



# *Coding Theory*

CO 331



Alfred John Menezes

# Preface

---

**Disclaimer** Much of the information on this set of notes is transcribed directly/indirectly from the lectures of CO 331 during Winter 2020 as well as other related resources. I do not make any warranties about the completeness, reliability and accuracy of this set of notes. Use at your own risk.

For any questions, send me an email via <https://notes.sibeliusp.com/contact/>.

You can find my notes for other courses on <https://notes.sibeliusp.com/>.

---

*Sibeliusp Peng*

# Contents

---

<b>Preface</b>	<b>1</b>
<b>0 Pre</b>	<b>3</b>
<b>1 Introduction &amp; Fundamentals</b>	<b>5</b>
1.1 Decoding Strategy . . . . .	7
1.1.1 Nearest Neighbour Decoding . . . . .	8
1.2 Error Correcting & Detecting Capabilities of a Code . . . . .	10
<b>2 Introduction to Finite Fields</b>	<b>13</b>
2.1 Non-existence of finite fields . . . . .	16
2.2 Constructing finite fields . . . . .	18
2.3 Properties of finite fields . . . . .	21
<b>3 Linear Codes</b>	<b>24</b>
3.1 Properties of Linear Codes . . . . .	24
3.2 Dual Codes . . . . .	28

# Pre

---

## Example: Replication code

source msgs		codewords
0	→	0
1	→	1

# of errors/codeword that be detected: 0  
 # errors/codeword that can be corrected: 0  
 Rate: 1

---

source msgs		codewords
0	→	00
1	→	11

# of errors/codeword that be detected: 1  
 # errors/codeword that can be corrected: 0  
 Rate: 1/2

---

source msgs		codewords
0	→	000
1	→	111

# of errors/codeword that be detected: 2  
 # errors/codeword that can be corrected: 1 (nearest neighbour decoding)  
 Rate: 1/3

---

source msgs		codewords
0	→	00000
1	→	11111

# of errors/codeword that be detected: 4  
 # errors/codeword that can be corrected: 2 (nearest neighbour decoding)  
 Rate: 1/5

**Goal of Coding Theory** Design codes so that:

1. High information rate
2. High error-correcting capability
3. Efficient encoding & decoding algorithms



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



# Introduction & Fundamentals

## alphabet, word, length...

An *alphabet*  $A$  is a finite set of  $q \geq 2$  symbols. E.g.  $A = \{0, 1\}$ .

A *word* is a finite sequence of symbols from  $A$ . (tuples or vectors)

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

A *code*  $C$  over  $A$  is a finite set of words over  $A$  (of size  $\geq 2$ ).

A *codeword* is a word in  $C$ .

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  $|C| = M$ . This is denoted by  $[n, M]$ .

### Example:

$A = \{0, 1\}$ .  $C = \{00000, 11100, 00111, 10101\}$  is a  $[5, 4]$ -code over  $\{0, 1\}$ .

Messages		Codewords
00	→	00000
10	→	11100
01	→	00111
11	→	10101

Encoding 1-1 map

The channel encoder transmits only codewords. But, what's received by the channel decoder might not be codeword.

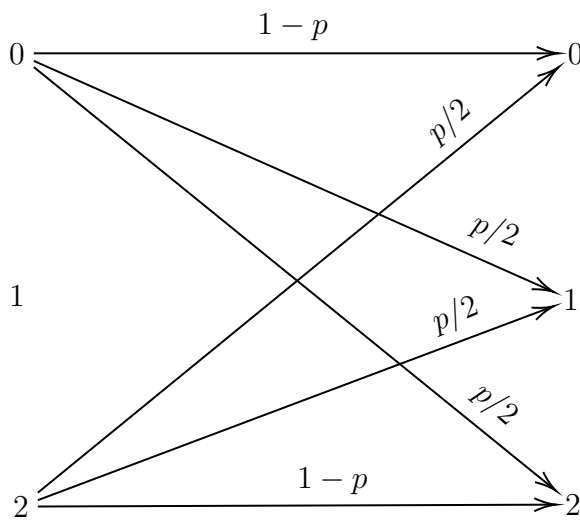
### Example:

Suppose the channel decoder receives  $r = 11001$ . What should it do?

Example:  $q = 2$  (Binary symmetric channel, BSC)



Example:  $q = 3$



Assumptions about the communications channel

- 1) The channel only transmits symbols from  $A$ .
- 2) No symbols are deleted, added, or transposed.
- 3) (Errors are “random”) Suppose the symbol transmitted are  $X_1, X_2, X_3, \dots$ . Suppose the symbols received are  $Y_1, Y_2, Y_3, \dots$ . Then for all  $i \geq 1$ , and all  $i \leq j, k \leq q$ ,

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

where  $p = \text{symbol error prob.}$

### Notes about BSC

- (i) If  $p = 0$ , the channel is perfect.
- (ii) If  $p = \frac{1}{2}$ , the channel is useless.
- (iii) If  $1 \geq p > \frac{1}{2}$ , then simply flip all bits that are received.

- (iv) WLOG, we will assume that  $0 < p < \frac{1}{2}$ .
- (v) Analogously, for a  $q$ -ary channel, we can assume that  $0 < p < \frac{q-1}{q}$ . (Optional exercise)

### Hamming distance

If  $x, y \in A^n$ , the *Hamming distance*  $d(x, y)$  is the # of coordinate positions in which  $x$  &  $y$  differ.

The *distance of a code*  $C$  is

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

**Example:**

$$d(10111, 01010) = 4$$

### Theorem 1.1

$d$  is a metric. For all  $x, y, z \in A^n$

- (i)  $d(x, y) \geq 0$ , and  $d(x, y) = 0$  iff  $x = y$ .
- (ii)  $d(x, y) = d(y, x)$
- (iii)  $\triangle$  inequality  $d(x, z) \leq d(x, y) + d(y, z)$

### rate

The *rate* of an  $[n, M]$ -code  $C$  over  $A$  with  $|A| = q$  is

$$R = \frac{\log_q M}{n}.$$

If the source messages are all  $k$ -tuples over  $A$ ,

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

**Example:**

$$C = \{00000, 11100, 00111, 10101\} \quad A = \{0, 1\}$$

Here  $R = \frac{2}{5}$  and  $d(C) = 2$ .

## 1.1 Decoding Strategy

Let  $C$  be an  $[n, M]$ -code over  $A$  of distance  $d$ . Suppose some codeword is transmitted, and  $r \in A^n$  is received. The channel decoder has to decide the following:



- (i) no errors have occurred, accept  $r$ .
- (ii) errors have occurred, and (decode) correct  $r$  to some codeword.
- (iii) errors has occurred, correction is not possible.

### 1.1.1 Nearest Neighbour Decoding

Incomplete Maximum Likelihood Decoding (IMLD). Correct  $r$  to the unique codeword  $c$  for which  $d(r, c)$  is smallest. If  $c$  is not unique, reject  $r$ . Complete MLD (CMLD). Same as IMLD, accept ties are broken arbitrarily.

**Question** Is IMLD a reasonable strategy?

#### Theorem 1.2

IMLD selects the codeword  $c$  that maximizes  $P(r|c)$  prob. that  $r$  is received given that  $c$  was sent.

**Proof:**

Suppose  $c_1, c_2 \in C$  with  $d(c_1, r) = d_1$  and  $d(c_2, r) = d_2$ . Suppose  $d_1 > d_2$ .

Now

$$P(r|c_1) = (1-p)^{n-d_1} \left( \frac{p}{q-1} \right)^{d_1}$$

and

$$P(r|c_2) = (1-p)^{n-d_2} \left( \frac{p}{q-1} \right)^{d_2}$$

So,

$$\frac{P(r|c_1)}{P(r|c_2)} = (1-p)^{d_2-d_1} \left( \frac{p}{q-1} \right)^{d_1-d_2} = \left( \frac{p}{(1-p)(q-1)} \right)^{d_1-d_2}$$

Recall

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

Hence

$$\frac{P(r|c_1)}{P(r|c_2)} < 1$$

and so

$$P(r|c_1) < P(r|c_2)$$

□

The ideal strategy is to correct  $r$  to  $c \in C$  that minimizes  $P(c|r)$ . This is Minimum

error decoding (MED).

**Example: (IMD is not the same as MED)**

Let  $C = \{\underbrace{000}_{c_1}, \underbrace{111}_{c_2}\}$ . (corresponding to 0, 1).

Suppose  $P(c_1) = 0.1, P(c_2) = 0.9$ . Suppose  $p = 1/4$  and  $r = 100$ .

**IMLD**  $r \rightarrow 000$

**MED**

$$\begin{aligned} P(c_1|r) &= \frac{P(r|c_1) \cdot P(c_1)}{P(r)} \\ &= p(1-p)^2 \times 0.1 / P(r) \\ &= \frac{9}{640 \cdot P(r)} \end{aligned}$$

Similarly

$$\begin{aligned} P(c_2|r) &= \frac{P(r|c_2) \cdot P(c_2)}{P(r)} \\ &= p(1-p)^2 \times 0.9 / P(r) \\ &= \frac{27}{640 \cdot P(r)} \end{aligned}$$

So MED:  $r \rightarrow 111$

**Note**

1. IMLD: Select  $c$ . s.t.  $P(r|c)$  is maximum  
MED: Select  $c$ . s.t.  $P(c|r)$  is maximum
2. MED has the drawback that it requires knowledge of  $P(c_i)$ ,  $1 \leq i \leq M$
3. Suppose source messages are equally likely, so  $P(c_i) = \frac{1}{M}$ , for each  $1 \leq i \leq M$ . Then

$$P(r|c_i) = P(c_i|r) \cdot P(c_i) / P(r) = P(c_i|r) \cdot \underbrace{\left[ \frac{1}{M \cdot P(r)} \right]}_{\text{does not depend on } i}$$

So IMLD is the same as MED.

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

## 1.2 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:

If  $r \notin C$ , then reject  $r$ ; otherwise accept  $r$ .

### e-error correcting code

A code  $C$  is called an *e-error correcting code* if the decoding always makes the correct decision if at most  $e$  errors per codeword are introduced. (Similarly: *e-error detecting code*)

#### Example:

$C = \{0000, 1111\}$  is 1-error correcting code, but not a 2-error correcting code.

$C = \{\underbrace{0 \dots 0}_m, \underbrace{1 \dots 1}_m\}$  is a  $\lfloor \frac{m-1}{2} \rfloor$ -error correcting code.

$C = \{0000, 1111\}$  is a 3-error detecting code.

### Theorem 1.3

Suppose  $d(C) = d$ . Then  $C$  is a  $(d - 1)$ -error detecting code.

#### Proof:

Suppose  $c \in C$  is transmitted and  $r$  is received.

- If no error occur, then  $r = c \in C$  and the decoder accepts  $r$ .
- If  $\geq 1$  and  $\leq (d - 1)$  errors occur, then  $1 \leq d(r, c) \leq d - 1$ . So,  $r \notin C$ , and hence the decoder rejects  $r$ .

□

### Theorem 1.4

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

#### Proof:

Since  $d(C) = d$ , there exist  $c_1, c_2 \in C$  with  $d(c_1, c_2) = d$ . If  $c_1$  is sent, it is possible that  $d$  errors occur and  $c_2$  is received. In this case, the decoder accepts  $c_2$ . □

### Theorem 1.5

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

**Proof:**

Suppose  $c \in C$  is transmitted, at most  $\frac{d-1}{2}$  errors are introduced, and  $r$  is received. Let  $c_1 \in C, c_1 \neq c$ .

By  $\triangle$  ineq,  $d(c, c_1) \leq d(c, r) + d(r, c_1)$ . So

$$d(r, c_1) \geq d(c, c_1) - d(c, r) \geq d - \frac{d-1}{2} = \frac{d+1}{2} \geq \frac{d-1}{2}$$

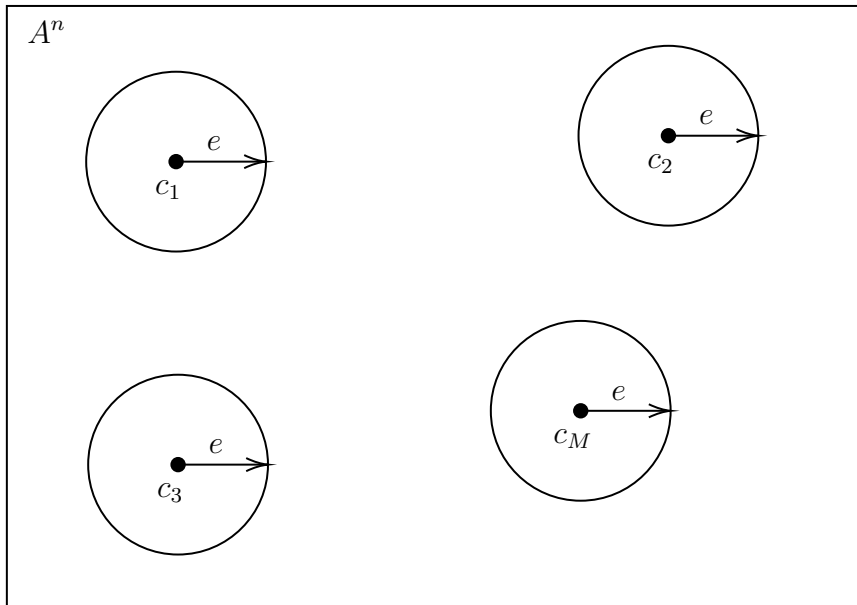
So  $c$  is the unique codeword closest to  $r$ .

So IMLD/CMLD will decode  $r$  to  $c$ . □

**Theorem 1.6**

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

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



$C = \{c_1, c_2, \dots, c_M\}$ . Let  $e = \lfloor \frac{d-1}{2} \rfloor$ . For  $c \in C$ , let  $S_c$  = sphere of radius  $e$  centered at  $c = \{r \in A^n : d(r, c) \leq e\}$ . We proved: If  $c_1, c_2 \in C, c_1 \neq c_2$ , then  $S_{c_1} \cap S_{c_2} = \emptyset$ . The question can be viewed as a *sphere packing problem*: Can we place  $M$  spheres of radius  $e$  in  $A^n$  (such that no 2 spheres overlap)? This is purely combinatorial problem.

**Example:**

Take  $q = 2, n = 128, M = 2^{64}, d \geq 22$ . Does a code with these parameters exist?

**Answer** YES.

**Question** What are the codewords?

**Question** How do we encode and decode efficiently?

**Preview** We'll view  $\{0, 1\}^{128}$  as a vector space of dimension 128 over  $\mathbb{Z}_2$ . We'll choose  $C$  to be a 64-dimensional subspace of this vector space.

# Introduction to Finite Fields

---

## field

A *field*  $(F, +, \cdot)$  consists of a set  $F$  and two operations

$$+ : F \times F \rightarrow F$$

and

$$\cdot : F \times F \rightarrow F,$$

such that

- (i)  $a + (b + c) = (a + b) + c \quad \forall a, b, c \in F.$
- (ii)  $a + b = b + a, \quad \forall a, b \in F.$
- (iii)  $\exists 0 \in F$  such that  $a + 0 = a, \forall a \in F.$
- (iv)  $\forall a \in F, \exists -a \in F$  such that  $a + (-a) = 0.$
- (v)  $a \cdot (b \cdot c) = (a \cdot b) \cdot c, \quad \forall a, b, c \in F.$
- (vi)  $a \cdot b = b \cdot a, \quad \forall a, b \in F.$
- (vii)  $\exists 1 \in F, 1 \neq 0$ , such that  $a \cdot 1 = a \quad \forall a \in F.$
- (viii)  $\forall a \in F, a \neq 0, \exists a^{-1} \in F$  such that  $a \cdot a^{-1} = 1.$
- (ix)  $a \cdot (b + c) = a \cdot b + a \cdot c, \quad \forall a, b, c \in F.$

## infinite, finite, order

A field  $F$  is *infinite* if  $|F|$  is infinite.  $F$  is *finite* if  $|F|$  is finite, in which case  $|F|$  is the *order* of  $F$ .

**Example:**

$\mathbb{Q}, \mathbb{R}, \mathbb{C}$  are infinite fields.  $\mathbb{Z}$  is *not* a field.

**Q** For what integers  $n \geq 2$  do there exist finite fields of order  $n$ ? if a field of order  $n$  exists, how do we “construct”?

**Recall** Let  $n \geq 2$ , the integers modulo  $n$ ,  $\mathbb{Z}_n$ , is the set of all equivalent classes mod  $n$ ,

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

where  $[a] + [b] = [a + b]$ ,  $[a] \cdot [b] = [a \cdot b]$ .

More simply  $\mathbb{Z}_n = \{0, 1, \dots, n-1\}$  with addition & multiplication performed mod  $n$ .

**Example:**

$\mathbb{Z}_9 = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$ .

In  $\mathbb{Z}_9$ ,  $5 + 7 = 3$ ,  $5 \cdot 7 = 8$ .

**Fact**  $\mathbb{Z}_n$  is a *commutative ring*. (i.e. field axioms (i)-(ix) are satisfied, except possibly (viii)).

**Theorem 2.1**

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

**Proof:**

( $\Leftarrow$ ) Suppose  $n$  is prime. Let  $a \in \mathbb{Z}_n, a \neq 0$  (so  $1 \leq a \leq n-1$ ). 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

$$as \equiv 1 \pmod{n}$$

So  $a^{-1} = s$ . So (viii) is satisfied, so  $\mathbb{Z}_n$  is a field (of order  $n$ ).

( $\Rightarrow$ ) Suppose  $n$  is composite, say  $n = a \cdot b$ , where  $2 \leq a, b \leq n-1$ . Suppose  $a^{-1}$  exists,  $a^{-1} = s$ . Then  $as \equiv 1 \pmod{n}$ . So

$$abs \equiv b \pmod{n},$$

so

$$ns \equiv b \pmod{n},$$

so  $0 \equiv b \pmod{n}$ , so  $n|b$  which is impossible.

$\therefore a^{-1}$  does not exist, so  $\mathbb{Z}_n$  is not a field.

□

**Q** Do there exist finite fields of orders 4 and 6?

### characteristic

The *characteristic* of a field denoted  $\text{char}(F)$ , is the smallest positive integer  $m$  such that

$$\underbrace{1 + 1 + 1 + \dots + 1}_m = 0.$$

If no such  $m$  exists, then  $\text{char}(F) = 0$ .

#### Example:

$\text{char}(\mathbb{Q}) = 0$ ,  $\text{char}(\mathbb{R}) = 0$ ,  $\text{char}(\mathbb{C}) = 0$ .

$\text{char}(\mathbb{Z}_p) = p$  ( $p$  is prime)

### Theorem 2.2

If  $\text{char}(F) = 0$ , then  $F$  is infinite.

#### Proof:

Consider  $1, 1+1, 1+1+1, 1+1+1+1, \dots$

Then no 2 elements in this list are equal, because if

$$\underbrace{1 + 1 + 1 + \dots + 1}_a = \underbrace{1 + 1 + 1 + \dots + 1}_b \quad \text{where } a < b$$

then  $0 = \underbrace{1 + 1 + 1 + \dots + 1}_{b-a}$  which contradicts  $\text{char}(F) = 0$ .

So  $F$  is infinite. □

### Theorem 2.3

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

#### Proof:

Suppose  $\text{char}(F) = m$ , which is composite. Say,  $m = a \cdot b$ , where  $2 \leq a, b \leq m-1$ .

Now  $\underbrace{(1 + 1 + 1 + \dots + 1)}_a \cdot \underbrace{(1 + 1 + 1 + \dots + 1)}_b = \underbrace{1 + 1 + 1 + \dots + 1}_m = 0$  since  $\text{char}(F) = m$ .

Let  $\underbrace{1 + \dots + 1}_a = s$  and  $\underbrace{1 + \dots + 1}_b = t$ , so  $s \cdot t = 0$ .

But  $s \neq 0$ , and so  $s^{-1}$  exists, thus  $s^{-1} \cdot s \cdot t = 0$ , therefore  $t = 0$ , which contradicts  $\text{char}(F) = m$ . □

**Next class** Let  $F$  be a finite field of order  $n$ . Then  $\text{char}(F) = p$  (prime). Then  $\mathbb{Z}_p$  is a “subfield” of  $F$ . And  $F$  is a vector space over  $\mathbb{Z}_p$  say of dimension  $k$ . Then order of  $F$  is  $p^k$ .



## 2.1 Non-existence of finite fields

Let  $F$  be a finite field of characteristic  $p$ . Consider

$$E = \{0, 1, 1+1, 1+1+1, \dots, \underbrace{1+1+1+\dots+1}_{p-1}\} \subseteq F$$

Check:  $E$  is a field w.r.t the field operations of  $F$ . Also,  $E$  has order  $p$ . If we label the elements of  $E$  in a natural way

$$1+1 \leftrightarrow 2, 1+1+1 \leftrightarrow \dots, \underbrace{1+1+1+\dots+1}_{p-1} \leftrightarrow p-1,$$

then  $E$  is really just  $\mathbb{Z}_p$ . ( $E$  is *isomorphic* to  $\mathbb{Z}_p$ ).

### Theorem 2.4

If  $F$  be a finite field of order  $n$ , then  $\text{char}(F) = p$  (prime). Then  $\mathbb{Z}_p$  is a “subfield” of  $F$ .

So let's identify:

elements of  $F \leftrightarrow$  vectors  
 elements of  $\mathbb{Z}_p \leftrightarrow$  scalars  
 addition in  $F \leftrightarrow$  vector addition  
 multiplication in  $F \leftrightarrow$  scalar multiplication

### Theorem 2.5

If  $F$  is a finite char  $P$ , then  $F$  is a vector space over  $\mathbb{Z}_p$ .

**Proof:**

Read Appendix A (of the textbook). □

### Theorem 2.6

If  $F$  is a finite field of char  $P$ , then order of  $F$  is  $p^n$  for some  $n \geq 1$ .

**Proof:**

Let  $n$  be the dimension of (the vector space)  $F$  over  $\mathbb{Z}_p$ . Let  $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$  be a basis. Then every element in  $F$  can be written uniquely as

$$c_1\alpha_1 + c_2\alpha_2 + \dots + c_n\alpha_n, \tag{*}$$

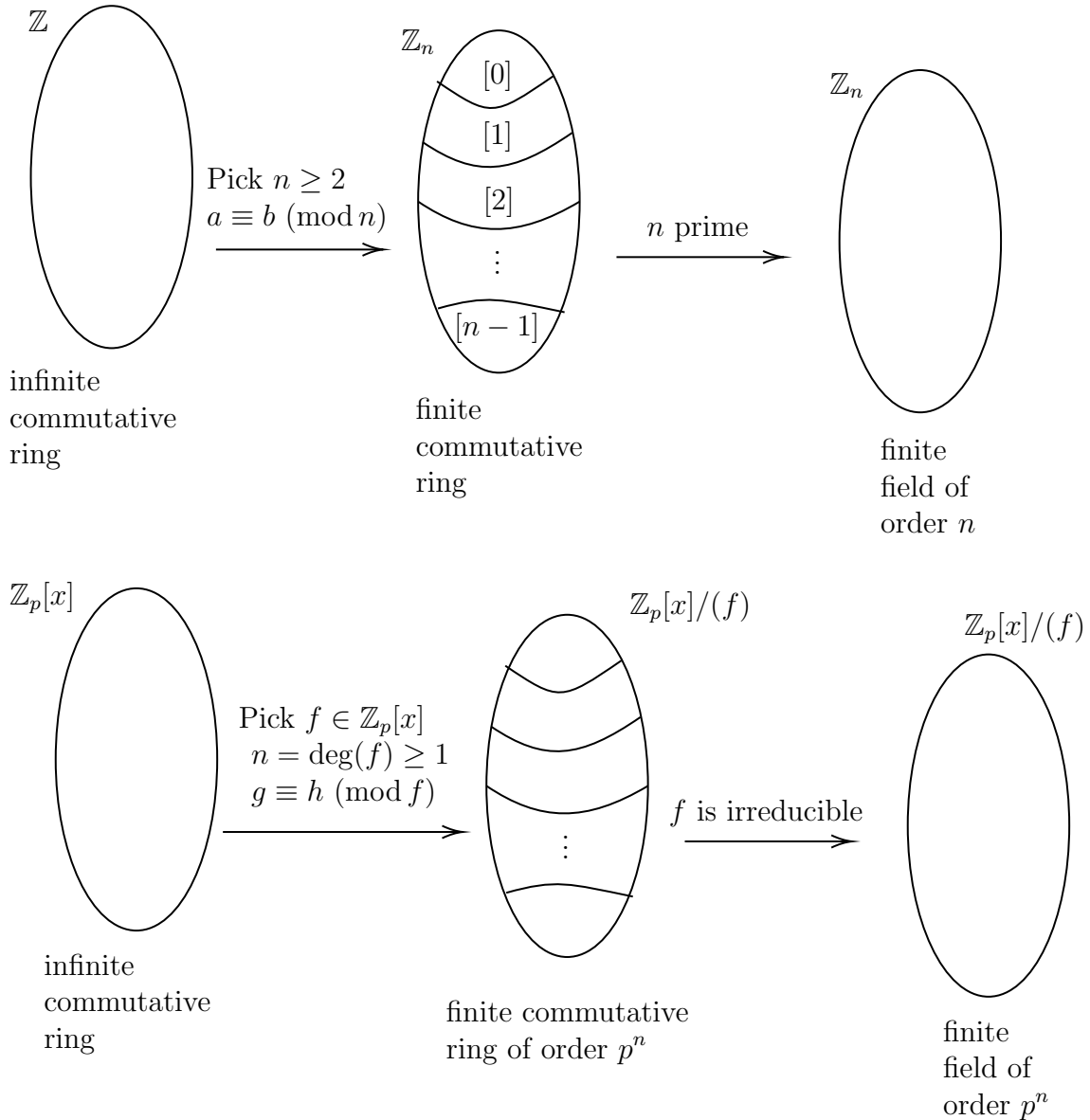
where  $c_i \in \mathbb{Z}_p$ .

Also every element  $(*)$  is in  $F$ . Hence  $\text{ord}(F) = p^n$ . □

**Example:**

There is no field of order 6.

**Q** Is there a finite field of order 4? 8? 9? Yes.



**$F[x]$**

If  $F$  is a field, then  $F[x]$  is the set of all polynomials in  $x$  with coefficients from  $F$ .

Addition and multiplication is done in the usual way, with coefficient arithmetic in  $F$ .

**Example:**

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

### Theorem 2.7

$F[x]$  is an infinite commutative ring.

### Some notations

Let  $f \in F[x]$ ,  $\deg(f) \geq 1$ .

If  $g, h \in F[x]$ , we write  $g \equiv h \pmod{f}$ .

If  $g - h = \ell f$  for some  $\ell \in F[x]$ , we write  $(f|g - h)$ .

### Facts

1.  $\equiv$  is an equivalence relation.

2. The equivalence class containing  $g \in F[x]$  is

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

3. We define  $[g_1] + [g_2] = [g_1 + g_2]$        $[g_1] \cdot [g_2] = [g_1 \cdot g_2]$

4. The set of all equivalence classes, denoted  $F[x]/(f)$  (where  $f \in F[x]$ ,  $\deg(f) \geq 1$ ) is a commutative ring.

5. The polynomials in  $F[x]$  of degree  $< \deg(f)$  are a system of distinct representatives of the equivalence classes in  $F[x]/(f)$ .

**Justification** Let  $g \in F[x]$ . By division algorithm for polynomials, we can write  $g = \ell f + r$  where  $\deg(r) < \deg(f)$ . [Convention:  $\deg(0) = -\infty$ ]

Then  $g - r = \ell f$ . So  $g \equiv r \pmod{f}$ . So  $[g] = [r]$ .

Also if  $r_1, r_2 \in F[x]$ ,  $r_1 \neq r_2$  and  $\deg(r_1), \deg(r_2) < \deg(f)$ , then  $f \nmid r_1 - r_2$ , so  $r_1 \not\equiv r_2 \pmod{f}$ . Hence  $[r_1] \neq [r_2]$ .

## 2.2 Constructing finite fields

**We proved** A system of distinct representatives for  $\mathbb{Z}_p[x]/(f)$  is  $[r(x)] : r \in \mathbb{Z}_p[x], \deg(r) < \deg(f)$ . Therefore,  $|\mathbb{Z}_p[x]/(f)| = p^n$ .

### irreducible

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

**Example:**

$x^2 + 1$  is irreducible over  $\mathbb{R}$ .

$x^2 + 1$  is reducible over  $\mathbb{C}$ , since  $(x^2 + 1) = (x + i)(x - i)$ .

$x^2 + 1$  is reducible over  $\mathbb{Z}_2$ , since  $x^2 + 1 = (x + 1)^2$ .

$x^2 + 1$  is irreducible over  $\mathbb{Z}_3$ .

**Theorem 2.8**

Let  $F$  be a field, and  $f \in F[x]$  of degree  $n \geq 1$ . Then  $F[x]/(f)$  is a field if and only if  $f$  is irreducible over  $F$ .

**Proof:**

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

( $\Leftarrow$ ) Suppose  $g \in F[x]/(f)$ ,  $g \neq 0$ , (and  $\deg(g) < \deg(f)$ ). Then  $\gcd(g, f) = 1$ , and by the EEA for polynomials, there exist  $s, t \in F[x]$  such that  $gs + ft = 1$ . Reducing both sides mod  $f$  gives  $gs \equiv 1 \pmod{f}$ . So  $g^{-1} = s$ . Hence  $F[x]/(f)$  is a field.

( $\Rightarrow$ ) Exercise.

□

So, to construct a finite field of order  $p^n$  ( $n \geq 2$ ), we need an irreducible polynomial  $f \in \mathbb{Z}_p[x]$  of degree  $n$ . Then  $\mathbb{Z}_p[x]/(f)$  is a finite field of order  $p^n$ .

**Fact** For any prime  $p$ , integer  $n \geq 2$ , there exists an irreducible polynomial degree  $n$  in  $\mathbb{Z}_p[x]$ .

**Theorem 2.9**

There exists a finite field of order  $q$  iff  $q$  is a prime power.

**Example: Construct a finite field of order 4.**

Take  $f(x) = x^2 + x + 1 \in \mathbb{Z}_2[x]$ , which is irreducible over  $\mathbb{Z}_2$ . So, the field is  $\mathbb{Z}_2[x]/(x^2 + x + 1) = \{0, 1, x, x + 1\}$ .

- $x + (x + 1) = 1$ .
- $x \cdot (x + 1) = x^2 + x = 1$ .
- So,  $x^{-1} = x + 1$ .
- $1^{-1} = 1$
- $x^{-1} = x + 1$
- $(x + 1)^{-1} = x$

**Example: Field of order  $8 = 2^3$** 

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

$$F_1 = \mathbb{Z}_2[x]/(x^3 + x + 1) = \{0, 1, x, x + 1, x^2, x^2 + 1, x^2 + x, x^2 + x + 1\}$$

- $x^2 + (x^2 + x + 1) = x + 1$
- $x^2 \cdot (x^2 + x + 1) = x^4 + x^3 + x^2 = 1$ .

$$\begin{array}{r} x^3 + x + 1 \overline{) \begin{array}{r} x^4 + x^3 + x^2 \\ - x^4 \phantom{+ x^3} - x^2 \phantom{+ x} \\ \hline x^3 \phantom{+ x^2} - x \\ - x^3 \phantom{+ x^2} - x - 1 \\ \hline - 2x - 1 \end{array}} \end{array}$$

- $(x^2)^{-1} = x^2 + x + 1$
- $x^{-1} = x^2 + 1$

**Example: Finite field of order 8**

Take  $f_2(x) = f(x) = x^3 + x^2 + 1$ . Then  $F_2 = \mathbb{Z}_2[x]/(x^3 + x^2 + 1)$  is a finite field of order 8. Its elements are  $F_2 = \{0, 1, x, x + 1, x^2, x^2 + 1, x^2 + x, x^2 + x + 1\}$ .

- $x^{-1} = x^2 + x$ .

**Note**

$F_1$  and  $F_2$  are two different field of order 8. In fact, they are “essentially the same”, i.e., they are *isomorphic*, i.e., there is a bijection  $\alpha : F_1 \rightarrow F_2$  such that  $\alpha(a + b) = \alpha(a) + \alpha(b)$  and  $\alpha(a \cdot b) = \alpha(a) \cdot \alpha(b)$ ,  $\forall a, b \in F$ .

**Fact** Any two fields of order  $q$  are isomorphic.

**GF( $q$ )**

We will denote *the* finite field of order  $q$  by  $\text{GF}(q)$ .

We saw two different representations of  $\text{GF}(2^3)$ .

**Recall** A finite field of order  $q$  exists iff  $q = p^n$  for some prime  $p$  and  $n \geq 1$ . ( $p = \text{characteristic}$ )

- Also  $\text{GF}(q) = \mathbb{Z}_p[x]/(f)$ , where  $f \in \mathbb{Z}_p[x]$  is irreducible and has degree  $n$ .

**Example: Construct  $\text{GF}(16)$** 

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 factor.

$\therefore f$  is irreducible over  $\mathbb{Z}_2$ . So  $\text{GF}(16) = \mathbb{Z}_2[x]/(x^4 + x + 1)$ .

## 2.3 Properties of finite fields

### Theorem 2.10: Frosh's Dream

Let  $\alpha, \beta \in \text{GF}(q)$ , where  $\text{char}(\text{GF}(q)) = p$ . Then  $(\alpha + \beta)^p = \alpha^p + \beta^p$ .

**Proof:**

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

$$\text{Now, } \binom{p}{i} = \frac{p(p-1) \cdots (p-i+1)}{1 \cdot 2 \cdots i} \in \mathbb{N}.$$

If  $1 \leq i \leq p-1$ , then  $p \mid$  numerator; but  $p \nmid$  denominator.  $\therefore p \nmid \binom{p}{i}$ . So,

$$\begin{aligned} \binom{p}{i} \alpha^i \beta^{p-i} &= \underbrace{\alpha^i \beta^{p-i} + \cdots + \alpha^i \beta^{p-i}}_{\binom{p}{i}} \\ &= \alpha^i \beta^{p-i} \underbrace{(1 + 1 + 1 + \cdots + 1)}_{\binom{p}{i}} \\ &= \alpha^i \beta^{p-i} \cdot 0 \quad \text{since char} = p \text{ and } p \mid \binom{p}{i} \\ &= 0 \end{aligned}$$

□

### Note

More generally,

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

for all  $m \geq 1$ .

### Theorem 2.11

Let  $\alpha \in \text{GF}(q)$ . Then  $\alpha^q = \alpha$ .

**Proof:**

- If  $\alpha = 0$ , then of course  $\alpha^q = \alpha$ .
- Suppose  $\alpha \neq 0$ . Let  $\alpha_1, \dots, \alpha_{q-1}$  be the nonzero elements in  $\text{GF}(q)$ . Consider  $\alpha\alpha_1, \dots, \alpha\alpha_{q-1}$ . The elements in this list are pairwise distinct because

if  $\alpha\alpha_i = \alpha\alpha_j$  ( $i \neq j$ ), then  $\alpha^{-1}\alpha\alpha_i = \alpha^{-1}\alpha\alpha_j$ , so  $\alpha_i = \alpha_j$ . Also

$$\alpha\alpha_i \neq 0, \quad \forall 1 \leq i \leq q-1.$$

Hence

$$\{\alpha_1, \alpha_2, \dots, \alpha_{q-1}\} = \{\alpha\alpha_1, \dots, \alpha\alpha_{q-1}\}$$

$$\therefore \alpha_1 \dots \alpha_{q-1} = (\alpha\alpha_1) \dots (\alpha\alpha_{q-1})$$

$$\therefore \alpha^{q-1} = 1$$

$$\therefore \alpha^q = \alpha$$

□

### $\text{GF}(q)^*$

Let  $\text{GF}(q)^* = \text{GF}(q) \setminus \{0\}$ .

### $\text{ord}(\alpha)$

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

#### Example:

How many elements of order 1 are there in  $\text{GF}(q)$ ?

$$\alpha = 1$$

#### Example:

Find  $\text{ord}(x)$  in  $\text{GF}(16) = \mathbb{Z}_2[x]/(x^4 + x + 1)$ .

$$x^1 = 1, x^2 = x^2, x^3 = x^3, x^4 = x + 1, x^5 = x^2 + x, \dots, x^{15} = 1.$$

Since  $\text{ord}(x) \neq 1, 3, 5$ ,  $\text{ord}(x) | 15$ , we have  $\text{ord}(x) = 15$ .

### Lemma 2.12

Let  $\alpha \in \text{GF}(q)^*$ ,  $\text{ord}(\alpha) = t$ ,  $s \in \mathbb{Z}$ .  $\alpha^s = 1 \iff t | s$ .

#### Proof:

Let  $s \in \mathbb{Z}$ . Long division  $g$  gives  $s = \ell t + r$ , where  $0 \leq r \leq t-1$ .

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

So

$$\begin{aligned} \alpha^s = 1 &\iff \alpha^r = 1 \\ &\iff r = 0 \quad \text{since } 0 \leq r \leq t-1 \\ &\iff t | s \end{aligned}$$

□

**Corollary 2.13**

If  $\alpha \in \text{GF}(q)^*$ , then  $\text{ord}(\alpha) | q - 1$ .

**Proof:**

We know that  $\alpha^{q-1} = 1$ . So  $\text{ord}(\alpha) | q - 1$  by previous lemma.  $\square$

**generator**

An element  $\alpha \in \text{GF}(q)$  is a *generator of  $\text{GF}(q)^*$*  (primitive element in  $\text{GF}(q)$ ).  
If  $\text{ord}(\alpha) = q - 1$ .

**Lemma 2.14**

If  $\alpha$  is a generator of  $\text{GF}(q)^*$  then  $\{\alpha^0, \alpha^1, \alpha^2, \dots, \alpha^{q-2}\} = \text{GF}(q)^*$ .

**Lemma 2.15**

If  $\alpha \in \text{GF}(q)^*$  has order  $t$ , then  $\alpha^0, \alpha^1, \dots, \alpha^{t-1}$  are pairwise distinct.

**Proof:**

Suppose  $\alpha^i = \alpha^j$ , where  $0 \leq i < j \leq t - 1$ . Then  $\alpha^{j-i} = 1$  which contradicts  $\text{ord}(\alpha) = t$  since  $1 \leq j - i \leq t - 1$ .  $\square$

So, if  $\alpha$  is a generator of  $\text{GF}(q)^*$  then  $\{\alpha^0, \alpha^1, \alpha^2, \dots, \alpha^{q-2}\} = \text{GF}(q)^*$ .

**Theorem 2.16**

$\text{GF}(q)^*$  has at least one generator.

**Proof:**

See LEARN (optional).  $\square$

**Example:**

Find a generator of  $\text{GF}(8) = \mathbb{Z}_2[x]/(x^3 + x + 1)$ .

$x$  is a generator.



# Linear Codes

---

Let  $F = \text{GF}(q)$ .

Let  $V_n(F) = F \times F \times \dots \times F = F^n$

Then  $V_n(F)$  is an  $n$ -dimensional vector space over  $F$ .

We have  $|V_n(F)| = q^n$ .

## linear $(n, k)$ -code over $F$

A *linear  $(n, k)$ -code over  $F$*  is a  $k$ -dimensional subspace of  $V_n(F)$ .

## subspace

A subspace of a vector space  $V$  over  $F$  is a subset  $S \subseteq V$  such that

- (i)  $S \neq \emptyset$ .
- (ii)  $v_1 + v_2 \in S \quad \forall v_1, v_2 \in S$ .
- (iii)  $\lambda v \in S, \quad \forall v \in S, \lambda \in F$ .

### Note

$S$  is also a vector space over  $F$ .

$0 \in S$ .

## 3.1 Properties of Linear Codes

Let  $C$  be an  $(n, k)$ -code over  $F$ . Let  $v_1, v_2, \dots, v_k$  be an ordered basis for  $C$ .

- 1) The codewords in  $C$  are precisely:

$$mv_1 + m_2v_2 + \dots + m_kv_k,$$

where  $m_i \in F$ .

So  $|C| = M = q^k$ .

- 2) The rate of  $C$  is  $R = \frac{\log_q M}{n} = \frac{k}{n}$ ,

- 3) Distance

#### weight

The (Hamming) *weight* of  $v \in V_n(F)$ ,  $\omega(v)$ , is the number of nonzero coordinate positions in  $vv$ .

The weight of  $C$  is  $\omega(C) = \min\{\omega(c) : c \in C, c \neq 0\}$ .

#### Theorem 3.1

If  $C$  is a linear code, then  $d(C) = \omega(C)$ .

**Proof:**

$$\begin{aligned} d(C) &= \min\{d(x, y) : x, y \in C, x \neq y\} \\ &= \min\{\omega(x - y) : x, y \in C, x \neq y\} \\ &= \min\{\omega(c) : c \in C, c \neq 0\} \\ &= \omega(C) \end{aligned}$$

□

- 4) Encoding.

Since  $M = q^k$ , there are  $q^k$  source messages. We'll assume that the source messages are elements of  $V_k(F)$ . A natural encoding rule is: Given  $(m_1, m_2, \dots, m_k) \in V_k(F)$ . We will encode it as  $c = m_1v_1 + m_2v_2 + \dots + m_kv_k$ .

#### Note

The encoding rule depends on the basis chosen for  $C$ .

- 5) Note if  $m = (m_1, \dots, m_k)$ , then the encoding rule can be written as follows.

$$c = (m_1, m_2, \dots, m_k) \begin{bmatrix} - & v_1 & - \\ & \vdots & \\ - & v_k & - \end{bmatrix}_{k \times n}$$

$$c = mG$$

**generator matrix**

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  $C$  w.r.t.  $G$  is  $c = mG$ .

**Note**

Performing elementary row operations on  $G$  gives a different matrix for the same code  $C$ .

**Example:** Consider a binary  $(5, 3)$ -code  $C$

where binary means “over  $F = \text{GF}(2) = \mathbb{Z}_2$ . 5 is  $n$ , length of code. 3 is  $k$ , dimension.

Then  $M = q^k = 2^3$  and  $R = \frac{k}{n} = \frac{3}{5}$ . and

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

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

indeed has rank 3 so  $G$  is a GM for  $C$ .

Encoding rule is  $c = mG$ .

$m$ source msgs	$\rightarrow$	$c$ 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 row equivalent to  $G$  is also a GM for  $C$ , but yields a different encoding rule.

**systematic, standard form**

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

**Example:**

$C = \langle 100011, 101010, 100110 \rangle$  is a non-systematic  $(6, 3)$ -code. A GM for  $C$  is

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

Another GM for  $C$  is

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

Another GM for  $C$ :

$$G_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

$C$  is not systematic.

However, if every codeword is permuted by moving the second bit to a new fourth bit, then we get a new code  $C'$  that is linear, and has the same  $n, k, d$  as  $C$ .

**equivalent**

Let  $C$  be an  $(n, k)$ -code. If  $\pi$  is a permutation on  $\{1, 2, \dots, n\}$ , Then  $\pi(C)$ <sup>a</sup> is an  $(n, k)$ -code and is said to be *equivalent* to  $C$ .

<sup>a</sup>i.e. apply  $\pi$  to each codeword

**Fact**

1. If  $C, C'$  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 GM for  $C$  in row reduced form. Then one can permute to columns of  $G$  to get a matrix  $G' = [I_k | A]$  in standard form.

Then  $G'$  is a GM for a code  $C'$  that is equivalent to  $C$ . □

## 3.2 Dual Codes

### inner product

Let  $x, y \in V_n(F)$ . The *inner product* of  $x$  and  $y$  is

$$x \cdot y = \sum_{i=1}^n x_i y_i \in F$$

**Properties** For all  $x, y, z \in V_n(F)$  and all  $\lambda \in F$

1.  $x \cdot y = y \cdot x$
2.  $x \cdot (y + z) = x \cdot y + x \cdot z$
3.  $(\lambda x) \cdot y = \lambda(x \cdot y)$
4.  $x \cdot x = 0$  does **not** imply that  $x = 0$ .

**Example:**

Consider  $V_2(\mathbb{Z}_2)$

Then  $(1, 1) \cdot (1, 1) = 0$ .

### dual code

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

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

### Theorem 3.2

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

**Proof:**

Let  $v_1, v_2, \dots, v_k$  be a basis for  $C$ .

**Claim** Let  $x \in V_n(F)$ . Then  $x \in C^\perp$  iff  $v_1 \cdot x = v_2 \cdot x = \dots = v_k \cdot x = 0$ .

( $\implies$ ) If  $x \in C^\perp$ , then  $x \cdot c \forall c \in C$ . In particular,  $x \cdot v_1 = 0, \dots, x \cdot v_k = 0$ .

( $\impliedby$ ) Suppose  $x \cdot v_1 = x \cdot v_2 = \dots = x \cdot v_k = 0$ . Let  $c \in C$ . We can write

$$c = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k, \quad v_i \in F$$

Then  $x \cdot c = \lambda_1(x \cdot v_1) + \dots + \lambda_k(x \cdot v_k) = 0$ . Hence  $x \in C^\perp$ .

Consider

$$G = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{bmatrix}_{k \times n}$$

Then  $x \in C^\perp$  iff  $Gx^T = 0$ . So  $C^\perp$  is the nullspace of  $G$ . Hence  $C^\perp$  is an  $(n - k)$ -dimensional subspace of  $V_n(F)$ .  $\square$

# Index

---

## A

alphabet, word, length..... 5

## C

characteristic ..... 15

## D

dual code ..... 28

## E

e-error correcting code..... 10

equivalent ..... 27

## F

$F[x]$  ..... 17

field ..... 13

## G

generator ..... 23

generator matrix..... 26

$GF(q)$  ..... 20

$GF(q)^*$  ..... 22

## H

Hamming distance ..... 7

## I

infinite, finite, order ..... 13

inner product ..... 28

irreducible ..... 18

## L

linear  $(n, k)$ -code over  $F$  ..... 24

## O

$\text{ord}(\alpha)$  ..... 22

## R

rate..... 7

## S

subspace..... 24

systematic, standard form ..... 26

## W

weight ..... 25