d-1)$-error detecting code.*

Proof. Suppose $\boldsymbol{c} \in C$ is transmitted $\boldsymbol{r}$ is received. Let e denote the amount of errors that have occurred in transmission.

- If $e = 0$, then $\boldsymbol{r} = \boldsymbol{c} \in C$, and the decoder accepts $\boldsymbol{r}$.
- If $e \geq d$, then the decoder can make the wrong choice since $d(C) = d$.
- If $e \in [1, d-1]$, then $1 \leqslant d(\boldsymbol{r}, \boldsymbol{c}) \leqslant d-1$. So, $\boldsymbol{r} \notin C$, hence the decoder rejects $\boldsymbol{r}$. Hence, C is a $(d-1)$-error detecting code.

□

THEOREM 1.7.4. *If $d(C) = d$, then C is not a d-error detecting code.*

Proof. Since $d(C) = d$, there exists codewords $\boldsymbol{c}_1, \boldsymbol{c}_2 \in C$ with $d(\boldsymbol{c}_1, \boldsymbol{c}_2) = d$. If $\boldsymbol{c}_1$ is sent and $\boldsymbol{r}$ is received with d errors, it is possible $\boldsymbol{r} = \boldsymbol{c}_2$ is received. In this case, the decoder accepts $\boldsymbol{c}_2$. Hence, C is not a d-error detecting code. □

THEOREM 1.7.5. *If $d(C) = d$, then C is a $\lfloor\frac{d-1}{2}\rfloor$-error correcting code.*

Proof. Suppose $\boldsymbol{c} \in C$ is transmitted, at most $\frac{d-1}{2}$ errors are introduced, and $\boldsymbol{r}$ is received. Let $\boldsymbol{z} \in C$ with $\boldsymbol{z} \neq \boldsymbol{c}$. By the triangle inequality, we have

$$\begin{aligned} d(\boldsymbol{c}, \boldsymbol{z}) \leqslant d(\boldsymbol{c}, \boldsymbol{r}) + d(\boldsymbol{r}, \boldsymbol{z}) \implies d(\boldsymbol{r}, \boldsymbol{z}) &\geqslant d(\boldsymbol{c}, \boldsymbol{z}) - d(\boldsymbol{c}, \boldsymbol{r}) \\ &\geqslant d - \frac{d-1}{2} \\ &= \frac{d+1}{2} \\ &> \frac{d-1}{2} \end{aligned}$$

So, $\boldsymbol{c}$ is the unique codeword closest to $\boldsymbol{r}$. Hence, IMLD/CMLD will decode $\boldsymbol{r}$ to $\boldsymbol{c}$. Thus, C is a $\lfloor\frac{d-1}{2}\rfloor$-error correcting code. □

THEOREM 1.7.6. *If $d(C) = d$, then C is not a $\left(\lfloor\frac{d-1}{2}\rfloor + 1\right)$- error correcting code.*

Proof. Exercise. □

Given q, n, M, d, does there exist an $[n, M]$-code over A with $|A| = q$ such that $d(C) = d$?

Let $C = \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_M\}$ and $e = \lfloor \frac{d-1}{2} \rfloor$. For any codeword $\boldsymbol{c} \in C$, let $S_{\boldsymbol{c}}$ be the sphere of radius e centered at $\boldsymbol{c}$; that is,

$$S_{\boldsymbol{c}} = \{\boldsymbol{r} \in A^n : d(\boldsymbol{r}, \boldsymbol{c}) \leqslant e\}$$

We proved that if $\boldsymbol{c}_i, \boldsymbol{c}_j \in C$ with $i \neq j$, then $S_{\boldsymbol{c}_i} \cap S_{\boldsymbol{c}_j} = \varnothing$ for each $i \neq j$. This question can be viewed as a **sphere packing problem**: Can we place M spheres of radius e in A^n such that no two spheres overlap? This is a purely combinatorial problem.

Given $A = \{0, 1\}$, $n = 128$, $M = 2^{64}$, determine if an $[n, M]$-code C over A with $d(C) = d$ exists.

The answer to this problem is yes and we will see this in the following lectures.

Roadmap: We'll view $\{0, 1\}^n$ as a vector space of dimension n over $\mathbb{Z}_q$ where $|A| = q$. We will chose the code C to be an M-dimensional subspace of this vector space and we will choose special subspaces that satisfy the $d(C) = d$ requirement.

Chapter 2

Finite Fields

2020-01-15

2.1 Introduction

DEFINITION 2.1.1. A **field** F is a set of elements under two binary operations, which we denote by $+$ and $\cdot$ such that $+ : F \times F \to F$ and $\cdot : F \times F \to F$ where all the following axioms are satisfied:
V1 $a + (b + c) = (a + b) + c$
V2 $a + b = b + a$
V3 $\exists\, 0 \in F$ such that $a + 0 = a$
V4 $\exists\, (-a) \in F$ such that $a + (-a) = 0$
V5 $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
V6 $a \cdot b = b \cdot a$
V7 $\exists\, 1 \in F$ such that $a \cdot 1 = a$
V8 $\forall\, a \neq 0$, $\exists (a^{-1}) \in F$ such that $a \cdot (a^{-1}) = 1$
V9 $a \cdot (b + c) = a \cdot b + a \cdot c$

DEFINITION 2.1.2. A field F is **infinite** if $|F|$ is infinite.

DEFINITION 2.1.3. A field F is **finite** if $|F|$ is finite.

DEFINITION 2.1.4. The **order** of a field F, denoted $\mathrm{ord}(F)$ is $|F|$.

EXAMPLE 2.1.5 (Infinite and Finite Fields)**.**
- $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are infinite fields.
- $\mathbb{Z}$ is **not** a field since $3 \in \mathbb{Z}$, but $\left(\frac{1}{3}\right) \notin \mathbb{Z}$.

Question: For what $n \in \mathbb{Z}_{\geqslant 2}$ do there exists finite fields of order n? If a field of order n exists, how do we "construct" it?

Recall: Let $n \geqslant 2$. The integers modulo n, $\mathbb{Z}_n$ is the set of all equivalence classes $\mod n$.

$$\mathbb{Z}_n = \{[0], [1], [2], \ldots, [n-1]\}$$

where $[a]+[b]=[a+b]$ and $[a][b]=[ab]$. More simply, $\mathbb{Z}_n=\{0,1,\ldots,n-1\}$ with addition and multiplication performed $\mod n$.

EXAMPLE 2.1.6 (Modulo). Let $\mathbb{Z}_9=\{0,1,\ldots,8\}$.
- $5+7=3$ (i.e. $5+7\equiv 3 \mod 9$)
- $5\cdot 7=8$ (i.e. $5+7\equiv 8 \mod 9$)

DEFINITION 2.1.7. A **commutative ring** satisfies field axioms V1-V9 except V8.

THEOREM 2.1.8. *$\mathbb{Z}_n$ is a commutative ring.*

THEOREM 2.1.9. *$\mathbb{Z}_n$ is a field if and only if n is prime.*

Proof. ($\Longleftarrow$) Suppose n is prime. Let $a\in\mathbb{Z}_n$, $a\neq 0$ (i.e. $1\leqslant a\leqslant n-1$). Since n is prime, $\gcd(a,n)=1$ so $\exists\, s,t\in\mathbb{Z}$ such that

$$as+nt=1$$

Reducing both sides $\mod n$ gives

$$as\equiv 1 \mod n$$

Define $a^{-1}=s$. Thus, V8 is satisfied and hence $\mathbb{Z}_n$ is a field of order n.

($\Longrightarrow$) Suppose for a contradiction that n is composite, say $n=ab$ where $2\leqslant a,b\leqslant n-1$. Suppose a^{-1} exists, and define $a^{-1}=s$. Then,

$$as\equiv 1 \mod n \implies abs\equiv b \mod n \implies ns\equiv b \mod n \implies 0\equiv b \mod n$$

So, $n\mid b$ which is impossible. Therefore, a^{-1} does not exist, and hence $\mathbb{Z}_n$ is not a field. □

Question: Do there exist finite fields of orders 4 and 6?

DEFINITION 2.1.10. The **characteristic** of a field, denoted $\mathrm{char}(F)$, is the smallest possible integer m such that

$$\underbrace{1+\cdots+1}_{m}=0$$

If no such m exists, then we define $\mathrm{char}(F)=0$

EXAMPLE 2.1.11 (Characteristic of Fields).
- $\mathrm{char}(\mathbb{Q})=0$
- $\mathrm{char}(\mathbb{R})=0$
- $\mathrm{char}(\mathbb{C})=0$
- $\mathrm{char}(\mathbb{Z}_p)=p$ where p is prime.

THEOREM 2.1.12. *If $\mathrm{char}(F)=0$, then F is infinite.*

Proof. Consider $1,1+1,\ldots,\underbrace{1+\cdots+1}_{a}\in F$. Suppose for a contradiction there exists distinct $a,b\in\mathbb{Z}$ such that

$$\underbrace{1+\cdots+1}_{a}=\underbrace{1+\cdots+1}_{b}$$

where $a > b$, then

$$\underbrace{1+\cdots+1}_{a} = \underbrace{1+\cdots+1}_{b} + \underbrace{1+\cdots+1}_{a-b} = \underbrace{1+\cdots+1}_{b}$$

Hence, $\underbrace{1+\cdots+1}_{a-b} = 0 \implies \text{char}(F) = (a-b)$ which contradicts $\text{char}(F) = 0$. Thus, F is infinite. □

THEOREM 2.1.13. *If F is a finite field, then $\text{char}(F)$ is prime.*

Proof. Suppose for a contradiction that $\text{char}(F) = m$ is composite, say $m = ab$ where $2 \leqslant a, b \leqslant m-1$. Now

$$(\underbrace{1+\cdots+1}_{a})(\underbrace{1+\cdots+1}_{b}) = \underbrace{1+\cdots+1}_{m} = 0$$

since $\text{char}(F) = m$. Let $\underbrace{1+\cdots+1}_{a} = s$ and $\underbrace{1+\cdots+1}_{b} = t$, so $st = 0$ where $s \neq 0$. Since $\text{char}(F) = m > a$, there exists $c \in F$ such that $cs = 1 \implies c = s^{-1}$. Therefore $s^{-1}st = 0$. Thus, $t = 0$ which is a contradiction to $\text{char}(F) = m$. □

Roadmap: Let F be a finite field of order n. Then, $\text{char}(F) = p$ where p is prime. Then, $\mathbb{Z}_p$ is a subfield of F. F is a vector space over $\mathbb{Z}_p$ of $\dim = k$. Then, order of F is p^k.

2020-01-17

DEFINITION 2.1.14. We say two fields F and S are **isomorphic** if they have the same binary operations and if there exists a bijection between them.

DEFINITION 2.1.15. Let F be a field. A subset $S \subseteq F$ is called a **subfield** of F if S is a field itself with respect to the same operations of F.

EXAMPLE 2.1.16 (Subfield)**.** Let F be a finite field where $\text{char}(F) = p$. Consider $E = \{0, 1, 1+1, \ldots, \underbrace{1+\cdots+1}_{p-1}\} \subseteq F$. We see that E is a field with the same field operations as F. Also, E has order p. If we label the elements of E in a natural way such that $\underbrace{1+\cdots+1}_{p-1} \longleftrightarrow p-1$, then

$$E = \{0, 1, 1+1, \ldots, \underbrace{1+\cdots+1}_{p-1}\} = \mathbb{Z}_p = \{0, 1, 2, \ldots, p-1\} \subseteq F$$

So E is isomorphic to $\mathbb{Z}_p$.

THEOREM 2.1.17. *If F is a finite field of characteristic p, then $\mathbb{Z}_p$ is a subfield of F.*

Proof. Exercise. □

DEFINITION 2.1.18. Let F be a finite field, and consider $\mathbb{Z}_p \subseteq F$.
- Each $v \in F$ is vector.
- Each $c \in \mathbb{Z}_p$ is a scalar.
- Addition in F is defined by vector addition.
- Multiplication in F by elements in $\mathbb{Z}_p$ is defined by scalar multiplication.

THEOREM 2.1.19. *If F is a finite field of characteristic p, then F is a vector space over $\mathbb{Z}_p$.*

Proof. Exercise. □

THEOREM 2.1.20. *If F is a finite field of characteristic p, then*

$$\text{ord}(F) = p^k$$

for some $k \in \mathbb{Z}_{\geqslant 1}$.

Proof. Let k be the dimension of the vector space F over $\mathbb{Z}_p$. Let $\{\alpha_1, \ldots, \alpha_k\}$ be a basis for F. Then, every element in F can be written as

$$c_1\alpha_1 + \cdots + c_k\alpha_k$$

where $c_i \in \mathbb{Z}_p$. For each α_i, there are p possible choices for c_i, hence $\text{ord}(F) = p^k$. □

EXAMPLE 2.1.21. There is no field of order 6.

Question: Is there a finite field of order 4, 8, 9?

2.2 Irreducible Polynomials

DEFINITION 2.2.1. Let F be a field. The **set of all polynomials in x over F** (polynomials with coefficients from F) is denoted $F[x]$. Addition and multiplication are both done in the usual way, with coefficient arithmetic in F.

EXAMPLE 2.2.2. In $\mathbb{Z}_{11}$, $(2 + 5x + 6x^2) + (3 + 9x + 5x^2) = 5 + 3x$.

THEOREM 2.2.3. *Let F be a field. $F[x]$ is a commutative ring.*

DEFINITION 2.2.4. Let F be a field and let $f \in F[x]$ with $\deg(f) \geqslant 1$. If $g, h \in F[x]$ with $f \mid (g - h)$, then we write

$$g \equiv h \mod f$$

or equivalently, we can write $g - h = \ell f$ for some $\ell \in F[x]$.

THEOREM 2.2.5. *Congruence is an equivalence relation.*

DEFINITION 2.2.6. For a given $f \in F[x]$, the **equivalence class containing $g \in F[x]$** is

$$[g] = \{h \in F[x] : h \equiv g \mod f\}$$

DEFINITION 2.2.7. For $g, h \in F[x]$, we define addition and multiplication as follows:
- Addition: $[g] + [h] = [g + h]$
- Multiplication: $[g][h] = [gh]$

THEOREM 2.2.8. *1. The set of all equivalence classes, denoted $F[x]/(f)$ where $f \in F[x]$ and $\deg(f) \geqslant 1$ is a commutative ring.*
2. The polynomials in $F[x]$ of degree less than degree of f are a system of distinct representatives of equivalence classes in $F[x]/(f)$.

Proof of 5:

Proof. Let $g \in F[x]$. By division algorithm for polynomials we can write $g = \ell f + r$ where $\deg(r) < \deg(f)$. So, $g - r = \ell f$. Hence, $g \equiv r \mod f$. Thus, $[g] = [r]$ and we have $\deg(r) < \deg(f)$. Also, if $r_1, r_2 \in F[x]$ with $r_1 \neq r_2$, and $\deg(r_1), \deg(r_2) < \deg(f)$, then

$$f \nmid (r_1 - r_2) \iff r_1 \not\equiv r_2 \mod f$$

Thus, $[r_1] \neq [r_2]$. □

2020-01-20

DEFINITION 2.2.9. Let F be a field, and $f \in F[x]$ of degree $n \geqslant 1$. f is **irreducible** over F if f cannot be written as $f = gh$, where $g, h \in F[x]$ and $\deg(g), \deg(n) \geqslant 1$.

EXAMPLE 2.2.10 (Irreducible)**.**
- $x^2 + 1$ is irreducible over $\mathbb{R}$
- $x^2 + 1$ is reducible over $\mathbb{C}$ since $(x + i)(x - i) = x^1 + 1$
- $x^2 + 1$ is reducible over $\mathbb{Z}_2$ since $(x + 1)^2 = x^1 + 1$
- $x^2 + 1$ is irreducible over $\mathbb{Z}_3$

THEOREM 2.2.11. *Let F be a field and $f \in F[x]$ of degree $n \geqslant 1$. $F[x]/(f)$ is a field if and only if f is irreducible over F.*

Proof. Note that $F[x]/(f)$ is a commutative ring.

($\Longleftarrow$) Suppose $g \in F[x]/(f)$ where $g \neq 0$ and $\deg(g) < \deg(f)$. Then, $\gcd(g, f) = 1$ and so by EEA for polynomials, there exists $s, t \in F[x]$ such that

$$gs + ft = 1$$

Reducing both sides modulo f gives

$$gs \equiv 1 \mod f$$

So, $g^{-1} = s$. Hence $F[x]/(f)$ is a field.

($\Longrightarrow$) Exercise. □

We need an irreducible polynomial $f \in \mathbb{Z}_p[x]$ of degree n. Then, $\mathbb{Z}[x]/(f)$ is a finite field of order p^n.

THEOREM 2.2.12. *For any prime p and $n \in \mathbb{Z}_{\geqslant 2}$, there exists an irreducible polynomial of degree n over $\mathbb{Z}_p$.*

The proof is beyond the scope of this course.

THEOREM 2.2.13. *There exists a finite field of order q if and only if q is a prime power.*

EXAMPLE 2.2.14. Construct a finite field of order $2^2 = 4$.
Solution: Take $f(x) = x^2 + x + 1 \in \mathbb{Z}_2[x]$ which is irreducible over $\mathbb{Z}_2[x]$. Thus, the field is

$$\mathbb{Z}_2[x]/(x^2 + x + 1) = \{0, 1, x, x + 1\}$$

Examples of operations:
- $x + (x + 1) = 1$
- $x(x + 1) = x^2 + x = 1$
- $x^{-1} = x + 1$
- $1^{-1} = 1$
- $x^{-1} = x + 1$
- $(x + 1)^{-1} = x$

EXAMPLE 2.2.15. Construct a field of order $2^3 = 8$.
Solution: We need an irreducible polynomial of degree 3 over $\mathbb{Z}_2$. Take $f_1(x) = x^3 + x + 1$ which is irreducible over $\mathbb{Z}_2$. Then a field of order 8 is

$$F_1 = Z_2[x]/(x^3 + x + 1) = \left\{0, 1, x, x + 1, x^2, x^2 + 1, x^2 + x, x^2 + x + 1\right\}$$

Examples of operations:
- $x^2 + (x^2 + x + 1) = x + 1$
- $x^2(x^2 + x + 1) = x^4 + x^3 + x^2 = 1$
- $(x^2)^{-1} = x^2 + x + 1$
- $x^{-1} = x^2 + 1$

EXAMPLE 2.2.16. Construct a field of order $2^3 = 8$.
Solution: Take $f_2(x) = x^3 + x^2 + 1$. Then a field of order 8 is

$$F_2 = \mathbb{Z}_2[x]/(x^3 + x^2 + 1) = \left\{0, 1, x, x + 1, x^2, x^2 + 1, x^2 + x, x^2 + x + 1\right\}$$

Examples of operations:
- $x^{-1} = x^2 + x$

Note: F_1 and F_2 are two different fields of order $2^3 = 8$, but they are isomorphic. That is, there is a bijection $\alpha : F_1 \to F_2$ such that

$$\alpha(a + b) = \alpha(a) + \alpha(b)$$
$$\alpha(ab) = \alpha(a)\alpha(b)$$

for all $a, b \in F_1$.

THEOREM 2.2.17. *Any two finite fields of order q are isomorphic.*

Proof. Exericse. □

DEFINITION 2.2.18. We will denote the **Galois field of order** q by $GF(q)$.

We saw two different representations of $GF(2^3)$.

2020-01-22

EXAMPLE 2.2.19. Construct $GF(2^4 = 16)$.
Solution: Take $f(x) = x^4 + x + 1 \in \mathbb{Z}_2[x]$.
- f has no roots in $\mathbb{Z}_2$ and hence no linear factors
- long division shows that $x^2 + x + 1 \nmid x^4 + x + 1$, so f has no irreducible quadratic factors
- f is irreducible over $\mathbb{Z}_2$.

Thus, $GF(16) = \mathbb{Z}_2[x]/(x^4 + x + 1)$.

2.3 Properties of Finite Fields

THEOREM 2.3.1 (Frosh's Dream)**.** *Let* $\alpha, \beta \in GF(q)$ *where* $\text{char}(GF(q)) = p$.

$$(\alpha + \beta)^p = \alpha^p + \beta^p$$

Proof.

$$(\alpha + \beta)^p = \alpha^p + \sum_{i=1}^{p-1} \binom{p}{i} \alpha^i \beta^{p-i} + \beta^p$$

Now,

$$\begin{aligned}\binom{p}{i} = \frac{p!}{i!(p-i)!} &= \frac{p(p-1)\cdots(p-i+1)(p-i)(p-i-1)\cdots(2)(1)}{[i(i-1)\cdots(2)(1)]\,[(p-i)(p-i-1)\cdots(2)(1)]} \\ &= p\left[\frac{(p-1)\cdots(p-i+2)}{i(i-1)\cdots(2)(1)}\right]\end{aligned}$$

If $1 \leqslant i \leqslant p-1$ then $p \mid$ numerator, but $p \nmid$ denominator. Thus,

$$p \mid \binom{p}{i} = p\lambda$$

where $\lambda \in \mathbb{N}$ with $\lambda \neq 0$ and $p \nmid \lambda$.

$$\begin{aligned}\sum_{i=1}^{p-1} \binom{p}{i} \alpha^i \beta^{p-i} &= \sum_{i=1}^{p-1} (p\lambda_i) \alpha^i \beta^{p-i} \\ &= \sum_{i=1}^{p-1} (\underbrace{1 + \cdots + 1}_{p}) \lambda_i \alpha^i \beta^{p-i} \\ &= 0\end{aligned}$$

Thus, $(\alpha + \beta)^p = \alpha^p + \beta^p$. □

COROLLARY 2.3.2.

$$(\alpha+\beta)^{p^m} = \alpha^{p^m} + \beta^{p^m}$$

for all $m \geqslant 1$.

Proof. Exercise. Hint: Induction on m. □

THEOREM 2.3.3. *Let* $\alpha \in GF(q)$. *Then*

$$\alpha^q = \alpha$$

Proof. If $\alpha = 0$, then $\alpha^q = 0 = \alpha$.

If $\alpha \neq 0$, let $\{\alpha_1, \ldots, \alpha_{q-1}\}$ be the non-zero elements in $GF(q)$. Consider

$$\{\alpha\alpha_1, \ldots, \alpha\alpha_{q-1}\}$$

Note that the elements in this list are pairwise distinct because if $\alpha\alpha_i = \alpha\alpha_j$ with $i \neq j$, then

$$\alpha^{-1}\alpha\alpha_i = \alpha^{-1}\alpha\alpha_j$$

which implies that $\alpha_i = \alpha_j$ which is a contradiction. Also $\alpha\alpha_i \neq 0$ for all $i \in [1, q-1]$. Hence, $\{\alpha_1, \ldots, \alpha_{q-1}\} = \{\alpha\alpha_1, \ldots, \alpha\alpha_{q-1}\}$. Therefore, $\alpha_1 \cdots \alpha_{q-1} = (\alpha\alpha_1) \cdots (\alpha\alpha_{q-1})$. Hence, $\alpha^{q-1} = 1$. Thus, $\alpha^q = \alpha$. □

DEFINITION 2.3.4. Let $GF(q)^* = GF(q)/\{0\}$.

DEFINITION 2.3.5. Let $\alpha \in GF(q)^*$. The **order of** α, denoted $\mathrm{ord}(\alpha)$ is the smallest positive integer t such that $\alpha^t = 1$.

EXAMPLE 2.3.6. How many elements of order 1 are there in $GF(q)$?
Solution: $\alpha = 1$

EXAMPLE 2.3.7. Find $\mathrm{ord}(x)$ in $GF(16) = \mathbb{Z}_2/(x^4 + x + 1)$.
Solution:

- $x^1 = x$
- $x^2 = x^2$
- $x^3 = x^3$
- $x^4 = x + 1$
- $x^5 = x^2 + x$
- $x^6 = x^3 + x^2$
- $x^7 = x^3 + x + 1$
- $x^8 = x^2 + 1$
- $x^9 = x^3 + x$
- $x^{10} = x^2 + x + 1$
- $x^{11} = x^3 + x^2 + x$
- $x^{12} = x^3 + x^2 + x + 1$
- $x^{13} = x^3 + x^2 + 1$
- $x^{14} = x^3 + 1$
- $x^15 \equiv 1 \mod x^4 + x + 1$

Since $\mathrm{ord}(x) \neq 1, 3, 5$ $\mathrm{ord}(x) \mid 15$, so we have $\mathrm{ord}(x) = 15$.

LEMMA 2.3.8. *Let $\alpha \in GF(q)^*$, $\operatorname{ord}(\alpha) = t$ and $s \in \mathbb{Z}$.*

$$\alpha^s = 1 \iff t \mid s$$

Proof. Let $s \in \mathbb{Z}$. By the division algorithm for integers,

$$s = \ell t + r$$

where $0 \leqslant r \leqslant t-1$. Then

$$\alpha^s = \alpha^{\ell t + r} = (\alpha^t)^\ell \alpha^r = \alpha^r$$

So,

$$\begin{aligned} \alpha^s = 1 &\iff a^r = 1 \\ &\iff r = 0 \qquad \text{since } 0 \leqslant 1 \leqslant t-1 \\ &\iff t \mid s \end{aligned}$$

□

COROLLARY 2.3.9. *If $\alpha \in GF(q)^*$, then $\operatorname{ord}(\alpha) \mid (q-1)$.*

Proof. We know $\alpha^{q-1} = 1$, so $\operatorname{ord}(\alpha) \mid (q-1)$ by the previous Lemma. □

DEFINITION 2.3.10. An element $\alpha \in GF(q)$ is a **generator** of $GF(q)^*$ if $\operatorname{ord}(\alpha) = q-1$.

THEOREM 2.3.11. *If α is a generator of $GF(q)^*$, then*

$$\{\alpha^1, \ldots, \alpha^{q-1}\} = GF(q)^*$$

2020-01-24

THEOREM 2.3.12. *If $GF(q)^*$ has order t, then*

$$\alpha^1, \ldots, \alpha^{t-1}$$

are pairwise distinct.

Proof. Suppose for a contradiction that $\alpha^i = \alpha^j$ where $0 \leqslant i, j \leqslant t-1$. WLOG suppose $j > i$, then $\alpha^{j-i} = 1$ which contradicts $\operatorname{ord}(\alpha) = t$ since $1 \leqslant j - i \leqslant t-1$. □

2.4 Existence of Generators (Optional)

LEMMA 2.4.1. *Let $\alpha \in GF(q)^*$ with $\operatorname{ord}(\alpha) = t$. Then $\operatorname{ord}(\alpha^i) = t/\gcd(t, i)$.*

Proof. Let $d = \gcd(t, i)$. The order of a^i is the smallest positive integer s such that $\alpha^{is} = 1$. Now,

$$\alpha^{is} = 1 \iff t \mid is \iff \frac{t}{d} \mid \frac{i}{d}s \iff \frac{t}{d} \mid s$$

Since the smallest positive integer s satisfying $\frac{t}{d} \mid s$ is $s = \frac{t}{d}$, we have $\mathrm{ord}(\alpha^i) = \frac{t}{d}$. □

LEMMA 2.4.2. *Let $\alpha, \beta \in GF(q)^*$, with $\mathrm{ord}(\alpha) = m$ and $\mathrm{ord}(\beta) = n$. If $\gcd(m, n) = 1$ then $\mathrm{ord}(\alpha\beta) = mn$.*

Proof. Let $t = \mathrm{ord}(\alpha\beta)$. Now,

$$(\alpha\beta)^{mn} = \alpha^{mn}\beta^{mn} = 1,$$

so $t \mid mn$. Also,

$$1 = (\alpha\beta)^{tn} = \alpha^{tn}\beta^{tn} = \alpha^{tn},$$

so $m \mid tn$. And, since $\gcd(m, n) = 1$, we have $m \mid t$. Similarly,

$$1 = (\alpha\beta)^{tm} = \alpha^{tm}\beta^{tm} = \beta^{tm},$$

so $n \mid tm$. And, since $\gcd(m, n) = 1$, we have $n \mid t$. Hence, since $\gcd(m, n) = 1$, we have $mn \mid t$. Thus $t = mn$. □

THEOREM 2.4.3. *Every finite field $GF(q)$ has a generator.*

Proof. Let α be an element of highest order in $GF(q)^*$; say $\mathrm{ord}(\alpha) = t$. Suppose that $t < (q - 1)$.

If the order of every element in $GF(q)^*$ were to divide t then the equation $y^t - 1 = 0$ would have $q - 1$ roots in $GF(q)$, which is impossible since $(q - 1) > t$. Hence there exists an element $\beta \in GF(q)^*$ whose order b does not divide t.

Now, let ℓ be a prime such that the highest power of ℓ which divides b (say ℓ^e) is greater than the highest power of ℓ which divides t (say ℓ^f) — such a prime ℓ must exist since b does not divide t.

Consider the field elements $\alpha' = \alpha^{\ell^f}$ and $\beta' = \beta^{b/\ell^e}$. We have

$$\mathrm{ord}(\alpha') = \frac{t}{\gcd(t, \ell^f)} = \frac{t}{\ell^f}$$

and

$$\mathrm{ord}(\beta') = \frac{b}{\gcd(b, \ell^e)} = \frac{b}{b/\ell^e} = \ell^e$$

Since $\gcd(t/\ell^f, \ell^e) = 1$, we have $\mathrm{ord}(\alpha'\beta') = (t/\ell^f)(\ell^e) = t\ell^{e-f} > t$. This contradicts the hypothesis that the highest order of any element in $GF(q)^*$ is t. Hence the hypothesis that $t < (q-1)$ is wrong, and so $t = q - 1$. Thus α is a generator of $GF(q)^*$. □

Chapter 3

Linear Codes

3.1 Introduction

Let $F = GF(q)$. Let $V_n(F) = \underbrace{F \times \cdots \times F}_{n} = F^n$. Then, $V_n(F)$ is an n-dimensional vector space over F and we have $|V_n(F)| = q^n$.

DEFINITION 3.1.1. Let $F = GF(q)$. A **linear** (n, k)**-code** over F is an n-dimensional subspace of $V_n(F)$.

DEFINITION 3.1.2. A subspace of a vector space V over F is a subset $S \subseteq V$ such that
V1 $0 \in S \implies S \neq \varnothing$
V2 $v_1 + v_2 \in S, \forall v_1, v_2 \in S$
V3 $\lambda v \in S, \forall \lambda \in F$ and $v \in S$
Note that $S \subseteq V$ is also a vector space over F.

Let C be an (n, k)-code over F. Let $v_1, \ldots, v_k$ be an ordered basis for C.

(1) The codewords in C are precisely:

$$m_1 v_1 + \cdots + m_k v_k$$

where $m_i \in F$. So, $|C| = M = q^k$ since there are q choices for each m. The length of C is n and has dimension k.

(2) The rate of C is

$$R = \frac{\log_q(M)}{n} = \frac{k}{n}$$

DEFINITION 3.1.3. The **Hamming weight** of $v \in V_n(F)$, denoted $w(v)$ is the number of non-zero coordinate positions in V.

DEFINITION 3.1.4. The **Hamming weight of an** (n, k)**-code** C is:

$$w(C) = \min \{w(c) : c \in C,\ c \neq 0\}$$

THEOREM 3.1.5. *If C is a linear code, then $d(C) = w(C)$.*

Proof.

$$\begin{aligned} d(C) &= \min\{d(x,y) : x,y \in C,\ x \neq y\} \\ &= \min\{w(x-y) : x,y \in C,\ x \neq y\} \qquad \text{by (A2Q1a)} \\ &= \min\{w(c) : c \in C,\ x \neq y\} \qquad \text{since } C \text{ is a vector space} \\ &= w(C) \end{aligned}$$

□

3.2 Generator Matrices and the Dual Code

Since $M = q^k$, there are q^k source messages. We'll assume that the source messages are elements of $V_k(F)$. Then, a natural encoding rule is, given $(m_1, \ldots, m_k) \in V_k(F)$ we'll encode the message as

$$c = m_1 v_1 + \cdots + m_k v_k$$

The encoding rule depends on the basis chosen for C.

If $m = (m_1, \ldots, m_k)$, then the encoding rule can be written as follows:

$$\begin{aligned} C &= (m_1, \ldots, m_k) \begin{bmatrix} -v_1- \\ -v_2- \\ \vdots \\ -v_k- \end{bmatrix}_{k \times n} \\ &= mG \end{aligned}$$

Note that v_i are row vectors in this course.

DEFINITION 3.2.1. Let C be an (n,k)-code. A **generator matrix** G for C is a $k \times n$ matrix whose rows form a basis for C.

Note: An encoding rule for C with respect to G is $C = mG$. Performing elementary row operations on G gives a different matrix for the same code C due to the order of the basis.

2020-01-27

EXAMPLE 3.2.2. Consider a $\underbrace{\text{binary}}_{F=GF(2)=\mathbb{Z}_2}$ $(\underbrace{5}_{n}, \underbrace{3}_{k})$-code C. Then $M = q^k = 2^3$ and $R = \frac{k}{n} = \frac{3}{5}$.

$C = \langle \underbrace{10010}_{v_1}, \underbrace{01011}_{v_2}, \underbrace{00101}_{v_3} \rangle$.

$$G = \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{array}\right]_{3 \times 5}$$

$\operatorname{rank}(G) = 3$, thus G is a generator matrix for C.

M (source messages)	$\rightarrow$	C (codewords)
000	$\rightarrow$	00000
001	$\rightarrow$	00101
010	$\rightarrow$	01011
011	$\rightarrow$	01110
100	$\rightarrow$	10010
101	$\rightarrow$	10111
110	$\rightarrow$	11001
111	$\rightarrow$	11100

$d(C) = 2$, $e = 0$.

Note: Any matrix equivalent to G is also a generator matrix for C, but yields a different encoding rule.

DEFINITION 3.2.3. Let $[\, I_k \mid A \,]_{k \times n}$ be a generator matrix for an (n, k)-code C. If an (n, k)-code has a generator matrix of this form, then C is **systematic**, and the generator matrix is in **standard form**.

EXAMPLE 3.2.4. $C = \langle 100011, 101010, 100110 \rangle$ is a non-systematic $(6, 3)$-code. Some generator matrices are:

$$G_1 = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{array}\right]$$

$R_2 + R_1$:

$$G_2 = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{array}\right]$$

$R_3 + R_1$:

$$G_3 = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{array}\right]$$

Clearly C is not systematic. However, if every codeword is permuted by moving the second bit to the fourth bit, we get C' that is linear and has the same length, dimension, and distance as C.

DEFINITION 3.2.5. Let C be an (n, k)-code. If π is a permutation on $\{1, \ldots, n\}$. Then $\pi(C)$ (that is, apply π to each codeword) is an (n, k)-code which is said to be an **equivalent code** for C.

THEOREM 3.2.6. *(1) If C and C' are equivalent codes, then*

$$d(C) = d(C')$$

(2) Every linear code is equivalent to a systematic code.

Proof. Let C be an (n, k) code. Let G be a generator matrix for C in RREF. Them, one can permute the columns of G to get a matrix $G' = [\, I_k \mid A \,]$ in standard form. Then, G' is a generator matrix for a code C' that is equivalent to C. □

DEFINITION 3.2.7. Let $x, y \in V_n(F)$. The **inner product** of x and y is $x \cdot y = \sum_{i=1}^{n} x_i y_i \in F$

THEOREM 3.2.8. *If $x, y, z \in V_n(F)$ and $\lambda \in F$, then*
(1) $x \cdot y = y \cdot x$
(2) $x \cdot (y + z) = x \cdot y + x \cdot y$
(3) $(\lambda x) \cdot y = \lambda(x \cdot y)$
(4) $x \cdot x = 0$ *does* ***not*** *imply* $x = 0$

EXAMPLE 3.2.9. Consider $V_2(\mathbb{Z}_2)$. Then, $(1,1) \cdot (1,1) = 0$.

DEFINITION 3.2.10. Let C be an (n,k)-code over F. The **dual code** of C is

$$C^\perp = \{x \in V_n(F) : x \cdot c = 0 \, \forall c \in C\}$$

THEOREM 3.2.11. *Let $x \in V_n(F)$.*

$$x \in C^\perp \iff v_1 \cdot x = \cdots = v_k \cdot x = 0$$

Proof. ($\implies$) If $x \in C^\perp$, then $x \cdot c = 0$ for all $c \in C$. In particular,

$$x \cdot v_1 = \cdots = x \cdot v_k = 0$$

($\impliedby$) Suppose $x \cdot v_1 = \cdots = x \cdot v_k = 0$. Let $c \in C$. We can write

$$c = \lambda_1 v_1 + \cdots + \lambda_k v_k$$

for all $\lambda_i \in F$. Then,

$$x \cdot c = \lambda_1(x \cdot v_1) + \cdots + \lambda_k(x \cdot v_k) = 0$$

Hence, $x \in C^\perp$. □

THEOREM 3.2.12. *If C is an (n,k)-code over F, then $C^\perp$ is an $(n, n-k)$-code over F.*

Proof. Consider

$$G = \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix}_{k \times n}$$

Then, $x \in C^\perp$ if and only if $Gx^\top = 0$. So, $C^\perp$ is the nullspace of G. Hence, $C^\perp$ is an $(n-k)$-dimensional subspace of $V_n(F)$. □

2020-01-29

DEFINITION 3.2.13. If $x, y \in V_n(F)$ and $x \cdot y = 0$, then x and y are **orthogonal**.

THEOREM 3.2.14. *If C is a linear code, then $(C^\perp)^\perp = C$.*

Proof. Let C be an (n,k)-code. Then $C^\perp$ is an $(n, n-k)$-code. So, $(C^\perp)^\perp$ is an (n,k)-code. But $C \subseteq (C^\perp)^\perp$ by definition of $C^\perp$. Suppose C is a code over $F = GF(q)$. Then $|C| = q^k$ and $|(C^\perp)^\perp| = q^k$. Thus, $C = (C^\perp)^\perp$. □

THEOREM 3.2.15. *Let C be an (n,k)-code with standard form $k \times n$ generator matrix. Then, a generator matrix for $C^\perp$ is*

$$H = \left[-A^\top \mid I_{n-k} \right]_{(n-k)\times n}$$

Proof. $\operatorname{rank}(H) = n - k$, so H is indeed a generator matrix for some $(n, n-k)$-code $\overline{C}$. Now,

$$\begin{aligned} GH^\top &= [\, I_k \mid A \,] \begin{bmatrix} -A \\ I_{n-k} \end{bmatrix} \\ &= -A + A \\ &= 0 \end{aligned}$$

Since $GH^\top = 0$, every row of H is orthogonal to every row of G, so every vector in the row space of H is orthogonal to every vector in the row space of G. Hence, $\overline{C} \subseteq C$. Since $\dim(\overline{C}) = \dim(C^\perp)$ we have $\overline{C} = C^\perp$. □

3.3 The Parity-Check Matrix

DEFINITION 3.3.1. A generator matrix for $C^\perp$ is called a **parity-check matrix** (PCM) for C.

EXAMPLE 3.3.2. Consider a $(5,2)$-code C over $\mathbb{Z}_3$ with generator matrix

$$G = \begin{bmatrix} 2 & 0 & 2 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} \leftarrow c_1 \\ \leftarrow c_2 \end{matrix}$$

Find the length, dimension, order, number of codewords, codewords, distance, weight and errors that can be corrected for C.

Solution:

- Length: $n = 5$ ((n,k)-code)
- Dimension: $k = 2$ ((n,k)-code)
- Order: $q = 3$ ($\mathbb{Z}_3$)
- Number of codewords: $M = q^k = 3^2 = 9$
- Codewords: $C = \{00000, 20210, 10120, 11001, 22002, 01211, 12212, 21121, 02122\}$
- Distance: $d(C) = w(C) = 3$
- Error-correcting capability: $e = 1$

Find a generator matrix for $C^\perp$.

Solution:

$$\begin{bmatrix} 2 & 0 & 2 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 1 & 1 \end{bmatrix}$$

So,

$$H = \left[\begin{array}{cc|ccc} 2 & 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 1 \end{array}\right]$$

is a generator matrix for $C^\perp$ which is a $(5,3)$-code over $\mathbb{Z}_3$. $M = 3^3 = 27$.

THEOREM 3.3.3. *Let C be an (n,k)-code over F, and let H be a PCM for C. Then $d(C) \geqslant s$ if and only if every $(s-1)$ columns of H are linearly independent over F.*

Proof. Let $h_1, \ldots, h_n$ be the columns of H.

($\Longleftarrow$) Suppose $d(C) \leqslant s-1$, so $w(C) \leqslant s-1$. Let $c \in C$, with $1 \leqslant w(c) \leqslant s-1$. WLOG, suppose $c_j = 0$ for all $s \leqslant j \leqslant n$. Since $c \in C$, we have $Hc^\top = 0$. Therefore, $c_1h_1 + \cdots + c_{s-1}h_{s-1} = 0$. Since $w(C) \geqslant 1$, this is a non-trivial linear combination of $h_1, \ldots, h_{s-1}$ that equal 0. So, $h_1, \ldots, h_{s-1}$ are linearly dependent over F.

($\Longrightarrow$) Suppose there are $s-1$ columns of H that are linearly dependent over F, say $h_1, \ldots, h_{s-1}$. So, we can write

$$c_1h_1 + \cdots + c_{s-1}h_{s-1}$$

where $c_j \in F$ not all zero. Let $c = (c_1, \ldots, c_{s-1}, \underbrace{0 \cdots 0}_{n-s+1}) \in V_n(F)$. Then, $Hc^\top = 0$. So, c_iC and $1 \leqslant w(c) \leqslant s-1$, so $d(C) \leqslant s-1$. □

COROLLARY 3.3.4. *Let C be an (n,k)-code over F with PCM H. Then, $d(C)$ is the smallest number of columns of H that are linearly dependent over F.*

EXAMPLE 3.3.5. Recall, we found a PCM

$$H = \left[\begin{array}{cc|ccc} 2 & 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 1 \end{array}\right]$$

for a $(5,2)$-code C over $\mathbb{Z}_3$. Find $d(C)$.
Solution:

- No 0 column in $H \implies d(C) \geqslant 2$
- No two linearly dependent columns in H (since no repeated columns, and no column is two times another column $\implies d(C) \geqslant 2$)

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

Therefore $d(C) \ngeqslant 4$, therefore $d(C) = 3$.

EXAMPLE 3.3.6. Let C be a binary code with PCM H.

- $d(C) = 1 \iff H$ has a 0 column.
- $d(C) = 2 \iff$ the columns of H are non-zero and two are the same.
- $d(C) = 3 \iff$ the columns of H are non-zero, distinct, and one column is the sum of two other (distinct) columns.

3.4 Hamming Codes and Perfect Codes

EXAMPLE 3.4.1. Construct a $(7, 4, 3)$-binary code C.
Solution: Consider a PCM for C:

$$H = \left[\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{array}\right]_{3\times 7}$$

This is a **Hamming Code** of order 3 over $GF(2)$.

DEFINITION 3.4.2. Let C be an $[n, M]$-code with distance d over an alphabet A of size q. Let $e = \lfloor \frac{d-1}{2} \rfloor$. The **sphere packing bound** or **Hamming bound** is:

$$M \sum_{i=0}^{e} \binom{n}{i} (q-1)^i \leqslant q^n$$

DEFINITION 3.4.3. Let C be an $[n, M]$-code over A of distance d. Then, C is perfect if

$$M \sum_{i=0}^{e} \binom{n}{i} (q-1)^i = q^n$$

Note: If C is perfect, then IMLD=CMLD.

2020-01-03

For fixed n, q, d, a perfect code maximizes

$$R = \frac{log_q(M)}{n}$$

EXAMPLE 3.4.4. $GF(q)^n$ is a trivial perfect code with $d = 1$.
$C = \{\underbrace{0\cdots 0}_{n}, \underbrace{1\cdots 1}_{n}\}$ over $\mathbb{Z}_2$ is a perfect code if n is odd.

Proof.

$$\begin{aligned} 2\left(\sum_{i=0}^{e}\binom{n}{i}\right) &= 2\left(\binom{n}{0} + \binom{n}{e}\right) \\ &= \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{e+1} + \cdots \binom{n}{n-1} + \binom{n}{n} \\ &= (1+1)^n \\ &= 2^n \end{aligned}$$

□

Exercise: Prove that every perfect code must have odd distance (without using the theorem below)

THEOREM 3.4.5 (Tietäväinen, 1973). *The only perfect codes are:*
(1) $V_n(GF(q))$.
(2) The binary replication code of odd length.
(3) The $(23, 12, 7)$*-binary Golay code and all codes equivalent to it.*
(4) The $(11, 6, 5)$*-ternary Golay code and all codes equivalent to it. A generator matrix for this code is:*

$$G = \left[\begin{array}{c|ccccc} & 1 & 1 & 1 & 1 & 1 \\ & 0 & 1 & 2 & 2 & 1 \\ I_6 & 1 & 0 & 1 & 2 & 2 \\ & 2 & 1 & 0 & 1 & 2 \\ & 2 & 2 & 1 & 0 & 1 \\ & 1 & 2 & 2 & 1 & 0 \end{array}\right]_{6\times 11}$$

(5) The Hamming codes and all codes of the same $[n, M, d]$ *parameters as them with* $d = 3$.

EXAMPLE 3.4.6. A Hamming code of order $r = 3$ over $GF(3)$ is a $(13, 10, 3)$-code over $GF(3)$ with PCM:

$$H = \left[\begin{array}{ccc|ccc|ccc|ccc|c} 1 & 0 & 0 & 1 & 0 & 1 & 2 & 2 & 0 & 2 & 1 & 2 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 2 & 2 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 2 & 1 & 2 & 2 & 1 \end{array}\right]_{3\times 13}$$

Observations:

(i) For every non-zero vector $v \in V_r(GF(q))$, exactly one scalar multiple of v must be a column of a PCM (for the Hamming code of order r over $GF(q)$)

(ii) The dimension of the code is indeed k since $\mathrm{rank}(\text{PCM}) = r = n - k$ since $\lambda_i e_i$ are columns of the PCM.

(iii) The Hamming codes have distance 3.

THEOREM 3.4.7. *Hamming codes are perfect.*

Proof. Recall that Hamming codes have $e = 1$ and $n = \frac{q^r-1}{q-1}$ with $r = n - k$.

$$\begin{aligned} M\sum_{i=0}^{e}\binom{n}{i}(q-1)^i &= q^{n-r}(1 + n(q-1)) \\ &= q^{n-r}\left(1 + \frac{q^r-1}{q-1}(q-1)\right) \\ &= q^n \end{aligned}$$

□

DEFINITION 3.4.8. Suppose $\boldsymbol{c} \in C$ is transmitted. Suppose $\boldsymbol{r} \in V_n(F)$ is received. Then, the **error vector** is $\boldsymbol{e} = \boldsymbol{r} - \boldsymbol{c}$.

EXAMPLE 3.4.9 (Error Vector). Over $\mathbb{Z}_3$, if $\boldsymbol{c} = (120212)$ is sent, and $\boldsymbol{r} = (122102)$ is received, then the error vector is $\boldsymbol{e} = (002220)$.