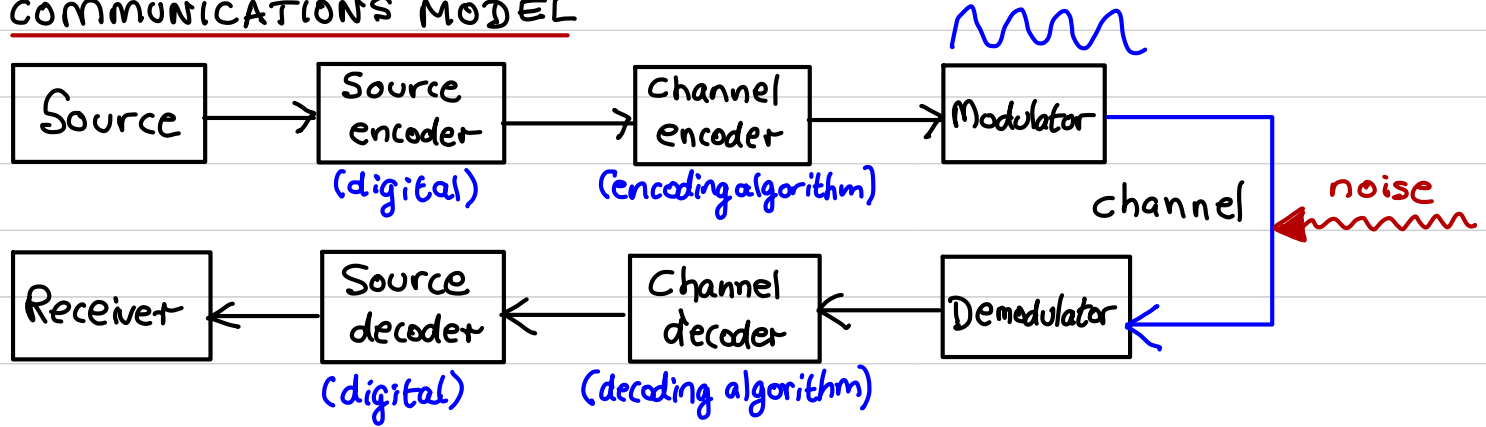


VO

INTRODUCTION

- Coding theory is about clever ways of adding redundancy to messages to allow for (efficient) error detection and correction.
- COMMUNICATIONS MODEL



- EXAMPLE (parity code)

Encoding algorithm: Add a parity bit to the (binary) message.

Decoding algorithm: If the number of 1's in the received message t is even, then accept t ; else reject t .

EXAMPLE (binary replication code)

Source messages	Codeword	# errors/codeword that can always be detected	# errors/codeword that can always be corrected *	Information rate
0 1	0 1	0	0	1
0 1	00 11	1	0	$\frac{1}{2}$
0 1	000 111	2	1	$\frac{1}{3}$
0 1	0000 1111	3	1	$\frac{1}{4}$
0 1	00000 11111	4	2	$\frac{1}{5}$

• →
encoding algorithm

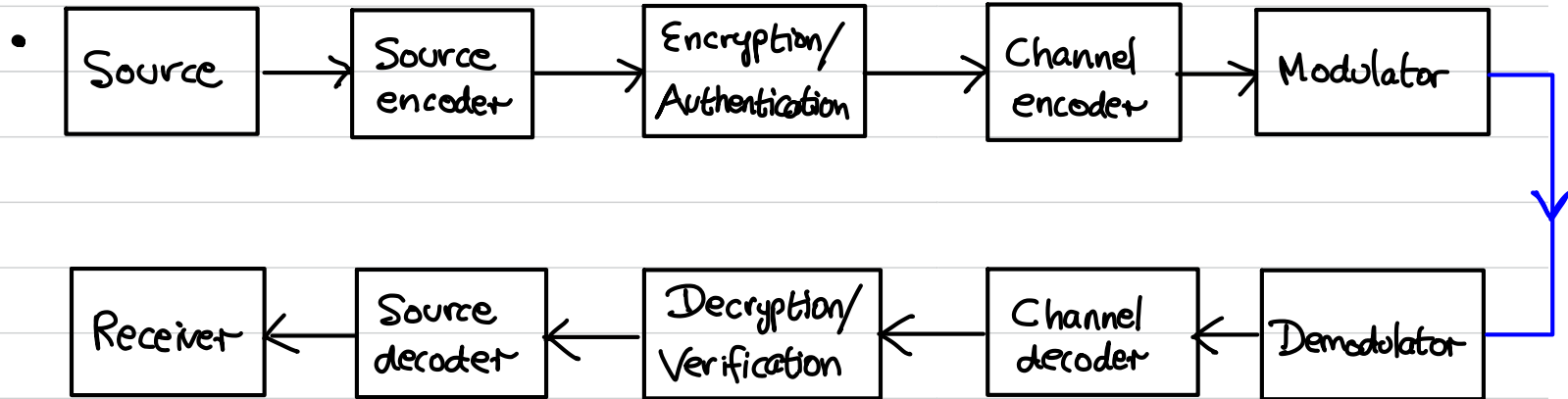
* using "nearest neighbour decoding"

GOALS OF CODING THEORY

- Design codes with (i) High error-correcting capability.
(ii) High information rate.
(iii) Efficient encoding and decoding algorithms.
- "Efficient" = efficient in hardware and software.
- COURSE OVERVIEW This course deals with algebraic methods for designing good (block) codes. These codes are used for reliable communications and storage: space probes, CD/DVD/BlueRay, wireless, disk storage, QR codes, etc.
- NOT COVERED LDPC codes, Turbo codes, Raptor codes, Polar codes,
And newer applications: Distributed storage, compressed sensing, group testing, Learning theory, algorithm design, complexity theory,

THE BIG PICTURE

- Coding theory in its broadest sense deals with techniques for the efficient, secure and reliable transmission of data over communications channels that might be subject to non-malicious errors (noise) and adversarial intrusion. The latter includes passive intrusion (eavesdropping) and active intrusion (injection/deletion/modification).



COURSE OUTLINE

-5-

LEARNING OUTCOMES

- 1) Demonstrate a fundamental understanding of the binary symmetric channel, decoding strategies, and the challenges in designing good codes.
- 2) Construct codes, and devise efficient encoding and decoding algorithms for them as a means of gaining exposure to the applications of linear algebra and abstract algebra.
- 3) Analyze the properties of major families of algebraic codes including linear codes, Hamming codes, Golay codes, cyclic codes, BCH codes, and Reed-Solomon codes.

COURSE PREREQUISITES

- Linear Algebra: vector spaces, linear independence, basis, Gaussian elimination, null space of a matrix, rank of a matrix.
- Elementary Number Theory: congruences, integers modulo n ($\mathbb{Z}_n$), primes, Fermat's Little Theorem.

COURSE INGREDIENTS

- Linear Algebra
- Abstract Algebra
(groups, rings, fields)

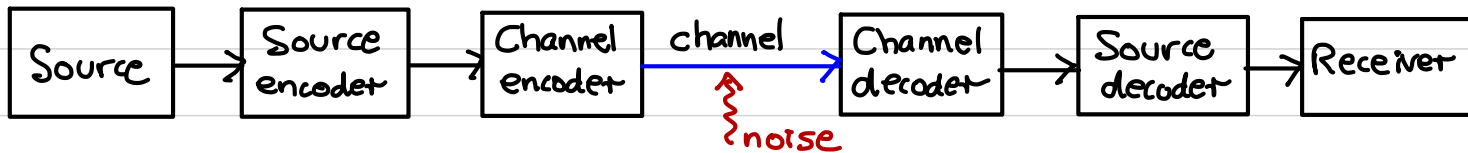
EXERCISES

- Problem sets for each chapter (with solutions).

REFERENCES

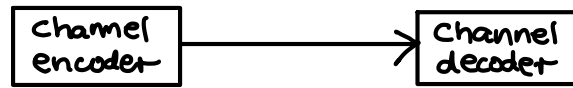
- "An Introduction to Error Correcting Codes",
Paul van Oorschot & Scott Vanstone (1989).
- "Modern Coding Theory",
Tom Richardson & Rüdiger Urbanke (2008)
- "Essential Coding Theory",
Venkatesan Guruswami, Atri Rudra & Madhu Sudan (2023)
- "Algebraic Coding Theory",
Mary Wootters (2021)

V1a BASIC DEFINITIONS AND CONCEPTS



DEFINITIONS

- An alphabet A is a finite set of $q \geq 2$ symbols.
- A word is a finite sequence of symbols from A (also: vector, tuple).
- The length of a word is the number of symbols it has.
- A code C over A is a finite set of words (with $|C| \geq 2$).
- A codeword is a word in the code C .
- A block code is a code in which all codewords have the same length.
- A block code of length n containing M codewords over A is a subset $C \subseteq A^n$ with $|C| = M$. C is called an $[n, M]$ -code over A .



-9-

- EXAMPLE

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

<u>Source messages</u>	<u>Codewords</u>
------------------------	------------------

00	00000
----	-------

01	11100
----	-------

10	00111
----	-------

11	10101
----	-------

- The channel encoder only transmits codewords.
However, what is received may not be a codeword.

- EXAMPLE

$r = 11000$ is received.

What should the channel decoder do?

ASSUMPTIONS ABOUT THE CHANNEL

(1) Only symbols from A are transmitted ("hard decision coding").

(2) No symbols are lost/added/interchanged during transmission.

(3) The channel is a q -symmetric channel:

- Let $A = \{a_1, a_2, \dots, a_q\}$.

- Let $X_i = i^{\text{th}}$ symbol sent.

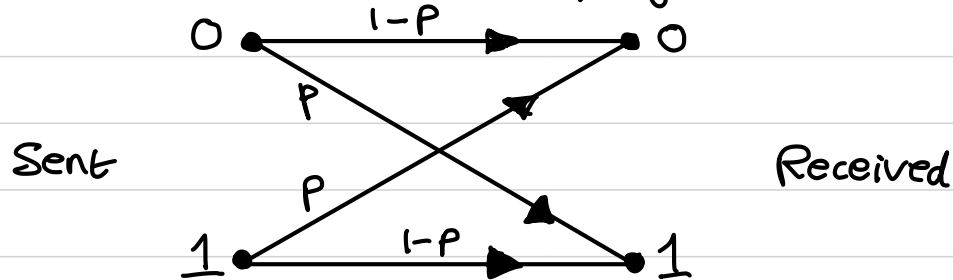
- Let $Y_i = i^{\text{th}}$ symbol received.

- For all $i \geq 1$, $1 \leq j, k \leq q$, $P_r(Y_i = a_k | X_i = a_j) = \begin{cases} 1-p, & \text{if } j=k, \\ \frac{p}{q-1}, & \text{if } j \neq k. \end{cases}$

p is called the symbol error probability of the channel ($0 \leq p \leq 1$).

BINARY SYMMETRIC CHANNEL (BSC)

- A 2-symmetric channel is called a binary symmetric channel.



- For a BSC:

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

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

(3) If $1/2 < p < 1$, then flipping all received bits converts the channel to a BSC with $0 < p < 1/2$.

(4) Henceforth, we will assume that $0 < p < 1/2$ for a BSC.

EXERCISE

- For a q -symmetric channel, show that one can take

$$0 < p < \frac{q-1}{q}$$

without loss of generality.

Hint: First consider the case $q=3$.

- In the remainder of the course, we shall assume that

$$0 < p < \frac{q-1}{q}$$

INFORMATION RATE

DEFINITION The information rate (or rate) R of an $[n, M]$ -block code C over an alphabet A of size q is $R = (\log_q M)/n$.

• If C encodes messages that are the k -tuples over A (so $|C| = q^k$), then $R = k/n$.

• Note: $0 < R \leq 1$. Ideally, R should be close to 1.

• EXAMPLE The rate of the binary code $C = \{00000, 11100, 00111, 10101\}$ is $R = 2/5$.

HAMMING DISTANCE

-14-

DEFINITION The Hamming distance (or distance) between two n -tuples in A^n is the number of coordinate positions in which they differ.

THEOREM (properties of Hamming distance)

For all $x, y, z \in A^n$: (1) $d(x, y) \geq 0$, with $d(x, y) = 0$ iff $x = y$.

(2) $d(x, y) = d(y, x)$.

(3) $d(x, y) + d(y, z) \geq d(x, z)$ (Δ inequality)

DEFINITION The Hamming distance (or distance) of an $[n, M]$ -code C is $d(C) = \min \{d(x, y) : x, y \in C, x \neq y\}$.

EXAMPLE The distance of $C = \{00000, 11100, 00111, 10101\}$ is $d(C) = 2$.

V1b DECODING STRATEGIES

-15-

EXAMPLE Let $C = \{00000, 11100, 00111, 10101\}$.

C is a $[5,4]$ -code over $\{0,1\}$ (a binary code) with $R = 2/5$ and $d(C) = 2$.

ERROR DETECTION If C is used for error detection only, the strategy is the following: A received word $t \in A^n$ is accepted iff $t \in C$.

ERROR CORRECTION Let C be an $[n,M]$ -code over A with distance d .

Suppose that $c \in C$ is transmitted and $t \in A^n$ is received.

The (channel) decoder must decide one of the following:

- (i) No errors have occurred; accept t .
- (ii) Errors have occurred; correct (or decode) t to a codeword $c \in C$.
- (iii) Errors have occurred; correction is not possible.

NEAREST NEIGHBOUR DECODING

INCOMPLETE MAXIMUM LIKELIHOOD DECODING (IMLD)

If there is a unique codeword $c \in C$ such that $d(r, c)$ is a minimum, then decode r to c . If no such c exists, then reject r (ask for retransmission or disregard the information).

COMPLETE MAXIMUM LIKELIHOOD DECODING (CMLD)

Same as IMLD, except that if there are two or more $c \in C$ for which $d(r, c)$ is minimum, decode r to an arbitrary one of these.

IS IMLD A REASONABLE STRATEGY?

THEOREM IMLD chooses the codeword $c \in C$ for which the conditional probability $P(r|c) = P(r \text{ is received} | c \text{ is sent})$ is largest.

PROOF Suppose $c_1, c_2 \in C$ with $d_1 = d(c_1, r)$, $d_2 = d(c_2, r)$. Suppose that $d_1 > d_2$.
Now, $P(r|c_1) = (1-p)^{n-d_1} \left(\frac{p}{q-1}\right)^{d_1}$ and $P(r|c_2) = (1-p)^{n-d_2} \left(\frac{p}{q-1}\right)^{d_2}$.

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

$$\text{Now, } \frac{p}{(1-p)(q-1)} < 1 \Leftrightarrow p < (1-p)(q-1) \Leftrightarrow p < q - pq^{-1} + p \Leftrightarrow pq < q-1 \Leftrightarrow p < \frac{q-1}{q}.$$

Thus, $\frac{P(r|c_1)}{P(r|c_2)} < 1$, so $P(r|c_1) < P(r|c_2)$, and the result follows. $\square$

MINIMUM ERROR PROBABILITY DECODING (MED)

- An ideal strategy would be to decode r to a codeword $c \in C$ for which $P(c|r) = P(c \text{ is sent} | r \text{ is received})$ is largest. This is MED.

- EXAMPLE (IMLD/CMLD is not the same as MED)

Consider $C = \{c_1, c_2\}$, and suppose that $P(c_1) = 0.1$ and $P(c_2) = 0.9$.

Suppose $p = 1/4$ (for a BSC).

Suppose $r = 100$ is the received word.

$$\text{Then } P(c_1|r) = P(r|c_1) \cdot P(c_1) / P(r) = p(1-p)^2 \times 0.1 / P(r) = \frac{9}{640} \cdot \frac{1}{P(r)}.$$

$$\text{And } P(c_2|r) = P(r|c_2) \cdot P(c_2) / P(r) = p^2(1-p) \times 0.9 / P(r) = \frac{27}{640} \cdot \frac{1}{P(r)}.$$

So, MED decodes r to c_2 , whereas IMLD decodes r to c_1 .

BAYES THEOREM: $P(A \cap B) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B).$

IMLD/CMLD vs. MED

- IMLD maximizes $P(r|c)$.
- MED maximizes $P(c|r)$.

(i) MED has the drawback that the decoding algorithm depends on the probability distribution of source messages.

(ii) If all source messages are equally likely, then CMLD and MED are identical:
$$P(c_i|r) = \underbrace{P(r|c_i)}_{P(r)} \cdot \underbrace{P(c_i)}_{\substack{M \cdot P(r) \\ \text{does not depend on } i}} = P(r|c_i) \cdot \frac{1}{M \cdot P(r)}.$$

(iii) In practice, IMLD/CMLD is used.

In this course, we will use IMLD/CMLD.

V1C ERROR CORRECTING AND DETECTING CAPABILITIES

DETECTION ONLY

STRATEGY If r is received, then accept r iff $r \in C$.

DEFINITION A code C is an e -error detecting code if the decoder always makes the correct decision if e or fewer errors per codeword are introduced by the channel.

EXAMPLE Consider $C = \{000, 111\}$.

- C is a 2-error detecting code.
- C is not a 3-error detecting code.

THEOREM A code C of distance d is a $(d-1)$ -error detecting code, but is not a d -error detecting code.

PROOF Suppose $c \in C$ is sent.

- If no errors occur, then c is received and is accepted.
- Suppose that the number of errors introduced is ≥ 1 and $\leq d-1$; let t be the received word. Then $1 \leq d(t, c) \leq d-1$, so $t \notin C$. Thus, t is rejected.
- Since $d(C) = d$, there exist $c_1, c_2 \in C$ with $d(c_1, c_2) = d$.

If c_1 is sent and c_2 is received, the decoder accepts c_2 ; the d errors go undetected. $\square$

CORRECTION

STRATEGY : IMLD / CMLD

DEFINITION A code C is an e -error correcting code if the decoder always makes the correct decision if e or fewer errors per codeword are introduced by the channel.

EXAMPLE Consider $C = \{000, 111\}$.

- C is a 1-error correcting code.
- C is not a 2-error correcting code.

THEOREM A code C of distance d is an e -error correcting code, where $e = \lfloor \frac{d-1}{2} \rfloor$. ($\lfloor x \rfloor$ is the largest integer $\leq x$.)

PROOF Suppose that $c \in C$ is sent, at most $(d-1)/2$ errors are introduced, and r is received. Then $d(r, c) \leq (d-1)/2$. On the other hand, if c' is any other codeword, then

$$\begin{aligned} d(r, c') &\geq d(c, c') - d(r, c) \quad (\text{by } \Delta \text{ inequality}) \\ &\geq d - (d-1)/2 \\ &= (d+1)/2 \\ &> (d-1)/2 \geq d(r, c). \end{aligned}$$

Hence, c is the unique codeword at minimum distance from r , so the IMLD/CMLD decoder correctly concludes that c was sent. $\square$

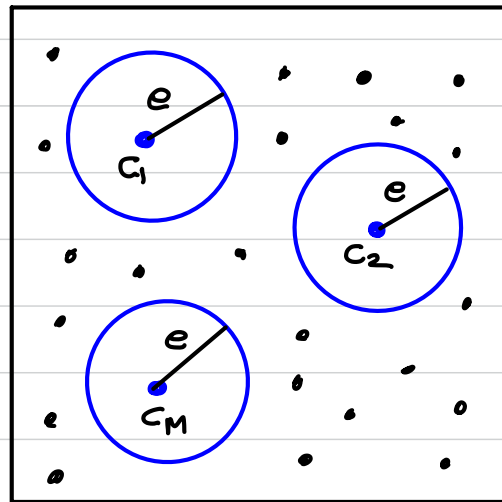
EXERCISE Suppose $d(C)=d$, and let $e = \lfloor \frac{d-1}{2} \rfloor$. Show that C is not an $(e+1)$ -error correcting code.

SPHERE PACKING A natural question to ask is: Given A, n, M, d , does there exist an $[n, M]$ -code C over A of distance $\geq d$?

This question can be phrased as an equivalent sphere packing problem:

Can we place M spheres of radius $e = \lfloor \frac{d-1}{2} \rfloor$ in A^n , so that no two spheres overlap?

A^n



$$C = \{c_1, c_2, \dots, c_M\}, \quad e = \lfloor \frac{d-1}{2} \rfloor.$$

S_c = sphere of radius e centered at $c \in C$
 = all words within distance e of c .

We proved: If $c_1, c_2 \in C, c_1 \neq c_2$, then $S_{c_1} \cap S_{c_2} = \emptyset$.

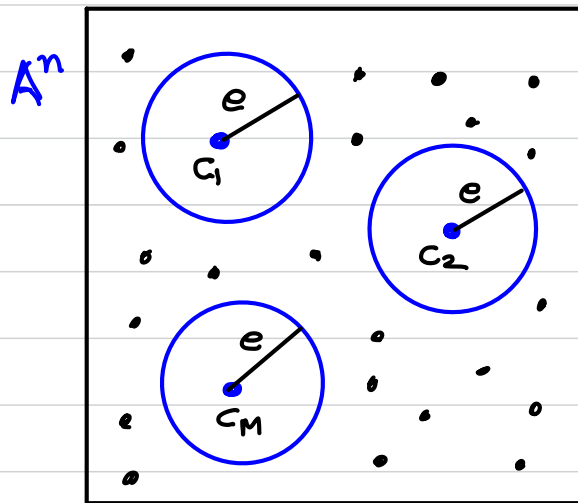
- QUESTION Let $q=2$, $n=127$, $M=2^{64}$.

Does there exist an $[n, M]$ -binary code with $d \geq 21$?

If so, can encoding and decoding be done efficiently?

- We will construct such a code in V6.

The main tools used will be linear algebra (over finite fields) and abstract algebra (rings and fields).



V2a INTRODUCTION TO FINITE FIELDS

-26-

DEFINITION A commutative ring $(R, +, \cdot)$ consists of a set R , together with two operations $+: R \times R \rightarrow R$ and $\cdot: R \times R \rightarrow R$, such that:

(i) $a + (b + c) = (a + b) + c \quad \forall a, b, c \in R.$

(ii) $a + b = b + a \quad \forall a, b \in R.$

(iii) $\exists 0 \in R$ with $0 + a = a \quad \forall a \in R.$

(iv) For each $a \in R$, $\exists -a \in R$,
such that $a + (-a) = 0.$

(v) $a \cdot (b \cdot c) = (a \cdot b) \cdot c \quad \forall a, b, c \in R.$

(vi) $a \cdot b = b \cdot a \quad \forall a, b \in R.$

(vii) $\exists 1 \in R$, $1 \neq 0$, such that
 $a \cdot 1 = a \quad \forall a \in R.$

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

• NOTATION We will denote $(R, +, \cdot)$ by R .

• EXAMPLE $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ are commutative rings.

DEFINITION A field $(F, +, \cdot)$ is a commutative ring with an additional property (ix): $\forall a \in F$ with $a \neq 0$, $\exists a^{-1} \in F$ such that $a \cdot a^{-1} = 1$.

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

DEFINITION A field $(F, +, \cdot)$ is a finite field if F is a finite set; otherwise it is an infinite field. If F is a finite field, its order is $|F|$.

- **EXAMPLE** $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ are infinite fields.

- **QUESTIONS** 1) For which integers $n \geq 2$ do finite fields of order n exist?
2) How does one construct such a field, i.e. what are the field elements, and how are field operations performed?

THE INTEGERS MODULO n

- Let $n \geq 2$. Recall that $\mathbb{Z}_n$ consists of the set of equivalence classes of integers modulo n , $\mathbb{Z}_n = \{[0], [1], [2], \dots, [n-1]\}$, with addition and multiplication done in the natural way: $[a] + [b] = [a+b]$, $[a] \cdot [b] = [a \cdot b]$.
- More simply, we write $\mathbb{Z}_n = \{0, 1, 2, \dots, n-1\}$, and perform addition and multiplication modulo n .
- EXAMPLE $\mathbb{Z}_9 = \{0, 1, 2, \dots, 8\}$. In $\mathbb{Z}_9$, $3+7=1$ and $3 \cdot 7=3$.
More precisely, $3+7 \equiv 1 \pmod{9}$ and $3 \cdot 7 \equiv 3 \pmod{9}$.
- FACT $\mathbb{Z}_n$ is a finite commutative ring.

QUESTION: When is $\mathbb{Z}_n$ a field?

THEOREM $\mathbb{Z}_n$ is a field iff n is prime.

PROOF ($\Leftarrow$) Suppose that n is prime. Let $a \in \mathbb{Z}_n, a \neq 0$ (so $1 \leq a \leq n-1$). Since n is prime, $\gcd(a, n) = 1$. Hence $\exists s, t \in \mathbb{Z}$ such that $as + nt = 1$. Reducing both sides modulo n gives $as \equiv 1 \pmod{n}$. Thus, $a^{-1} = s$, and so $\mathbb{Z}_n$ is a field.

($\Rightarrow$) Suppose that n is composite, say $n = ab$ where $2 \leq a, b \leq n-1$. Now, if a^{-1} exists, say $ac \equiv 1 \pmod{n}$, then $abc \equiv b \pmod{n}$, so $nc \equiv b \pmod{n}$. Thus, $b \equiv 0 \pmod{n}$, so $n \mid b$ which is absurd since $2 \leq b \leq n-1$. Hence, $\mathbb{Z}_n$ is not a field, $\square$

QUESTIONS

- We have established the existence of finite fields of order n , for each prime n .
- What about finite fields of order n , where n is composite?
- In particular, is there a finite field of order 4? 6? 8? 9? 10?

Vab NON-EXISTENCE OF FINITE FIELDS

DEFINITION Let F be a field. The characteristic of F , $\text{char}(F)$, is the smallest positive integer m such that $\underbrace{1+1+\dots+1}_m = 0$.
If no such m exists, then $\text{char}(F) = 0$.

EXAMPLE $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ have characteristic 0.

$\mathbb{Z}_p$ (p prime) has characteristic p .

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

PROOF The field elements $1, 1+1, 1+1+1, \dots$ are distinct, because if $\underbrace{1+1+\dots+1}_a = \underbrace{1+1+\dots+1}_b$ where $a < b$, then $\underbrace{(1+1+\dots+1)}_b - \underbrace{(1+1+\dots+1)}_a = \underbrace{1+1+\dots+1}_{b-a} = 0$, contradicting $\text{char}(F) = 0$. $\square$

THEOREM Let F be a field with $\text{char}(F) = m \neq 0$. Then m is prime.

PROOF Suppose m is composite, say $m = ab$ where $2 \leq a, b \leq m-1$.

Let $s = \underbrace{1+1+\dots+1}_a$ and $t = \underbrace{1+1+\dots+1}_b$, and note that $s \neq 0, t \neq 0$.

Then $s \cdot t = (\underbrace{1+1+\dots+1}_a) \cdot (\underbrace{1+1+\dots+1}_b) = \underbrace{1+1+\dots+1}_m = 0$.

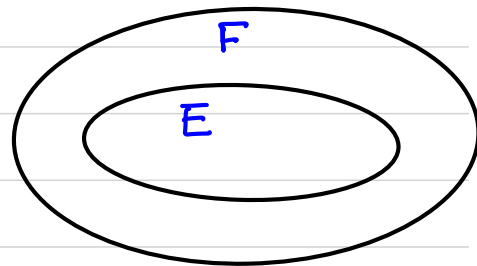
Thus, $s \cdot t \cdot t^{-1} = s \cdot 1 = s = 0$, a contradiction.

We conclude that m is prime. $\square$

SUBFIELDS

- Let F be a finite field of characteristic p .
- Consider the subset of elements of F :

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



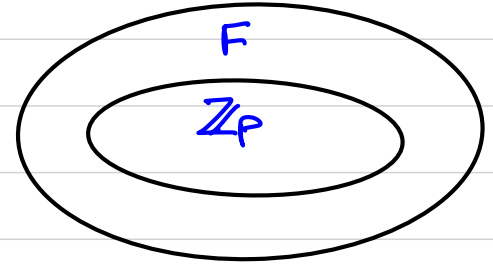
- The elements of E are distinct.
- One can verify that E is a field, using the same operations as F .
 E is said to be a subfield of F .
- If we identify the elements of E with the elements of $\mathbb{Z}_p$ in the natural way, then E is essentially the same field as $\mathbb{Z}_p$.
- We have proven:

THEOREM

Let F be a finite field of characteristic p . Then $\mathbb{Z}_p$ is a subfield of F .

FINITE FIELDS AS VECTOR SPACES

- Let F be a finite field of characteristic p .

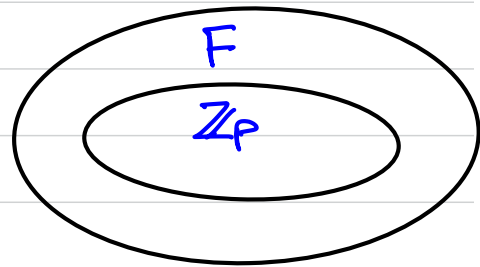


- Identify $\begin{array}{l} \text{vectors} \longleftrightarrow \text{elements of } F \\ \text{scalars} \longleftrightarrow \text{elements of } \mathbb{Z}_p \\ \text{vector addition} \longleftrightarrow \text{addition of } F \\ \text{scalar multiplication} \longleftrightarrow \text{multiplication of } F. \end{array}$

- Then F is a vector space over $\mathbb{Z}_p$
(i.e. the axioms that define a vector space are satisfied).

FINITE FIELDS: NON-EXISTENCE

THEOREM Let F be a finite field of characteristic p . Then the order of F is p^n , for some $n \geq 1$.



PROOF Let the dimension of F as a vector space over $\mathbb{Z}_p$ be n . Let $\alpha_1, \alpha_2, \dots, \alpha_n$ be a basis for F over $\mathbb{Z}_p$. Then each element $\beta \in F$ can be written uniquely in the form $\beta = c_1\alpha_1 + c_2\alpha_2 + \dots + c_n\alpha_n$, where $c_i \in \mathbb{Z}_p$. Thus, $F = \{c_1\alpha_1 + c_2\alpha_2 + \dots + c_n\alpha_n : c_i \in \mathbb{Z}_p\}$, so $|F| = p^n$. $\square$

EXAMPLE There do not exist finite fields of orders 6, 10, 12, 14, 15,

QUESTION Do finite fields of orders 4, 8, 9, 16, 25, 27, ... exist?

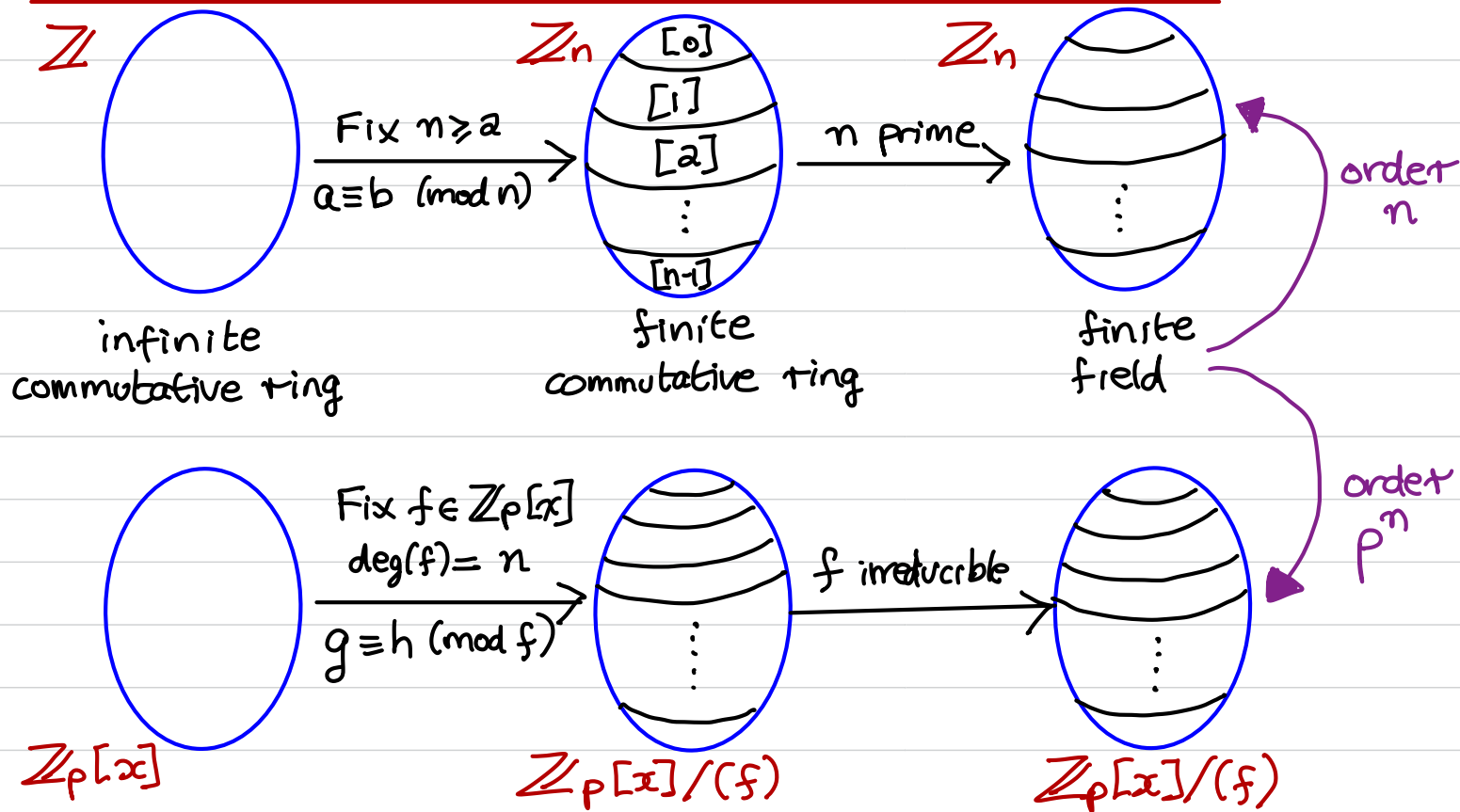
V2C EXISTENCE OF FINITE FIELDS

-36-

- POLYNOMIAL RINGS Let F be a field. Then $F[x]$ denotes the set of all polynomials in x with coefficients from F . Addition and multiplication of polynomials in $F[x]$ is done in the usual way, with coefficient arithmetic done in F .
- EXAMPLE In $\mathbb{Z}_5[x]$,
 - $(3x^4 + 2x^3 + x + 4) + (x^5 + 2x^4 + x^2 + 2x + 3) = x^5 + 2x^3 + x^2 + 3x + 2.$
 - $(3x^2 + 4x + 1) \cdot (2x^2 + x + 2) = x^4 + x^3 + 2x^2 + 4x + 2.$

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

CONSTRUCTION OF FINITE FIELDS: MAIN IDEA



POLYNOMIAL DIVISION

- Let $f, g \in F[x]$, with $g \neq 0$. Then there exist unique polynomials $r, s \in F[x]$ with $f = sg + r$ and $\deg(r) < \deg(g)$.

$\text{quotient} \nearrow$ $\nwarrow \text{remainder}$
- By convention, $\deg(0) = -\infty$.

• EXAMPLE Consider $f = 3x^4 + 2x^3 + 2x^2 + x + 1$, $g = 2x^2 + 3x + 4 \in \mathbb{Z}_5[x]$.

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

quotient
 $\searrow$

divisor
 $\searrow$

remainder
 $\searrow$

$$x^2 + x + 1$$

$$\underline{x^2 + 4x + 2}$$

$$2x + 4$$

So, $f = (4x^2 + 3)g + (2x + 4)$.

THE RING $F[x]/(f)$

DEFINITION Let $f \in F[x]$ with $\deg(f) \geq 1$. Let $g, h \in F[x]$. Then g is congruent to h modulo f , written $g \equiv h \pmod{f}$, if $g - h = lf$ for some $l \in F[x]$. Equivalently, $f \mid (g - h)$, or g, h leave the same remainder upon division by f .

FACTS The relation $g \equiv h \pmod{f}$ is an equivalence relation and partitions $F[x]$ into equivalence classes: $[g] = \{h \in F[x] : g \equiv h \pmod{f}\}$. Addition and multiplication is well defined: $[g] + [h] = [g + h]$, $[g] \cdot [h] = [g \cdot h]$.

DEFINITION The set of equivalence classes is denoted $F[x]/(f)$.

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

REPRESENTATIVES OF EQUIVALENCE CLASSES OF $F[x]/(f)$

- Suppose that $\deg(f) = n$.
- Let $g \in F[x]$. Long division of g by f yields $g = sf + r$, where $s, r \in F[x]$ with $\deg(r) < n$. Thus $g \equiv r \pmod{f}$, so $[g] = [r]$.
- Also, if $r_1, r_2 \in F[x]$, $r_1 \neq r_2$, $\deg(r_1) < n$, $\deg(r_2) < n$, then $f \nmid (r_1 - r_2)$, so $r_1 \not\equiv r_2 \pmod{f}$. Thus $[r_1] \neq [r_2]$.

• Hence, the polynomials in $F[x]$ of degree $< n$ are a complete set of representatives of the equivalence classes of $F[x]/(f)$.

• Now, let $F = \mathbb{Z}_p$. Then $\mathbb{Z}_p[x]/(f) = \{[r] : r \in \mathbb{Z}_p[x], \deg(r) < n\}$
 $= \{[r_0 + r_1x + \dots + r_{n-1}x^{n-1}] : r_i \in \mathbb{Z}_p\}$, so $|\mathbb{Z}_p[x]/(f)| = p^n$.

Hence, $\mathbb{Z}_p[x]/(f)$ is a finite commutative ring of order p^n .

- QUESTION When is $F[x]/(f)$ a field?

DEFINITION Let $f \in F[x]$ with $\deg(f) \geq 1$. Then f is irreducible over F if f cannot be written as $f = g \cdot h$, $g, h \in F[x]$, $\deg(g) \geq 1$, $\deg(h) \geq 1$.

- EXAMPLE • $x^2 + 1$ is irreducible over $\mathbb{R}$, since it has no roots in $\mathbb{R}$.
- $x^2 + 1$ is reducible over $\mathbb{C}$, since $x^2 + 1 = (x + i)(x - i)$.
- $x^2 + 1$ is reducible over $\mathbb{Z}_2$, since $x^2 + 1 = (x + 1)(x + 1)$.
- $x^2 + 1$ is irreducible over $\mathbb{Z}_3$, since it has no roots in $\mathbb{Z}_3$.

THEOREM $F[x]/(f)$ is a field iff f is irreducible over F .

PROOF Analogous to the proof that $\mathbb{Z}_n$ is a field iff n is prime (see slide 29).

V2d CONSTRUCTION OF FINITE FIELDS

-42-

FACTS ABOUT POLYNOMIAL RINGS $F[x]$ (F =field)

1) FACTOR THEOREM: Let $f \in F[x]$ and $c \in F$.

Then $x-c$ is a factor of $f(x)$ iff $f(c)=0$.

2) NOT TOO MANY ROOTS: Let $f \in F[x]$ with $\deg(f)=n \geq 0$.

Then f has at most n roots in F .

3) UNIQUE FACTORIZATION Let $f \in F[x]$ with $\deg(f)=n \geq 0$. Then, up to rearrangement of terms, f has a unique factorization over F :

$f = c f_1^{e_1} f_2^{e_2} \dots f_t^{e_t}$, where $c \in F$, $f_i \in F[x]$ with f_i monic and irreducible over F , and $e_i \geq 1$.

[A polynomial is monic if its leading coefficient is 1, eg. $f(x)=x^4-3x^2+19$]

THEOREM Let $f \in \mathbb{Z}_p[x]$ be an irreducible polynomial of degree $n \geq 1$. Then $\mathbb{Z}_p[x]/(f)$ is a finite field of order p^n and characteristic p . The field elements are the polynomials in $\mathbb{Z}_p[x]$ of degree $< n$.

EXAMPLE (finite field of order $4=2^2$) Here, $p=2$ and $n=2$.

Let $f(x) = x^2 + x + 1 \in \mathbb{Z}_2[x]$. Then $f(0) = 1$, $f(1) = 1$, so f has no roots in $\mathbb{Z}_2$. Thus f is irreducible over $\mathbb{Z}_2$.

So, $F = \mathbb{Z}_2[x]/(x^2 + x + 1)$ is a finite field of order $2^2 = 4$.

- The elements of F are $\{0, 1, x, x+1\}$. ($[\]$ is omitted.)
- Example of addition: $x + (x+1) = 2x+1 = 1$.
- Example of multiplication: $x \cdot (x+1) = x^2 + x = 1$.

EXAMPLE (finite field of order $8=2^3$) Here, $p=2$ and $n=3$.

We need an irreducible polynomial of degree 3 over $\mathbb{Z}_2$.

Candidates: $x^3, x^3+1, x^3+x, x^3+x+1, x^3+x^2, x^3+x^2+1, x^3+x^2+x, x^3+x^2+x+1$.

Try $f(x)=x^3+x+1$. Since $f(0)=1$ and $f(1)=1$, f has no roots in $\mathbb{Z}_2$, and thus no linear factors in $\mathbb{Z}_2[x]$. Thus, f is irreducible over $\mathbb{Z}_2$, and $F_1 = \mathbb{Z}_2[x]/(x^3+x+1)$ is a finite field of order $2^3=8$.

- The elements of F_1 are $\{0, 1, x, x+1, x^2, x^2+1, x^2+x, x^2+x+1\}$.
- Example of addition: $(x^2+x) + (x^2+x+1) = 1$.
- Example of multiplication: $(x^2+x) \cdot (x^2+x+1) = x^4+x = x^2$.
- Example of inversion: $x^{-1} = x^2+1$, since $x \cdot (x^2+1) = x^3+x = 1$.

EXAMPLE (finite field of order $8 = 2^3$) Here, $p=2$ and $n=3$.

$f(x) = x^3 + x^2 + 1$ is irreducible over $\mathbb{Z}_2$, and so

$F_2 = \mathbb{Z}_2[x]/(x^3 + x^2 + 1)$ is a finite field of order $2^3 = 8$.

• The elements of F_2 are $\{0, 1, x, x+1, x^2, x^2+1, x^2+x, x^2+x+1\}$.

• Note that F_1 and F_2 are not the same field.

eg. In F_1 , $x \cdot x^2 = x+1$, whereas $x \cdot x^2 = x^2+1$ in F_2 .

• However, F_1 and F_2 are isomorphic (essentially the same).

Formally, there is a bijection $\phi: F_1 \rightarrow F_2$ such that

$\phi(a+b) = \phi(a) + \phi(b)$ and $\phi(a \cdot b) = \phi(a) \cdot \phi(b) \quad \forall a, b \in F_1$.

EXISTENCE AND UNIQUENESS OF FINITE FIELDS

FACT Let p be prime and $n \geq 1$. Then $\exists$ irred. poly. of degree n over $\mathbb{Z}_p$.

THEOREM There exists a finite field of order q iff $q = p^n$ for some prime p and integer $n \geq 1$.

FACT Any two finite fields of the same order are isomorphic.

- We will denote the finite field of order q by $\text{GF}(q)$
"the Galois Field of order q ".
- In slides 44 and 45, we saw two ways of representing the finite field $\text{GF}(2^3)$.

V2e PROPERTIES OF FINITE FIELDS

-47-

THEOREM (Frosh's Dream) Let F be a finite field of characteristic p , and let $\alpha, \beta \in F$. Then $(\alpha + \beta)^{p^m} = \alpha^{p^m} + \beta^{p^m}$ for all $m \geq 1$.

PROOF ($m=1$) By the Binomial Theorem,

$$(\alpha + \beta)^p = \binom{p}{0} \beta^p + \sum_{i=1}^{p-1} \binom{p}{i} \alpha^i \beta^{p-i} + \binom{p}{p} \alpha^p.$$

Now, for $1 \leq i \leq p-1$, $\binom{p}{i} = \frac{p(p-1)(p-2)\dots(p-i+1)}{1 \cdot 2 \cdot 3 \dots i} \equiv 0 \pmod{p}$, since p divides the numerator but not the denominator, and $\binom{p}{i}$ is an integer.

$$\text{Thus, } \binom{p}{i} \alpha^i \beta^{p-i} = \underbrace{\alpha^i \beta^{p-i} + \dots + \alpha^i \beta^{p-i}}_{\binom{p}{i}} = \underbrace{(1+1+\dots+1)}_{\binom{p}{i}} \alpha^i \beta^{p-i} = 0.$$

$$\text{Hence, } (\alpha + \beta)^p = \alpha^p + \beta^p. \quad \square$$

THE MULTIPLICATIVE GROUP $\text{GF}(q)^*$

-48-

DEFINITION The multiplicative group of $\text{GF}(q)$ is $\text{GF}(q)^* = \text{GF}(q) \setminus \{0\}$.

THEOREM Let $\alpha \in \text{GF}(q)^*$. Then $\alpha^{q-1} = 1$.

NOTE: If $\text{GF}(q) = \mathbb{Z}_p$, then this is Fermat's Little Theorem.

PROOF Let the distinct elements of $\text{GF}(q)^*$ be $\alpha_1, \alpha_2, \dots, \alpha_{q-1}$.

Consider the (nonzero) elements $\alpha\alpha_1, \alpha\alpha_2, \dots, \alpha\alpha_{q-1}$. These elements are distinct because if $\alpha\alpha_i = \alpha\alpha_j$ for some $i \neq j$, then $\alpha^{-1}(\alpha\alpha_i) = \alpha^{-1}(\alpha\alpha_j)$, so $\alpha_i = \alpha_j$, a contradiction.

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

$$(\alpha\alpha_1)(\alpha\alpha_2)\cdots(\alpha\alpha_{q-1}) = \alpha_1\alpha_2\cdots\alpha_{q-1}.$$

Cancelling gives $\alpha^{q-1} = 1$. $\square$

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

ORDERS OF FIELD ELEMENTS

-49-

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

THEOREM Let $\alpha \in \text{GF}(q)^*$, $\text{ord}(\alpha) = t$. Then $\alpha^s = 1$ iff $t \mid s$.

PROOF Let $s \in \mathbb{Z}$. Then long division of s by t yields

$$s = lt + r, \text{ where } 0 \leq r < t.$$

$$\text{Now, } \alpha^s = \alpha^{lt+r} = (\alpha^t)^l \cdot \alpha^r = \alpha^r,$$

$$\text{Hence, } \alpha^s = 1 \iff \alpha^r = 1 \iff r = 0 \iff t \mid s. \quad \square$$

COROLLARY Let $\alpha \in \text{GF}(q)^*$. Then $\text{ord}(\alpha) \mid (q-1)$.

EXAMPLE There is only one element in $\text{GF}(q)$ of order 1, namely $\alpha = 1$.

- EXAMPLE Consider $\text{GF}(2^3) = \mathbb{Z}_2[x]/(x^3+x+1)$.

The order of $\alpha = x^2+1$ is 7.

- EXAMPLE Consider $\text{GF}(2^4) = \mathbb{Z}_2[x]/(x^4+x+1)$.

$f(x) = x^4+x+1$ has no roots in $\mathbb{Z}_2$, thus no linear factors.

Also, $f(x)$ has no irreducible quadratic factors since $(x^2+x+1) \nmid f(x)$.

Hence, f is irreducible over $\mathbb{Z}_2$.

Find $\text{ord}(\alpha)$ in $\text{GF}(2^4)$.

SOLUTION We have $\alpha^1 = \alpha$, $\alpha^2 = \alpha^2$, $\alpha^3 = \alpha^3$, $\alpha^4 = \alpha+1$, $\alpha^5 = \alpha^2+\alpha$.

Thus, $\text{ord}(\alpha) \neq 1, 3, 5$. Since $\text{ord}(\alpha) \mid 15$, we must have $\text{ord}(\alpha) = 15$.

FACT Let $\alpha \in \text{GF}(q)^*$ with $\text{ord}(\alpha) = t$. Then the elements $\alpha^0, \alpha^1, \alpha^2, \dots, \alpha^{t-1}$ are distinct. In particular, if $\text{ord}(\alpha) = q-1$, then $\{\alpha^0, \alpha^1, \dots, \alpha^{q-2}\} = \text{GF}(q)^*$.

DEFINITION A generator of $\text{GF}(q)^*$ is an element in $\text{GF}(q)$ of order $q-1$.

EXAMPLE $\alpha = x$ is a generator of $\text{GF}(2^4)^*$, where $\text{GF}(2^4) = \mathbb{Z}_2[x]/(x^4+x+1)$.

Let's verify that $\text{ord}(x) = 15$.

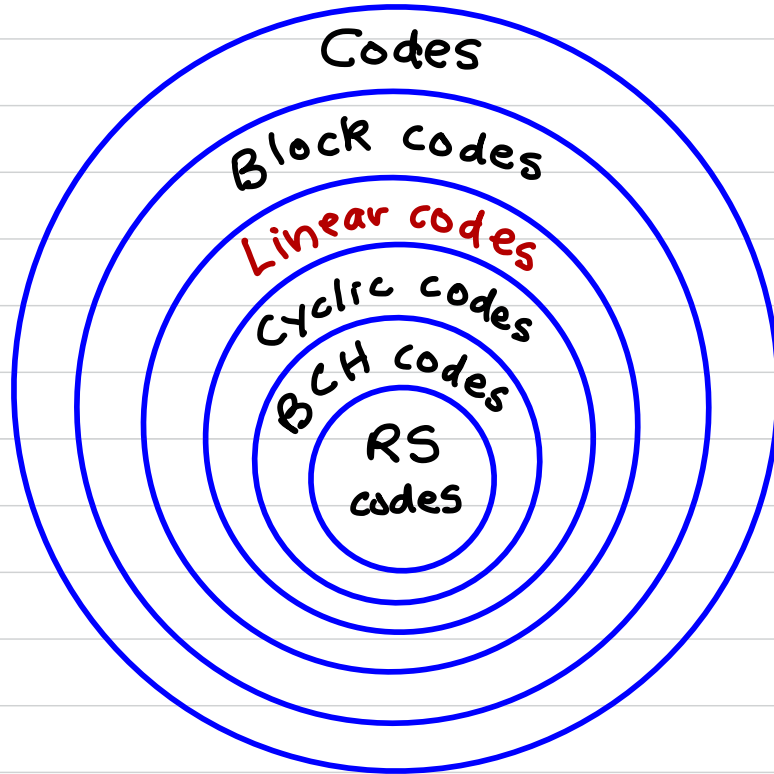
$$\begin{aligned} x^0 &= 1, & x^1 &= x, & x^2 &= x^2, & x^3 &= x^3, & x^4 &= x+1, & x^5 &= x^2+x, & x^6 &= x^3+x^2, \\ x^7 &= x^3+x+1, & x^8 &= x^2+1, & x^9 &= x^3+x, & x^{10} &= x^2+x+1, & x^{11} &= x^3+x^2+x, \\ x^{12} &= x^3+x^2+x+1, & x^{13} &= x^3+x^2+1, & x^{14} &= x^3+1, & x^{15} &= 1. \end{aligned}$$

THEOREM $\text{GF}(q)^*$ has a generator.

PROOF omitted. $\square$

V3a LINEAR CODES

-52-



Linear codes

- Weight
- Generator matrix
- Dual code
- Parity-check matrix
- Distance
- Perfect codes
- Hamming codes
- Syndrome decoding

DEFINITION OF A LINEAR CODE

• Let $F = GF(q)$, and let $V_n(F) = F^n = \overbrace{F \times F \times \dots \times F}^n$.

$V_n(F)$ is an n -dimensional vector space over F , with $|V_n(F)| = q^n$.

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

RECALL A subspace S of a vector space V over F is a non-empty subset $S \subseteq V$ such that (i) $a, b \in S \Rightarrow a + b \in S$, and (ii) $a \in S, \lambda \in F \Rightarrow \lambda a \in S$.

If S is a subspace of V , then S is itself a vector space over F ; in particular $0 \in S$. A basis of S is a linearly independent, spanning subset of S . All bases of S have the same size, called the dimension of S .

PROPERTIES OF LINEAR CODES

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

1) NUMBER OF CODEWORDS The codewords in C are precisely $m_1 v_1 + m_2 v_2 + \dots + m_k v_k$, where $m_i \in F$.

Thus, $|C| = M = q^k$.

2) RATE The rate of C is $R = \frac{\log_q M}{n} = \frac{\log_q q^k}{n} = \frac{k}{n}$.

3) WEIGHT

DEFINITION The weight $w(v)$ of a vector $v \in V_n(F)$ is the number of nonzero coordinates in v . The weight of a linear code C is $w(C) = \min \{w(c) : c \in C, c \neq 0\}$.

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

PROOF We have $d(C) = \min \{d(x, y) : x, y \in C, x \neq y\}$
 $= \min \{w(x - y) : x, y \in C, x \neq y\}$ (since $d(x, y) = w(x - y)$)
 $= \min \{w(c) : c \in C, c \neq 0\}$ (since C is linear, $x - y \in C$)
 $= w(C)$. $\square$

NOTATION An (n, k, d) -code C over F is a linear code of length n , dimension k , and distance d .

4) ENCODING

- Since there are q^k codewords, there are also q^k source messages.
- We shall assume that the source messages are the elements of F^k .
- Then, a convenient and natural bijection between F^k and C , i.e. an encoding rule, is defined by

$$m = (m_1, m_2, \dots, m_k) \mapsto c = m_1 v_1 + m_2 v_2 + \dots + m_k v_k.$$
- NOTE Different ordered bases for C yield different encoding rules.

5) GENERATOR MATRIX: A convenient way to describe C .

DEFINITION A generator matrix (GM) for an (n, k) -code C is a $k \times n$ matrix whose rows form a basis for C : $G =$

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{bmatrix}$$

NOTE The encoding rule is $c = mG$.

EXAMPLE (linear code) Consider the $(5,3)$ -binary code

$$C = \langle \underbrace{10011}_{c_1}, \underbrace{01001}_{c_2}, \underbrace{00110}_{c_3} \rangle. \quad (\text{NOTE: } c_1, c_2, c_3 \text{ are l.i. over } \mathbb{Z}_2.)$$

• A generator matrix for C is $G = \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{array} \right]_{3 \times 5}.$

• The encoding rule w.r.t. the GM G is $c = mG$.

$$000 \longrightarrow 00000$$

$$100 \longrightarrow 10011$$

$$001 \longrightarrow 00110$$

$$101 \longrightarrow 10101$$

$$010 \longrightarrow 01001$$

$$110 \longrightarrow 11010$$

$$\underbrace{011}_m \longrightarrow \underbrace{0111}_c$$

$$\underbrace{111}_m \longrightarrow \underbrace{11100}_c$$

• $M = |C| = 2^3 = 8$, $R = 3/5$, $d(C) = w(C) = 2$.

So, C is a $(5,3,2)$ -binary code.

STANDARD FORM G_M

DEFINITION Let C be an (n, k) -code over F . A G_M G for C of the form $G = [I_k | A]_{k \times n}$ is said to be in standard form.

If C has a G_M in standard form, then C is a systematic code.

EXAMPLE $C = \langle 100011, 001001, 000110 \rangle$ is a non-systematic $(6, 3)$ -binary code.

However, $C' = \langle 100011, 001001, 010010 \rangle$ obtained by swapping the 2nd and 4th coordinates of every codeword in C , is systematic.

A G_M for C' is

$$G' = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right]_{3 \times 6}$$

EQUIVALENT CODES

DEFINITION Two codes C, C' of length n are equivalent if C' can be obtained from C by choosing a permutation of coordinate positions $\{1, 2, \dots, n\}$, and then consistently rearranging every codeword in C according to this permutation.

FACTS (equivalent codes)

1. If C is linear and C' is equivalent to C , then C' is linear.
2. Equivalent linear codes have the same length, dimension, distance.
3. Every linear code is equivalent to a systematic code.

V3b DUAL CODE AND PARITY CHECK MATRICES

-60-

DEFINITION Let $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n) \in V_n(F)$.

The inner product of x and y is $x \cdot y = x_1 y_1 + x_2 y_2 + \dots + x_n y_n \in F$.

x, y are orthogonal if $x \cdot y = 0$.

PROPERTIES For all $x, y, z \in V_n(F)$ and $\lambda \in F$:

(i) $x \cdot y = y \cdot x$.

(ii) $x \cdot (y + z) = x \cdot y + x \cdot z$.

(iii) $(\lambda x) \cdot y = \lambda(x \cdot y)$.

(iv) $x \cdot x = 0$ does not imply that $x = 0$.

EXAMPLE Let $x = 111100 \in V_6(\mathbb{Z}_2)$. Then $x \cdot x = 0$, but $x \neq 0$.

More generally, if $x \in V_n(\mathbb{Z}_2)$, then $x \cdot x = 0$ iff $\omega(x)$ is even.

THE DUAL CODE

DEFINITION Let C be an (n, k) -code over F . The dual code of C is $C^\perp = \{x \in V_n(F) : x \cdot y = 0 \ \forall y \in C\}$. " $C^\perp = C^{\text{perp}}$ "

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

PROOF Let G be a G.M. for C , and let the rows of G be $v_1, v_2, \dots, v_k$.

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

PROOF OF CLAIM ($\Rightarrow$) is clear since $v_1, v_2, \dots, v_k \in C$.

($\Leftarrow$) Suppose $v \in C$. Then we can write $v = \lambda_1 v_1 + \dots + \lambda_k v_k$, where $\lambda_i \in F$. Then $v \cdot x = (\lambda_1 v_1 + \dots + \lambda_k v_k) \cdot x = \lambda_1 (v_1 \cdot x) + \dots + \lambda_k (v_k \cdot x) = 0$. $\square$

Thus, $C^\perp = \{x \in V_n(F) : Gx^T = 0\} = \text{null space of } G$.

Since G has rank k , its null space is a subspace of $V_n(F)$ of dim $n-k$. $\square$

PARITY-CHECK MATRIX

THEOREM 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.

Furthermore, $(C^\perp)^\perp$ is an (n, k) -code, and $C \subseteq (C^\perp)^\perp$.

Since $\dim(C) = \dim((C^\perp)^\perp)$, it follows that $C = (C^\perp)^\perp$. $\square$

DEFINITION If C is a linear code, then a generator matrix H for $C^\perp$ is called a parity-check matrix (PCM) for C .

NOTES (i) H is an $(n-k) \times n$ matrix.

(ii) C has many PCMs.

CONSTRUCTING A PCM FOR C (A GM FOR $C^\perp$)

THEOREM Let C be an (n, k) -code with GM $G = [I_k | A]$.

Then $H = [-A^T | I_{n-k}]$ is a GM for $C^\perp$.

NOTE: A is a $k \times (n-k)$ matrix.

PROOF Since $\text{rank}(H) = n-k$, H is a GM for an $(n, n-k)$ -code $\bar{C}$.

$$\text{Also, } GH^T = [I_k | A] \begin{bmatrix} -A \\ I_{n-k} \end{bmatrix} = -A + A = 0.$$

Thus, $\bar{C} \subseteq C^\perp$. Since $\dim(\bar{C}) = \dim(C^\perp) = n-k$, we have $\bar{C} = C^\perp$.

Hence, H is a GM for $C^\perp$. $\square$

EXAMPLE Consider the $(5,2)$ -code C over $\mathbb{Z}_3$ with
 G.M $G = \begin{bmatrix} 2 & 0 & 2 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix}$. Find a PCM for C .

SOLUTION First find a G.M for C in standard form.

$$G \xrightarrow{R_1 \leftarrow 2R_1} \begin{bmatrix} 1 & 0 & 1 & 2 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_2 \leftarrow R_2 - R_1} \begin{bmatrix} 1 & 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 1 & 1 \end{bmatrix}$$

A

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

- We have $C = \{00000, 20210, 10120, 11001, 22002, 01211, 02122, 21101, 12212\}$.
- $d(C) = w(C) = 3$, and $R = 2/5$.

NOTES ON PCMS

Let C be an (n, k) -code over F with G.M. G .

1) An $(n-k) \times n$ matrix H over F is a PCM for C iff $GH^T = 0$ and $\text{rank}(H) = n-k$.

2) G is a PCM for $C^\perp$ (since $(C^\perp)^\perp = C$).

3) Let H be a PCM for C . Then H is a G.M. for $C^\perp$, so $C = \text{null space of } H$.

4) Let H be a PCM for C , and let $x \in V_n(F)$. Then $x \in C$ iff $Hx^T = 0$.

V3C DISTANCE OF A LINEAR CODE

-66-

C	$C^\perp$
(n, k) -code over F	$(n, n-k)$ -code over F
G : G.M for C	H : G.M for $C^\perp$
C = row space of G	$C^\perp$ = row space of H
H : PCM for C	G : PCM for $C^\perp$
C = null space of H	$C^\perp$ = null space of G
$x \in C$ iff $Hx^T = 0$	$x \in C^\perp$ iff $Gx^T = 0$

$$(C^\perp)^\perp = C$$

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

PROOF Let $H = [h_1 | h_2 | \dots | h_n]$.

($\Leftarrow$) Suppose $d(C) \leq s-1$. Let $c = (c_1, c_2, \dots, c_n) \in C$ with $c \neq 0$ and $w(c) \leq s-1$. Without loss of generality, suppose that $c_j = 0$ for $s \leq j \leq n$. Then, since $Hc^T = 0$, we have

$$c_1 h_1 + c_2 h_2 + \dots + c_{s-1} h_{s-1} + c_s h_s + \dots + c_n h_n = 0,$$

so $c_1 h_1 + \dots + c_{s-1} h_{s-1} = 0$. Since at least one of $c_1, c_2, \dots, c_{s-1}$ is nonzero, the $s-1$ columns $h_1, h_2, \dots, h_{s-1}$ of H are linearly dependent.



($\Rightarrow$) Suppose there is a set of $s-1$ columns of H that are linearly dependent over F . Without loss of generality, let them be $h_1, h_2, \dots, h_{s-1}$. So, $\exists \lambda_i \in F$, not all 0, such that

$$\lambda_1 h_1 + \lambda_2 h_2 + \dots + \lambda_{s-1} h_{s-1} = 0.$$

Let $c = (\lambda_1, \lambda_2, \dots, \lambda_{s-1}, 0, \dots, 0) \in V_n(F)$. Then $c \in C$ since $Hc^T = \lambda_1 h_1 + \lambda_2 h_2 + \dots + \lambda_{s-1} h_{s-1} = 0$. But $1 \leq \omega(c) \leq s-1$, so $1 \leq d(C) \leq s-1$. $\square$

COROLLARY Let H be a PCM for a linear code C over F . Then $d(C)$ is the smallest number of columns of H that are linearly dependent over F .

EXAMPLE In the example on slide 64, $F = \mathbb{Z}_3$ and $H = \begin{bmatrix} 2 & 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 1 \end{bmatrix}$.

- No single column of H is l.d. over $\mathbb{Z}_3$ (i.e. no zero column), so $d(C) \geq 2$.
- No two columns of H are l.d. over $\mathbb{Z}_3$ (i.e. scalar multiples of each other), so $d(C) \geq 3$.
- There exist three columns of H that are l.d. over $\mathbb{Z}_3$:

$$\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \text{ so } d(C) \neq 4.$$
- Hence $d(C) = 3$.

EXAMPLE Let H be a PCM for a binary linear code C .

1) $d(C)=1$ iff H has a zero column.

2) $d(C)=2$ iff H has no zero column, and two columns of H are identical.

3) $d(C)=3$ iff the columns of H are nonzero and distinct, and some column of H is the sum of two other columns.

4) etc.

EXAMPLE Find a (n, k, d) -binary code C .

SOLUTION Construct a PCM for C : $H = \begin{bmatrix} 1 & 0 & 0 & | & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & | & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & | & 0 & 1 & 1 & 1 \end{bmatrix}_{3 \times 7}$.

Then a GM for C is $\begin{bmatrix} 1 & 1 & 0 & | & \\ 0 & 1 & 1 & | & \\ 1 & 0 & 1 & | & \\ 1 & 1 & 1 & | & \end{bmatrix}_{4 \times 7} I_4$.

C has $n=7$, $k=4$, $d=3$, $q=2$, $M=16$.

NOTE C is a Hamming code of order 3 over $\mathbb{Z}_2$.

V3d HAMMING CODES

-72-

- Hamming codes are an infinite family of single-error correcting codes discovered by Richard Hamming in 1950.

- EXAMPLE Recall that $H = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{array} \right]$ is a PCM for a

$(7,4,3)$ -binary code C , called the Hamming code of order 3 over $\mathbb{Z}_2$.

DEFINITION A Hamming code of order r over $F = \text{GF}(q)$ is an (n,k) -code over F with $n = \frac{q^r - 1}{q - 1}$ and $k = n - r$, and with PCM H_r ,

an $r \times n$ matrix whose columns are nonzero, and no two of whose columns are scalar multiples of each other.

EXAMPLE A PCM for a Hamming code of order 3 over $\mathbb{Z}_3$ is:

$$H_3 = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 & 2 & 1 & 0 & 1 & 1 & 2 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 2 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 2 & 1 & 2 & 1 & 1 \end{bmatrix}_{3 \times 13}$$

$$n=13$$

$$k=10$$

$$d=3$$

NOTES (Hamming codes)

1) If $v \in V_r(F)$, $v \neq 0$, then exactly one nonzero scalar multiple of v must be column of H_r , giving $n = (q^r - 1) / (q - 1)$ columns in total.

2) H_r has rank r , since among the columns of H_r are scalar multiples of the unit vectors. Hence Hamming codes of order r over $GF(q)$ do indeed have dimension $n - r$.

3) By design, Hamming codes of order r over $GF(q)$ have distance 3, and so are single-error correcting codes.

DECODING SINGLE-ERROR CORRECTING CODES

- Let H be a PCM for an (n, k, d) -code C over F with $d \geq 3$.

DEFINITION Suppose $c \in C$ is sent and $r \in V_n(F)$ is received. The error vector is $e = r - c$ (so $r = c + e$).

EXAMPLE Over $\mathbb{Z}_3$: If $c = 100101$ and $e = 020000$, then $r = 120101$.

KEY OBSERVATIONS

- 1) $Hr^T = H(c+e)^T = Hc^T + He^T = He^T$ (since $Hc^T = 0$).
- 2) If $e = 0$, then $He^T = 0$. (The converse is not true.)
- 3) If $w(e) = 1$, say $e = (0, \dots, 0, \alpha, 0, \dots, 0)$, then $He^T = \alpha h_i$, where h_i is the i^{th} column of H . (The converse is not true.)

DECODING ALGORITHM FOR SINGLE-ERROR CORRECTING CODES

INPUT PCM H and a received word $r \in V_n(F)$.

- 1) Compute $S = Hr^T$.
- 2) If $S = 0$, then accept r as the transmitted word (so $e = 0$); STOP.
- 3) Compare S with the columns of H . If $S = \alpha h_i$ for some i , then set $e = (0, \dots, 0, \alpha, 0, \dots, 0)$, and decode to $c = r - e$; STOP.
 i th position 
- 4) Report that more than one error has occurred.

CORRECTNESS If $w(e) = 0$ or $w(e) = 1$, the decoding algorithm is guaranteed to make the correct decision.

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

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

Suppose that $r = 0111110$.

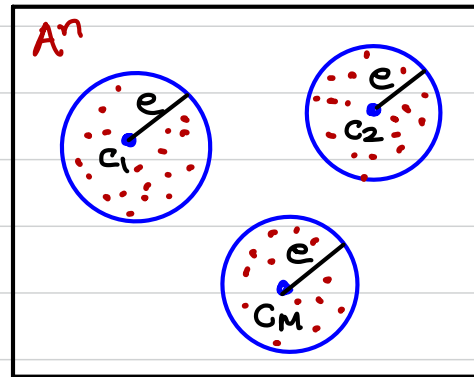
We compute $s = Hr^T = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, which is the 6th column of H .

So, we set $e = (000010)$ and decode r to $c = r - e = \underline{0111100}$.
 6th position 

CHECK Verify that $Hc^T = 0$.

V3e PERFECT CODES

DEFINITION Let C be an $[n, M]$ -code of distance d over A , with $|A|=q$ and $e = \lfloor \frac{d-1}{2} \rfloor$. Then C is perfect if each $x \in A^n$ is in the sphere of radius e centered at some $c \in C$.



• Equivalently, C is perfect if
$$M \cdot \sum_{i=0}^e \binom{n}{i} (q-1)^i = q^n.$$

• For fixed q, n, d , a perfect code has maximum possible M . In other words, a perfect code has maximum possible rate $R = \frac{\log_q M}{n}$ for fixed q, n, d .

- EXAMPLE $C = A^n$ is a (trivial) perfect code with distance $d=1$.
- EXAMPLE Let n be odd. Then $C = \{\underbrace{000\cdots 0}_n, \underbrace{111\cdots 1}_n\}$ is a perfect binary code with distance n .

PROOF Let $e = (n-1)/2$. Then

$$\begin{aligned}
 M \cdot \sum_{i=0}^e \binom{n}{i} (q-1)^i &= 2 \left[\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{e} \right] \\
 &= \left[\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{e} \right] + \left[\binom{n}{e+1} + \binom{n}{e+2} + \cdots + \binom{n}{n} \right] \quad \left[\text{since } \binom{n}{k} = \binom{n}{n-k} \right] \\
 &= (1+1)^n = 2^n. \quad \square
 \end{aligned}$$

- EXERCISE Prove that every perfect code has odd distance.
- EXERCISE Show that $\text{IMLD} = \text{CMLD}$ for perfect codes.

EXAMPLE All Hamming codes of order r over $GF(q)$ are perfect.

PROOF We have $n = \frac{q^r - 1}{q - 1}$, $k = n - r$, $d = 3$, $e = 1$.

$$\text{Now, } M \cdot \sum_{i=0}^e \binom{n}{i} (q-1)^i = q^k \left[\binom{n}{0} (q-1)^0 + \binom{n}{1} (q-1)^1 \right]$$

$$= q^k [1 + n(q-1)]$$

$$= q^{n-r} \left[1 + \frac{q^r - 1}{q - 1} (q-1) \right]$$

$$= q^n. \quad \square$$

CLASSIFICATION OF PERFECT CODES

THEOREM (Tietäväinen, 1973) The only perfect codes are:

- 1) $V_n(\mathbb{GF}(q))$ [trivial codes].
- 2) The binary replication codes of odd lengths.
- 3) The Hamming codes, and all codes with the same $[n, M, d]$ parameters.
- 4) The $(23, 12, 7)$ -binary Golay code C_{23} , and all codes equivalent to it [see V4a].
- 5) The $(11, 6, 5)$ -ternary Golay code and all codes equivalent to it.

A G.M for this code is

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

V3f SYNDROME DECODING

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

DEFINITION Let $x, y \in V_n(F)$. We write $x \equiv y \pmod{C}$ if $x - y \in C$.

FACTS 1) $\equiv \pmod{C}$ is an equivalence relation.

2) The set of equivalence classes partitions $V_n(F)$.

3) The equivalence class containing $x \in V_n(F)$ is
 $C + x = \{y \in V_n(F) : y \equiv x \pmod{C}\} = \{c + x : c \in C\}$,
 and is called a coset of C .

$V_n(F)$

C	$C + x_1$	$C + x_2$	$C + x_3$	$\dots$	
-----	-----------	-----------	-----------	---------	--

EXAMPLE (cosets) Consider the (5,2)-binary code C with G.M $G = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix}$.

Find all cosets of C .

SOLUTION The cosets of C are:

$$C = C + 00000 = \{00000, 10111, 01110, 11001\} = C + 10111 = C + 01110 = C + 11001.$$

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

$$C + 01000 = \{01000, 11111, 00110, 10001\}.$$

$$C + 00100 = \{00100, 10011, 01010, 11101\}.$$

$$C + 00010 = \{00010, 10101, 01100, 11011\}.$$

$$C + 00001 = \{00001, 10110, 01111, 11000\}.$$

$$C + 10100 = \{10100, 00011, 11010, 01101\}.$$

$$C + 10010 = \{10010, 00101, 11100, 01011\}.$$

FACTS (cosets)

1) $C + 0 = C$.

2) If $y \in C + x$, then $C + y = C + x$.

3) All cosets of C have the same size, namely $|C| = q^k$.

4) The number of distinct cosets is q^{n-k} .

DEFINITION Let H be a PCM for an (n, k) -code C over F .
 Let $x \in V_n(F)$. The syndrome of x (w.r.t. H) is $s = Hx^T$.

NOTES 1) $s \in V_{n-k}(F)$.

2) All codewords have syndrome 0.

THEOREM Let $x, y \in V_n(F)$. Then $x \equiv y \pmod{C}$ iff $Hx^T = Hy^T$.
So, cosets are characterized by their syndromes.

PROOF We have $x \equiv y \pmod{C}$ iff $x - y \in C$ iff $H(x - y)^T = 0$
iff $Hx^T = Hy^T$. $\square$

DECODING Recall that $c \in C$ is sent and $r \in V_n(F)$ is received.
The (unknown) error vector is $e = r - c$. Since $r - e = c$, we have
 $r \equiv e \pmod{C}$. Thus, r and e are in the same coset of C .

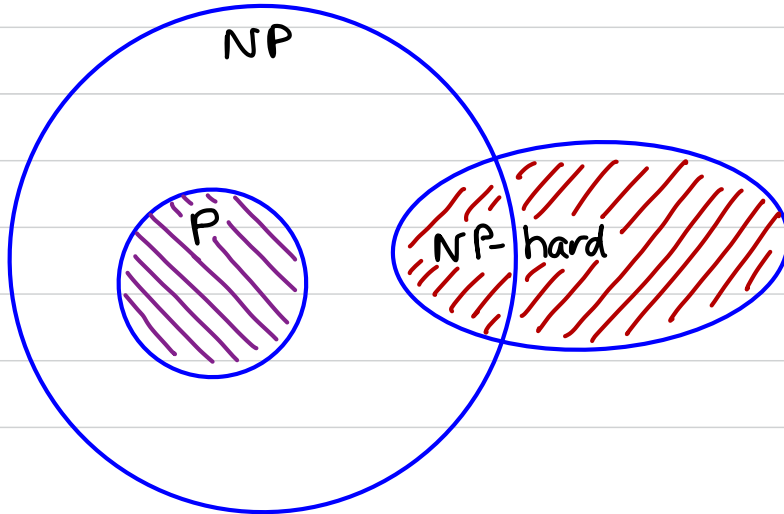
DECODING STRATEGY Given r , find a vector e of smallest weight
that has the same syndrome as r : $He^T = Hr^T$.

CMLD: Decode r to $c = r - e$.

IMLD: If e is unique, decode r to $c = r - e$; else reject r .

QUESTION Given H and τ , can one efficiently find a vector e of smallest weight such that $He^T = H\tau^T$? [syndrome decoding problem]

FACT This problem is NP-hard, which strongly suggests that no general-purpose efficient algorithm exists.



If any NP-hard problem can be solved efficiently, then all problems in NP can be solved efficiently, so " $P = NP$ ".

SYNDROME DECODING ALGORITHM (CMLD)

SETUP For each coset of C , select an arbitrary vector of smallest weight in that coset, and call it the coset leader of that coset.

Store a table of coset leaders and their syndromes.

coset leader	syndrome
--------------	----------

} q^{n-k} rows

DECODING ALGORITHM (CMLD)

Given $\mathbf{r} \in V_n(F)$, compute $\mathbf{s} = \mathbf{H}\mathbf{r}^T$. Let the corresponding coset leader be $\mathbf{e}$. Then decode $\mathbf{r}$ to $\mathbf{c} = \mathbf{r} - \mathbf{e}$.

NOTE The decoding algorithm is guaranteed to make the correct decision if the error vector is a coset leader; otherwise, it is guaranteed to make an incorrect decision.

SELECTING COSET LEADERS

THEOREM Let C be an (n, k, d) -code over F . Let $x \in V_n(F)$ be a vector of weight $\leq \lfloor \frac{d-1}{2} \rfloor$. Then x is a coset leader.

PROOF Suppose y is in the same coset as x , with $y \neq x$ and $w(y) \leq w(x) \leq \lfloor \frac{d-1}{2} \rfloor$. Then $x \equiv y \pmod{C}$, so $x - y \in C$ and $x - y \neq 0$. But

$$w(x - y) = w(x + (-y)) \leq w(x) + w(-y) = w(x) + w(y) \leq \lfloor \frac{d-1}{2} \rfloor + \lfloor \frac{d-1}{2} \rfloor \leq d-1.$$

This contradicts $d(C) = d$, so no such y exists.

Hence, x is the unique vector of smallest weight in its coset, so must be a coset leader. $\square$

EXAMPLE (syndrome decoding for the code on slide 82)

- For the $(5,2)$ -binary code C with G.M. $G = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix}$, we have

a PCM $H = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}$. Note that $d(C) = 3$.

• SYNDROME TABLE

<u>coset leader</u>	<u>syndrome</u>
00000	000
10000	111
01000	110
00100	100
00010	010
00001	001
10100	011
10010	101

DECODING

- Suppose $r = 11011$.
- Compute $s = Hr^T = 010$.
- The corresponding coset leader is $e = 00010$.
- Decode r to $c = r - e = \underline{11001}$.

NOTE Syndrome decoding is not efficient in general since the syndrome table is exponentially large.

For an (n, k) -binary code, the syndrome table has size

$$2^{n-k} (n + (n-k)) = 2^{n-k} (2n-k) \text{ bits.}$$

cosets
coset leader
syndrome

[Actually, only $2^{n-k}n$ bits are needed since the table can be sorted by syndrome, and then the syndromes do not need to be stored.]

V4a THE BINARY GOLAY CODE (1949)

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

} each row is the left cyclic shift of the previous row.

• Let $\hat{G} = [I_{12} | \hat{B}]_{12 \times 23}$. $\hat{G}$ is a G_M for a (23,12)-binary code called the (binary) Golay code C_{23} . In fact, $d(C_{23}) = 7$. [proof later]

CLAIM C_{23} is a perfect code.

PROOF We have $e = \lfloor (d-1)/2 \rfloor = 3$, and

$$M \cdot \sum_{i=0}^e \binom{n}{i} (q-1)^i = 2^{12} \left[\binom{23}{0} + \binom{23}{1} + \binom{23}{2} + \binom{23}{3} \right] = 2^{23}. \quad \square$$

THE EXTENDED GOLAY CODE C_{24}

- C_{24} is the binary code with GM $G = [I_{12} | B]_{12 \times 24}$, where

$$B = \left[\begin{array}{c|c} 0 & \hat{B} \\ \vdots & \\ 1 & \end{array} \right]_{12 \times 12} \quad \left(\text{so } B = \left[\begin{array}{c|c} 0 & 111111111111 \\ 1 & 11011100010 \\ \vdots & \vdots \\ 1 & 01101110001 \end{array} \right]_{12 \times 12} \right).$$

PROPERTIES OF C_{24}

- C_{24} is a $(24, 12)$ -binary code.
- $GG^T = 0$ [check this]. Hence $C_{24} \subseteq C_{24}^\perp$, so C_{24} is self-orthogonal. ↗ $C \subseteq C^\perp$
- Since $\dim(C_{24}) = \dim(C_{24}^\perp) = 12$, we have $C_{24} = C_{24}^\perp$. So, C_{24} is self-dual. ↘ $C = C^\perp$
- $B = B^T$ (so B is symmetric).
- A PCM for C_{24} is $H = [-B^T | I_{12}] = [B | I_{12}]$.
- Since $C_{24} = C_{24}^\perp$, H is also a GM for C_{24} .

THEOREM $d(C_{24}) = 8$.

$$G = \begin{bmatrix} I_{12} & \begin{matrix} 0111111111 \\ 111011100010 \\ 110111000101 \\ \vdots \\ 101101110001 \end{matrix} \end{bmatrix} \begin{matrix} r_1 \\ r_2 \\ r_3 \\ \vdots \\ r_{12} \end{matrix}$$

-92-

PROOF Let the rows of G be $r_1, r_2, \dots, r_{12}$.

i) Note that $w(r_1) = 12$ and $w(r_i) = 8$ for $2 \leq i \leq 12$.

Hence $4 \mid w(r_i)$ for all i . Now, $r_i \cdot r_j = 0$ since $GG^T = 0$, and so the number of coordinates of r_i and r_j that are both 1 is even. Hence, $4 \mid w(r_i + r_j)$, and so every codeword has weight divisible by 4.

Thus, $d(C_{24}) = 4$ or 8 .

ii) Next, we'll show that no codeword has weight 4 (so $d(C_{24}) = 8$).

- Each row of G has weight ≥ 8 .

- Adding two rows of G : $w(r_1 + r_j) = 8$ for $2 \leq j \leq 12$. [Check this]

$w(r_i + r_j) = 8$ for $2 \leq i < j \leq 12$. [Check this]



- Adding 3 rows of G_1 : Let $c = r_i + r_j + r_k$, where $1 \leq i < j < k \leq 12$.
Let $c = (x, y)$, where x, y have length 12. Suppose $w(c) = 4$.
Since $w(x) = 3$, we have $w(y) = 1$. Since $H = [B \mid I_{12}]$ is also a GM for C_{24} , c must be a single row of H . This is impossible since each row of H has weight 8 or 12. Hence, $w(c) \neq 4$.
- Adding 4 rows of G_1 : Let $c = r_i + r_j + r_k + r_l$, where $1 \leq i < j < k < l \leq 12$.
Let $c = (x, y)$, and suppose $w(c) = 4$. Then $w(x) = 4$ and $w(y) = 0$.
But H does not have any such vector in its row space. Thus, $w(c) \neq 4$.
- Adding ≥ 5 rows of G_1 : If $c = (x, y)$ is the sum of 5 or more rows of G_1 , then $w(x) \geq 5$, so $w(c) \neq 4$. $\square$

COROLLARY $d(C_{23}) = 7$.

V4b A DECODING ALGORITHM FOR C_{24}

RECALL $n=24$, $k=12$, $d=8$, $e=3$.

$G = [I_{12} | B]$ and $H = [B | I_{12}]$ are both GIMs and PCMs for C_{24} .

DECODING STRATEGY (IMLD) Compute a syndrome S of the received word r . Find a vector e of weight ≤ 3 that has syndrome S . If such a vector e exists, then decode r to $c = r - e$; else reject r .

CORRECTNESS If the error vector has weight ≤ 3 , then the decoder always make the correct decision. If the error vector has weight ≥ 4 , the decoder will either reject r , or will decode r to a codeword different from the transmitted one.

DECODING ALGORITHM FOR C24 (WITH JUSTIFICATION)

- Let $r = (x, y)$ and $e = (e_1, e_2)$, where x, y, e_1, e_2 have length 12.
- There are 5 cases (not mutually exclusive) in the event $w(e) \leq 3$:

(A) $w(e_1) = 0$ and $w(e_2) = 0$.

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

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

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

(E) $w(e_1) = 1$ and $w(e_2) = 1$ or 2 .

1) Compute the syndrome $s_1 = [I_{12} | B] r^T$.

If $s_1 = 0$, then accept r and STOP. [case A]

2) [Note: $s_1 = [I_{12} | B] r^T = [I_{12} | B] e^T = e_1^T + B e_2^T = e_1^T$. So, if we are in case B, then $1 \leq w(s_1) \leq 3$.] If $w(s_1) \leq 3$, then set $e = (s_1^T, 0)$. [case B]

Correct x in the positions corresponding to the 1's in s_1 and STOP.



3) [Recall: $s_i = e_i^T + Be_i^T$. If we are in case (C), then s_i is equal to a column of B with one or two bits flipped (depending on which bits of e_i are 1)]. Compare s_i with the columns (or rows) of B . [Case C]

If any column of B , say column i , differs in exactly one (say j) or two (say j and k) positions from s_i , then decode $r = (x, y)$ as follows:

Correct x in position j , or in positions j, k .

Correct y in position i , and STOP.

4) Compute the syndrome $s_2 = [B | I_{12}] r^T = [B | I_{12}] e^T = Be_1^T + e_2^T = e_2^T$.

If $w(s_2) \leq 3$, then correct y in the positions corresponding to the 1's in s_2 and STOP. [Case D]

5) Analogous to step 3. [Case E]

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

DECODING ALGORITHM FOR C24

Suppose $\mathbf{r} = (x, y)$ is received.

- 1) Compute $\mathbf{s}_1 = [\mathbf{I}_{12} | \mathbf{B}] \mathbf{r}^T$. If $\mathbf{s}_1 = \mathbf{0}$, then accept $\mathbf{r}$ and STOP. (A)
- 2) If $\omega(\mathbf{s}_1) \leq 3$, then set $\mathbf{e} = (\mathbf{s}_1^T, \mathbf{0})$ and decode $\mathbf{r}$ to $\mathbf{c} = \mathbf{r} - \mathbf{e}$ and STOP. (B)
- 3) Compare $\mathbf{s}_1$ to the rows of $\mathbf{B}$. If row i of $\mathbf{B}$ differs in exactly one position (say j) or two positions (say j and k), then:
Correct x in position j , or positions j and k ; Correct y in position i ; STOP. (C)
- 4) Compute $\mathbf{s}_2 = [\mathbf{B} | \mathbf{I}_{12}] \mathbf{r}^T$.
If $\omega(\mathbf{s}_2) \leq 3$, then set $\mathbf{e} = (\mathbf{0}, \mathbf{s}_2^T)$ and decode $\mathbf{r}$ to $\mathbf{c} = \mathbf{r} - \mathbf{e}$ and STOP. (D)
- 5) Compare $\mathbf{s}_2$ to the rows of $\mathbf{B}$. If row i of $\mathbf{B}$ differs in exactly one position (say j) or two positions (say j and k), then:
Correct y in position j , or positions j and k ; Correct x in position i ; STOP. (E)
- 6) Reject $\mathbf{r}$.

- EXAMPLE Decode $\mathbf{r} = (1006 \ 1000 \ 0000 \ 1001 \ 0001 \ 1101)$.

SOLUTION Compute $\mathbf{S}_1 = [\mathbf{I}_{12} | \mathbf{B}] \mathbf{r}^T = (0100 \ 1000 \ 0000)^T$.

Since $w(\mathbf{S}_1) = 2$, we set $\mathbf{e} = (\mathbf{S}_1^T, 0)$ and decode $\mathbf{r}$ to

$$\mathbf{c} = \mathbf{r} - \mathbf{e} = (\underline{1100} \ \underline{0000} \ 0000 \ 1001 \ 0001 \ 1101). \quad [\text{Check: } \mathbf{H}\mathbf{c}^T = \mathbf{0}]$$

- EXAMPLE Decode $\mathbf{r} = (1000 \ 0010 \ 0000 \ 1000 \ 1101 \ 0010)$.

SOLUTION Compute $\mathbf{S}_1 = [\mathbf{I}_{12} | \mathbf{B}] \mathbf{r}^T = (1011 \ 1110 \ 1011)^T$.

Then $w(\mathbf{S}_1) > 3$. We see that $\mathbf{S}_1$ differs in positions 6 and 7

from row 4 of $\mathbf{B}$. So, we set $\mathbf{e} = (0000 \ 0\underline{11}0 \ 0000 \ 00\underline{01} \ 0000 \ 0000)$

and decode $\mathbf{r}$ to $\mathbf{c} = \mathbf{r} - \mathbf{e} = (1000 \ 0\underline{1}00 \ 0000 \ 100\underline{1} \ 1101 \ 0010)$.

[Check: $\mathbf{H}\mathbf{c}^T = \mathbf{0}$]

B =

0	1	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	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	
1	0	0	1	0	1	1	0	1	1	1	0	
1	0	1	0	1	1	0	1	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	12x12

NOTES (decoding algorithm for C_{24})

1) The decoding algorithm only needs to store B (144 bits).

In contrast, a syndrome table has size $2^{12} \times 24 = 98,304$ bits.

2) Decoding is efficient and simple $\Rightarrow$ good for hardware implementation.

3) C_{24} was used in the Voyager space mission to transmit photos of Jupiter and Saturn to earth.

V4C RELIABILITY OF C24

- QUESTION Is C24 better than simpler codes such as the binary replication codes and the Hamming codes?

- Let p = symbol error probability, and $C = \{c_1, c_2, \dots, c_M\}$.
- Let w_i = prob. that the decoding algorithm makes an incorrect decision or rejects if c_i is sent.
- $P_C = \frac{1}{M} \sum_{i=1}^M w_i = w_1 = \text{error probability of } C$.
- $1 - P_C = \text{reliability of } C = \text{prob. that } r \text{ is decoded correctly.}$

(1) For C24, $1 - P_{C24} = (1-p)^{24} + \binom{24}{1} p(1-p)^{23} + \binom{24}{2} p^2(1-p)^{22} + \binom{24}{3} p^3(1-p)^{21}$.

- (2) If no channel encoding is used, the prob. that a 12-bit message is transmitted with no errors is $(1-p)^{12}$.

(3) Suppose the binary triplication code T is used to encode 12-bit messages. Then $1 - P_T = [(1-p)^3 + 3p(1-p)^2]^{12}$.

(4) Suppose that a (15,11)-binary Hamming code H is used to encode 11-bit messages. Then $1 - P_H = (1-p)^{15} + 15p(1-p)^{14}$.

p	$(1-p)^{12}$	$1 - P_T$	$1 - P_H$	$1 - P_{C24}$
0.1	0.282429	0.711206	0.549043	0.785738
0.01	0.886385	0.996480	0.990378	0.999909
0.001	0.988066	0.999964	0.999896	0.999999895
Rate	1	$\frac{1}{3} \approx 0.33$	$\frac{11}{15} \approx 0.73$	$\frac{1}{2} = 0.5$

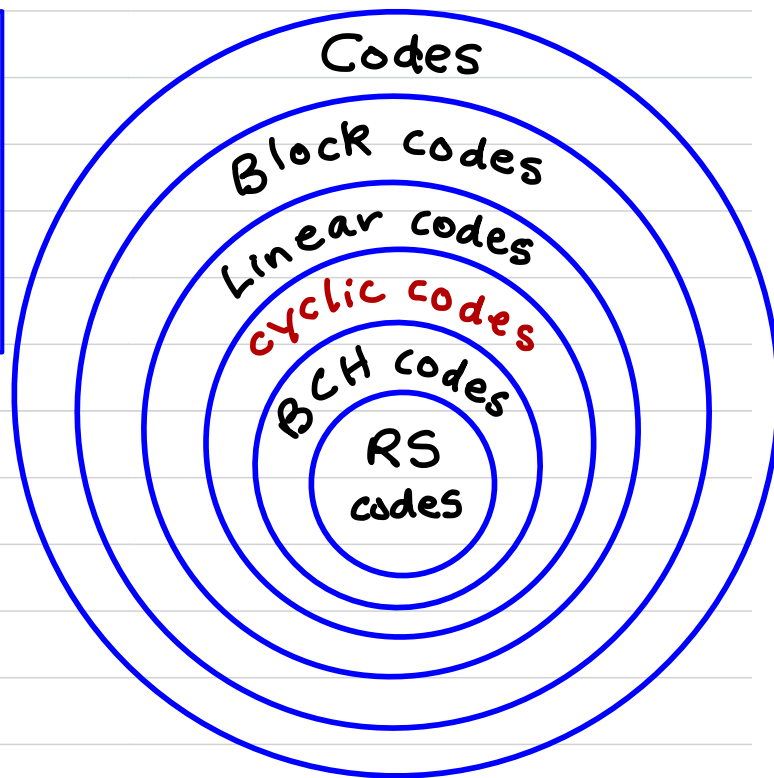
V5a CYCLIC CODES

DEFINITION A subspace S of $V_n(F)$ is cyclic if $(a_0, a_1, \dots, a_{n-1}) \in S$ implies that $(a_{n-1}, a_0, a_1, \dots, a_{n-2}) \in S$.

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

NEXT GOAL:

An algebraic characterization of cyclic subspaces of $V_n(F)$ as ideals of the polynomial ring $R = F[x]/(x^n - 1)$.



THE POLYNOMIAL RING $R = F[x]/(x^n - 1)$

• Let $R = F[x]/(x^n - 1)$ where $F = GF(q)$. Then R is a commutative ring (but not a field since $x^n - 1$ is reducible over F).

• We have the following bijection between $V_n(F)$ and R :

$$a = (a_0, a_1, a_2, \dots, a_{n-1}) \longleftrightarrow a(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1}.$$

• This bijection preserves vector addition and scalar multiplication:

$$\text{If } a, b \in V_n(F) \text{ and } \lambda \in F, \text{ then } a + b \longleftrightarrow a(x) + b(x)$$

$$\text{and } \lambda a \longleftrightarrow \lambda a(x).$$

• We can use this bijection to define a natural multiplication on $V_n(F)$:

DEFINITION Let $a, b \in V_n(F)$. Then $a \cdot b = c \in V_n(F)$, where
 $c \longleftrightarrow c(x) = a(x) \cdot b(x) \bmod (x^n - 1).$

WHY CHOOSE x^{n-1} AS THE MODULUS IN R ?

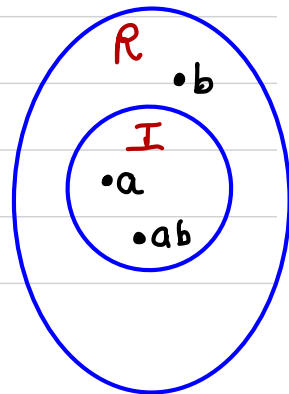
- Let $a = (a_0, a_1, \dots, a_{n-1}) \in V_n(F)$, and let $a(x)$ be the associated polynomial in R . Then

$$\begin{aligned} x \cdot a(x) &= a_0x + a_1x^2 + \dots + a_{n-2}x^{n-1} + a_{n-1}x^n \\ &\equiv a_{n-1} + a_0x + a_1x^2 + \dots + a_{n-2}x^{n-1} \pmod{x^{n-1}} \\ &\longleftrightarrow (a_{n-1}, a_0, a_1, \dots, a_{n-2}). \end{aligned}$$

- So, multiplication by x of a polynomial in R corresponds to a (right) cyclic shift of the associated vector in $V_n(F)$.

DEFINITION Let R be a finite commutative ring. A non-empty subset I of R is an ideal of R if (i) $a, b \in I \Rightarrow a+b \in I$, and (ii) $a \in I, b \in R \Rightarrow a \cdot b \in I$.

EXAMPLE $\{0\}$ and R are (trivial) ideals of R .



THEOREM (algebraic characterization of cyclic subspaces of $V_n(F)$)
 Let S be a non-empty subset of $V_n(F)$, and let I be the set of associated polynomials in $R = F[x]/(x^n - 1)$. Then S is a cyclic subspace of $V_n(F)$ iff I is an ideal of R .

PROOF ($\Rightarrow$) Suppose S is a cyclic subspace of $V_n(F)$. Since S is non-empty and closed under addition, so is I . Now, let $a(x) \in I$ and $b(x) = b_0 + b_1x + \dots + b_{n-1}x^{n-1} \in R$. Since S is a cyclic subspace, $x \cdot a(x) \in I$. Hence, $x^i a(x) \in I$ for all $0 \leq i \leq n-1$. Since S is closed under scalar multiplication, we have $b_i x^i a(x) \in I$ for all $0 \leq i \leq n-1$. Finally, since S is closed under addition, $\sum_{i=0}^{n-1} b_i x^i a(x) = b(x)a(x) \in I$. Thus, I is an ideal of R .



($\Leftarrow$) Suppose I is an ideal of R . Since I is non-empty and closed under addition, so is S . Since I is closed under multiplication by constant polynomials (i.e., elements of F), S is closed under scalar multiplication. Thus, S is a subspace of $V_n(F)$. Finally, since I is closed under multiplication by x , S is closed under right cyclic shifts. Thus, S is a cyclic subspace of $V_n(F)$. $\square$

So, to study cyclic subspaces of $V_n(F)$, we proceed to study ideals of $R = F[x]/(x^n - 1)$.

V56 IDEALS OF $R = F[x]/(x^n - 1)$

-108-

DEFINITIONS Let R be a (finite commutative) ring, and let $g \in R$.

- Let $\langle g \rangle = \{g \cdot r : r \in R\}$. Then $\langle g \rangle$ is an ideal of R , called the ideal generated by g .
- An ideal I of R is principal if $I = \langle g \rangle$ for some $g \in I$.
- R is a principal ideal ring if every ideal of R is principal.

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

PROOF Let I be an ideal of R .

- If $I = \{0\}$, then $I = \langle 0 \rangle$.
- If $I \neq \{0\}$, then let $g(x)$ be a nonzero polynomial of smallest degree in I . We now show that $I = \langle g \rangle$.

Let $h \in I$. We can write $h(x) = l(x)g(x) + r(x)$, where $l, r \in F[x]$ and

$\deg(r) < \deg(g)$. Now, $h(x), l(x)g(x) \in I$ implies that

$h(x) - l(x)g(x) \in I$, so $r(x) \in I$. But $\deg(r) < \deg(g)$, so we must have $r(x) = 0$. Thus, $h(x) = l(x)g(x)$, so $h \in \langle g \rangle$. Thus, $I \subseteq \langle g \rangle$.

And, since $g \in I$, we have $\langle g \rangle \subseteq I$. Thus, $I = \langle g \rangle$. $\square$

NOTE In the previous proof, we can take $g(x)$ to be monic, i.e. a polynomial whose leading coefficient is 1. This is because if $g(x) = g_0 + g_1x + \dots + g_tx^t \in I$ where $g_t \neq 0$, then $g_t^{-1}g(x) = g_0g_t^{-1} + g_1g_t^{-1}x + \dots + x^t \in I$ is monic.

DEFINITION Let I be an ideal of $R = F[x]/(x^n - 1)$.

- If $I = \{0\}$, then x^{n-1} is the canonical generator of I .
- If $I \neq \{0\}$, then the monic polynomial of smallest degree in I is the canonical generator of I .

- The following theorem justifies the qualifier "the" in the above definition.

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

- 1) There is a unique monic poly. $g(x)$ of smallest degree in I , and $I = \langle g(x) \rangle$.
- 2) $g(x) \mid (x^n - 1)$ in $F[x]$.

PROOF 1) Let $g(x), h(x)$ be monic polynomials of (the same) smallest degree in I . Then $g(x) - h(x) \in I$. But $\deg(g - h) < \deg(g)$. Hence $g(x) - h(x) = 0$, so $g(x) = h(x)$. This proves uniqueness of $g(x)$.

2) Write $x^n - 1 = l(x)g(x) + r(x)$, where $l, r \in F[x]$, $\deg(r) < \deg(g)$.

Then $r(x) = -l(x)g(x) + x^n - 1 \equiv -l(x)g(x) \pmod{x^n - 1}$.

Thus, $r(x) \in I = \langle g \rangle$, and so we must have $r(x) = 0$ since $\deg(r) < \deg(g)$. Hence $g(x) \mid (x^n - 1)$. $\square$

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

PROOF • If $h(x) = x^n - 1$, then $I = \{0\}$.

• Suppose that $h(x) \neq x^n - 1$, so $I \neq \{0\}$. Let $g(x)$ be the monic polynomial of smallest degree in I . Since $h(x)$ generates I , we can write $g(x) = a(x)h(x) \bmod (x^n - 1)$ for some $a(x) \in F[x]$, $\deg(a) < n$.

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

Since $h(x) \mid (x^n - 1)$, we have $h(x) \mid g(x)$, so $\deg(h) \leq \deg(g)$.

But $\deg(g) \leq \deg(h)$, so $\deg(g) = \deg(h)$.

Finally, since g and h are both monic, we have $g(x) = h(x)$.

Hence, $h(x)$ is the canonical generator of $\langle h(x) \rangle$. $\square$

COROLLARY There is a 1-1 correspondence between ideals of R and monic divisors of $x^n - 1$ over F , and thus also a 1-1 correspondence between cyclic subspaces of $V_n(F)$ and monic divisors of $x^n - 1$ over F .

NOTE (monic divisors of $x^n - 1$ over F)

Let $x^n - 1 = p_1(x)^{e_1} p_2(x)^{e_2} \cdots p_t(x)^{e_t}$ be the complete factorization of $x^n - 1$ over F , where $p_1, p_2, \dots, p_t$ are monic irreducible polynomials in $F[x]$, and $e_i \geq 1$. Then the set of all monic divisors of $x^n - 1$ over F is

$$\{ p_1(x)^{f_1} p_2(x)^{f_2} \cdots p_t(x)^{f_t} : 0 \leq f_i \leq e_i \}.$$

Hence, the number of monic divisors of $x^n - 1$ over F is

$$(e_1 + 1)(e_2 + 1) \cdots (e_t + 1).$$

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

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

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

So, the monic divisors of x^3-1 over $\mathbb{Z}_2$ are

$$g_1(x)=1, \quad g_2(x)=x+1, \quad g_3(x)=x^2+x+1, \quad g_4(x)=(x+1)(x^2+x+1).$$

Hence, $V_3(\mathbb{Z}_2)$ has 4 cyclic subspaces.

<u>CYCLIC SUBSPACE</u>	<u>DIMENSION</u>
$\langle g_1(x) \rangle = S_1 = \{000, 001, 010, 011, 100, 101, 110, 111\} = V_3(\mathbb{Z}_2)$	3
$\langle g_2(x) \rangle = S_2 = \{000, 110, 011, 101\}$	2
$\langle g_3(x) \rangle = S_3 = \{000, 111\}$	1
$\langle g_4(x) \rangle = S_4 = \{000\}$	0

V5C DIMENSION AND A GM OF A CYCLIC CODE

-115-

THEOREM Let $g(x)$ be a monic divisor of $x^n - 1$ over F , where $F = GF(q)$. Suppose $\deg(g) = n - k$. Then the cyclic subspace S of $V_n(F)$ generated by $g(x)$ has dimension k .

PROOF Recall that $\langle g(x) \rangle = \{a(x)g(x) \bmod (x^n - 1) : a(x) \in F[x], \deg(a) < n\}$. We claim that $\langle g(x) \rangle = \{b(x)g(x) : b(x) \in F[x], \deg(b) < k\}$. To see this, let $h(x) = a(x)g(x) \bmod (x^n - 1)$ for some $a(x) \in F[x], \deg(a) < n$. Then we can write $a(x)g(x) = h(x) + l(x)(x^n - 1)$ for some $l(x) \in F[x]$. Since $g(x) \mid (x^n - 1)$, we have $g(x) \mid h(x)$, so $h(x) = b(x)g(x)$ for some $b(x) \in F[x], \deg(b) < k$. This proves the claim. Finally since there are q^k polynomials of degree $< k$ in $F[x]$, $\langle g(x) \rangle$ has size q^k . Thus, S has dimension k . $\square$

EXAMPLE Construct a $(7,4)$ -cyclic code over $F = \mathbb{Z}_2$.

SOLUTION We need a monic divisor of x^7-1 over $\mathbb{Z}_2$ of degree 3.

Now, the complete factorization of x^7-1 over $\mathbb{Z}_2$ is

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

Choose $g(x) = x^3+x^2+1$.

Then $C = \langle x^3+x^2+1 \rangle$ is a $(7,4)$ -cyclic code over $\mathbb{Z}_2$.

Let's find a GM for C . We need a basis for C , i.e. 4 linearly independent codewords in C . We can choose $g(x)$, $xg(x)$, $x^2g(x)$, $x^3g(x)$ to get

$$G = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} \begin{array}{l} \leftarrow g(x) \\ \leftarrow xg(x) \\ \leftarrow x^2g(x) \\ \leftarrow x^3g(x) \end{array}$$

NOTE: C is systematic

EXAMPLE (cont'd) Encode the message $m = 1001$.

SOLUTION

$$c = mG = [1001] \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix} = \underline{1010011}.$$

$$\begin{aligned} \text{Equivalently, } c(x) &= m(x)g(x) \\ &= (1+x^3)(1+x^2+x^3) \\ &= 1+x^2+x^5+x^6 \\ &\leftrightarrow \underline{1010011}. \end{aligned}$$

A GM FOR A CYCLIC CODE

THEOREM Let $g(x)$ be the canonical generator of an (n,k) -cyclic code C over F (so $g(x)$ is a monic divisor of x^n-1 over F of degree $n-k$). Then a (non-standard) GM for C is

$$G = \begin{bmatrix} g(x) \\ xg(x) \\ x^2g(x) \\ \vdots \\ x^{k-1}g(x) \end{bmatrix}_{k \times n}.$$

ENCODING Source messages are the polynomials in $F[x]$ of degree $< k$. If $m(x) = m_0 + m_1x + \dots + m_{k-1}x^{k-1} \in F[x]$, then the encoding of m w.r.t. G is $c = [m_0, m_1, \dots, m_{k-1}]G = m_0g(x) + m_1xg(x) + \dots + m_{k-1}x^{k-1}g(x)$, so $c(x) = m(x)g(x)$. [NOTE: No reduction by x^n-1 is needed.]

SUMMARY

- $V_n(F) \longleftrightarrow R = F[x]/(x^n - 1)$
- $a = (a_0, a_1, \dots, a_{n-1}) \in V_n(F) \longleftrightarrow a(x) = a_0 + a_1x + \dots + a_{n-1}x^{n-1} \in R$
- C cyclic subspace of $V_n(F) \longleftrightarrow I$ ideal of R
- $\dim(C) = k \longleftrightarrow$ The canonical generator of I has degree $n-k$.
- Encoding: $c = mG \longleftrightarrow c(x) = m(x)g(x)$, where $G = \begin{bmatrix} g(x) \\ xg(x) \\ \vdots \\ x^{k-1}g(x) \end{bmatrix}_{k \times n}$
- $C^\perp$ is cyclic $\xleftrightarrow{\text{v5d}}$ The canonical generator is $h^*(x)$,
where $h(x) = (x^n - 1)/g(x)$.
- Syndrome w.r.t. $a \xleftrightarrow{\text{v5e}}$ $s(x) = a(x) \bmod g(x)$
particular PCM
- MISSING: Distance of a cyclic code.

V5d THE DUAL CODE OF A CYCLIC CODE

-120-

- Let C be an (n, k) -cyclic code over F with canonical generator $g(x)$.
- Let $g(x) = \underbrace{g_0 + g_1 x + \dots + g_{n-k} x^{n-k}}_{\neq 0} + \underbrace{g_{n-k+1} x^{n-k+1} + \dots + g_{n-1} x^{n-1}}_{=0}$.

DEFINITION The parity-check polynomial of C is $h(x) = (x^n - 1)/g(x)$.

- Let $h(x) = \underbrace{h_0 + h_1 x + \dots + h_k x^k}_{\neq 0} + \underbrace{h_{k+1} x^{k+1} + \dots + h_{n-1} x^{n-1}}_{=0}$.

- Let $a(x) = a_0 + a_1 x + \dots + a_{n-1} x^{n-1} = g(x) h(x) \bmod (x^n - 1)$.

Note that $a(x) = 0$. Equating coefficients of x^i for $0 \leq i \leq n-1$, gives:

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

- Thus, the vector $\mathbf{g} = (g_0, g_1, \dots, g_{n-1})$ is orthogonal to the vector $\bar{\mathbf{h}} = (h_{n-1}, h_{n-2}, \dots, h_1, h_0)$ and all its cyclic shifts.

A PCM FOR C

- It follows that all cyclic shifts of g are orthogonal to all cyclic shifts of $\bar{h}$, where $\bar{h}(x) = h_{n-1} + h_{n-2}x + \dots + h_2x^{n-3} + h_1x^{n-2} + h_0x^{n-1}$.
- Recall the following GM for C :

• DEFINE:

$$G = \begin{bmatrix} g_0 & g_1 & \dots & g_{n-k} & 0 & \dots & 0 \\ 0 & g_0 & g_1 & \dots & g_{n-k-1} & g_{n-k} & \dots & 0 \\ \vdots & \vdots & & & \vdots & & & \\ 0 & 0 & \dots & g_0 & g_1 & \dots & g_{n-k} \end{bmatrix}_{k \times n}$$

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

- From the above observation, we have $GH^T = 0$.

Thus, $C' \subseteq C^\perp$, where C' is the code generated by the rows of H . But $\text{rank}(H) = n-k$ (since $h_k = 1$), so $\dim(C') = n-k = \dim(C^\perp)$.

Hence, $C' = C^\perp$, and H is a (non-standard) PCM for C .

$C^\perp$ is CYCLIC

-122-

DEFINITION Let $h(x) = h_0 + h_1x + \dots + h_kx^k$ be a polynomial of degree k (so $h_k \neq 0$). The reciprocal polynomial of $h(x)$ is

$$h_R(x) = x^k h(1/x) = h_k + h_{k-1}x + \dots + h_1x^{k-1} + h_0x^k.$$

If $h_0 \neq 0$, we define $h^*(x) = h_0^{-1} h_R(x)$. [$h^*(x)$ is monic]

THEOREM Let C be an (n, k) -cyclic code over F with canonical generator $g(x)$. Let $h(x) = (x^n - 1)/g(x)$. Then $C^\perp$ is cyclic, with canonical generator $h^*(x)$.

PROOF We have $g(x)h(x) = x^n - 1$, so $g(\frac{1}{x})h(\frac{1}{x}) = (\frac{1}{x})^n - 1$. Multiplying both sides by x^n gives $x^{n-k}g(\frac{1}{x})x^k h(\frac{1}{x}) = -(x^n - 1)$. Hence $g_R(x)h_R(x) = -(x^n - 1)$, so $h_R(x) \mid (x^n - 1)$. Thus, $h^*(x)$ is a monic divisor of $x^n - 1$. We saw on slide 121 that the $(n, n-k)$ -cyclic code generated by $h_R(x)$ (and thus also by $h^*(x)$) is $C^\perp$. $\square$

V5e COMPUTING SYNDROMES

-123-

- Let C be an (n,k) -cyclic code over F with canonical generator $g(x)$.
- We will find a "nice" PCM for C .

1) Find a GM for C of the form $[R | I_k]$.

For $0 \leq i \leq k-1$, long division gives $x^{n-k+i} = l_i(x)g(x) + t_i(x)$, $\deg(t_i) < n-k$, $\deg(l_i) < k$.

Then, $x^{n-k+i} - t_i(x) = l_i(x)g(x) \in C$.

Thus, a GM for C is:

$$G_1 = \begin{bmatrix} \underbrace{-t_0(x)}_{n-k} + \underbrace{x^{n-k}}_k & \\ \underbrace{-t_1(x)}_{n-k} + \underbrace{x^{n-k+1}}_k & \\ \vdots & \\ \underbrace{-t_{k-1}(x)}_{n-k} + \underbrace{x^{n-1}}_k \end{bmatrix}_{k \times n} = \begin{bmatrix} \underbrace{-x^{n-k} \bmod g(x)}_{n-k} & \underbrace{I_k}_k \\ \underbrace{-x^{n-k+1} \bmod g(x)}_{n-k} & \\ \vdots & \\ \underbrace{-x^{n-1} \bmod g(x)}_{n-k} & \end{bmatrix}_{k \times n} = [R | I_k].$$

NOTE: $\text{rank}(G_1) = k$.

a) A (standard form) PCM for C is $H = [I_{n-k} \mid -R^T]$.

Note that $H^T = \begin{bmatrix} I_{n-k} \\ -R \end{bmatrix}$, so the rows of H^T (the columns of H)

are $x^0 \bmod g(x)$, $x^1 \bmod g(x)$, $\dots$, $x^{n-1} \bmod g(x)$.

THEOREM (computing syndromes) The syndrome of $r \in V_n(F)$ w.r.t. the above PCM is $s \in V_{n-k}(F)$, where $s(x) = r(x) \bmod g(x)$.

PROOF Let $r = (r_0, r_1, \dots, r_{n-1}) \in V_n(F)$. The syndrome of r is $s = Hr^T$.
Hence $s(x) = [r_0 x^0 \bmod g(x)] + [r_1 x^1 \bmod g(x)] + \dots + [r_{n-1} x^{n-1} \bmod g(x)]$
 $= (r_0 + r_1 x + \dots + r_{n-1} x^{n-1}) \bmod g(x)$
 $= r(x) \bmod g(x). \quad \square$

EXAMPLE Consider the $(15,9)$ -binary cyclic code C with canonical generator $g(x) = 1 + x + x^2 + x^3 + x^6$. Compute the syndrome of $r = (1100\ 1000\ 1110\ 000)$.

SOLUTION $r(x) = 1 + x + x^4 + x^8 + x^9 + x^{10}$.

$$\begin{array}{r}
 x^4 + x^3 + x^2 + x \quad r(x) \\
 \hline
 x^6 + x^3 + x^2 + x + 1 \quad g(x) \left\{ \begin{array}{l} x^{10} + x^9 + x^8 + x^4 + x + 1 \\ x^{10} + x^7 + x^6 + x^5 + x^4 \\ x^9 + x^8 + x^7 + x^6 + x^5 + x + 1 \\ x^9 + x^6 + x^5 + x^4 + x^3 \\ x^8 + x^7 + x^4 + x^3 + x + 1 \\ x^8 + x^5 + x^4 + x^3 + x^2 \\ x^7 + x^5 + x^2 + x + 1 \\ x^7 + x^4 + x^3 + x^2 + x \\ \hline x^5 + x^4 + x^3 + 1 \end{array} \right. \\
 \hline
 \end{array}$$

$s(x)$ →

$$\begin{array}{r}
 1001111 \left\{ \begin{array}{l} 11100010011 \\ 1001111 \\ \hline 1111100011 \\ 1001111 \\ \hline 110011011 \\ 1001111 \\ \hline 10100111 \\ 1001111 \\ \hline 111001 \end{array} \right. \\
 \hline
 \end{array}$$

Hence, $s = 100111$.

- The syndromes of a vector and its cyclic shifts are closely related.

THEOREM Let $\mathbf{r} \in V_n(F)$ with syndrome polynomial $S(x) = S_0 + S_1x + \dots + S_{n-k-1}x^{n-k-1}$.
 The syndrome of $x\mathbf{r}(x)$ is (i) $xS(x)$, if $S_{n-k-1} = 0$
 (ii) $xS(x) - S_{n-k-1}g(x)$, if $S_{n-k-1} \neq 0$.
 (i) is labeled "cyclic shift of $\mathbf{r}$ " with an arrow pointing to $x\mathbf{r}(x)$.
 (ii) is labeled "not cyclic shifts" with an arrow pointing to $x\mathbf{r}(x)$.

PROOF Since $\mathbf{r}(x)$ has syndrome $S(x)$, we have $\mathbf{r}(x) = l(x)g(x) + S(x)$ for some $l \in F[x]$. Hence, $x\mathbf{r}(x) = xl(x)g(x) + xS(x)$. Since $g(x) \mid (x^n - 1)$, $x\mathbf{r}(x)$ and $x\mathbf{r}(x) \bmod (x^n - 1)$ leave the same remainder upon division by $g(x)$.

(i) If $S_{n-k-1} = 0$, then $\deg(S) < n-k-1$, so $\deg(xS) < n-k$. Hence, $xS(x)$ is the (unique) remainder upon dividing $x\mathbf{r}(x)$ by $g(x)$.

(ii) If $S_{n-k-1} \neq 0$, then $x\mathbf{r}(x) = xl(x)g(x) + xS(x) - S_{n-k-1}g(x) + S_{n-k-1}g(x)$
 $= [xl(x) + S_{n-k-1}]g(x) + [xS(x) - S_{n-k-1}g(x)]$. Notice that $\deg(\bar{S}) < n-k$,
 so $\bar{S}(x)$ is the unique remainder $\bar{S}(x)$ upon dividing $x\mathbf{r}(x)$ by $g(x)$. $\square$

- So, given the syndrome s of t , we can easily compute the syndromes of cyclic shifts of t .

• EXAMPLE (continuing the example on slide 125)

$$g(x) = x^6 + x^3 + x^2 + x + 1 \leftrightarrow 1111001, \quad t = 1100\ 1000\ 1110\ 000, \quad s = 100111.$$

i	<u>$S_i(x) = \text{syndrome of } x^i t(x)$</u>
0	100111
1	101111
2	101011
3	101001
4	101000
5	010100
6	001010
7	000101
$\vdots$	$\vdots$

$$\begin{array}{r}
 0100111 \\
 1111001 \\
 \hline
 1011110
 \end{array}$$

~~1011110~~

V5f BURST ERROR CORRECTION

- Cyclic codes are good for correcting burst errors.

DEFINITION Let $e \in V_n(F)$. The cyclic burst length of e is the length of the shortest cyclic block of e that contains all its nonzero components.

EXAMPLE The cyclic burst length of $e = 0110100010$ is 7.

DEFINITION A linear code C is a t -cyclic burst error correcting code if all cyclic burst errors of length $\leq t$ are in different cosets of C , i.e. have different syndromes. The largest such t is called the cyclic burst error correcting capability of C .

EXAMPLE $g(x) = x^6 + x^3 + x^2 + x + 1$ is the canonical generator for a $(15, 9)$ -binary cyclic code C . In fact, C is a 3-cyclic burst error correcting code. To check this, we verify that all cyclic bursts of length ≤ 3 have different syndromes.

cyclic burst error	Syndrome	Integer representation	cyclic burst error	Syndrome	Integer representation
0	000006	0	x^{12}	010110	22
x^0	100006	32	x^{13}	001011	11
x^1	010000	16	x^{14}	111001	57
x^2	001000	8	$1+x$	110000	48
x^3	000106	4	$x(1+x)$	011000	24
x^4	000010	2	$\vdots$	$\vdots$	$\vdots$
x^5	000001	1	$x^{14}(1+x)$	011001	25
x^6	111100	60	$1+x+x^2$	111000	56
x^7	011110	30	$x(1+x+x^2)$	011100	28
x^8	001111	15	$\vdots$	$\vdots$	$\vdots$
x^9	111011	59	$x^{14}(1+x+x^2)$	001001	9
x^{10}	100001	33	$1+x^2$	101000	40
x^{11}	101100	44	$x(1+x^2)$	010100	20
			$\vdots$	$\vdots$	$\vdots$
			$x^{14}(1+x^2)$	101001	41

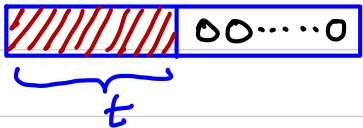
EXAMPLE $g(x) = x^8 + x^7 + x^6 + x^4 + 1$ is the canonical generator for a $(15, 7)$ -binary cyclic code C . C is a 4-cyclic burst error-correcting code.

THEOREM (bounds on t , the cyclic burst error correcting capability of C)
Let C be an (n, k, d) -code over $F = GF(q)$. Then $\lfloor (d-1)/2 \rfloor \leq t \leq n-k$.

PROOF • Recall that the vectors of weight $\leq \lfloor (d-1)/2 \rfloor$ lie in different cosets of C . In particular, all cyclic burst errors of length $\leq \lfloor (d-1)/2 \rfloor$ lie in different cosets of C . Thus, $t \geq \lfloor (d-1)/2 \rfloor$.

• No two cyclic burst errors of length $\leq t$ lie in the same coset of C .

In particular, no two vectors in which all the nonzero components are in the first t coordinate positions can lie in the same coset of C . Since there are q^t such vectors and q^{n-k} cosets, we must have $q^t \leq q^{n-k}$. Hence, $t \leq n-k$. $\square$



The diagram shows a vector represented as a sequence of components. The first t components are enclosed in a red hatched box, with a blue bracket underneath labeled t . The remaining components are zeros, represented as $00 \dots 0$, and are enclosed in a blue box.

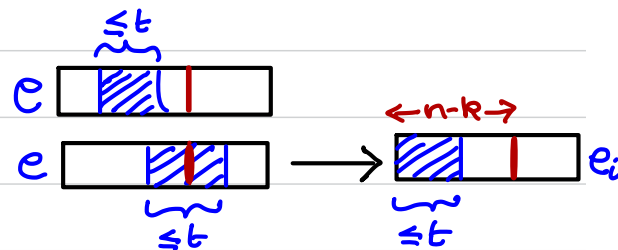
V5g DECODING ALG. FOR CYCLIC BURST ERROR CORRECTING CODES

- Let C be an (n, k) -cyclic code over F with canonical generator $g(x)$.
Let t be the cyclic burst error correcting capability of C , so $t \leq n - k$.
- Recall that $H = [I_{n-k} | -R^T]$ is a standard form PCM for C .
The syndrome of $r \in V_n(F)$ w.r.t. this PCM is $s(x) = r(x) \bmod g(x)$.

- IDEA OF DECODING ALGORITHM Suppose the error vector e is a cyclic burst of length $\leq t$. Then, some cyclic shift of e , say $e_i \leftrightarrow x^i e(x)$, has all its nonzero components in the first $n - k$ coordinate positions.

Then $s_i = H e_i^T$ has (non)-cyclic burst length $\leq t$, and $x = (s_i, 0)$ satisfies

$H x^T = s_i$. Thus, $e_i = (s_i, 0)$ and $e(x) = x^{n-i} e_i(x)$.



QUESTION How to compute s_i ? Let $r = c + e$, so $x^i r - x^i e = x^i c \in C$.

Thus, $x^i r$ and $x^i e$ have the same syndrome. So, we compute syndromes of $r_i \iff x^i r(x)$ for $0 \leq i \leq n-1$.

ERROR TRAPPING ALGORITHM FOR CYCLIC BURST ERROR CORRECTING CODES

- Let r be the received word.
- For i from 0 to $n-1$ do:
 - Compute $s_i(x)$, the syndrome of $x^i r(x)$.
 - If $s_i(x)$ has non-cyclic burst length $\leq t$ then
 - Let $e(x) = x^{n-i}(s_i, 0)$.
 - Decode r to $c = r - e$ and STOP.
- Reject r .

CORRECTNESS

If the error vector is in fact a cyclic burst error of length $\leq t$, then the algorithm will make the correct decision.

EXAMPLE Recall that $g(x) = x^6 + x^3 + x^2 + x + 1$ is the canonical generator for a $(15, 9)$ -binary cyclic code with cyclic burst error correcting capability $t=3$. Decode $r = 1110\ 1110\ 1100\ 000$ using error trapping.

SOLUTION

<u>i</u>	<u>$s_i(x) = \text{syndrome of } x^i r(x)$</u>
0	110011 $\leftarrow s(x) = r(x) \bmod g(x)$
1	100101
2	101110
3	010111
4	110111
5	100111
6	101111
7	101011
8	101001
9	<u>101000</u>

$\rightarrow$ burst of length ≤ 3

• So, $e(x) = x^{15-9} s_9(x)$
 $= x^6 (1+x^2)$
 $= \underline{0000\ 0010\ 1000\ 000}.$

• Decode r to $c = r - e$
 $= \underline{1110\ 1100\ 0100\ 000}$

• Check: $g(x) \mid c(x).$

INTERLEAVING

PURPOSE: Increase the cyclic burst error correcting capability of a code.

- Let C be an (n, k) -code with cyclic burst error correcting capability t .

Suppose $C_1 = (C_{11}, C_{12}, \dots, C_{1n}) \in C$,

$C_2 = (C_{21}, C_{22}, \dots, C_{2n}) \in C$,

$\vdots$

$C_s = (C_{s1}, C_{s2}, \dots, C_{sn}) \in C$.

- Interleaving to a depth s : Instead of sending $C_1, C_2, \dots, C_s$ in that order, transmit the columns of the above array:

$$C^* = (C_{11}, C_{21}, \dots, C_{s1} \mid C_{12}, C_{22}, \dots, C_{s2} \mid \dots \mid C_{1n}, C_{2n}, \dots, C_{sn}).$$

Then, any cyclic burst error of length $\leq st$ in C^* results in cyclic burst errors of length $\leq t$ in each of the original codewords $C_1, C_2, \dots, C_s$ (and these errors can be corrected).

THEOREM (interleaving) Let C be an (n, k) -code over F with cyclic burst error correcting capability t . Let C^* be the code obtained by interleaving C to a depth s .

- 1) C^* is an (ns, ks) -code over F with cyclic burst error correcting capability ts .
- 2) Suppose C is cyclic with canonical generator $g(x)$.
Then C^* is cyclic with canonical generator $g(x^s)$.

PROOF (sketch) 1) Show that C^* is a vector subspace, has length ns , size q^{ks} (so dimension ks), and cyclic burst error correcting ts .

2) Show that $g(x^s)$ is a monic divisor of $x^{ns}-1$ of degree $ns-ks$, and that $g(x^s) \mid c^*(x)$ for all $c^* \in C^*$. $\square$

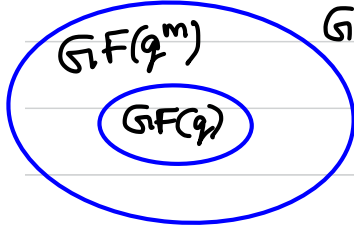
EXAMPLE (cf. slide 133) $g(x^{100}) = x^{600} + x^{300} + x^{200} + x^{100} + 1$ is the canonical generator of a $(1500, 900)$ -binary cyclic code with cyclic burst error correcting capability $t=300$.

V6a BCH CODES

-136-

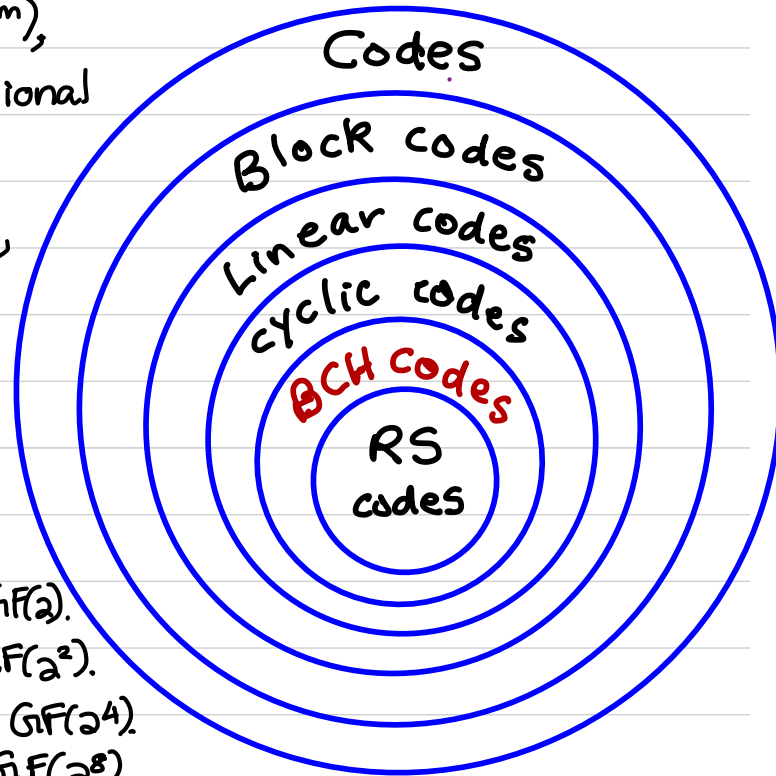
- Recall that $\mathbb{Z}_p$ is a subfield of $\text{GF}(p^m)$, and we can view $\text{GF}(p^m)$ as an m -dimensional vector space over $\mathbb{Z}_p$.

- More generally, for any prime power q , $\text{GF}(q)$ is a subfield of $\text{GF}(q^m)$, and we can view $\text{GF}(q^m)$ as an m -dimensional vector space over $\text{GF}(q)$.



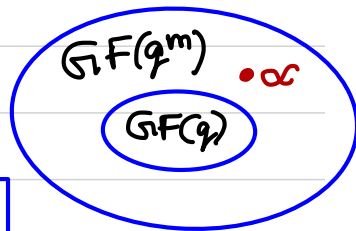
• EXAMPLE

- $\text{GF}(2^{16})$ is a 16-dimensional v.s. over $\text{GF}(2)$.
-)) » an 8-dimension v.s. over $\text{GF}(2^2)$.
-)) » a 4-dimensional v.s. over $\text{GF}(2^4)$.
-)) » a 2-dimensional v.s. over $\text{GF}(2^8)$.
-)) » a 1-dimensional v.s. over $\text{GF}(2^{16})$.



MINIMAL POLYNOMIALS

• We call $\text{GF}(q^m)$ an extension field and $\text{GF}(q)$ a subfield.



DEFINITION Let $\alpha \in \text{GF}(q^m)$. The minimal polynomial of α over $\text{GF}(q)$, denoted $m_\alpha(y)$, is the monic polynomial of smallest degree in $\text{GF}(q)[y]$ that has α as a root.

NOTES 1) If $m(y) \in \text{GF}(q)[y]$ is a nonzero polynomial with $m(\alpha) = 0$, and c is its leading coefficient, then $\bar{m}(y) = c^{-1}m(y)$ is a monic polynomial in $\text{GF}(q)[y]$ with $\bar{m}(\alpha) = 0$ and $\deg(\bar{m}) = \deg(m)$.

2) More generally, multiplying a polynomial by a nonzero constant does not change the roots of the polynomial.

3) We have $m_0(y) = y$.



4) If $\alpha \neq 0$, then let t be the order of α in $\text{GF}(q^m)$, and recall that $t \mid (q^m - 1)$. Then, α is a root of $y^t - 1 \in \text{GF}(q)[y]$. Hence, there does indeed exist a monic polynomial of smallest degree in $\text{GF}(q)[y]$ having α as a root.

EXAMPLE Let's find the minimal polynomials over $\text{GF}(2)$ of elements in $\text{GF}(2^2) = \mathbb{Z}_2[x]/(x^2 + x + 1) = \{0, 1, x, x+1\}$.

SOLUTION

- $m_0(y) = y.$
- $m_1(y) = y + 1,$
- $m_x(y) = y^2 + y + 1.$
- $m_{x+1}(y) = y^2 + y + 1.$

PROPERTIES OF MINIMAL POLYNOMIALS

THEOREM Let $\alpha \in \text{GF}(q^m)$.

- 1) The minimal polynomial $m_\alpha(y)$ of α over $\text{GF}(q)$ is unique.
- 2) $m_\alpha(y)$ is irreducible over $\text{GF}(q)$.
- 3) $\deg(m_\alpha) \leq m$.
- 4) If $f \in \text{GF}(q)[y]$, then $f(\alpha) = 0 \iff m_\alpha(y) \mid f(y)$.

PROOF 1) Suppose $m_1(y), m_2(y) \in \text{GF}(q)[y]$ are two monic polynomials of the same smallest degree with $m_1(\alpha) = m_2(\alpha) = 0$. Consider $r(y) = m_1(y) - m_2(y)$. Then $r(\alpha) = m_1(\alpha) - m_2(\alpha) = 0$. But $\deg(r) < \deg(m_1)$, so we must have $r(y) = 0$. Hence, $m_1(y) = m_2(y)$. $\square$

- PROOF OF 2) Suppose that $m_\alpha(y)$ is reducible over $\text{GF}(q)$, say $m_\alpha(y) = s(y)t(y)$ where $s, t \in \text{GF}(q)[y]$ and $1 \leq \deg(s), \deg(t) < \deg(m_\alpha)$. Then $m_\alpha(\alpha) = s(\alpha)t(\alpha) = 0$, so either $s(\alpha) = 0$ or $t(\alpha) = 0$. In either case we have a contradiction of the minimality of $\deg(m_\alpha)$. We conclude that $m_\alpha(y)$ is irreducible over $\text{GF}(q)$. $\square$

- PROOF OF 3) Recall that $\text{GF}(q^m)$ is an m -dimensional vector space over $\text{GF}(q)$. So, the field elements $1, \alpha, \alpha^2, \dots, \alpha^m$ are linearly dependent over $\text{GF}(q)$. Thus, we can write $a_0 + a_1\alpha + a_2\alpha^2 + \dots + a_m\alpha^m = 0$ for some $a_0, a_1, \dots, a_m \in \text{GF}(q)$ that are not all 0. Hence, α is a root of the nonzero polynomial $m(y) = a_0 + a_1y + a_2y^2 + \dots + a_my^m \in \text{GF}(q)[y]$ of degree $\leq m$. It follows that $\deg(m_\alpha) \leq m$. $\square$

• PROOF OF 4) Let $f(y) \in GF(q)[y]$. By long division, we can write $f(y) = l(y)m_\alpha(y) + r(y)$, where $l, r \in GF(q)[y]$ and $\deg(r) < \deg(m_\alpha)$.

Now, $f(\alpha) = l(\alpha)m_\alpha(\alpha) + r(\alpha) = r(\alpha)$ (since $m_\alpha(\alpha) = 0$).

Hence, $f(\alpha) = 0 \iff r(\alpha) = 0$

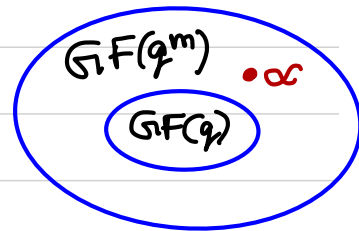
$\iff r(y) = 0$ (since $\deg(r) < \deg(m_\alpha)$)

$\iff m_\alpha(y) \mid f(y)$. □

V6b COMPUTING MINIMAL POLYNOMIALS

-142-

- We will show that the roots of $m_\alpha(y)$ are precisely the "conjugates" of α over $\text{GF}(q)$.
- We'll need the following result.



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

PROOF Since $\beta^q = \beta$ for all $\beta \in \text{GF}(q)$, the elements of $\text{GF}(q)$ are all roots of the polynomial $Y^q - Y \in \text{GF}(q)[Y]$. Since this polynomial has degree q , it can't have any other roots in $\text{GF}(q^m)$. Thus, $\alpha^q = \alpha$ iff $\alpha \in \text{GF}(q)$. $\square$

DEFINITION Let $\alpha \in \text{GF}(q^m)$. Let t be the smallest positive integer such that $\alpha^{q^t} = \alpha$ (note: $t \leq m$). Then the set of conjugates of α w.r.t. $\text{GF}(q)$ is $C(\alpha) = \{\alpha, \alpha^q, \alpha^{q^2}, \dots, \alpha^{q^{t-1}}\}$.

NOTE The t conjugates in $C(\alpha)$ are distinct. This is because if $\alpha^{q^i} = \alpha^{q^j}$ where $0 \leq i < j \leq t-1$, then $\alpha^{q^j} - \alpha^{q^i} = 0$, so $(\alpha^{q^{j-i}} - \alpha)^{q^i} = 0$. Hence $\alpha^{q^{j-i}} - \alpha = 0$, so $\alpha^{q^{j-i}} = \alpha$, which contradicts the minimality of t .

THEOREM Let $\alpha \in \text{GF}(q^m)$. Then the minimal polynomial of α over $\text{GF}(q)$ is $m(y) = \prod_{\beta \in C(\alpha)} (y - \beta) = (y - \alpha)(y - \alpha^q)(y - \alpha^{q^2}) \cdots (y - \alpha^{q^{t-1}})$.

PROOF i) Clearly, $m(y)$ is monic and $m(\alpha) = 0$.

ii) Let $f \in \text{GF}(q)[y]$, $f \neq 0$, with $f(\alpha) = 0$. Let's prove that $\deg(f) \geq t$.

Let $f(y) = \sum_{i=0}^d f_i y^i$. Then $f(\alpha^q) = \sum_{i=0}^d f_i \alpha^{iq} = \left(\sum_{i=0}^d f_i \alpha^i \right)^q = f(\alpha)^q = 0$.

So, $\alpha, \alpha^q, \alpha^{q^2}, \dots, \alpha^{q^{t-1}}$ are roots of f . Hence $\deg(f) \geq t$.



PROOF (cont'd)

iii) Let $m(y) = \prod_{\beta \in C(\alpha)} (y - \beta) = \sum_{i=0}^t m_i y^i$. Then $m(y) \in \text{GF}(q^n)[y]$.

We need to prove that $m(y) \in \text{GF}(q)[y]$.

$$\begin{aligned} \text{Now, } m(y)^q &= \prod_{\beta \in C(\alpha)} (y - \beta)^q = \prod_{\beta \in C(\alpha)} (y^q - \beta^q) = \prod_{\beta \in C(\alpha)} (y^q - \beta) \\ &= m(y^q) = \sum_{i=0}^t m_i y^{iq}. \quad (*) \end{aligned}$$

$$\text{Also, } m(y)^q = \left(\sum_{i=0}^t m_i y^i \right)^q = \sum_{i=0}^t m_i^q y^{iq}. \quad (**)$$

Comparing coefficients of y^{iq} of $(*)$ and $(**)$ yields $m_i^q = m_i$ for $0 \leq i \leq t$. Thus, $m_i \in \text{GF}(q)$, and so $m(y) \in \text{GF}(q)[y]$.

iv) We conclude that $m(y)$ is a monic polynomial of smallest degree in $\text{GF}(q)[y]$ that has α as a root. $\square$

EXAMPLE Consider $\text{GF}(2^4) = \mathbb{Z}_2[x]/(x^4+x+1)$. Find the minimal polynomial of $\beta = x^3+x^2$ over $\mathbb{Z}_2$. (Here, $q=2, m=4$).

SOLUTION When doing computations by hand, it's useful to have a generator α of $\text{GF}(2^4)^*$, and a table of powers of α . It turns out that $\alpha=x$ is a generator of $\text{GF}(2^4)^*$. Now, $\beta = \alpha^6$. Hence $C(\beta) = \{\alpha^6, \alpha^{12}, \alpha^9, \alpha^3\}$ (so $t=4$).

$$\begin{aligned} \text{Thus, } m_\beta(y) &= (y-\alpha^6)(y-\alpha^{12})(y-\alpha^9)(y-\alpha^3) \\ &= [y^2 - (\alpha^6 + \alpha^{12})y + \alpha^{18}][y^2 - (\alpha^9 + \alpha^3)y + \alpha^{12}] \\ &= [y^2 + \alpha^4 y + \alpha^3][y^2 + \alpha y + \alpha^{12}] \\ &= y^4 + (\alpha + \alpha^4)y^3 + (\alpha^{12} + \alpha^5 + \alpha^3)y^2 + (\alpha^{16} + \alpha^4)y + \alpha^{15} \\ &= \underline{y^4 + y^3 + y^2 + y + 1}. \end{aligned}$$

$$\begin{aligned} \alpha^0 &= 1 \\ \alpha^1 &= \alpha \\ \alpha^2 &= \alpha^2 \\ \alpha^3 &= \alpha^3 \\ \alpha^4 &= \alpha + 1 \\ \alpha^5 &= \alpha^2 + \alpha \\ \alpha^6 &= \alpha^3 + \alpha^2 \\ \alpha^7 &= \alpha^3 + \alpha + 1 \\ \alpha^8 &= \alpha^2 + 1 \\ \alpha^9 &= \alpha^3 + \alpha \\ \alpha^{10} &= \alpha^2 + \alpha + 1 \\ \alpha^{11} &= \alpha^3 + \alpha^2 + \alpha \\ \alpha^{12} &= \alpha^3 + \alpha^2 + \alpha + 1 \\ \alpha^{13} &= \alpha^3 + \alpha^2 + 1 \\ \alpha^{14} &= \alpha^3 + 1 \\ \alpha^{15} &= 1 \end{aligned}$$

V6C FACTORIZING $x^n - 1$ OVER $\text{GF}(q)$ [Part 1]

-146-

GOAL Describe the factorization of $x^n - 1$ over $\text{GF}(q)$. From this, we will see how canonical generators $g(x)$ can be chosen so that we have a non-trivial lower bound on the distance of the cyclic code generated by $g(x)$.

PRELIMINARIES Let p be the characteristic of $\text{GF}(q)$. If $\gcd(n, q) \neq 1$, then write $n = \bar{n} p^l$ where $l \geq 1$ and $\gcd(\bar{n}, q) = 1$.
Then $x^n - 1 = x^{\bar{n} p^l} - 1 = (x^{\bar{n}} - 1)^{p^l}$. So, wLOG, we shall assume $\gcd(n, q) = 1$.

NOTATION Let m be the smallest positive integer such that $q^m \equiv 1 \pmod{n}$, i.e. $n \mid (q^m - 1)$. [FACT: such an m exists.]

Let α be a generator of $\text{GF}(q^m)^*$.

Let $\beta = \alpha^{(q^m - 1)/n}$, and note that $\beta \in \text{GF}(q^m)$.

- Also, $\text{ord}(\beta) = n$, and so $1, \beta, \beta^2, \dots, \beta^{n-1}$ are distinct.

Furthermore, $(\beta^i)^n = (\beta^n)^i = 1^i = 1$ for each $0 \leq i \leq n-1$.

Hence, $1, \beta, \beta^2, \dots, \beta^{n-1}$ are roots of $x^n - 1$. So, the complete factorization of $x^n - 1$ over $\text{GF}(q^m)$ is $x^n - 1 = (x-1)(x-\beta)(x-\beta^2) \cdots (x-\beta^{n-1})$.

- However, we seek the factorization of $x^n - 1$ over $\text{GF}(q)$.

- Consider β^i , where $0 \leq i \leq n-1$.

Since β^i is a root of $x^n - 1$, we have $m_{\beta^i}(x) \mid (x^n - 1)$.

Also, the roots of $m_{\beta^i}(x)$ are $C(\beta^i) = \{\beta^i, \beta^{iq}, \beta^{iq^2}, \dots, \beta^{iq^{t-1}}\}$, where t is the smallest positive integer such that

$$iq^t \equiv i \pmod{n}.$$

CYCLOTOMIC COSETS

- The discussion on the previous slides motivates the following definition.

DEFINITION Suppose that $\gcd(n, q) = 1$, and let $0 \leq i \leq n-1$. The cyclotomic coset of $q \bmod n$ containing i is

$$C_i = \{i, iq \bmod n, iq^2 \bmod n, \dots, iq^{t-1} \bmod n\},$$

where t is the smallest positive integer such that $iq^t \equiv i \pmod{n}$.

Also, $C = \{C_i : 0 \leq i \leq n-1\}$ is the set of cyclotomic cosets of $q \bmod n$.

- EXAMPLE** The cyclotomic cosets of $2 \bmod 15$ ($q=2, n=15$) are:

$$C_0 = \{0\}, \quad C_1 = \{1, 2, 4, 8\} = C_2 = C_4 = C_8,$$

$$C_3 = \{3, 6, 12, 9\} = C_6 = C_{12} = C_9, \quad C_5 = \{5, 10\} = C_{10},$$

$$C_7 = \{7, 14, 13, 11\} = C_{14} = C_{13} = C_{11}. \quad \text{Hence } C = \{C_0, C_1, C_3, C_5, C_7\}.$$

• As the example suggests, if $j \in C_i$, then $C_j = C_i$.

• NOTE: $m_{\beta^i}(x) = (x - \beta^i)(x - \beta^{iq})(x - \beta^{iq^2}) \cdots (x - \beta^{iq^{t-1}}) = \prod_{j \in C_i} (x - \beta^j)$

is a monic irreducible factor of $x^n - 1$ over $\text{GF}(q)$ of degree $|C_i|$.

• This proves the following theorem.

THEOREM Suppose that $\gcd(n, q) = 1$.

- 1) The number of monic irreducible factors of $x^n - 1$ over $\text{GF}(q)$ is equal to the number of (distinct) cyclotomic cosets of $q \bmod n$.
- 2) The number of monic irreducible factors of degree d is equal to the number of (distinct) cyclotomic cosets of $q \bmod n$ of size d .

V6d FACTORIZING $x^n - 1$ OVER $GF(q)$ [Part 2]

THEOREM Suppose $\gcd(n, q) = 1$. Let m be the smallest positive integer such that $q^m \equiv 1 \pmod{n}$, and let $\beta \in GF(q^m)$ be an element of order n . Then the monic irreducible factors of $x^n - 1$ over $GF(q)$ are $\{m_{\beta^i}(x) : 0 \leq i \leq n-1\}$, where

$$m_{\beta^i}(x) = \prod_{j \in C_i} (x - \beta^j).$$

NOTE: If $j \in C_i$, then $m_{\beta^j}(x) = m_{\beta^i}(x)$.

EXAMPLE Factor $x^{15}-1$ over $\mathbb{Z}_2$. (Here $q=2$, $n=15$)

SOLUTION • The cyclotomic cosets of 2 mod 15 are $C_0=\{0\}$, $C_1=\{1,2,4,8\}$, $C_3=\{3,6,12,9\}$, $C_5=\{5,10\}$, $C_7=\{7,14,13,11\}$. So, $x^{15}-1$ has 5 irreducible factors over $\mathbb{Z}_2$, one of degree 1, one of degree 2, and three of degree 4.

• The smallest m for which $2^m \equiv 1 \pmod{15}$ is $m=4$.

• We need an element $\beta \in \text{GF}(2^4)$ of order 15. ↗ see slide 145

Let's take $\beta = \alpha$, since α is a generator for $\text{GF}(2^4) = \mathbb{Z}_2[\alpha]/(\alpha^4 + \alpha + 1)$.

• We then compute: $m_{\beta^0}(x) = x+1$ $C_0 = \{0\}$

$$m_{\beta}(x) = x^4 + x + 1 \quad \text{color: blue} \quad C_1 = \{1, 2, 4, 8\}$$

See slide 145 ↗ $m_{\beta^3}(x) = x^4 + x^3 + x^2 + x + 1$ $C_3 = \{3, 6, 12, 9\}$

$$m_{\beta^5}(x) = x^2 + x + 1 \quad \text{color: blue} \quad C_5 = \{5, 10\}$$

$$m_{\beta^7}(x) = x^4 + x^3 + 1 \quad \text{color: blue} \quad C_7 = \{7, 14, 13, 11\}.$$

• Thus, $x^{15}-1 = (x+1)(x^2+x+1)(x^4+x+1)(x^4+x^3+1)(x^4+x^3+x^2+x+1)$.

EXAMPLE Determine the number of cyclic subspaces of $V_{90}(\mathbb{Z}_3)$.

SOLUTION • First, we observe that $x^{90}-1 = (x^{10}-1)^9$.

• To determine the factorization pattern of $x^{10}-1$ over $\mathbb{Z}_3$, we find the cyclotomic cosets of $q=3$ modulo $n=10$:

$$C_0 = \{0\}, \quad C_1 = \{1, 3, 9, 7\}, \quad C_2 = \{2, 6, 8, 4\}, \quad C_5 = \{5\}.$$

• Thus, $x^{90}-1 = [f_0 \cdot f_1 \cdot f_2 \cdot f_5]^9$, where $f_0, f_1, f_2, f_5 \in \mathbb{Z}_3[x]$ are monic irred. polynomials over $\mathbb{Z}_3$, and $\deg(f_0)=1$, $\deg(f_1)=4$, $\deg(f_2)=4$, $\deg(f_5)=1$.

• Hence, the number of monic factors of $x^{90}-1$ over $\mathbb{Z}_3$ is $10 \times 10 \times 10 \times 10 = \underline{10000}$.

This is also the number of cyclic subspaces of $V_{90}(\mathbb{Z}_3)$.

NOTE $f_i(x) = m_{\beta^i}(x)$, where β is an element of order 10 in $\text{GF}(3^4)$.

In fact, $x^{10}-1 = (x+1)(x+2)(x^4+x^3+x^2+x+1)(x^4+2x^3+x^2+2x+1)$.

V6e BCH CODES: DEFINITION

- Discovered in 1960 by R.C. Bose and D. Ray-Chaudhuri, and independently in 1959 by A. Hocquenghem.

- A BCH code is a cyclic code that is constructed so that a non-trivial lower bound is known on its distance.

SETUP Suppose $\gcd(n, q) = 1$.

- Let m be the smallest positive integer such that $q^m \equiv 1 \pmod{n}$.
- Let α be a generator of $\text{GF}(q^m)^*$, and let $\beta = \alpha^{(q^m - 1)/n}$ (so $\text{ord}(\beta) = n$).
- Let $m_{\beta^i}(x)$ denote the minimal polynomial of β^i over $\text{GF}(q)$, for $0 \leq i \leq n-1$. Recall that $m_{\beta^i}(x) \mid (x^n - 1)$.
- We will let $m_{\beta^i}(x) = m_{\beta^{i \bmod n}}(x)$ for $i \geq n$ (since $\beta^i = \beta^{i \bmod n}$).

DEFINITION A BCH code C over $GF(q)$ of length n and designed distance δ is a cyclic code of length n over $GF(q)$ with canonical generator

$$g(x) = \text{lcm} \{ m_{\beta^i}(x) : a \leq i \leq a + \delta - 2 \}, \text{ for some integer } a.$$

NOTES 1) $\text{lcm} \{ 3, 3, 5, 7, 7, 7, 11, 11 \} = 3 \times 5 \times 7 \times 11.$

2) Since each $m_{\beta^i}(x)$ is a monic irreducible factor of $x^n - 1$, it follows that $g(x)$ is a monic divisor of $x^n - 1$. Hence $g(x)$ is indeed the canonical generator for a cyclic code of length n over $GF(q)$.

3) Among the roots of $g(x)$ are the $\delta - 1$ consecutive powers of β :

$$\beta^a, \beta^{a+1}, \beta^{a+2}, \dots, \beta^{a+\delta-2}.$$

4) BCH bound : $d(C) \geq \delta$. [Proof in V6f]

5) If $a=1$, the BCH code is narrow-sense.

EXAMPLE (BCH code) Let $q=3$, $n=13$. Then $m=3$ since $3^3 \equiv 1 \pmod{13}$.

- Consider $\mathbb{GF}(3^3) = \mathbb{Z}_3[x]/(x^3 + 2x^2 + 1)$.
- Then α is a generator of $\mathbb{GF}(3^3)^*$ (see next slide).
- Also, $\beta = \alpha^2$ has order 13.
- Compute the cyclotomic cosets of $q=3 \pmod{n=13}$:

$$C_0 = \{0\} \text{ ————— } m_{\beta^0}(x) = x + 2.$$

$$C_1 = \{1, 3, 9\} \text{ ————— } m_{\beta^1}(x) = x^3 + 2x^2 + 2x + 2.$$

$$C_2 = \{2, 6, 5\} \text{ ————— } m_{\beta^2}(x) = x^3 + 2x + 2.$$

$$C_4 = \{4, 12, 10\} \text{ ————— } m_{\beta^4}(x) = x^3 + x^2 + x + 2.$$

$$C_7 = \{7, 8, 11\} \text{ ————— } m_{\beta^7}(x) = x^3 + x^2 + 2.$$



EXAMPLE (cont'd)

$$\begin{aligned}
 \bullet \ m_{\beta^2}(x) &= (x - \beta^2)(x - \beta^6)(x - \beta^5) \\
 &= (x - \alpha^4)(x - \alpha^{12})(x - \alpha^{10}) \\
 &= [(x^2 - (\alpha^4 + \alpha^{12})x + \alpha^{16})][x - \alpha^{10}] \\
 &= [x^2 + \alpha^{10}x + \alpha^{16}][x + \alpha^{23}] \\
 &= x^3 + (\alpha^{10} + \alpha^{23})x^2 + (\alpha^{16} + \alpha^{33})x + \alpha^{39} \\
 &= \underline{x^3 + 2x + 2}.
 \end{aligned}$$

Sample
minimal
polynomial
calculation

$$\begin{aligned}
 \bullet \ \text{Let } g(x) &= m_{\beta^0}(x) \cdot m_{\beta^1}(x) \cdot m_{\beta^2}(x) \\
 &= \underline{x^7 + x^6 + 2x^5 + x^4 + 2x + 2}.
 \end{aligned}$$

• The roots of $g(x)$ are $\beta^0, \beta^1, \beta^3, \beta^9, \beta^2, \beta^6, \beta^5$.

Among these roots are $\beta^0, \beta^1, \beta^2, \beta^3$, so

$$\delta = 5 \Rightarrow d \geq 5.$$

• Thus, $g(x)$ is the canonical generator for a (13, 6)-BCH code over $GF(3)$ of distance at least 5.

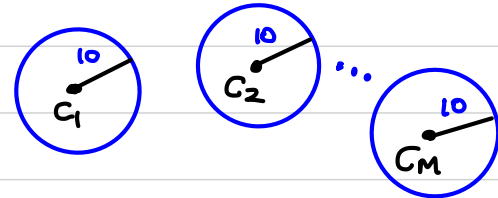
$\alpha^0 = 1$	$\alpha^{13} = 2$
$\alpha^1 = \alpha$	$\alpha^{14} = 2\alpha$
$\alpha^2 = \alpha^2$	$\alpha^{15} = 2\alpha^2$
$\alpha^3 = 2 + \alpha^2$	$\alpha^{16} = 1 + 2\alpha^2$
$\alpha^4 = 2 + 2\alpha + \alpha^2$	$\alpha^{17} = 1 + \alpha + 2\alpha^2$
$\alpha^5 = 2 + 2\alpha$	$\alpha^{18} = 1 + \alpha$
$\alpha^6 = 2\alpha + 2\alpha^2$	$\alpha^{19} = \alpha + \alpha^2$
$\alpha^7 = 1 + \alpha^2$	$\alpha^{20} = 2 + 2\alpha^2$
$\alpha^8 = 2 + \alpha + \alpha^2$	$\alpha^{21} = 1 + 2\alpha + 2\alpha^2$
$\alpha^9 = 2 + 2\alpha + 2\alpha^2$	$\alpha^{22} = 1 + \alpha + \alpha^2$
$\alpha^{10} = 1 + 2\alpha + \alpha^2$	$\alpha^{23} = 2 + \alpha + 2\alpha^2$
$\alpha^{11} = 2 + \alpha$	$\alpha^{24} = 1 + 2\alpha$
$\alpha^{12} = 2\alpha + \alpha^2$	$\alpha^{25} = \alpha + 2\alpha^2$
	$\alpha^{26} = 1$

V6f BCH BOUND

EXAMPLE Does there exist a block code with parameters $q=2$, $n=127$, $M=2^{64}$, $d \geq 21$? [slide 25]

The corresponding sphere packing problem is: Can we place $M=2^{64}$ spheres of radius $e = \lfloor \frac{21-1}{2} \rfloor = 10$ in $V_{127}(\mathbb{Z}_2)$ so that no two spheres overlap?

$V_{127}(\mathbb{Z}_2)$



SOLUTION YES! We will describe a BCH code with parameters $q=2$, $n=127$, $k=64$, $d=21$.

We have $m=7$ since $2^7 \equiv 1 \pmod{127}$.

EXAMPLE (cont'd) The cyclotomic cosets of 2 mod 127 are:

$$C_0 = \{0\}$$

$$C_1 = \{1, 2, 4, 8, 16, 32, 64\}$$

$$C_3 = \{3, 6, 12, 24, 48, 96, 65\}$$

$$C_5 = \{5, 10, 20, 40, 80, 33, 66\}$$

$$C_7 = \{7, 14, 28, 56, 112, 97, 67\}$$

$$C_9 = \{9, 18, 36, 72, 17, 34, 68\}$$

$$C_{11} = \{11, 22, 44, 88, 49, 98, 69\}$$

$$C_{13} = \{13, 26, 52, 104, 81, 35, 70\}$$

$$C_{15} = \{15, 30, 60, 120, 113, 99, 71\}$$

$$C_{19} = \{19, 38, 76, 25, 50, 100, 73\}$$

⋮

- Let β be an element of order 127 in $\text{GF}(2^7)^*$.
- Then $g(x) = m_\beta(x) \cdot m_{\beta^3}(x) \cdot m_{\beta^5}(x) \cdot m_{\beta^7}(x) \cdot m_{\beta^9}(x) \cdot m_{\beta^{11}}(x) \cdot m_{\beta^{13}}(x) \cdot m_{\beta^{15}}(x) \cdot m_{\beta^{19}}(x)$ is a degree-63 monic divisor of $x^{127} - 1$ over $\text{GF}(2)$.
- The roots of $g(x)$ include β^i , $1 \leq i \leq 20$.
- Thus, $g(x)$ is the canonical generator for a $(127, 64)$ -binary BCH code with designed distance $\delta = 21$ (so distance ≥ 21).

VANDER MONDE MATRICES

DEFINITION A Vandermonde matrix over a field F is a matrix of the form $A(x_1, x_2, \dots, x_t) = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_t \\ x_1^2 & x_2^2 & \dots & x_t^2 \\ \vdots & \vdots & & \vdots \\ x_1^{t-1} & x_2^{t-1} & \dots & x_t^{t-1} \end{bmatrix}_{t \times t}$, where $x_1, x_2, \dots, x_t \in F$.

THEOREM $\det(A(x_1, x_2, \dots, x_t)) \neq 0$ iff $x_1, x_2, \dots, x_t$ are distinct.

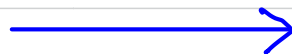
PROOF Perform the following row operations on A :

$$\left\{ \begin{array}{l} R_t \leftarrow R_t - x_1 R_{t-1} \\ \vdots \\ R_3 \leftarrow R_3 - x_1 R_2 \\ R_2 \leftarrow R_2 - x_1 R_1 \end{array} \right.$$

to get $A_1 =$

$$\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & x_2 - x_1 & x_3 - x_1 & \dots & x_t - x_1 \\ 0 & x_2^2 - x_1 x_2 & x_3^2 - x_1 x_3 & \dots & x_t^2 - x_1 x_t \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & x_2^{t-1} - x_1 x_2^{t-2} & x_3^{t-1} - x_1 x_3^{t-2} & \dots & x_t^{t-1} - x_1 x_t^{t-2} \end{bmatrix}.$$

Now compute $\det(A_1)$ by expanding along the first column:



PROOF (cont'd)

$$\det(A) = \det(A_1) = (x_2 - x_1)(x_3 - x_1) \cdots (x_t - x_1) \cdot \det$$

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_2 & x_3 & \cdots & x_t \\ x_2^2 & x_3^2 & \cdots & x_t^2 \\ \vdots & \vdots & & \vdots \\ x_2^{t-2} & x_3^{t-2} & \cdots & x_t^{t-2} \end{bmatrix}.$$

(t-1) × (t-1)

$$\text{By induction, } \det(A) = \prod_{1 \leq i < j \leq t} (x_j - x_i).$$

Thus, $\det(A) \neq 0$ iff $x_1, x_2, \dots, x_t$ are distinct. $\square$

COROLLARY The Vandermonde matrix $A(x_1, x_2, \dots, x_t)$ is non-singular iff $x_1, x_2, \dots, x_t$ are distinct.

THEOREM (BCH bound)

Let C be an (n, k) -BCH code over $GF(q)$ with designed distance δ . Then $d(C) \geq \delta$.

PROOF • Let $g(x)$ be the canonical generator for C . For simplicity, we'll assume that C is narrow-sense (so $a=1$). Hence,

$$g(x) = \text{lcm} \{ m_{\beta^i}(x) : 1 \leq i \leq \delta-1 \},$$

where $\beta \in GF(q^m)$ has order n .

• Now, let $r \in V_n(GF(q))$.

$$\text{Then } r \in C \iff g(x) \mid r(x) \iff m_{\beta^i}(x) \mid r(x) \iff r(\beta^i) = 0 \quad \forall 1 \leq i \leq \delta-1.$$



PROOF (cont'd)

• Let $H_1 = \begin{bmatrix} 1 & \beta & \beta^2 & \dots & \beta^{n-1} \\ 1 & \beta^2 & (\beta^2)^2 & \dots & (\beta^2)^{n-1} \\ 1 & \beta^3 & (\beta^3)^2 & \dots & (\beta^3)^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \beta^{\delta-1} & (\beta^{\delta-1})^2 & \dots & (\beta^{\delta-1})^{n-1} \end{bmatrix}_{(\delta-1) \times n}$.

• Now, $r \in C \Leftrightarrow r(\beta^i) = 0 \forall 1 \leq i \leq \delta-1 \Leftrightarrow H_1 r^T = 0$.

• Furthermore, no $t = \delta-1$ columns of H_1 are linearly dependent over $GF(q^m)$ since

$$\det \begin{bmatrix} \beta^{i_1} & \beta^{i_2} & \dots & \beta^{i_t} \\ (\beta^2)^{i_1} & (\beta^2)^{i_2} & \dots & (\beta^2)^{i_t} \\ \vdots & \vdots & \ddots & \vdots \\ (\beta^{\delta-1})^{i_1} & (\beta^{\delta-1})^{i_2} & \dots & (\beta^{\delta-1})^{i_t} \end{bmatrix} = \beta^{i_1} \beta^{i_2} \dots \beta^{i_t} \cdot \det \begin{bmatrix} 1 & 1 & \dots & 1 \\ \beta^{i_1} & \beta^{i_2} & \dots & \beta^{i_t} \\ \vdots & \vdots & \ddots & \vdots \\ (\beta^{i_1})^{\delta-2} & (\beta^{i_2})^{\delta-2} & \dots & (\beta^{i_t})^{\delta-2} \end{bmatrix}$$

$$= \left(\prod_{j=1}^t \beta^{i_j} \right) \cdot \det \underbrace{(A(\beta^{i_1}, \beta^{i_2}, \dots, \beta^{i_t}))}_{\text{Vandermonde matrix}} \neq 0, \text{ since } \beta^{i_1}, \beta^{i_2}, \dots, \beta^{i_t} \text{ are distinct.}$$

Vandermonde matrix



PROOF (cont'd)

- Since $\text{GF}(q) \subseteq \text{GF}(q^m)$, we also have that no $\delta-1$ columns of H_1 are linearly dependent over $\text{GF}(q)$.
- Now, if $c \in C$, $c \neq 0$, $w(c) < \delta$, then $H_1 c^T = 0$ gives 0 as a non-trivial linear combination over $\text{GF}(q)$ of $\delta-1$ (or fewer) columns of H_1 , contradicting what we just proved.
- Hence, every nonzero codeword in C has weight $\geq \delta$.
- Thus, $d(C) \geq \delta$. $\square$

V6g EXAMPLES OF BCH CODES

-164-

- EXAMPLE #1 • Let $q=2$, $n=2^t-1$ where $t \geq 2$. Then $\gcd(n, q)=1$ and $m=t$.
- Let β be a generator of $GF(2^t)^*$.
 - The cyclotomic cosets of 2 mod n are $C_0=\{0\}$, $C_1=\{1, 2, 4, \dots, 2^{t-1}\}$,
 - Let $g(x) = m_p(x) = (x-\beta)(x-\beta^2)(x-\beta^4) \dots (x-\beta^{2^{t-1}})$.
 - Then $g(x)$ is the canonical generator for a $(2^t-1, 2^t-1-t)$ -binary BCH code C with designed distance $\delta=3$. So, $d(C) \geq 3$.

- A PCM for C must look like $H = \left[\begin{array}{c|c|c|c|c} | & | & | & | & \dots & | \end{array} \right]_{t \times (2^t-1)}$.
all nonzero vectors in $V_t(\mathbb{Z}_2)$

So, C is a cyclic binary Hamming code (and $d(C)=3$).

- Hence, all binary Hamming codes are cyclic, up to equivalence.

EXAMPLE #2 • Let $q=2$ and $n=23$. Then $m=11$ [$2^{11} \equiv 1 \pmod{23}$].

• The cyclotomic cosets of 2 mod 23 are:

$$C_0 = \{0\}, C_1 = \{1, 2, 4, 8, 16, 9, 18, 13, 3, 6, 12\}, C_5 = \{5, 10, 20, 17, 11, 22, 21, 19, 15, 7, 14\}.$$

$$\text{In fact } x^{23}-1 = (x+1)(x^{11}+x^9+x^7+x^6+x^5+x+1)(x^{11}+x^{10}+x^6+x^5+x^4+x^2+1).$$

• Let β be an element of order 23 in $GF(2^{11})^*$. Let $g(x) = m_\beta(x)$.

Then $g(x)$ is the canonical generator for a $(23, 12)$ -binary BCH code C of designed distance $\delta=5$. Hence $d(C) \geq 5$.

Furthermore, $g(x) \in C$, so $d(C) \leq 7$.

• FACT C is equivalent to the binary Golay code C_{23} .

So, C_{23} is equivalent to a cyclic code.

EXAMPLE #3 • Let $q=2$, $n=2^{16}-1=65535$. Then $\gcd(n,q)=1$ and $m=16$.

• Let $GF(2^{16}) = \mathbb{Z}_2[\alpha] / (\alpha^{16} + \alpha^5 + \alpha^3 + \alpha^2 + 1)$.

• FACT: α is a generator of $GF(2^{16})^*$. Let $\beta = \alpha$.

• The cyclotomic cosets of 2 mod 65535 are: $C_0 = \{0\}$,
 $C_1 = \{1, 2, 4, 8, 16, \dots\}$, $C_3 = \{3, 6, 12, 24, \dots\}$, $C_5 = \{5, 10, 20, \dots\}$, $C_7 = \{7, 14, \dots\}$,
 $C_9 = \{9, 18, \dots\}$, $C_{11} = \{11, 22, \dots\}$, $C_{13} = \{13, \dots\}$, $C_{15} = \{15, \dots\}$,
 $C_{17} = \{17, \dots\}$, $C_{19} = \{19, \dots\}$, $C_{21} = \{21, \dots\}$, $C_{23} = \{23, \dots\}$,

• FACT $C_1, C_3, C_5, \dots, C_{23}$ are distinct and have size 16.

• Let $g(x) = \prod_{i \in \{1, 3, 5, \dots, 23\}} m_{\beta^i}(x)$. Then $\deg(g) = 12 \times 16 = 192$.

• $g(x)$ is the canonical generator for a (65535, 65343)-binary
BCH code C with $\delta=25$, so $d(C) \geq 25$.



EXAMPLE #3 (cont'd)

THEOREM (Shortening a code) Let C be a systematic (n, k, d) -code over $\text{GF}(q)$, and let $t < k$. Let C' be the code obtained by "shortening" C in its first t coordinate positions, i.e. taking all codewords in C that have 0 in the first t coordinate positions, and then deleting those coordinates. Then C' is an $(n-t, k-t, d')$ -code over $\text{GF}(q)$ with $d' \geq d$.

PROOF Let $G = \left[\begin{array}{c|c} \begin{matrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{matrix} & A \end{array} \right]_{k \times n}$ be a standard-form G.M for C .

Let G' be the matrix obtained by deleting the first t rows of G , and then deleting the first t columns. Then G' is a $(k-t) \times (n-t)$ G.M for C' with $d' \geq d$. $\square$



EXAMPLE #3 (cont'd)

- C : $(65535, 65343)$ -binary BCH code with $d \geq 25$.

- C is systematic, since C has a G.M. $G =$

$$G = \begin{bmatrix} \boxed{g(x)} & & & \\ & \boxed{g(x)} & & \\ & & \boxed{g(x)} & \\ & & & \ddots \\ & & & & \boxed{g(x)} \end{bmatrix}.$$

- Consider the shortened code C' obtained by shortening C by $t=33135$.

- Then C' is a $(32400, 32208)$ -binary code with distance $d' \geq 25$.

- C' is used (together with an LDPC code) in the DVB-S2 standard for digital video broadcasting—satellite.

EXAMPLE #4 (QR codes)

- 1) • Let $q=2$ and $n=15$, so $\gcd(n,q)=1$ and $m=4$.
 • Let $GF(2^4) = \mathbb{Z}_2[\alpha]/(\alpha^4 + \alpha + 1)$ and $\beta = \alpha$. Then $\text{ord}(\beta) = 15$.
 • Let $g(x) = m_\beta(x) \cdot m_{\beta^3}(x) \cdot m_{\beta^5}(x) = x^{10} + x^8 + x^5 + x^4 + x^2 + x + 1$. [see slide 151]
 • Then $g(x)$ is the canonical generator for a (15,5)-binary BCH code C with $\delta=7$ (since $\beta, \beta^2, \beta^3, \beta^4, \beta^5, \beta^6$ are roots of $g(x)$). In fact, $d(C)=7$.
 • C is used in QR codes to encode the "format data". [There are 32 formats]

- 2) • $g(x) = x^{11} + x^9 + x^7 + x^6 + x^5 + x + 1$ is the canonical generator for a $(23,12,7)$ -BCH code B_{23} that is equivalent to C_{23} .
 • Let E_{23} be the set of even-weight codewords in B_{23} . Then E_{23} is a $(23,11,8)$ -binary cyclic code with canonical generator $g(x) \cdot (x+1)$.
 • Let S_{23} be the (18,6,8)-binary code obtained by shortening E_{23} by $t=5$.
 • S_{23} is used to encode the "version data". [There are 34 versions]



EXAMPLE #4 (cont'd)

3) A $(255, 231, 25)$ -RS code over $\text{GF}(2^8)$ is used for the payload, more precisely a shortened $(36, 12, 25)$ -code over $\text{GF}(2^8)$ ($t=219$), and a shortened $(37, 13, 25)$ -code over $\text{GF}(2^8)$ ($t=218$). [see slide 175]

DECODING BCH CODES

Several efficient algorithms have been designed for decoding BCH codes. We don't have the time to study them. Instead, we'll study a decoding algorithm for a family of BCH codes called Reed-Solomon (RS) codes.

V7a Reed-Solomon Codes

- Invented by Irving Reed and Gustave Solomon in 1960.

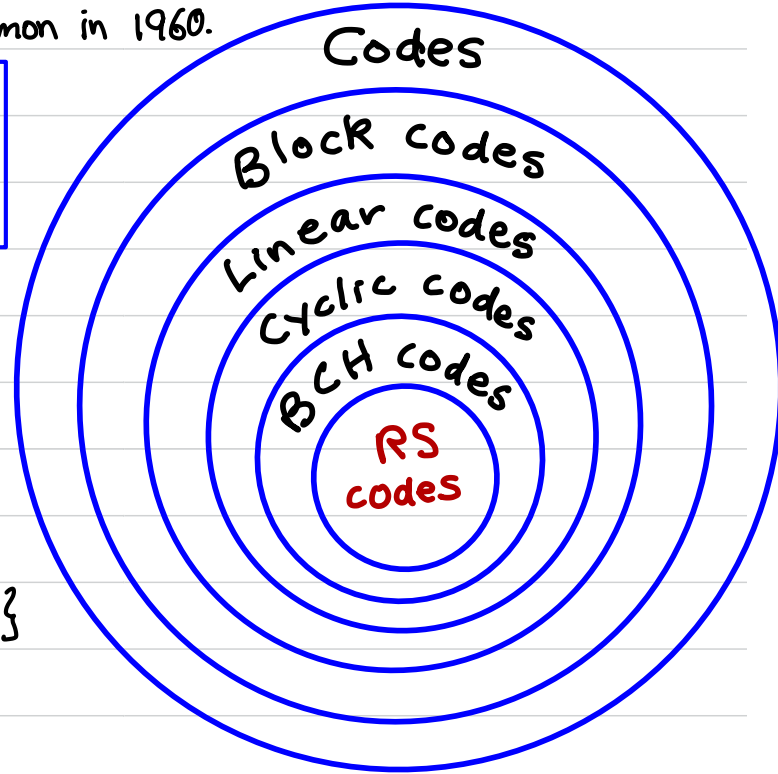
DEFINITION A Reed-Solomon (RS) code is a BCH code of length n over $\text{GF}(q)$ where $n \mid (q-1)$.

- **NOTE** Since $q^1 \equiv 1 \pmod{n}$, we have $m=1$.

- **EXAMPLE** Let $q=2^4$ and consider $\text{GF}(2^4) = \mathbb{Z}_2[\alpha]/(\alpha^4 + \alpha + 1)$. Recall that α is a generator of $\text{GF}(2^4)^*$. Let $\beta = \alpha^3$.

Then $\text{ord}(\beta) = 5$, so $q=16$, $n=5$, $m=1$.

$$\begin{aligned} \text{Let } g(x) &= \text{lcm}\{m_\beta(x), m_{\beta^2}(x), m_{\beta^3}(x)\} \\ &= (x - \beta)(x - \beta^2)(x - \beta^3) \\ &= x^3 + \alpha^{11}x^2 + \alpha^2x + \alpha^3. \end{aligned}$$



EXAMPLE (cont'd)

• Then $g(x)$ is the canonical generator for a $(5,2)$ -RS code C over $GF(2^4)$ with $\delta=4$. Since $w(g(x))=4$, we have $d(C)=4$.

• A GIM for C is $G = \begin{bmatrix} \alpha^3 & \alpha^2 & \alpha'' & 1 & 0 \\ 0 & \alpha^3 & \alpha^2 & \alpha'' & 1 \end{bmatrix}_{2 \times 5}$.

• Consider the code C' obtained from C by replacing each symbol in codewords in C by its binary representation.

eg. $(\alpha^3, \alpha^2, \alpha'', 1, 0) \leftrightarrow (0001 \ 0010 \ 0111 \ 1000 \ 0000)$.

• It is not difficult to see that C' is closed under addition and scalar multiplication over $\mathbb{Z}_2$. So, C' is a $(20,8)$ -binary code.

MORE GENERALLY....

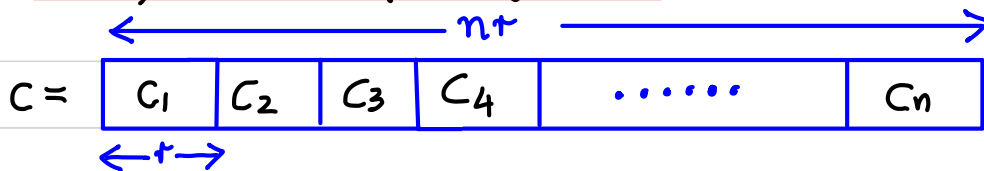
- Suppose $n \mid (q-1)$, and let $\beta \in \text{GF}(q)$ be an element of order n .
Then $m_{\beta^i}(x) = x - \beta^i$ for all i .
- ARS code C of length n over $\text{GF}(q)$ with designed distance δ is a BCH code of length n over $\text{GF}(q)$ with canonical generator $g(x) = (x - \beta^a)(x - \beta^{a+1})(x - \beta^{a+2}) \cdots (x - \beta^{a+\delta-2})$ for some a .
- Since $\deg(g) = \delta - 1$, we have $w(g) \leq \delta$, and so $d(C) \leq \delta$.
By the BCH bound, $d(C) \geq \delta$. Hence, $d(C) = \delta$.
- Since $\dim(C) = k = n - \deg(g) = n - \delta + 1$, we have $k = n - d + 1$, so $d = n - k + 1$.
Now, $d \leq n - k + 1$ for any (n, k, d) -code. Thus, RS codes are optimal in the sense that, for any fixed n, k, q with $n \mid (q-1)$, they achieve maximum distance among all (n, k) -codes over $\text{GF}(q)$.

RS CODES HAVE GOOD (CYCLIC) BURST ERROR CORRECTING CAPABILITY

- Let C be a RS code of length n over $GF(2^r)$ and designed distance δ .
- Consider $c = (c_1, c_2, \dots, c_n) \in C$, and note that $c_i \in GF(2^r)$.

Let $e = \lfloor (d-1)/2 \rfloor = \lfloor (n-k)/2 \rfloor$.

- By writing each c_i as a binary vector of length r , we can view c as a binary vector of length nr .



- Now, if c is transmitted and a cyclic burst of length $\leq 1 + (e-1)r$ bits is introduced, then at most e $GF(2^r)$ symbols of c are received incorrectly. Thus, the received word can be decoded correctly.

THEOREM Let C be an (n, k) -RS code over $GF(2^r)$. Then C' , the code obtained by replacing each symbol in codewords in C by its r -bit binary representation, is an (nr, kr) -binary code with cyclic burst error correcting capability $t = 1 + \left(\left\lfloor \frac{n-k}{2} \right\rfloor - 1\right)r$.

EXAMPLE Consider $GF(2^8) = \mathbb{Z}_2[\alpha]/(\alpha^8 + \alpha^4 + \alpha^3 + \alpha^2 + 1)$, so $q = 256$. Then $\beta = \alpha$ has order $n = 255$. Let $g(x) = \prod_{i=1}^{24} (x - \beta^i)$. Then $g(x)$ is

the canonical generator for a $(255, 231, 25)$ -RS code C over $GF(2^8)$ with error correcting capability $e = 12$. The related code C' is a $(2040, 1848)$ -binary code with cyclic burst error correcting capability $t = 89$. The code C , and others derived from it, have been widely deployed in practice, including in CDs, DVDs, Blu-rays, and QR codes.

V7b ALTERNATIVE VIEW OF NARROW-SENSE RS CODES

- Let $F = \text{GF}(q)$, let α be a generator of $\text{GF}(q)^*$, and let $2 \leq \delta < q$.

Let $g(x) = (x - \alpha)(x - \alpha^2)(x - \alpha^3) \cdots (x - \alpha^{\delta-1})$.

Then $g(x)$ is the canonical generator for a BCH-code C over $\text{GF}(q)$ of length $n = q - 1$, dimension $k = n - \delta + 1$, and distance $d = \delta = n - k + 1$.

- C is a narrow-sense RS code.

DEFINITION Let $f \in F[x]$. Then $\underline{c}(f) = (f(\alpha), f(\alpha^2), f(\alpha^3), \dots, f(\alpha^{q-2})) \in V_n(F)$.

DEFINITION Let $\underline{D} = \{ \underline{c}(f) : f \in F[x], \deg(f) \leq k-1 \} \subseteq V_n(F)$.

THEOREM (alternative view of RS codes) $C = \mathcal{D}$.

PROOF 1) $\mathcal{D}$ is a vector space over F . [exercise]

2) CLAIM $\dim(\mathcal{D}) = k$.

PROOF OF CLAIM Let $f, g \in F[x]$, $\deg(f) \leq k-1$, $\deg(g) \leq k-1$.

Suppose $c(f) = c(g)$. Then $f(\alpha^i) = g(\alpha^i)$ for $0 \leq i \leq q-2$, so $(f-g)(\alpha^i) = 0$ for $0 \leq i \leq q-2$. Thus, $f-g$ is a polynomial of degree $\leq k-1$ with $\geq q-1$ roots. But $k < q$ (recall: $k+d=q$).

So, we must have $f-g=0$, so $f=g$.

It follows that $|\mathcal{D}| = q^k$, so $\dim(\mathcal{D}) = k$. $\square$

PROOF (cont'd)

3) Let's prove that $D \subseteq C$. Let $c(f) \in D$, where $f(x) = \sum_{j=0}^{k-1} f_j x^j$.
 Then $c(f) = (f(1), f(\alpha), \dots, f(\alpha^{q-2})) \leftrightarrow C_f(x) = \sum_{i=0}^{n-1} f(\alpha^i) x^i$.

Now, for each $t, 1 \leq t \leq \delta-1$, we have

$$C_f(\alpha^t) = \sum_{i=0}^{n-1} \left(\sum_{j=0}^{k-1} f_j \alpha^{ij} \right) \alpha^{it} = \sum_{j=0}^{k-1} f_j \left(\sum_{i=0}^{n-1} \alpha^{(j+t)i} \right) = \sum_{j=0}^{k-1} f_j \left(\frac{\alpha^{(j+t)n} - 1}{\alpha^{j+t} - 1} \right).$$

Note that $\alpha^{j+t} \neq 1$, since $1 \leq j+t \leq k-1 + \delta-1 = n-1$, so $C_f(\alpha^t) = 0$.

Thus $g(x) \mid C_f(x)$, so $c(f) \in C$. Hence $D \subseteq C$.

4) Since $\dim(C) = \dim(D) = k$, we have $C = D$. $\square$

• Encoding $m = (m_0, m_1, \dots, m_{k-1}) \in F^k$: Let $f(x) = \sum_{j=0}^{k-1} m_j x^j$. Then $m \mapsto c = c(f)$.

V7C BERLEKAMP-WELCH DECODING ALGORITHM

-179-

RECALL $C = \{c(f) = (f(1), f(\alpha), \dots, f(\alpha^{n-1})) : f \in F[x], \deg(f) \leq k-1\}$.
 $n = q-1, d = n-k+1, e = \lfloor (n-k)/2 \rfloor$.

• Suppose $c = c(f)$ is transmitted and r is received. Suppose that $d(r, c) = t \leq e$. The objective is to determine c from r .

DEFINITION Let $\tilde{E}(x) = \left[\prod_{\substack{0 \leq i \leq n-1 \\ c_i \neq r_i}} (x - \alpha^i) \right] x^{e-t}$.

NOTE: $\tilde{E}$ is monic and $\deg(\tilde{E}) = e$.

Let $\tilde{Q}(x) = \tilde{E}(x)f(x)$.

NOTE: $\deg(\tilde{Q}) \leq e+k-1$.

THEOREM For all $0 \leq i \leq n-1$, $r_i \tilde{E}(\alpha^i) = \tilde{Q}(\alpha^i)$.

PROOF If $c_i \neq r_i$, then $\tilde{E}(\alpha^i) = 0$, so $r_i \tilde{E}(\alpha^i) = 0 = \tilde{Q}(\alpha^i)$.

If $c_i = r_i$, then $r_i = f(\alpha^i)$, so $r_i \tilde{E}(\alpha^i) = f(\alpha^i) \tilde{E}(\alpha^i) = \tilde{Q}(\alpha^i)$. $\square$

THEOREM Let $E(x), Q(x) \in F[x]$ with:

(i) E monic, $\deg(E) = e$,

(ii) $\deg(Q) \leq e+k-1$, and

(iii) $\tau_i E(\alpha^i) = Q(\alpha^i)$ for all $0 \leq i \leq n-1$.

Then $Q(x)/E(x) = \tilde{Q}(x)/\tilde{E}(x) = f(x)$.

PROOF Let $R(x) = \tilde{Q}(x)E(x) - Q(x)\tilde{E}(x)$. Then $\deg(R) \leq 2e+k-1$.

For all $0 \leq i \leq n-1$,

$$R(\alpha^i) = \tilde{Q}(\alpha^i)E(\alpha^i) - Q(\alpha^i)\tilde{E}(\alpha^i) = \tau_i \tilde{E}(\alpha^i)E(\alpha^i) - \tau_i E(\alpha^i)\tilde{E}(\alpha^i) = 0.$$

Thus, $R(x)$ has n roots, $1, \alpha, \alpha^2, \dots, \alpha^{n-1}$. But $e = \lfloor (n-k)/2 \rfloor < (n-k+1)/2$, so $\deg(R) \leq 2e+k-1 < (n-k+1) + k-1 = n$.

Hence, we must have $R(x) = 0$, so $\frac{Q(x)}{E(x)} = \frac{\tilde{Q}(x)}{\tilde{E}(x)} = f(x)$. $\square$

BERLEKAMP-WELCH DECODING (1986)INPUT: $\tau = (\tau_0, \tau_1, \tau_2, \dots, \tau_{n-1})$.

1. Find $E, Q \in F[x]$ with E monic, $\deg(E) = e$, $\deg(Q) \leq e + k - 1$, and $\tau_i E(\alpha^i) = Q(\alpha^i) \forall 0 \leq i \leq n-1$.
If no such E, Q exist, then reject τ and STOP.
2. Compute $f(x) = Q(x)/E(x)$.
(If $E(x) \nmid Q(x)$, then reject τ and STOP.)
3. Compute $c = c(f)$.
4. If $d(\tau, c) \leq e$ then return (c) ; else reject τ .

Q How to find E and Q ?A Linear algebra!

- There are $2e + k$ variables, the coefficients of E and Q .
- Each condition $\tau_i E(\alpha^i) = Q(\alpha^i)$ yields one linear equation, so there are n equations.

RUNNING TIME: Step 1 + Step 2 + Step 3 + Step 4 = $O(n^3)$.

$O(n^3)$ $O(n^2)$ $O(n^2)$ $O(n)$

The algorithm is efficient, i.e., polynomial time.

EXAMPLE • $q=11$, $n=10$, $\alpha=2$, $k=4$, $d=n-k+1=7$, $e=\lfloor (n-k)/2 \rfloor=3$.

• α is a generator of $\mathbb{Z}_{11}^*$:

$$\alpha^1=2, \alpha^2=4, \alpha^3=8, \alpha^4=5, \alpha^5=10, \alpha^6=9, \alpha^7=7, \alpha^8=3, \alpha^9=6, \alpha^{10}=1.$$

• $C = \{c(f) : f \in \mathbb{Z}_{11}[x], \deg(f) \leq 3\}$ is a $(10, 4, 7)$ -RS code over $\mathbb{Z}_{11}$.

• ENCODING Let $m = (3, 0, 7, 9) \longleftrightarrow f(x) = 3 + 7x^2 + 9x^3$.

Then $m \mapsto c = c(f) = (8, 4, 9, 10, 5, 1, 3, 1, 1, 10)$.

• TRANSMISSION c is sent, $r = (8, 4, \underline{0}, 10, \underline{6}, 1, \underline{3}, \underline{7}, 1, 10)$ is received.

• DECODING $E(x) = a_0 + a_1x + a_2x^2 + x^3$, $a_i \in \mathbb{Z}_{11}$.

$$\underline{Q(x)} = b_0 + b_1x + b_2x^2 + \dots + b_6x^6, \quad b_i \in \mathbb{Z}_{11}.$$

The conditions $r_i E(\alpha^i) = Q(\alpha^i)$, $0 \leq i \leq 9$, yield 10 equations:

$$i: \quad r_i a_0 + r_i \alpha^i a_1 + r_i \alpha^{2i} a_2 + \underbrace{r_i \alpha^{3i}}_t = b_0 + \alpha^i b_1 + \alpha^{2i} b_2 + \dots + \alpha^{6i} b_6.$$



EXAMPLE (cont'd)

i	a_0	a_1	a_2	t	b_0	b_1	b_2	b_3	b_4	b_5	b_6
0	8	8	8	8	1	1	1	1	1	1	1
1	4	8	5	10	1	2	4	8	5	10	9
2	0	0	0	0	1	4	5	9	3	1	4
3	10	3	2	5	1	8	9	6	4	10	3
4	6	8	7	2	1	5	3	4	9	1	5
5	1	10	1	10	1	10	1	10	1	10	1
6	3	5	1	9	1	9	4	3	5	1	9
7	7	5	2	3	1	7	5	2	3	10	4
8	1	3	9	5	1	3	9	5	4	1	3
9	10	5	8	4	1	6	3	7	9	10	5

• Solve to get $a_0=3, a_1=6, a_2=6,$
 $b_0=9, b_1=7, b_2=6, b_3=6,$
 $b_4=8, b_5=6, b_6=9.$

• $E(x) = x^3 + 6x^2 + 6x + 3.$

• $Q(x) = 9x^6 + 6x^5 + 8x^4 + 6x^3$
 $+ 6x^2 + 7x + 9.$

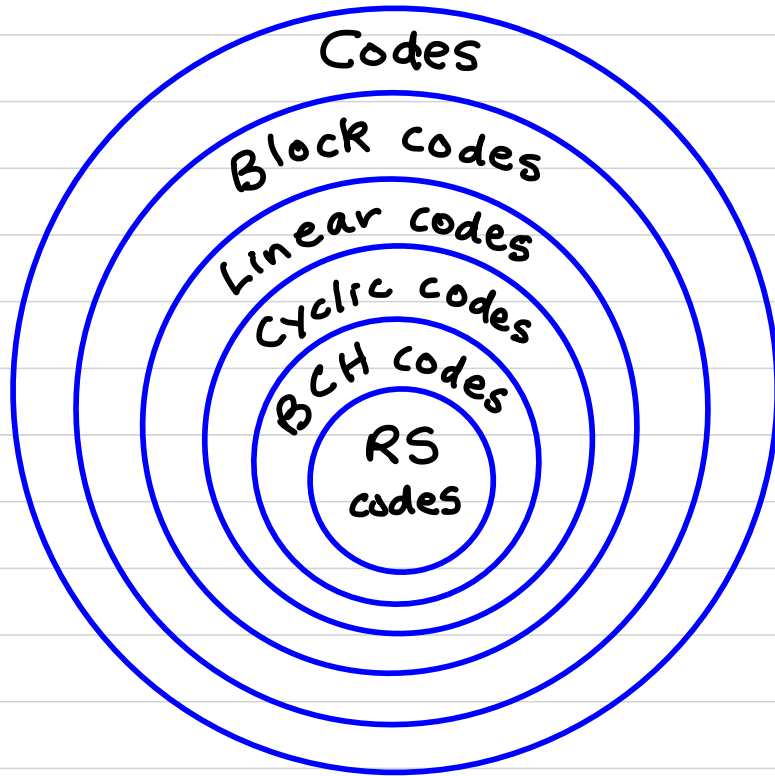
• $Q(x)/E(x) = 9x^3 + 7x^2 + 3.$

• $c = (8, 4, 9, 10, 5, 1, 3, 1, 1, 10),$
 and $d(r, c) = 3.$

• Thus, $m = (3, 0, 7, 9).$

V8 WRAP-UP

-184-



CODING THEORY 2

- LDPC codes
 - Low Density Parity Codes
 - Good for soft-decision decoding
 - Used in digital TV, wireless comm., optical communications, etc.
- Concatenation codes
 - BCH/RS + LDPC
- Turbo codes
- Raptor codes
- Polar codes
- Code-based cryptography

REED-SOLOMON CODES

- There is a faster decoding algorithm for RS (and BCH) codes, with running time $O(n \log n)$, discovered by Berlekamp and Massey in 1969.
- Berlekamp-Welch, discovered in 1986, is slower ($O(n^3)$), but can be adapted for L-list decoding of RS codes: Find all codewords c with $d(r, c) \leq L$. [Sudan 1996, Guruswami-Sudan 1998]
- RS codes are optimal erasure codes, so good for storage.
- Hardware architectures for encoding and decoding.

APPLICATIONS OF REED-SOLOMON CODES

- Communications
- Storage (CDs, DVDs, Blu-Ray, QR codes,....)
- Distributed storage
- Compressed sensing (signal processing)
- Group testing
- Learning theory
- Pseudorandomness
- Complexity theory
- Algorithm design

⋮

} newer
applications

"ESSENTIAL CODING THEORY" by Guruswami, Rudra and Sudan.

V9 LINEAR ALGEBRA REVIEW

- VECTOR SPACES (OVER FINITE FIELDS)
- MATRICES

VECTOR SPACES

- Let $F = GF(q)$.
- $V_n(F) = \underbrace{F \times F \times \dots \times F}_n = F^n = \{ (x_1, x_2, \dots, x_n) : x_i \in F \}$ is an n -dimensional vector space over F , where
 addition is $(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$,
 and scalar multiplication is $\lambda \cdot (x_1, x_2, \dots, x_n) = (\lambda x_1, \lambda x_2, \dots, \lambda x_n)$.
- A subset $C \subseteq V_n(F)$ is a vector subspace of $V_n(F)$ if:
 - (i) C is non-empty;
 - (ii) C is closed under addition: $\forall x, y \in C, x + y \in C$;
 - (iii) C is closed under scalar multiplication: $\forall x \in C, \forall \lambda \in F, \lambda \cdot x \in C$.
- Note: A vector subspace of $V_n(F)$ is also a vector space over F .

LINEAR INDEPENDENCE

- Let C be a vector space over $F = GF(q)$, and let $v_1, v_2, \dots, v_k \in C$.
- A linear combination of $v_1, v_2, \dots, v_k$ is a vector in C :

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k, \text{ where } \lambda_i \in F.$$

- The span of $\{v_1, v_2, \dots, v_k\}$ is $\text{Span}(v_1, \dots, v_k) = \{\lambda_1 v_1 + \dots + \lambda_k v_k : \lambda_i \in F\}$.

Note: $\text{Span}(v_1, \dots, v_k) \subseteq C$.

- The vectors $v_1, v_2, \dots, v_k$ are linearly independent over F if the only solution to $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k = 0$ is $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$; otherwise, the vectors are linearly dependent over F .

- If $v_1, v_2, \dots, v_k$ are l.i. over F , then $|\text{Span}(v_1, \dots, v_k)| = q^k$; otherwise, $|\text{Span}(v_1, \dots, v_k)| < q^k$.

BASES AND DIMENSION

- Let C be a vector space over $F = GF(q)$.
- A basis of C is a linearly independent set $\{v_1, v_2, \dots, v_k\} \subseteq C$, with $\text{Span}(v_1, v_2, \dots, v_k) = C$.
- FACTS
 - 1) C has a basis.
 - 2) All bases of C have the same size, called the dimension of C .
 - 3) If $\dim(C) = k$, then $|C| = q^k$.
- Note: $C = \{0\}$ is a 0-dimensional vector space over F , with basis $\{\}$.

MATRICES

• Addition and Subtraction: Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be $m \times n$ matrices over F . Then $A+B = [a_{ij}+b_{ij}]$ and $A-B = [a_{ij}-b_{ij}]$.

• Multiplication: Let $A_{m_1 \times n_1}$ and $B_{m_2 \times n_2}$ be matrices over F . Then $A \cdot B$ is only defined if $n_1 = m_2$, in which case $A \cdot B$ has dimensions $m_1 \times n_2$:

$$\begin{array}{c}
 \left[\begin{array}{c} \text{---} a_i \text{---} \end{array} \right] \\
 A: m_1 \times n_1
 \end{array}
 \begin{array}{c}
 \left[\begin{array}{c} | \\ b_j \\ | \end{array} \right] \\
 B: m_2 \times n_2
 \end{array}
 =
 \begin{array}{c}
 \left[\begin{array}{c} \boxed{c_{ij}} \end{array} \right] \\
 C: m_1 \times n_2
 \end{array}$$

EXAMPLE Let $C \subseteq V_n(F)$ be a vector space over F with basis $\{v_1, v_2, \dots, v_k\}$. Then C can be defined by the $k \times n$ matrix

$$G = \begin{bmatrix} \text{---} & v_1 & \text{---} \\ \text{---} & v_2 & \text{---} \\ & \vdots & \\ \text{---} & v_k & \text{---} \end{bmatrix}.$$

Note that $C = \{[m_1, m_2, \dots, m_k] \cdot G : [m_1, m_2, \dots, m_k] \in V_k(F)\}$
 $= \{m_1 v_1 + m_2 v_2 + \dots + m_k v_k : [m_1, m_2, \dots, m_k] \in V_k(F)\}$
 $= \text{row space of } G.$

RANK OF A MATRIX

- Let A be an $m \times n$ matrix over F .
- The row space of A is the vector subspace of $V_n(F)$ spanned by the rows of A . The row rank of A is the dimension of its row space.
- The column space of A is the vector subspace of $V_m(F)$ spanned by the columns of A . The column rank of A is the dimension of its column space.
- FACT : The row rank and column rank of A are equal, and is called the rank of A .

EXAMPLE Determine the dimension of the vector subspace C of $V_5(\mathbb{Z}_5)$ that is spanned by $v_1 = (3, 0, 2, 3, 1)$, $v_2 = (2, 1, 3, 4, 0)$, $v_3 = (0, 4, 3, 1, 1)$, and $v_4 = (2, 1, 1, 2, 2)$.

SOLUTION

$$\text{Let } G_1 = \begin{bmatrix} 3 & 0 & 2 & 3 & 1 \\ 2 & 1 & 3 & 4 & 0 \\ 0 & 4 & 3 & 1 & 1 \\ 2 & 1 & 1 & 2 & 2 \end{bmatrix}_{4 \times 5}.$$

Then $\dim(C) = \text{rank}(G_1)$.

Use elementary row operations to determine the (reduced) row echelon form of G_1 :

$$\begin{array}{c} R_1 \leftarrow 2R_1 \\ \hline \end{array} \rightarrow \begin{bmatrix} 1 & 0 & 4 & 1 & 2 \\ 2 & 1 & 3 & 4 & 0 \\ 0 & 4 & 3 & 1 & 1 \\ 2 & 1 & 1 & 2 & 2 \end{bmatrix} \quad \begin{array}{c} R_2 \leftarrow R_2 - 2R_1 \\ R_4 \leftarrow R_4 - 2R_1 \\ \hline \end{array}$$

$$\begin{bmatrix} 1 & 0 & 4 & 1 & 2 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 4 & 3 & 1 & 1 \\ 0 & 1 & 3 & 0 & 3 \end{bmatrix} \xrightarrow{\substack{R_3 \leftarrow R_3 - 4R_2 \\ R_4 \leftarrow R_4 - R_2}} \begin{bmatrix} 1 & 0 & 4 & 1 & 2 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 3 & 3 & 2 \\ 0 & 0 & 3 & 3 & 2 \end{bmatrix}$$

$$\xrightarrow{R_4 \leftarrow R_4 - R_3} \begin{bmatrix} 1 & 0 & 4 & 1 & 2 \\ 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 3 & 3 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Hence $\text{rank}(G) = 3$, so $\dim(C) = 3$.

NULL SPACE AND INVERTIBILITY

- Let A be an $m \times n$ matrix over F .
 - The null space of A is $\{x \in V_n(F) : Ax^T = 0\}$.
 - The null space of A is a vector subspace of $V_n(F)$; its dimension is called the nullity of A .
 - FACT $\text{rank}(A) + \text{nullity}(A) = n$.
-

- Let A be an $n \times n$ matrix over F .
 A is invertible (or non-singular) if there exists an $n \times n$ matrix B over F with $AB = BA = I_n$; we write $B = A^{-1}$.
- FACT A is invertible $\Leftrightarrow A$ has rank $n \Leftrightarrow \det(A) \neq 0$.