Coding Theory

CO 331

Alfred John Menezes

LaTeXed by Sibelius Peng

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/`.

Sibelius Peng

Contents

Preface **1**

0 Pre **3**

1 Introduction & Fundamentals **5**
1.1 Decoding Strategy 7
1.1.1 Nearest Neighbour Decoding 8
1.2 Error Correcting & Detecting Capabilities of a Code 10

2 Introduction to Finite Fields **13**
2.1 Non-existence of finite fields 16
2.2 Constructing finite fields 18
2.3 Properties of finite fields 21

3 Linear Codes **24**
3.1 Properties of Linear Codes 24
3.2 Dual Codes 28
3.3 Perfect Code 32
3.4 Error Correction (for Hamming Codes) 34
3.5 General Decoding Problem for binary linear codes 36
3.5.1 Syndrome Decoding 38

4 The binary Golay code **40**
4.1 The (binary) Golay code C_{23} (1949) 40
4.2 The Extended Golay Code C_{24} 41
4.2.1 Decoding Algorithm for C_{24} 41
4.2.2 Reliability of C_{24} 43

5 Cyclic Codes **45**
5.1 Dual Code of a Cyclic Code 51
5.2 Computing Syndromes 53
5.3 Burst Error Correcting 56
5.4 Decoding cyclic burst errors 58
5.5 Error trapping decoding (for cyclic burst errors) 59
5.6 Interleaving 60

6 BCH codes **61**
6.1 Minimal Polynomials 61

0

Pre

Example: Replication code

source msgs		codewords
0	$\to$	0
1	$\to$	1

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

---

source msgs		codewords
0	$\to$	00
1	$\to$	11

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

---

source msgs		codewords
0	$\to$	000
1	$\to$	111

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

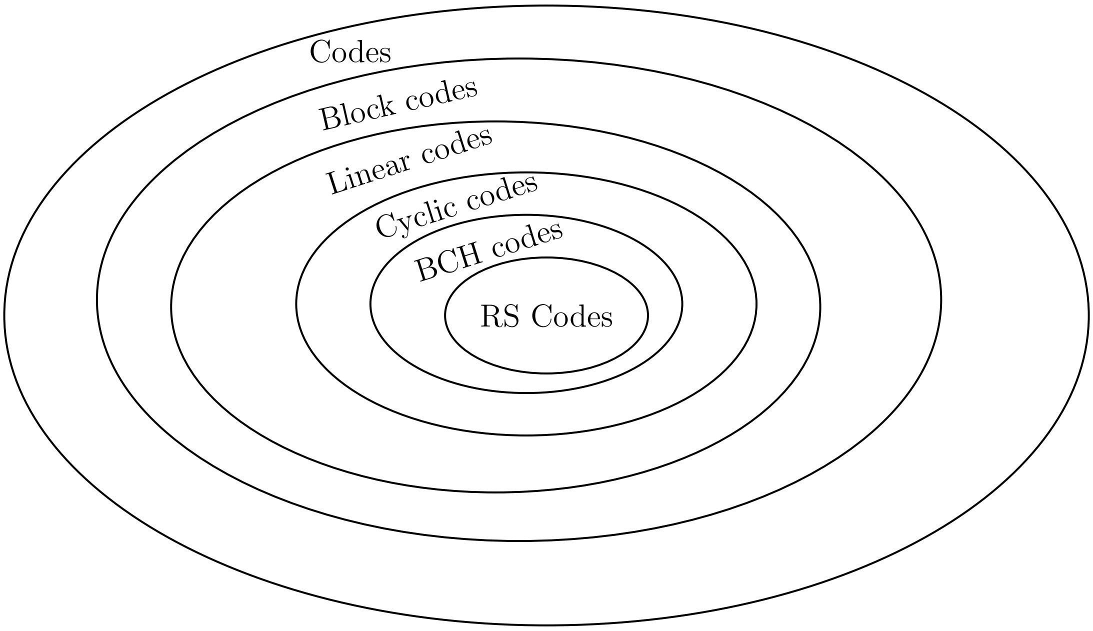
---

source msgs		codewords
0	$\to$	00000
1	$\to$	11111

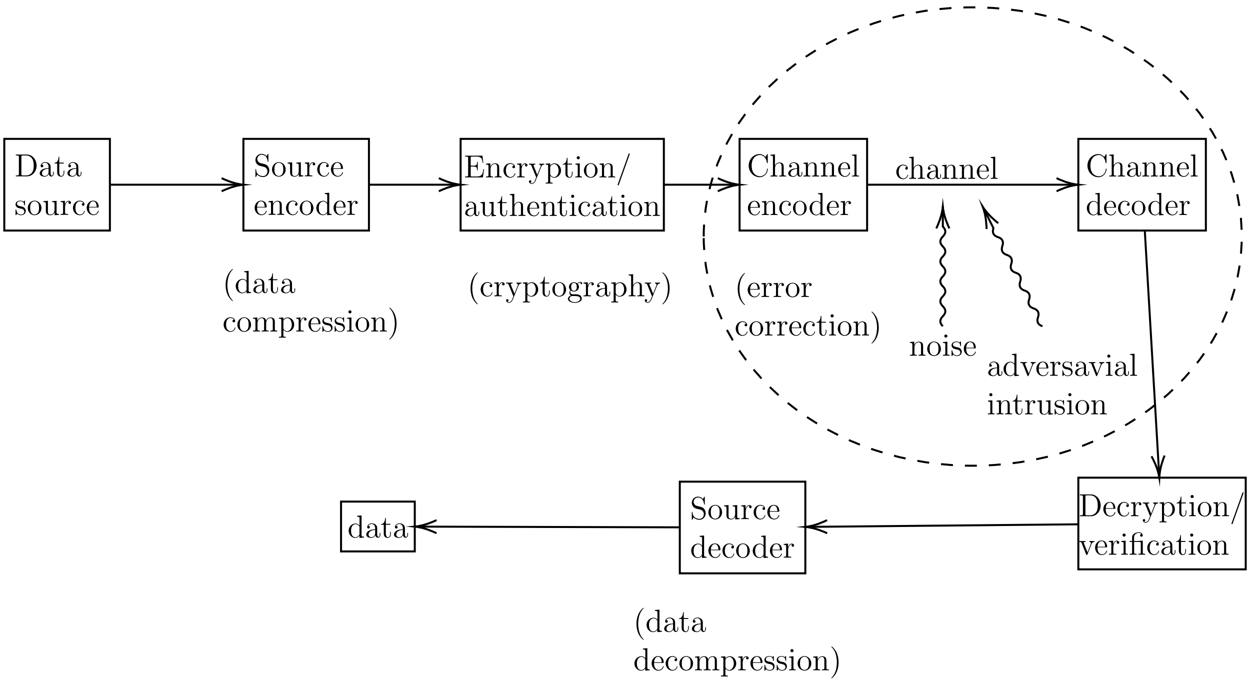
# of errors/codeword that be detected: 4
errors/codeword that can be corrected: 2 (nearest neighbour decoding)
Rate: 1/5

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.



1

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 ≥ 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	$\rightarrow$	00000
10	$\rightarrow$	11100
01	$\rightarrow$	00111
11	$\rightarrow$	10101

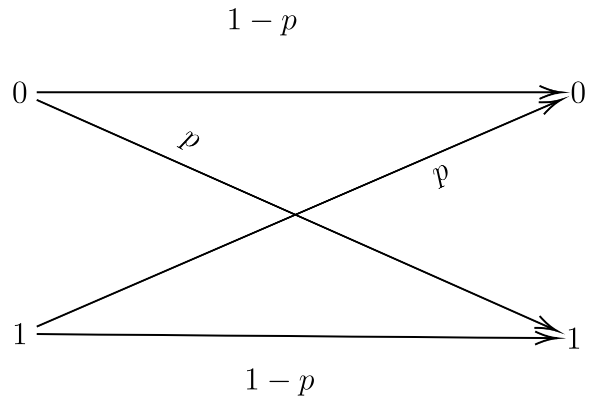
Encoding 1-1 map

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

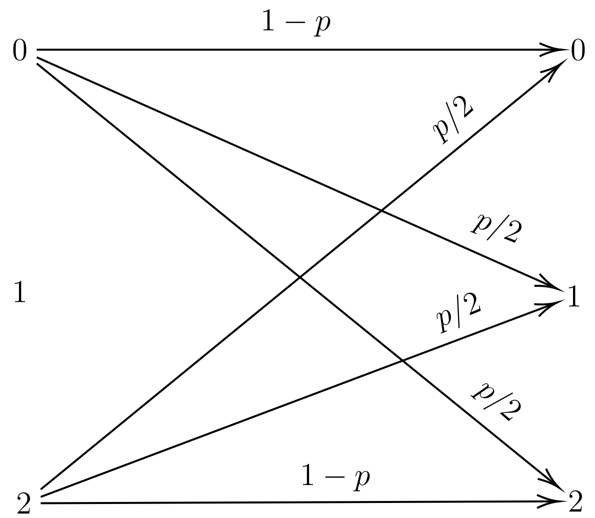
Example:

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 $X_1, X_2, X_3, \ldots$. Suppose the symbols received and $Y_1, Y_2, Y_3, \ldots$. 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 $p =$ 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:

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

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

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 $r \to 000$

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: $r \to 111$

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 $P(c_i)$, $1 \le i \le M$

3. Suppose source messages are equally likely, so $P(c_i) = \frac{1}{M}$, for each $1 \le i \le M$. 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:

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

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.

$C = \{\underbrace{0\ldots0}_{m}, \underbrace{1\ldots1}_{m}\}$ 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 ≥ 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

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

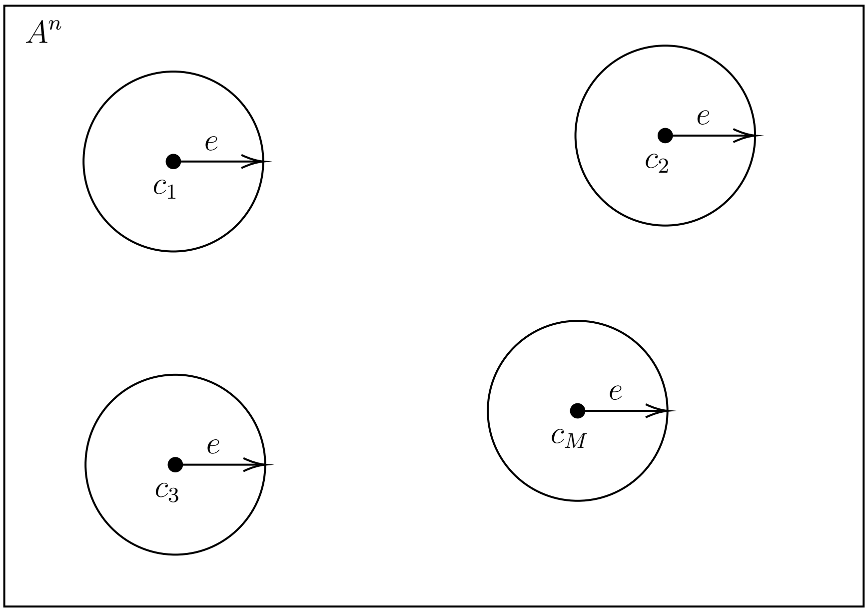
So c is the unique codeword closest to r.

So IMLD/CMLD will decode r to c. $\square$

Theorem 1.6

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

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



$C = \{c_1, c_2, \ldots, c_M\}$. Let $e = \left\lfloor \frac{d-1}{2} \right\rfloor$. 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 $S_{c_1} \cap S_{c_2} \neq \varnothing$. 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.

2

Introduction to Finite Fields

field

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

$$+ : F \times F \to F$$

and

$$\cdot : F \times F \to F,$$

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) $a \cdot (b + c) = a \cdot b + b \cdot c, \quad \forall a, b, c \in F.$

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

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:
($\Longleftarrow$) 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 sizes $[\bmod n]$, gives

$$as \equiv 1 \ [\bmod n]$$

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

($\Longrightarrow$) 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 $as \equiv 1 \ [\bmod n]$. So

$$abs \equiv b \ [\bmod n],$$

so

$$ns \equiv b \ [\bmod n],$$

so $0 \equiv b \ [\bmod n]$, so $n | b$ which is impossible.

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

$\square$

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

characteristic

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

$$\underbrace{1+1+1+\ldots+1}_{m} = 0.$$

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

Example:

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

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

Theorem 2.2

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

Proof:

Consider 1, 1+1, 1+1+1, 1+1+1+1,...

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

$$\underbrace{1+1+1+\ldots+1}_{a} = \underbrace{1+1+1+\ldots+1}_{b} \qquad \text{where } a < b$$

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

So F is infinite. $\square$

Theorem 2.3

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

Proof:

Suppose $\operatorname{char}(F) = m$, which is composite. Say, $m = a \cdot b$, where $2 \leq a, b \leq m-1$. Now $(\underbrace{1+1+1+\ldots+1}_{a}) \cdot (\underbrace{1+1+1+\ldots+1}_{b}) = \underbrace{1+1+1+\ldots+1}_{m} = 0$ since $\operatorname{char}(F) = m$.

Let $\underbrace{1+\ldots+1}_{a} = s$ and $\underbrace{1+\ldots+1}_{b} = t$, 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 $\operatorname{char}(F) = m$. $\square$

Next class Let F be a finite field of order n. Then $\operatorname{char}(F) = p$ (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

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

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

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

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:

$$\begin{aligned} \text{elements of } F &\leftrightarrow \text{vectors} \\ \text{elements of } \mathbb{Z}_p &\leftrightarrow \text{scalars} \\ \text{addition in } F &\leftrightarrow \text{vector addition} \\ \text{multiplication in } F &\leftrightarrow \text{scalar multiplication} \end{aligned}$$

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 $\{\alpha_1, \alpha_2, \ldots \alpha_n\}$ be a basis. Then every element in F can be written uniquely as

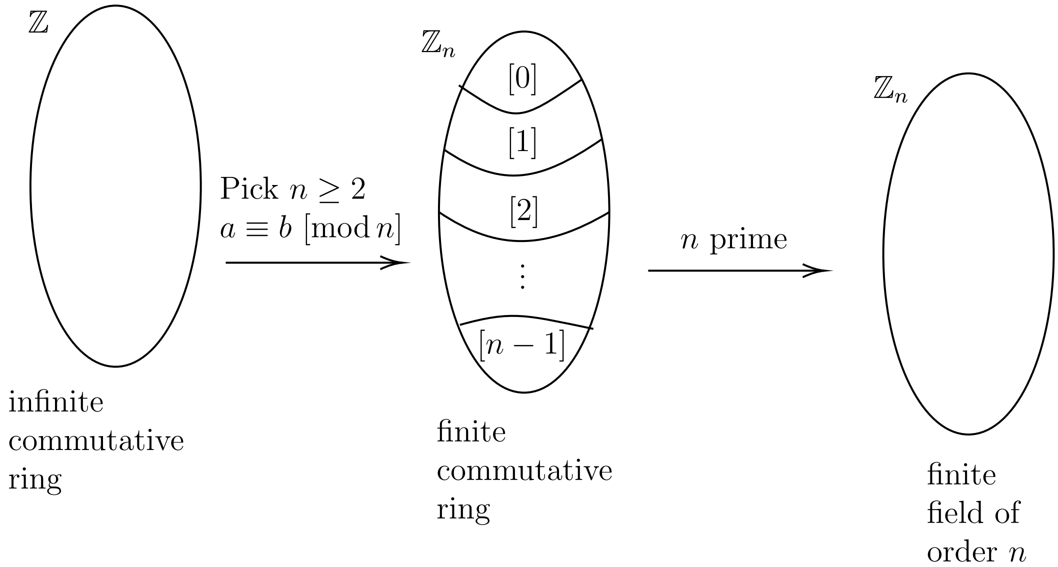
$$c_1\alpha_1 + c_2\alpha_2 + \ldots + c_n\alpha_n, \tag{*}$$

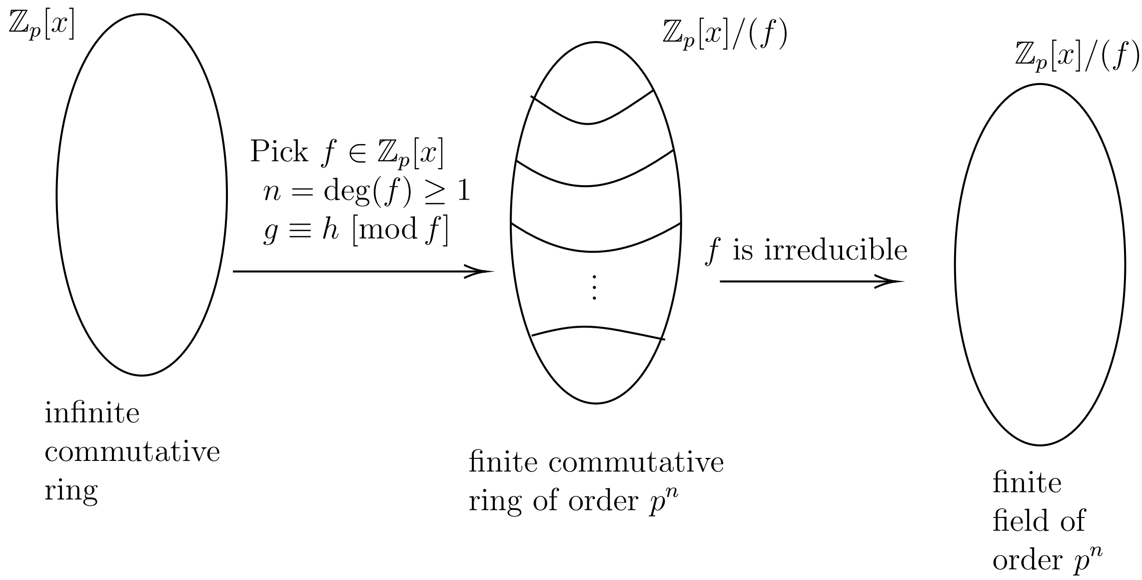
where $c_i \in \mathbb{Z}_p$.

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 $g \equiv h \ [\text{mod}\, f]$.

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
$$[g] = \{h \equiv g \ [\text{mod}\, f] : h \in F[x]\}$$
3. We define $[g_1] + [g_2] = [g_1 + g_2] \qquad [g_1] \cdot [g_2] = [g_1 \cdot g_2]$
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 $g \equiv r \ [\text{mod}\, f]$. 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 $r_1 \not\equiv r_2 \ [\text{mod}\, f]$. 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 irreducible over F.

Proof:

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

($\Longleftarrow$) 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 $gs \equiv \ [\bmod f]$. So $g^{-1} = s$. Hence $F[x]/(f)$ is a field.

($\Longrightarrow$) Exercise.

$\square$

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

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

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

$$\begin{array}{r|l} & \quad\quad\quad\quad\quad x + 1 \\ \hline x^3 + x + 1 & x^4 + x^3 + x^2 \\ & -x^4 \quad -x^2 \quad -x \\ \hline & x^3 \quad\quad -x \\ & -x^3 \quad\quad -x - 1 \\ \hline & \quad\quad -2x - 1 \end{array}$$

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

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 $\alpha : F_1 \to F_2$ such that $\alpha(a + b) = \alpha(a) + \alpha(b)$ and $\alpha(a \cdot b) = \alpha(a) \cdot \alpha(b), \quad \forall a, b \in F$.

Fact Any two fields of order q are isomorphic.

GF(q)

We will denote *the* finite field of order q by GF(q).

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

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

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

Example: Construct 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 $\mathrm{GF}(16) = \mathbb{Z}_2[x]/(x^4+x+1)$.

2.3 Properties of finite fields

Theorem 2.10: Frosh's Dream

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

Proof:

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

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

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

$$\begin{aligned}
\binom{p}{i}\alpha^i\beta^{p-i} &= \underbrace{\alpha^i\beta^{p-i} + \ldots + \alpha^i\beta^{p-i}}_{\binom{p}{i}} \\
&= \alpha^i\beta^{p-i}(\underbrace{1+1+1+\ldots+1}_{\binom{p}{i}}) \\
&= \alpha^i\beta^{p-i}\cdot 0 \qquad \text{since char} = p \text{ and } p \left| \binom{p}{i} \right. \\
&= 0
\end{aligned}$$

□

Note
More generally,

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

for all $m \ge 1$.

Theorem 2.11

Let $\alpha \in \mathrm{GF}(q)$. 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 $\mathrm{GF}(q)$. Consider $\alpha\alpha_1, \dots, \alpha\alpha_{q-1}$. The elements in this list are pairwise distinct because

if $\alpha\alpha_i = \alpha\alpha_j \quad (i \neq j)$, 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

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

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

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

$\therefore \alpha^q = \alpha$

□

GF$(q)^*$

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

ord(alpha)

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

Example:

How many elements of order 1 are there in GF(q)?

$\alpha = 1$

Example:

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

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

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

Lemma 2.12

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

Proof:

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

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

So

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

□

Corrollary 2.13

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

Proof:
We know that $\alpha^{q-1} = 1$. So $\mathrm{ord}(\alpha)|q-1$ by previous lemma. □

generator

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

Lemma 2.14

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

Lemma 2.15

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

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

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

Theorem 2.16

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

Proof:
See LEARN (optional). □

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

x is a generator.

3

Linear Codes

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

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

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

(i) $S \neq \varnothing$.

(ii) $v_1 + v_2 \in S \qquad \forall v_1, v_2 \in S$.

(iii) $\lambda v \in S, \qquad \forall v \in S, \lambda \in F$.

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 $v_1, v_2, \ldots, v_k$ be an ordered basis for C.

1) The codewords in C are precisely:

$$mv_1 + m_2v_2 + \ldots + m_kv_k,$$

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 v.

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 $(m_1, m_2, \ldots, m_k) \in V_n(F)$. We will encode it as $c = m_1v_1 + m_2v_2 + \ldots + m_kv_k$.

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

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

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

$$c = mG$$

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 basic 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 $F = \mathrm{GF}(2) = \mathbb{Z}_2$. 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	$\to$	c codewords
000	$\to$	00000
001	$\to$	00101
010	$\to$	01011
011	$\to$	01110
100	$\to$	10010
101	$\to$	10111
110	$\to$	11001
111	$\to$	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 π is a permutation on $\{1, 2, \ldots, n\}$, Then $\pi(C)$[a] is an (n,k)-code and is said to be *equivalent* to C.

[a] i.e. apply π 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. $\square$

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

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

orthogonal

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

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

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

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

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

$$c = \lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_k v_k, \qquad v_i \in F$$

Then $x \cdot c = \lambda_1(x \cdot v_1) + \ldots + \lambda_k(x \cdot v_k) = 0$. 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 $V_n(F)$. □

Theorem 3.3

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 $F = \mathrm{GF}(q)$.Then $|C| = q^k$ and $|(C^\perp)^\perp| = q^k$.

$\therefore C = (C^\perp)^\perp$. □

Theorem 3.4: Constructing a GM for $C^\perp$

Let C be an (n,k)-code with GM $G = \left[I_k | A_{k \times (n-k)}\right]_{k \times n}$. Then *a* GM for $C^\perp$ is

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

Proof:
$\mathrm{rank}(H) = n-k$, so H is indeed a GM for some $(n, n-k)$-code $\overline{C}$.

Now,

$$GH^T = [I_k | A] \left[\frac{-A}{I_{n-k}}\right] = -A + A = 0$$

Since $GH^T = 0$, 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 $\overline{C} \subseteq C^\perp$. Since $\dim(\overline{C}) = \dim(C^\perp)$, we have $\overline{C} = C^\perp$. □

parity-check matrix

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

Example:

Consider a $(5,2)$-code C over $\mathbb{Z}_3$ with GM

$$G = \begin{bmatrix} 2 & 0 & 2 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix}_{2\times 5}$$

For C: $q = 3, n = 5, k = 2, M = 3^2 = 9$.

$$C = \{00000, \underset{c_1}{20210}, \underset{2c_1}{10120}, \underset{c_2}{11001}, \underset{2c_2}{22002}, \underset{c_1+c_2}{01211}, \underset{c_1+2c_2}{12212}, \underset{2c_1+c_2}{21121}, \underset{2c_1+2c_2}{02122}\}$$

Now find a GM for $C^\perp$

$$\begin{bmatrix} 2 & 0 & 2 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{\text{row reductions}} \left[\begin{array}{cc|ccc} 1 & 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 1 & 1 \end{array}\right]$$

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 GM for $C^\perp$ which is an $(5,3)$-code over $\mathbb{Z}_3$.

Note

Let C be an (n,k)-code over F with GM G:

1. $C^\perp$ is the nullspace of G.
2. $C^\perp$ is an $(n, n-k)$-code over F.
3. $(C^\perp)^\perp = C$
4. Let H be a GM for $C^\perp$, then H is a PCM for C (by definition).
5. G is a PCM for $C^\perp$.
6. $GH^T = 0$.
7. $\boxed{\text{For } x \in V_n(F), x \in C \text{ iff } Hx^T = 0.}$

   [C is the nullspace of H.]

Theorem 3.5

Let C be an (n,k)-code over F, and let H be a PCM for C. Then $d(C) \geq s$ iff every $s-1$ cols of H are linearly independent over F.

Proof:
Let $h_1, h_2, \ldots, h_n$ be the cols of H.

$\Longleftarrow$) Suppose $d(C) \leq s-1$, so $\omega(C) \leq s-1$. Let $c \in C$, with $1 \leq \omega(C) \leq s-1$. WLOG, suppose $c_j = 0$, $\forall s \leq j \leq n$. Since $c \in C$, we have $Hc^T = 0$.

$\therefore c_1h_1 + c_2h_2 + \ldots + c_{s-1}h_{s-1} = 0$

Since $\omega(C) \geq 1$, this is a non-trivial linear combinations of $h_1, \ldots, h_{s-1}$ that equal 0. So $h_1, \ldots, h_{s-1}$ are linear dependent over F.

$\Longrightarrow$) Suppose there are $s-1$ cols of H that are linear dependent over F, say $h_1, \ldots, h_{s-1}$. So we can write $c_1h_1 + c_2h_2 + \ldots + c_{s_1}h_{s-1}$ where $c_j \in F$, not all zero.

Let $c = (c_1, c_2, \ldots, c_{s-1}, \underbrace{0, \ldots, 0}_{n-s+1}) \in V_n(F)$.

Then $Hc^T = 0$. So $c \in C$. And $1 \leq \omega(C) \leq s-1$, so $d(C) \leq s-1$.

□

Corrollary 3.6

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

Example:
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 $d(C)$

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

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

so columns 1, 3, 4 are linearly dependent over $\mathbb{Z}_3$. Then $d(C) \ngeq 4$, so $d(C) = 3$.

Example:
C be a binary code, with PCM H

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

Example: Construct a $(\underset{n}{7}, \underset{k}{4}, \underset{d}{3})$-binary code C

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* $\mathbb{Z}_2$.

3.3 Perfect Code

perfect code

Let C be an $[n, M]$-code C over A of distance d. Then

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

where $e = \left\lfloor \frac{d-1}{2} \right\rfloor$. [Sphere packing bound]

Then C is *perfect* if

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

Note
If C is perfect, then IMLD = CMLD.

For fixed n, q, d, a perfect code maximized $R = \frac{\log_q M}{n}$.

Example:
$C = \mathrm{GF}(q)^n$ is a (trivial) perfect code with $d = 1$.

Example:
$C = \{\underbrace{0\ldots0}_{n}, \underbrace{1\ldots1}_{n}\}$ over $\mathbb{Z}_2$ is a perfect code if n is odd. (distance $= n$).

Proof:

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

□

Exercise

Prove that every perfect code must have odd distance.

Theorem 3.7: Tietäräinen, 1973

The only perfect codes are

(i) $V_n(\mathrm{GF}(q))$

(ii) The binary replication code of odd length.

(iii) The $(23, 12, 7)$-binary Golay code and all codes equivalent to it.

(iv) The $(11, 6, 5)$-ternary[a] Golay and all codes equivalent to it.

A GM is

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

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

[a] over $\mathbb{Z}^3$

Hamming code of order r over GF(q)

A Hamming code of order r over $\mathrm{GF}(q)$ is a linear code over $\mathrm{GF}(q)$ with $n = \frac{q^r-1}{q-1}$, $k = n - r$ and PCM a $r \times n$ matrix whose columns are nonzero & no two are scalar multiples of each other.

Example: A Hamming code of order $r = 3$ over $\mathrm{GF}(2)$

is a $(7, 4, 3)$-binary code with PCM

$$H = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}_{3\times 7}$$

Example: A Hamming code of order $r = 3$ over $\mathrm{GF}(3)$

is a $(13, 10, 3)$-code over $\mathrm{GF}(3)$ 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

1. For every nonzero vector $v \in V_r(\mathrm{GF}(q))$, exactly one scalar of v must be a column of a PCM (for the Hamming code of order r over $\mathrm{GF}(q)$) $n = \frac{q^r-1}{q-1}$
2. The dimension of the code is indeed k, since $\mathrm{rank}(P(M)) = r = n - k$. Since $\lambda_i e_i$ ($e_i = i^{th}$ unit vector, λ_i's are non-zero scalars) are cols of PCM.
3. The Hamming codes have distance 3. (since $\lambda_1 e_1, \lambda_2 e_2$ and $\lambda_3(e_1 + e_2)$ are cols of H for some scalar multiples $\lambda_1, \lambda_2, \lambda_3$)
4. The Hamming codes are perfect:

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

3.4 Error Correction (for Hamming Codes)

error vector

Suppose $c \in C$ is transmitted. Suppose $r \in V_n(F)$ is received. The *error vector* is $e = r - c$. $(c + e = r)$

Example:

Over $\mathbb{Z}_3$, if $c = (120212)$ is sent and $r = (122102)$ is received, then $e = (00220)$.

Decoding algorithm for single-error correcting codes (e.g. Hamming codes)

Let H be a PCM for an (n, k)-code C over $\mathrm{GF}(q)$ with $d \geq 3$.

Recall $c \in C$ is sent, $r \in V_n(\mathrm{GF}(q))$ is received, the *error vector* is $e = r - c$.

Main idea $Hr^T = H(c+e)^T = Hc^T + He^T = He^T$

syndrome

If $r \in V_n(\mathrm{GF}(q))$, $s = Hr^T$ is called the *syndrome* of r.

Note

1) r and e have the same syndrome.

2) If $e = 0$, then $He^T = 0$

3) If $\omega(e) = 1$, say $e = (0, \dots, \underbrace{\alpha}_{i^{th} \text{ position}}, \dots, 0)$ where $\alpha \neq 0$.

   Then $He^T = \alpha h_i$ (nonzero), where $h_i = i^{th}$ col of H.

4) Note: The converses of 2) and 3) are false.

Algorithm 1: Decoding algorithm (for single error-correcting codes)

Given : H, r

Compute $s = Hr^T$

If $\omega(s) = 0$, then accept r. (STOP)

Compare r with the columns of H. If $s = \alpha h_i$ (where $\alpha \neq 0$), then $e = (0, \dots, \underbrace{\alpha}_{i^{th}}, \dots, 0)$, and correct r to $c = r - e$. (STOP)

Reject.] NOT NEEDED if H is a Hamming code (because it is perfect)

Claim If $\omega(e) \leq 1$, then the decoding algorithm always makes the correct decision.

Note

If H is a Hamming code & $\omega(e) \geq 2$, then this decoding algorithm always makes the wrong decision.

Example:

Consider the $(7, 4, 3)$-binary Hamming code with PCM

$$H = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}_{3\times 7}$$

Decode $r = (0111110)$.

1. Compute $s = Hr^T = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, which is the 6^{th} col of H.

2. So, $e = (0000010)$

3. Decode r to $c = (0111100)$. [Check $Hc^T = 0$]

3.5 General Decoding Problem for binary linear codes

Instance

- An $(n-k) \times n$ matrix H over GF(2) with $\text{rank}(H) = n-k$.
- $r \in V_n(\text{GF}(2))$

Find Find a vector $e \in V_n(\text{GF}(2))$ of minimum weight with $Hr^T = He^T$.

Fact This problem is **NP-hard**[1].

Decoding Linear Codes

Let C be an (n,k)-code over $F = \text{GF}(q)$ with PCM H.

$\equiv$ [mod C]

We write $x \equiv y$ [mod C], where $x, y \in V_n(F)$ if $x - y \in C$.

Note

1) $\equiv$ [mod C] is an equivalence relation. (Reflexive, Symmetric, Transitive)
2) So, the set of equivalence classes partitions $V_n(F)$.
3) The equivalence class containing $x \in V_n(F)$ is called a *coset* of $V_n(F)$.

   This class is $\{y \in V_n(F) : y \equiv x \text{ [mod } C]\} = \{x + c : c \in C\} = C + x$.

   We call $C = x$ the coset of C represented by x. (See the example below)
4) $C + 0 = C$
5) If $y \in C + x$, then $C + y = C + x$
6) Every coset has size q^k

[1] These ideas could be found in CS 341/466/666 ...

- P = problems solvable in "polynomial time" (i.e. efficiently)
- NP = a "certain" class of problems including many problems of strong practical interest which do not know to solve efficiently.
- NP-hard: If any single problem in this class of problems can be solved efficiently, then so can all problems in NP (in which, P=NP)

7) # of cosets is $q^n/q^k = q^{n-k}$.

Example:

Consider the $(5, 2)$-binary code C with GM

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

The cosets of C are:

$$\begin{aligned} C + 00000 &= \{00000, 10111, 01110, 11001\} \\ = C + 10111 &= \{10111, 00000, 11001, 01110\} \\ = C + 01110 &= \dots \\ = C + 11001 &= \dots \end{aligned}$$

and

$$C + 10000 = \{10000, 00111, 11110, 01001\} = C + 00111$$

which is disjoint from the previous equivalent class. And brand new cosets:

$$\begin{aligned} C + 01000 &= \{01000, 11111, 00110, 10001\} \\ C + 00100 &= \{00100, 10011, 01010, 11101\} \\ C + 00010 &= \{00010, 10101, 01100, 11011\} \\ C + 00001 &= \{00001, 10110, 01111, 11000\} \\ C + 00011 &= \{00011, 10100, 01101, 11010\} \\ C + 11100 &= \{11100, 01011, 10010, 00101\} \end{aligned}$$

Theorem 3.8

Let $x, y \in V_n(F)$. Then $x \equiv y \ [\bmod C]$ iff $Hx^T = Hy^T$.

Proof:

$$x \equiv y \ [\bmod C] \iff H(x-y)^T = 0 \iff Hx^T = Hy^T$$

□

So, cosets are characterized by their syndromes.

Decoding $c \in C$ is send, $r \in V_n(F)$ is received. $e = r - c \in V_n(F)$. $Hr^T = He^T$.

So, r and e belong to the same coset of C.

CMLD Given r, find a vector e of smallest weight in $C+r$, or, equivalently, find a vector e of smallest weight with the same syndrome as r. Then decode r to $c = r - e$.

IMLD Find the unique vector e of smallest weight having the same syndrome as r. If no such e exists, then reject r. Otherwise, decode r to $c = r - e$.

coset leader

A vector of smallest weight is a coset of C is distinguished and called a *coset leader* (of that coset).

3.5.1 Syndrome Decoding

Algorithm 2: Syndrome Decoding Algorithm

Given : A PCM H for an (n, k)-code C over $F = \mathrm{GF}(q)$

Create a table of coset leaders and their syndromes. Given r, do

Compute $s = Hr^T$

Look up the coset leader corresponding to s, say ℓ.

Decode r to $c = r - \ell$.

Example:

$$G = \left[\begin{array}{cc|ccc} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{array}\right]_{2\times 5}$$

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

$n = 5, k = 2, q = 2$

Coset Leaders		Syndromes
00000	—	000
10000	—	111
01000	—	110
00100	—	100
00010	—	010
00001	—	001
00011	—	011
10010	—	101

Suppose $r = 10111$.

Compute $s = Hr^T = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

Coset leader is $\ell = 00000$.

So $c = r - e = 10111$.

Recall There is a 1-1 correspondence between cosets of C and syndromes. Also, $s = Hr^T = He^T$.

For a binary (n, k)-code C, the syndrome table has size $2^{n-k} \times n$, which is exponentially large.

Goal Design decoding algorithm which require very little space.

Example:

Use only the PCH, H, which is $(n-k)n$ bits.

4

The binary Golay code

4.1 The (binary) Golay code C_{23} (1949)

Let

$$\hat{B} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ \vdots \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}_{12\times 11} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

All rows below the second are left cyclic shifts of second row.

Let $\hat{G} = [I_{12}|\hat{B}]_{12\times 23}$

Then $\hat{G}$ is a GM for a $(23, 12)$-binary code called C_{23}

Facts

(i) $d(C_{23}) = 7$ (proof later) So $e = 3$

(ii) C_{23} is perfect:

$$2^{12}\left(\binom{23}{0} + \binom{23}{1} + \binom{23}{2} + \binom{23}{3}\right) = 2^{23}$$

4.2 The Extended Golay Code C_{24}

Let

$$B = \left[\begin{array}{c|cc} 0 & & \\ 1 & & \\ 1 & & \\ 1 & & \\ 1 & & \\ 1 & & \\ 1 & \hat{B} & \\ 1 & & \\ 1 & & \\ 1 & & \\ 1 & & \\ 1 & & \end{array}\right]_{12\times 12} = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Let $G = [I_{12}|B]_{12\times 24}$

Note

(i) C_{24} is a $(24, 12, 8)$-binary code (so $e = 3$)

(ii) $G^T G = 0$ (Check this)

(iii) Hence $C_{24} \subseteq C_{24}^{\perp}$ (C_{24} is a self orthogonal code)

(iv) $\dim(C_{24}) = 12$ and $d(C_{24}^{\perp}) = 12$, so $C_{24} = C_{24}^{\perp}$ (C_{24} is a self dual code)

(v) B is symmetric (check this).

(vi) A PCM for C_{24} is $H = [-B^T|I_{12}] = [B|I_{12}]$

(vii) Since $C_{24} = C_{24}^{\perp}$, H is also a GM for C_{24}

(viii) G is also a PCM for $C_{24}^{\perp}$.

4.2.1 Decoding Algorithm for C_{24}

$s = Hr^T$ syndrome table has size $2^{12} \times 24 \approx 96,000$ bits.

Recall C_{24} is a $(24, 12, 8)$-binary codes with PCMs $[B|I_{12}]$ and $[I_{12}|B]$.

Decoding strategy (IMLD) Compute a syndrome of r. Find a vector e of weight ≤ 3 that has the same syndrome as r.

If no such vector e exists, then reject r; else decode r to $c = r - e$.

Let $r = (x, y), e = (e_1, e_2)$. [all are 12 bits in length][1]

There are cases 5 cases (not mutually exclusive) in the event $\omega(e) \leq 3$.

(A) $w(e_1) = 0, w(e_2) = 0$ [2]

(B) $1 \leq w(e_1) \leq 3, w(e_2) = 0$

(C) $w(e_1) = 1$ or $2, w(e_2) = 1$

(D) $w(e_1) = 0, 1 \leq w(e_2) \leq 3$

(E) $w(e_1) = 1, w(e_2) = 1$ or 2

Theorem 4.1

Let C be an (n, k, d)-code over $\mathrm{GF}(q)$. Let $x \in V_n(\mathrm{GF}(q))$ have weight $\leq \left\lfloor \frac{d-1}{2} \right\rfloor$. Then x is the unique vector of min weight in the coset of C containing x (so must be a coset leader).

Proof:
Let y be a vector in the same coset of C as x, with $y \neq x$ and

$$w(y) \leq w(x) \leq \left\lfloor \frac{d-1}{2} \right\rfloor$$

Then $y - x \neq 0$, and $x \equiv y \;[\bmod C]$, and $x - y \in C$. Now

$$\begin{aligned} w(x-y) &= w(x + (-y)) \\ &\leq w(x) + w(-y) \\ &= w(x) + w(y) \\ &\leq \left\lfloor \frac{d-1}{2} \right\rfloor + \left\lfloor \frac{d-1}{2} \right\rfloor \\ &\leq d - 1 \end{aligned}$$

contradicting $d(C) = d$. □

Recall $[I_{12}|B]$ and $[B|I_{12}]$ are both PCMs for C_{24}.

Note

$$\begin{aligned} S_1 &= [I_{12}|B]r^T \\ &= [I_{12}|B]e^T \\ &= [I_{12}|B] \begin{bmatrix} e_1^T \\ e_2^T \end{bmatrix} \\ &= e_1^T + Be_2^T \end{aligned}$$

Similarly, $s_2 = [B|I_{12}]r^T = Be_1^T + e_2^T$

[1] Different presentation from textbook
[2] Note that I use w to represent weight from now on...

Algorithm 3: Decoding Algorithm for C_{24}

Suppose: $r = (x, y)$ is received

1 Compute $s_1 = [I_{12}|B]r^T$. If $s_1 = 0$, then accept r; STOP.

2 If $1 \leq w(s_1) \leq 3$, then correct x in the positions corresponding to is in s_1; STOP.

3 Compare s_1 to the columns (or rows) of B. If ant column, say col i, differs in one position (say j) or differs in two position (j and k) from s_1, then correct r as follows:

- Correct x in the position k or position j and k
- Correct y in position i.

STOP.

4 Compute $s_2 = [B|I_{12}]r^T$. If $w(s_2) \leq 3$, then correct y in position corresponding to the pos in s_2. STOP.

5 Compute s_2 to the cols (or rows) of B. If any col, say col i, differs in one pos (say j) or 2 pos (j and k), then

- Correct y in pos j, or pos j and k
- Correct x in pos i.

STOP.

6 Reject (since $w(e) \geq 4$).

See examples in handouts on LEARN.

Note

1. If $w(e) \leq 3$, then the algorithm makes the correct decision.
2. No storage is needed:
$$s_1 = [I_{12}|B] \begin{bmatrix} x \\ y \end{bmatrix} = x + By$$
3. Algorithm is very simple and efficient (good for hardware)

4.2.2 Reliability of C_{24}

- $p =$ symbol error prob. (BSC)
- $C = \{c_1, c_2, \ldots, c_M\}$
- $w_i =$ prob. that decoding algorithm makes an incorrect decision if c_i is sent.
- Error prob of C is $P_C = \dfrac{1}{M} \sum w_i$
- $1 - P_C =$Reliability of C

p	(1) $(1-p)^{12}$	(2) 1-$P_{C_{24}}$	(3) $1-P_T$	(4) $1-P_H$
0.1	0.282429	0.7857377	0.7112056	0.5490430
0.01	0.8863848	0.99990946	0.9964298	0.9903702
0.001	0.9880657	0.9999999895	0.99996402	0.998959
Rate	1	1/2	1/3	$11/15 \approx 0.73$

(1) If no source is used, then the reliability for 12-bit messages is $(1-p)^{12}$.

(2) $w_i = 1 - [(1-p)^{24} + \binom{24}{1}p(1-p)^{23} + \binom{24}{2}p^2(1-p)^{22} + \binom{24}{3}(1-p)^{21}]$

$$P_{C_{24}} = \frac{1}{2^{12}} \sum_{i=1}^{2^{12}} w_i = w_i$$

(3) T =Triplication code

$$\underbrace{10110\ldots 0}_{12} \to \underbrace{111\ 000\ 111\ 111\ 000\ldots\ 111}_{36}$$

$$1 - P_T = [(1-p)^3 + 3p(1-p)^2]^{12}$$

(4) $(15, 11)$-binary Hamming code

$$1 - P_H = (1-p)^{15} + 15p(1-p)^{14}$$

5

Cyclic Codes

cyclic subspace

A *cyclic subspace* S of $V_n(F)$ is a subspace such that $(a_0, a_1, \ldots, a_{n-1}) \in S \implies (a_{n-1}, a_0, \ldots, a_{n-2}) \in S$.

cyclic code

A *cyclic code* is a cyclic subspace of $V_n(F)$.

Let $R = F[x]/(x^n - 1)$

Associate

$$\underbrace{(a_0, a_1, \ldots, a_{n-1})}_{\in V_n(F)} \leftrightarrow \underbrace{a_0 + a_1 x + \ldots + a_{n-1}x^{n-1}}_{\in R}$$

Addition is preserved

$$a + b \leftrightarrow a(x) + b(x)$$

Scalar multiplication is preserved

$$\lambda a \leftrightarrow \lambda a(x)$$

Why choose $x^n - 1$?

Let $a = (a_0, \ldots, a_{n-1}) \in V_n(F)$. Let $a(x)$ be its associated polynomial in R. Then

$$\begin{aligned} x \cdot a(x) &= a_0 x + a_1 x^2 + \ldots + a_{n-2}x^{n-1} + a_{n-1}x^n \\ &\equiv a_{n-1} + a_0 x + a_1 x^2 + \ldots + a_{n-2}x^{n-1} \; [\bmod x^n - 1] \\ &\leftrightarrow (a_{n-1}, a_0, \ldots, a_{n-2}) \end{aligned}$$

So multiplying a polynomial in R by x corresponds to a (right) cyclic shift of the associated vector.

We'll define $\cdot : V_n(F) \times V_n(F) \to V_n(F) : a \cdot b \leftrightarrow a(x) \cdot b(x)\ [\bmod x^n - 1]$

$$\text{Cyclic subspaces of } V_n(F) \leftrightarrow \text{Ideals in } R \leftrightarrow \text{monic divisors of} x^n - 1$$

ideal

Let R be a commutative finite ring. A non-empty subset I of R is an *ideal* of R if

(1) For all $a, b \in I$, $a + b \in I$.

(2) For all $a \in I, b \in R, a \cdot b \in I$.

Example:
$\{0\}$ and R are (trivial) ideal of R.

Theorem 5.1

Let $S \subseteq V_n(F)$, non-empty. Let I be the associated polynomials.

Then S is a cyclic subspace of $V_n(F)$ iff I is an ideal of $R = F[x]/(x^n - 1)$.

Proof:

$\Rightarrow$) Suppose S is a cyclic subspace of $V_n(F)$.

Since S is closed under addition, so is I. Let $a(x) \in I$, $b(x) = b_0 + b_1x + \ldots + b_{n-1}x^{n-1} \in R$. Then $xa(x) \in I$ since S is a cyclic subspace.

So, $x^i a(x) \in I$, for all $0 \leq i \leq n - 1$.

Also, $b_i x^i a(x) \in I$, since S is closed under scalar multiplication.

Finally, $a(x) \cdot b(x) = a(x)(b_0 + b_1x + \ldots + b_{n-1}x^{n-1})$ which is in I, since I is closed under addition.

$\therefore$ I is an ideal.

$\Leftarrow$) Suppose I is an ideal of R. Since I is closed under addition, so is S. Since I is closed under multiplication by constant polynomials, S is closed under scalar multiplication. Since I is closed under multiplication by x, S is closed under (right) cyclic shifts.

$\therefore$ S is a cyclic subspace.

$\square$

Remark:
So, we have a 1-1 correspondence between cyclic subspaces of $V_n(F)$ and ideals of $R = F[x]/(x^n - 1)$.

ideal generated by $g(x)$

Let $g(x) \in R$. Then $\langle g(x) \rangle = \{g(x) \cdot a(x) : a(x) \in R\}$. Then $\langle g(x) \rangle$ is an ideal of R. Called the *ideal generated by* $g(x)$.

principal ideal & principal ideal ring

If I is an ideal of R, then I is *principal ideal* if $\exists g(x) \in I$ such that $I = \langle g(x) \rangle$.

R is called a *principal ideal ring* if every ideal of R is principal.

Theorem 5.2

$R = F[x]/(x^n - 1)$ is a principal ideal ring.

Proof:
Let I be an ideal of R.

Suppose first that $I = \{0\}$. Then $I = \langle 0 \rangle$ is principle.

Suppose $I \neq \{0\}$. Let $g(x)$ ne a polynomial of smallest degree in I. Let $a(x) \in I$.

Long division gives

$$a(x) = \ell(x)g(x) + r(x)$$

where $\ell, r \in F[x]$, and $\deg(r) < \deg(g)$.

But $\ell(x)g(x) \in I$. (since I is closed under multiplication by R).

And $a(x) - \ell(x)g(x) \in I$.

$\therefore$ $r(x) \in I$. Since $\deg(r) < \deg(g)$, we must have $r(x) = 0$. Hence $a(x) = \ell(x)g(x)$.

$\therefore I = \langle g(x) \rangle$

$\therefore$ R is a principal ideal ring. $\square$

Note
We can take $g(x)$ to be *monic* (i.e. $g(x) = x^\ell + g_{\ell-1}x^{\ell-1} + \ldots + g_0$)

If $g(x)$ were not monic, say $g_\ell x^\ell + \ldots + g_0$ where $g_\ell \neq 0, 1$, then

$$g_\ell^{-1} g(x) = x^\ell + \ldots + g_\ell^{-1} g_0$$

is monic and is also in I. We will call this process "making $g(x)$ monic".

the generator polynomial of I

Let I be an ideal in $R = F[x]/(x^n - 1)$. If $I = \{0\}$, then *the generator polynomial* of I is $x^n - 1$ (since $x^n - 1 \equiv 0 \; [\bmod\, x^n - 1]$).

If $I \neq \{0\}$, the monic polynomial of least degree in I is called *the generator polynomial of I*.

Theorem 5.3

Let I be a nonzero ideal in $R = F[x]/(x^n - 1)$.

(i) There is *unique* monic polynomial $g(x)$ of smallest degree in I.

(ii) $g(x) | x^n - 1$.

Proof:

(i) Suppose $g(x), h(x)$ are two monic polynomials of (the same) smallest degree in I. Then $g(x) - h(x) \in I$ and $\deg(g - h) < \deg(g)$. Hence we must have $g - h = 0$, so $g(x) = h(x)$.

(ii) We can write $x^n - 1 = \ell(x)g(x) + r(x)$ where $\ell, r \in F[x]$, and $\deg(r) < \deg(g)$. So
$$0 \equiv \ell(x)g(x) + r(x) \; [\bmod\, x^n - 1],$$
so
$$r(x) \equiv -\ell(x)g(x) \; [\bmod\, x^n - 1].$$
Since $\langle g(x) \rangle = I$, we have $r(x) \in I$. Hence $\deg(r) < \deg(g)$, we have $r(x) = 0$.

$\therefore g(x) | x^n - 1$.

$\square$

Theorem 5.4

Let $h(x)$ be a monic divisor of $x^n - 1$ in $F[x]$. Then the generator polynomial of $\langle h(x) \rangle$ is $h(x)$.

Proof:

If $h(x) = x^n - 1$, then $I = \{0\}$ and, by defn, its generator polynomial is $x^n - 1$.

If $\deg(h) < n$, so $I \neq \{0\}$, then let $g(x)$ be the monic polynomial of smallest degree in I. Since $g(x)$ is a generator of I, we can write $h(x) \equiv a(x)g(x) \; [\bmod\, x^n - 1]$.

So, $g(x) = a(x)h(x) + \ell(x)(x^n - 1)$ for some $\ell(x) \in F[x]$.

Since $h | x^n - 1$ and $h | ah$, we have $h(x) | g(x)$.

So, $\deg(h) \leq \deg(g)$. Since g is the monic polynomial of smallest degree in I, we have $\deg(g) \leq \deg(h)$, so $\deg(g) = \deg(h)$. Since g, h are both monic, we have $g(x) = h(x)$. □

Corollary 5.5

There is 1-1 correspondence between monic divisors of $x^n - 1$ in $F[x]$ and ideals in R. There is a 1-1 correspondence between monic divisors of $x^n - 1$ in $F[x]$ and cyclic subspaces of $V_n(F)$.

Example:

Find all cyclic subspaces of $V_3(\mathbb{Z}_2)$. ($n = 3$)

The complete factorization of $x^3 - 1$ over $\mathbb{Z}_2$ is:

$$x^3 - 1 = (1 + x)(1 + x + x^2) \qquad R = \mathbb{Z}_2[x]/(x^3 - 1)$$

Monic divisor of $x^3 - 1$	$\langle g_i(x) \rangle$ [a]	$\dim \langle g_i(x) \rangle$
$g_1(x) = 1$	$\underbrace{\{000, 001, \ldots, 111\}}_{8}$	3
$g_2(x) = 1 + x$	$\{000, 110, 011, 101\}$	2
$g_3(x) = 1 + x + x^2$	$\{000, 111\}$	1
$g_4(x) = 1 + x^4$	$\{0\}$	0

[a] Cyclic subspace generated by $g_i(x)$

We have one to one associations:

$$V_n(F) \leftrightarrow R = F[x]/(x^n - 1)$$

$$a = (a_0, a_1, \ldots, a_{n-1}) \in V_n(F) \leftrightarrow a_0 + a_1x^1 + \ldots + a_{n-1}x^{n-1} \in R = F[x]/(x^n - 1)$$

C : (cyclic subspace)		$g(x)$ [monic divisor of x^n-1]
$\dim C = k$		$\deg g = n - k$
GM for C in terms of $g(x)$		$h(x) = (x^n - 1)/g(x)$
Encoding: $mG \leftrightarrow m(x)g(x)$	$\leftrightarrow I$, (ideal in R) $\leftrightarrow$	$\deg(h) = k$
$C^\perp$: dual code of C is cyclic		$h_R(x)$: reciprocal poly of $h(x)$
PCM H for C:		$h^*(x) = h_R$ by making h_R monic
$s(x) \equiv r(x)$ [mod $g(x)$]		$h^*(x)$ the gen polynomial of $C^\perp$

Distance of C?

C: BCH code: $g(x)$ is specially selected to give a lower bound of C.

Lemma 5.6

Let $g(x)$ be a monic divisor of deg $n - k$ of $x^n - 1$ in $F[x]$.

Recall $\langle g(x) = \{g(x)a(x) : a(x) \in R\}\rangle$ [a]

In fact, $\langle g(x)\rangle = \{g(x)\overline{a}(x) : \deg(\overline{a}) < k\}$ [b]

[a] $a(x)$ with mod

[b] $\overline{a}(x)$ no mod

Proof:
Let $h(x) = g(x)a(x)$ [mod $x^n - 1$] where $\deg(a) < n$. So, $h(x) - g(x) = \ell(x)(x^n - 1)$ for some $\ell \in F[x]$.

$\therefore g(x) | h(x)$

So $h(x) = g(x)\overline{a}(x)$, for some $\overline{a} \in F[x]$ with $\deg(\overline{a}) \leq k - 1$. $\square$

Theorem 5.7

Let $g(x)$ be a monic divisor of $x^n - 1$ of deg $n - k$ in $F[x]$. Then the cyclic code C generated by $g(x)$ has $\dim k$.

Proof:
We'll show that

$$B = \{g(x), xg(x), \ldots, x^{k-1}g(x)\}$$

is a basis of C.

B is lin indep over F.

Suppose

$$\underbrace{\lambda_0 g(x)}_{\deg = n-k} + \underbrace{\lambda_1 x g(x)}_{\deg = n-k+1} + \ldots + \underbrace{\lambda_{k-1} x^{k-1} g(x)}_{\deg = n-k+k-1 = n-1} = 0$$

where $\lambda_i \in F$.

The coeff of x^{n-1} in the LHS is λ_{k-1}. The coeff of x^{n-1} in RHS is 0. Hence $\lambda_{k-1} = 0$. Similarly, $\lambda_0 = \lambda_1 = \ldots = \lambda_{k-2} = 0$.

Claim B spans C.

Let $h(x) \in \langle g(x) \rangle$. By Lemma, we can write

$$h(x) = \underbrace{\underbrace{g(x)}_{n-k}\underbrace{a(x)}_{k-1}}_{n-1 \text{ no mod}}$$

for some $a(x) \in F[x], \deg(a) \leq k - 1$.

Let $a(x) = \sum_{i=0}^{k-1} a_i x^i$, where $a_i \in F$.

Then $h(x) = g(x)a(x) = \sum_{i=0}^{k-1} a_i[x^i g(x)]$

Hence $\dim C = k$. $\square$

So a GM for C is

$$G = \begin{bmatrix} g(x) \\ xg(x) \\ \vdots \\ x^{k-1}g(x) \end{bmatrix}_{k\times n} = \begin{bmatrix} \boxed{g(x)} & 0 & \cdots & 0 \\ 0 & \boxed{xg(x)} & \cdots & 0 \\ \vdots & & & \\ 0 & \cdots & 0 & \boxed{x^{k-1}g(x)} \end{bmatrix}$$

Note
G is a non-systematic GM for C.

Encoding

$$\begin{aligned} c &= mG \\ &= (m_0, \cdots, m_{k-1}) \begin{bmatrix} g(x) \\ xg(x) \\ \vdots \\ x^{k-1}g(x) \end{bmatrix} \\ &= m_0 g(x) + m_{k-1}x^{k-1}g(x) \\ &= g(x)\left(m_0 + \cdots + m_{k-1}x^{k-1}\right) \\ &\implies c(x) = m(x)g(x) \end{aligned}$$

Example: Construct a cyclic $(7,4)$-code over $\mathbb{Z}_2$

We need a monic divisor of deg 3 of $x^7 - 1$ in $\mathbb{Z}_2[x]$. Table 3 in page 157 of textbook:

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

Let's take $g(x) = 1 + x + x^3$. Then $\langle g(x)\rangle = 1 + x + x^3$. Then $\langle g(x)\rangle$ is a $(7,4)$-cyclic code over $\mathbb{Z}_2$. A GM for C is

$$G = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}_{4\times 7}$$

(hamming code cyclic)

Encode $m = \begin{pmatrix} 1 & 0 & 1 & 1 \end{pmatrix}$

$$c = mG = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

$$c(x) = m(x)g(x) = \left(1 + x + x^3\right)\left(1 + x + x^3\right) = \left(1 + x + \cdots + x^6\right) = c$$

5.1 Dual Code of a Cyclic Code

Let C be an (n,k)-cyclic code over F with generated polynomial $g(x)$.

Let $g(x) = \underset{\neq 0}{g_0} + g_1 x + g_2 x^2 + \ldots + \underset{=1}{g_{n-k}} x^{n-k} + \underbrace{g_{n-k+1}x^{n-k+1} + \ldots + g_{n-1}x^{n-1}}_{0}$

Let $h(x) = (x^n - 1)/g(x) = \underset{\neq 0}{h_0} + h_1 x + \ldots + h_{k-1}x^{k-1} + \underset{=1}{h_k} x^j + \underbrace{\ldots + h_{n-1}x_0^{n-1}}_{0}$

Let $a(x) = a_0 + a_1 x + \ldots + a_{n-1}x^{n-1}$

and

$$a(x) = g(x)h(x) \;[\mathrm{mod}\, x^n - 1] \qquad (*)$$

Note $a(x) = 0$. Equating coeffs of $x^i, 0 \leq i \leq n-1$, of $(*)$:

$$a_i = 0 = g_0 h_i + g_1 h_{i-1} + \ldots + g_i h_0 + g_{i+1}h_{n-1} + g_{i+1}h_{n-2} + \ldots + g_{n-1}h_{i+1}$$

Let $g = (g_0, g_1, \ldots, g_{n-1})$ and $\overline{h} = (h_{n-1}, h_{n-2}, \ldots, h_1, h_0)$

Then g is orthogonal to $\overline{h}$ and all the cyclic shifts of $\overline{h}$. So every cyclic shift of g is orthogonal to every cyclic shift of $\overline{h}$.

Recall a GM for C is

$$\begin{bmatrix} g_0 & g_1 & \ldots & g_{n-k} & 0 & \ldots & 0 & 0 \\ 0 & g_0 & g_1 & \ldots & g_{n-k} & 0 & \ldots & 0 \\ \vdots \\ 0 & \ldots & 0 & 0 & g_0 & g_1 & \ldots & g_{n-k} \end{bmatrix}_{k\times n}$$

Consider

$$H = \begin{bmatrix} h_k & h_{k-1} & \ldots & h_0 & 0 & \ldots & 0 & 0 \\ 0 & h_k & h_{k-1} & \ldots & h_0 & 0 & \ldots & 0 \\ \vdots \\ 0 & \ldots & 0 & 0 & h_k & h_{k-1} & \ldots & h_0 \end{bmatrix}_{(n-k)\times n}$$

We have observed that $GH^T = 0$. Let C' be the code spanned by the rows of H. Then $C' \subseteq C^\perp$. But $\mathrm{rank}(H) = n - k$ (since $h_k = 1$). Since $\dim(C') = n - k$, we have $C' = C^\perp$.

Hence H is a PCM for C.

reciprocal of h

Let $h(x) = h_0 + h_1 x + \ldots + h_k x^k$ be a degree k polynomial. The reciprocal of h is $h_R(x) = h_k x^0 + h_{k-1}x + \ldots + h_1 x^{k-1} + h_0 x^k$.

Note

$h_R(x) = x^k h\left(\frac{1}{k}\right)$

If $h_0 \neq 0$, then $h^*(x) = h_0^{-1} \cdot h_R(x)$.

Theorem 5.8

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

Proof:
$g(x)h(x) = x^n - 1$

So,

$$g\left(\frac{1}{x}\right) h\left(\frac{1}{x}\right) = \frac{1}{x^n} - 1$$

So,

$$\left(x^{n-k} g\left(\frac{1}{x}\right)\right)\left(x^k h\left(\frac{1}{x}\right)\right) = 1 - x^n$$

Then

$$g_R(x) \cdot h_R(x) = -(x^n - 1)$$

Then

$$h_R(x) | x^n - 1$$

So, $h_R(x)$ is a degree k divisor pf $x^n - 1$. So, the matrix H is a GM for the cyclic code generated by $h^*(x)$.

Hence $C^\perp$ is cyclic with generator polynomial $h^*(x)$. □

5.2 Computing Syndromes

Let's find a more convenient PCM for C.

(i) Find a GM for C of the form $[R|I_k]_{k\times n}$. (Essentially systematic)

For $0 \le i \le k-1$, long division gives:

$$x^{n-k+i} = \underbrace{\ell_i(x)}_{\deg \le k-1} \cdot \underbrace{g(x)}_{\deg = n-k} + \underbrace{r_i(x)}_{\deg \le n-k-1}$$

Then $-r_i(x) + x^{n-k+i} = \ell_i(x) \cdot g(x) \in C$

Let

$$G = \begin{bmatrix} -r_0(x) + x^{n-k} \\ -r_1(x) + x^{n-k+1} \\ \vdots \\ -r_{k-1}(x) + x^{n-1} \end{bmatrix}_{k\times n} = \begin{bmatrix} \overbrace{\boxed{-r_0(x)}}^{n-k} & 1 & 0 & \dots & 0 \\ \boxed{-r_1(x)} & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ \boxed{-r_{k-1}(x)} & 0 & 0 & \dots & 1 \end{bmatrix} = [R|I_k]$$

Then G has rank k, so G is a GM for C.

(ii) Construct a PCM for C

This is $H = [I_{n-k}| - R^T]_{(n-k)\times n}$

Then $Hr^T = r(x)\ [\bmod g(x)]$

Then

$$H^T = \begin{bmatrix} I_{n-k} \\ \hline -R \end{bmatrix}_{n\times(n-k)} = \begin{array}{l} \left[\begin{array}{c} x^0 \ [\bmod g(x)] \\ x^1 \ [\bmod g(x)] \\ \vdots \\ x^{n-k-1} \ [\bmod g(x)] \\ \hline x^{n-k} \ [\bmod g(x)] \\ \vdots \\ x^{n-1} \ [\bmod g(x)] \end{array}\right] \begin{array}{l} \\ \\ \\ \\ \longrightarrow r_0(x) \\ \\ \longrightarrow r_{k-1}(x) \end{array} \end{array}$$

So, columns of H are

$$x^0 \ [\bmod g(x)], \ldots, x^{n-1} \ [\bmod g(x)]$$

Hence if $r = (r_0, r_1, \ldots, r_{n-1}) \in V_n(F)$, then

$$\begin{aligned} s &= Hr^T \\ &= (r_0 \cdot x^0 \ [\bmod g(x)]) + \ldots + (r_{n-1} \cdot x^{n-1} \ [\bmod g(x)]) \\ &= (r_0 x^0 + r_1 x + \ldots r_{n-1} x^{n-1}) \ [\bmod g(x)] \\ &= r(x) \ [\bmod g(x)] \end{aligned}$$

Theorem 5.9

Let C be a cyclic code with g.p. $g(x)$, and $r \in V_n(F)$. Then the syndrome of r (w.r.t. the previous PCM) is

$$s(x) = r(x) \ [\bmod g(x)]$$

Example:

$g(x) = 1 + x + x^2 + x^3 + x^6$ is the g.p. for a $(15, 9)$-binary cyclic code. (Check: $g(x) | x^{15} - 1$ over GF(2))

Compute the syndrome of $r = (1110\ 1110\ 1100\ 000)$

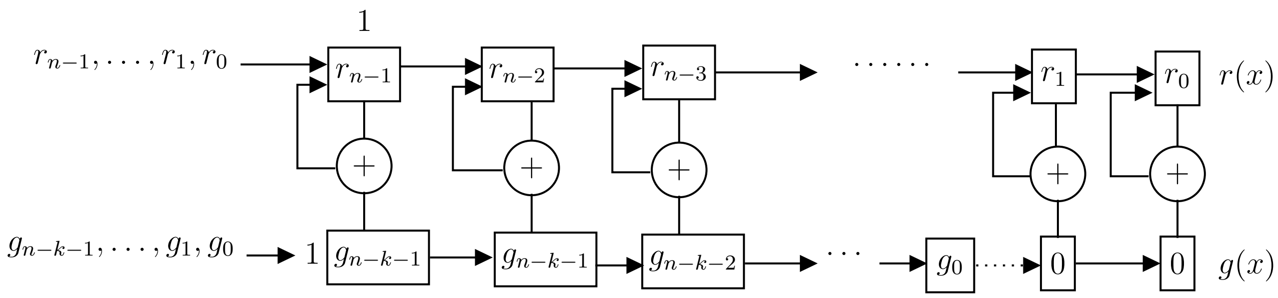
So $r(x) = 1 + x + x^2 + x^4 + x^5 + x^6 + x^8 + x^9$.

$$\begin{array}{r|l} & x^3 + x^2 \\ \hline x^6 + x^3 + x^2 + x + 1 & x^9 + x^8 + x^6 + x^5 + x^4 \qquad + x^2 + x + 1 \\ & -x^9 \qquad - x^6 - x^5 - x^4 - x^3 \\ \hline & x^8 \qquad\qquad - x^3 + x^2 \\ & -x^8 \qquad - x^5 - x^4 - x^3 - x^2 \\ \hline & - x^5 - x^4 - 2x^3 \qquad + x + 1 \end{array}$$

$s(x) = 1 + x + x^4 + x^5, \quad s = (110011)$

$$
\begin{array}{r|l}
g(x) & r(x) \\
1\,0\,0\,1\,1\,1\,1 & 1\,1\,0\,1\,1\,1\,0\,1\,1\,1 \\
 & 1\,0\,0\,1\,1\,1 \\
\hline
 & \quad 1\,0\,0\,0\,0\,1\,1\,1\,1 \\
 & \quad 1\,0\,0\,1\,1\,1\,1 \\
\hline
 & \qquad\quad 1\,1\,0\,0\,1\,1
\end{array}
$$

Hardware: $r(x)$ [mod $g(x)$]



So, $r(x)$ [mod $g(x)$] can be implemented in hardware using a very simple and fast circuit.

Note
Given the syndrome s of r, the syndromes of cyclic shifts of r can be easily computed.

Theorem 5.10

Let $r \in V_n(F)$ and $s = r(x)$ [mod $g(x)$] $= s_0 + s_1 x + \ldots + s_{n-k-1}x^{n-k-1}$.

Then the syndrome of $\underbrace{xr(x)}_{\text{cyclic shift}}$ is

(i) $xs(x)$, if $s_{n-k-1} = 0$

(ii) $xs(x) + s_{n-k-1}g(x)$, if $s_{n-k-1} \neq 0$

These two above are not cyclic shifts.

Proof:
We have $r(x) = \ell(x) + g(x) + s(x)$

Multiply by x:

$$xr(x) = x\ell(x) \underbrace{g(x)}_{\deg = n-k} + xs(x)$$

If $s_{n-k-1} = 0$, then $\deg(s) \leq n-k-2$, so $\deg(xs(x)) \leq n-k-1$.

So $xs(x)$ is the remainder upon dividing $xr(x)$ by $g(x)$.

So, $xs(x)$ is the syndrome of $r(x)$.

If $s_{n-k-1} \neq 0$, then $\deg(s) = n-k-1$. Then

$$\begin{aligned} xr(x) &= x\ell(x)g(x) + xs(x) + s_{n-k-1}g(x) - s_{n-k-1}g(x) \\ &= (x\ell(x) + s_{n-k-1})g(x) + \underbrace{[xs(x) - s_{n-k-1}g(x)]}_{\deg \leq n-k-1} \end{aligned}$$

Because

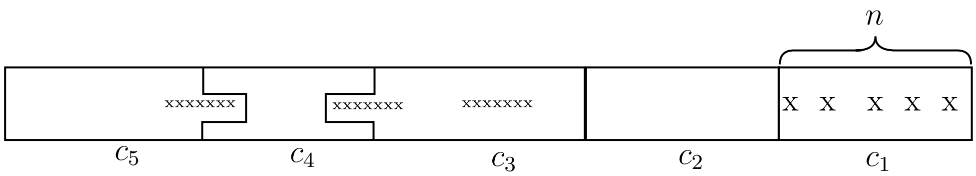
$$\begin{array}{rl} xs(x) & = s_0 x + \ldots + s_{n-k-1}x^{n-k} \\ - \quad s_{n-k-1}g(x) & = \ldots + \ldots + s_{n-k-1}x^{n-k} \\ \hline \multicolumn{2}{c}{xr(x)} \end{array}$$

The last term canceled.

So, $xs(x) - s_{n-k-1}g(x)$ is the syndrome of $xr(x)$. □

5.3 Burst Error Correcting

- "cyclic codes are good for (cyclic) burst error correcting."
- $C : (n,k,d)$-binary code, say $e = \left\lfloor \frac{d-1}{2} \right\rfloor = 5$



cyclic burst length of e

Let $e \in V_n(F)$. The *cyclic burst length of* e is the length of the smallest cyclic block that contain all the nonzero entries of e.

Example:

$e = \quad 0\;1\;0\;0\;0\;0\;0\;1$

has cyclic burst length 4. Say e is a cyclic burst error of length t if its cyclic burst length is t.

t-cyclic burst error correcting code

A linear code C is a t-cyclic burst error correcting code if every cyclic burst error of length at most t lies unique coset of C. The largest such t is called the cyclic burst error capability of C

Example:

$g(x) = 1 + x + x^2 + x^3 + x^6$ generates a $(15,9)$-binary cyclic code C that is a 3-cyclic burst error correcting code.

Note

$d(C) \leq 5$, so $e \leq 2$. Verify by checking that each cyclic burst error of length ≤ 3 has a unique syndrome.

Cyclic burst errors	Syndromes
0	000000
x^0	100000
x^1	010000
x^2	001000
$\vdots$	
x^5	000001
x^6	111100
x^7	011110
x^8	001111
x^9	111011
$\vdots$	
x^{14}	111001
$1+x$	110000
$x(1+x)$	011000
$\vdots$	
$x^{14}(1+x)$	011001
$1+x+x^2$	111000
$x(1+x+x^2)$	01100
$\vdots$	
$x^{14}(1+x+x^2)$	001001
$1+x^2$	101000
$x(1+x^2)$	010100
$\vdots$	
$x^{14}(1+x^2)$	101001

All syndromes are unique.

# of cyclic bursts of length ≤ 3 is 61. # of syndromes is 64.

Example:

$g(x) = 1 + x^4 + x^6 + x^7 + x^8$ generates a $(15, 7)$-binary cyclic code that is 4-cyclic burst error correcting. Distance ≤ 5, so $e \leq 2$.

Q How to construct codes with high cyclic burst error, correcting capability?

A 1. Use computer search. 2. RS codes. 3. Interleaving.

Theorem 5.11

Let C be an (n, k, d)-code over $\mathrm{GF}(q)$.

Let t be its cyclic burst error correcting capability.

$$\left\lfloor \frac{d-1}{2} \right\rfloor \leq t \leq n-k$$

Proof:
Every cyclic burst of length $\leq t$ has weight $\leq t$. Since every vector of weight $\leq \left\lfloor \frac{d-1}{2} \right\rfloor$ has a unique syndrome. We have $\left\lfloor \frac{d-1}{2} \right\rfloor \leq t$.

To prove the upper bound. Note that the # of cyclic burst errors where all the nonzero entries lie in the first t coordinate positions is q^t. Each has unique coset, and the total # of cosets is q^{n-k}. So $q^t \leq q^{n-k}$, $t \leq n-k$. □

Exercise

Prove that $t \leq \dfrac{n-k}{2}$.

5.4 Decoding cyclic burst errors

Let C be a t-cyclic burst e.c.c.[1] generated by $g(x)$ which is a degree-k monic divisor of $x^n - 1$ over $\mathrm{GF}(q)$.

Recall A PCM for C is $H = [I_{n-k} | -R^T]$ whose columns are x^0 [mod $g(x)$], $\ldots$, x^{n-1} [mod $g(x)$]. The syndrome of $r(x)$ is $s(x) = r(x)$ [mod $g(x)$].

Idea Suppose e is a cyclic burst of length $\leq t$.

Compute $s = Hr^T = r(x)$ [mod $g(x)$]

$e = \boxed{\text{x o —— o x x x}}$

$xe = \boxed{\text{x x o —— o x x}}$

$x^2e = \boxed{\text{x x x o —— o x}}$

$x^3e = \boxed{\text{x x x x o —— o}}$

$$\begin{aligned} s &= Hr^T = He^T \\ s_1 &= H(xr)^T = H(xe)^T \\ s_2 &= H(x^2r)^T = H(x^2e)^T \\ s_3 &= H(x^3r)^T = H(x^3e)^T \end{aligned}$$

[1] error correcting code

Idea (cont'd) Suppose e is a cyclic burst of length $\leq t$. Compute shifts of e, say $e_i = x^i$, has all its nonzero entries if the first $n-k$ positions. Then $s_i(x) = e_i(x)$ $[\bmod g(x)]$, and we can recognize such an $s_i(x)$ since it is a (non-cyclic) burst of length $\leq t$. Then $e = x^{n-i}e_i$.

How to compute $s(x)$? Recall $r = c + e$.

So $x^i r = x^i c + x^i e$, so $x^i r$ and $x^i e$ have the same syndrome.

5.5 Error trapping decoding

- Let $r(x)$ = received poly
- Let $s_i(x)$ = syndrome of $x^i r(x)$, $0 \leq i \leq n-1$

  So $s_0(x) = r(x)$ $[\bmod g(x)]$

Algorithm 4: Error trapping decoding

```
for i ← 0 to n − 1 do
    Compute s_i(x)
    if s_i(x) is a (non-cyclic) burst of length ≤ t then
        Let e_i(x) = (s_i(x), 0)
        Let e(x) = x^{n−i} e_i(x)
        Decode r(x) to r(x) − e(x)
    end
end
Reject r
```

Example:

$g(x) = 1 + x + x^2 + x^3 + x^6$ is the g.p. for $(15, 9)$-binary cyclic code with c.b.e.c capability 3. Decode $r = (1110\ \ 1110\ \ 1100\ \ 000)$

Solution Compute $s_0(x) = r(x)$ $[\bmod g(x)] = x^5 + x^4 + x + 1$

i	$s_i(x)$
0	110011
1	100101
2	101110
3	010111
4	110111
5	100111
6	101111
7	101011
8	101001
9	$\underbrace{101}_{\substack{\text{non-cyclic} \\ \text{burst of} \\ \text{length } \leq 3}}\ 000$

So $e_9 = (101000\ 000000000)$

So $e = x^6 e_9 = (000000101000000)$

So $c = r - e = (1110\ 1100\ 0100\ 000)$

Check $Hc^T = 0$, or $g(x)|c(x)$

5.6 Interleaving

Goal Improve c.b.e.c capability of a code

Suppose C is an (n, k)-code with c.b.e.c capability t.

Suppose the following codewords are transmitted.

$$\begin{aligned} v_1 &= (v_{11}, v_{12}, \ldots, v_{1n}) \in C \\ v_2 &= (v_{21}, v_{22}, \ldots, v_{2n}) \in C \\ &\vdots \\ v_s &= (v_{s1}, v_{s2}, \ldots, v_{sn}) \in C \end{aligned}$$

Suppose $v_1, v_2, \ldots, v_s$ are transmitted in that order. If a cyclic error of length $\leq t$ occurs in any codeword, that error can be corrected.

Instead we transmit the columns in order:

$$[v_{11}, v_{21}, \ldots, v_{s1}, v_{12}, v_{22}, \ldots, v_{s2}, \ldots, v_{1n}, \ldots, v_{sn}]$$

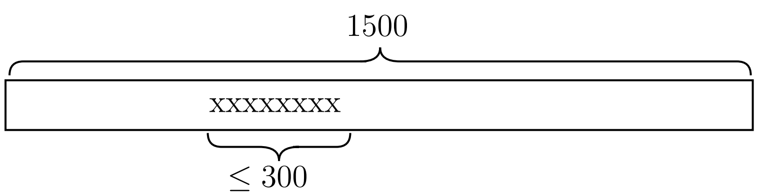
Now, if a cyclic burst error of length $\leq st$ occurs in this (flat) codeword, this means each original codeword suffered a cyclic burst error of length $\leq t$.

Theorem 5.12

Suppose C is an (n, k)-cyclic code with g.p. $g(x)$ and c.b.e.c. capability t. Then C^*, the code obtained by interleaving C to a depth s_1 is an (ns, ks)-cyclic code with g.p. $g^*(x) = g(x^s)$.

Example:

Interleave C to depth $s = 100$, to get code C^*. Then C^* is a $(1500, 900)$, 300-c.b.e.c.c, binary cyclic code with g.p. is $g(x^{100}) = 1 + x^{100} + x^{200} + x^{300} + x^{600}$,

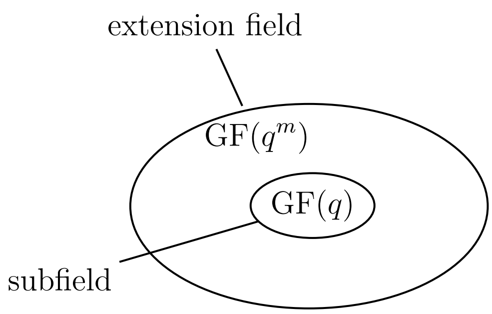


6 BCH codes

6.1 Minimal Polynomials

Recall We can view $F = \mathrm{GF}(p^m)$ as a vector space of dim m over $\mathbb{Z}_p$ and $\mathbb{Z}_p$ is a **subfield** of F.

More generally, for any prime power q, we can view the finite field $\mathrm{GF}(q^m)$ as a vector space of dim m over $\mathrm{GF}(q)$, and $\mathrm{GF}(q)$ is a subfield of $\mathrm{GF}(q^m)$.



Example:

$\mathrm{GF}(2^{16})$ is vector space of dim 16 over $\mathrm{GF}(2)$.

$\mathrm{GF}(2^{16})$ is vector space of dim 8 over $\mathrm{GF}(2^2)$.

$\mathrm{GF}(2^{16})$ is vector space of dim 4 over $\mathrm{GF}(2^4)$.

$\mathrm{GF}(2^{16})$ is vector space of dim 2 over $\mathrm{GF}(2^8)$.

$\mathrm{GF}(2^{16})$ is vector space of dim 1 over $\mathrm{GF}(2^{16})$.

minimal polynomial

Let $\alpha \in \mathrm{GF}(q^m)$. The minimal polynomial of α over $\mathrm{GF}(q)$, denoted $m_\alpha(x)$, is the monic polynomial of least degree in $\mathrm{GF}(q)[x]$ such that $m(\alpha) = 0$.

Example:

Let $0 \in \mathrm{GF}(q^m)$. Then $m_0(x) = x$.

Example:

Consider $\mathrm{GF}(2^2) = \mathbb{Z}_2[x]/(x^2+x+1) = \{0, 1, x, x+1\}$. Find min poly of $\alpha \in \mathrm{GF}(2^2)$ over $\mathrm{GF}(2)$:

$$\begin{aligned} m_0(y) &= y \\ m_1(y) &= y+1 \\ m_x(y) &= y^2+y+1 \\ m_{x+1}(y) &= y^2+y+1 \end{aligned}$$

Q Why does $m_\alpha(x)$ exist? ($\alpha \neq 0$)

The $\mathrm{ord}(\alpha) = t | (q^m - 1)$. So $\alpha^t = 1$. So, α is a root of $x^t - 1 \in \mathrm{GF}(q)[x]$. Also, if $f(x) \in \mathrm{GF}(q)[x]$ with $f(\alpha) = 0$, then if $c \in \mathrm{GF}(q)$ is the leading coeff of f, then $f' = c^{-1}f \in c^{-1}f \in \mathrm{GF}(q)[x]$ and $f'(\alpha) = 0$.

$\therefore m_\alpha(x)$ exists.

Theorem 6.1: Properties of $m_\alpha(x)$

Let $\alpha \in \mathrm{GF}(q^m)$, and $m_\alpha(x)$ ⓐ minimal poly of α over $\mathrm{GF}(q)$,
the

(1) $m_\alpha(x)$ is unique.

(2) $m_\alpha(x)$ is irreducible over $\mathrm{GF}(q)$.

(3) $\deg(m_\alpha) \leq m$

Proof:

(1) Suppose not. Let $s(x), t(x) \in \mathrm{GF}(q)[x]$ be two monic polys of (the same) least degree having α as a root. Consider $r(x) = s(x) - t(x) \in \mathrm{GF}(q)[x]$. Then $r(\alpha) = s(\alpha) - t(\alpha) = 0$ and $\deg(r) < \deg(s)$. So we must have $r(x) = 0$. Hence $s(x) = t(x)$.

(2) Suppose not. Let $m_\alpha(x) = s(x)t(x)$, where $s, t \in \mathrm{GF}(q)[x]$, and $\deg(s), \deg(t) < \deg(m_\alpha)$. So $m_\alpha(\alpha) = s(\alpha)t(\alpha) = 0$, so $s(\alpha) = 0$ or $t(\alpha) = 0$. This contradicts minimality of $\deg(m_\alpha)$.

(3) Recall $\mathrm{GF}(q^m)$ is a vector space of dim m over $\mathrm{GF}(q)$. So, $1, \alpha, \alpha^2, \alpha^3, \ldots, \alpha^m$ are linearly dependent over $\mathrm{GF}(q)$. So we can write

$$a_0 + a_1\alpha + \ldots + a_m\alpha^m = 0, \qquad \text{where } a_i \in \mathrm{GF}(q), \text{ not all } 0$$

So, α is a root of $f(x) = a_0 + a_1x + \ldots + a_mx^m$ which is nonzero and $\deg(f) \leq m$.

□

Index

A

alphabet, word, length... 5

C

characteristic 15
coset leader 38
cyclic burst length of e 56
cyclic code 45
cyclic subspace 45

D

dual code 28

E

e-error correcting code 10
equivalent 27
error vector 34

F

F[x] 17
field 13

G

generator 23
generator matrix 26
GF(q) 20
GF(q)* 22

H

Hamming code of order r over GF(q) 33
Hamming distance 7

I

ideal 46
ideal generated by $g(x)$ 47
infinite, finite, order 13
inner product 28
irreducible 18

L

linear (n, k)-code over F 24

M

minimal polynomial 61

O

ord(alpha) 22
orthogonal 28

P

parity-check matrix 30
perfect code 32
principal ideal & principal ideal ring 47

R

rate 7
reciprocal of h 52

S

subspace 24
syndrome 35
systematic, standard form 26

T

t-cyclic burst error correcting code 56
the generator polynomial of I 48

W

weight 25