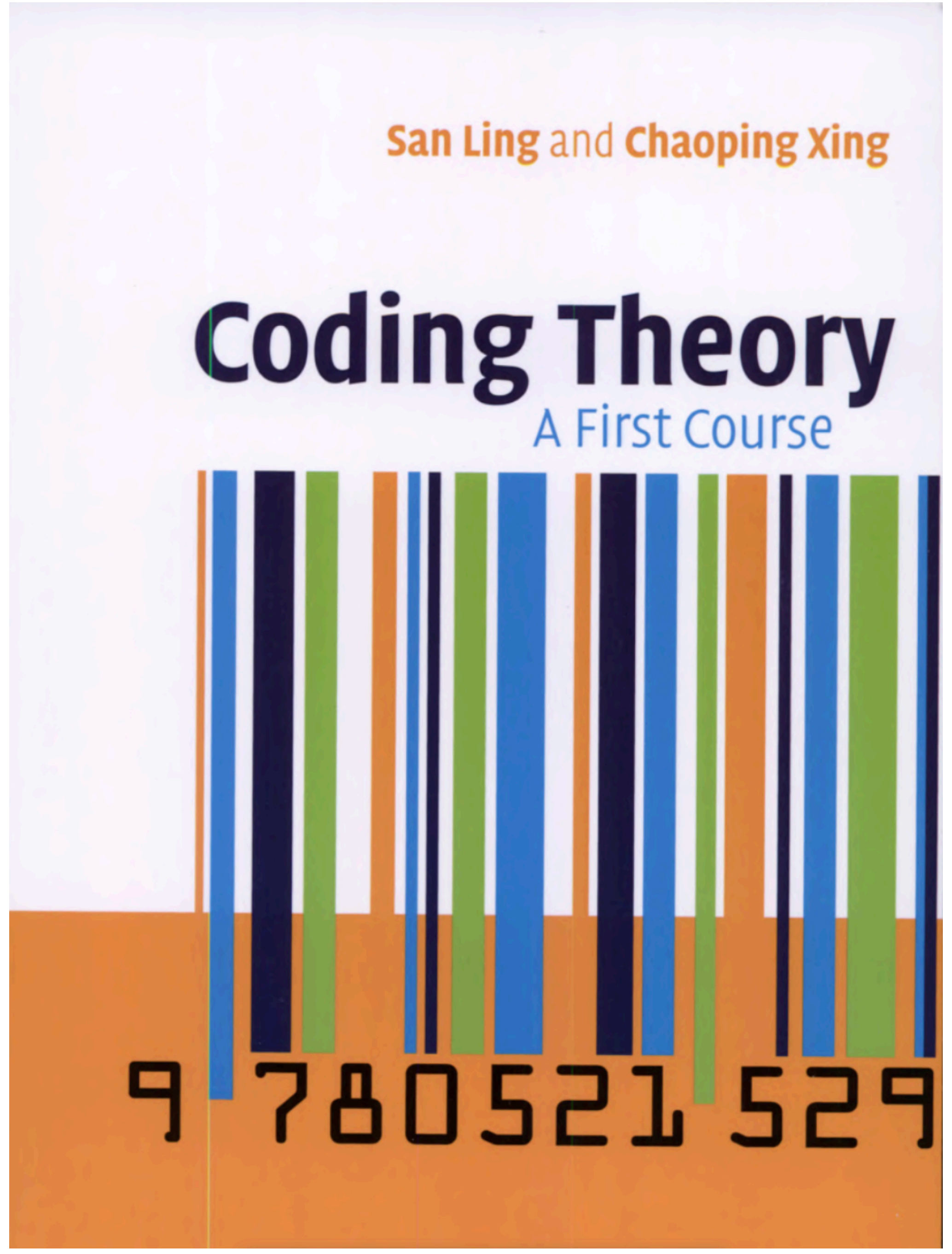
Coding Theory

Sahel Torkamani

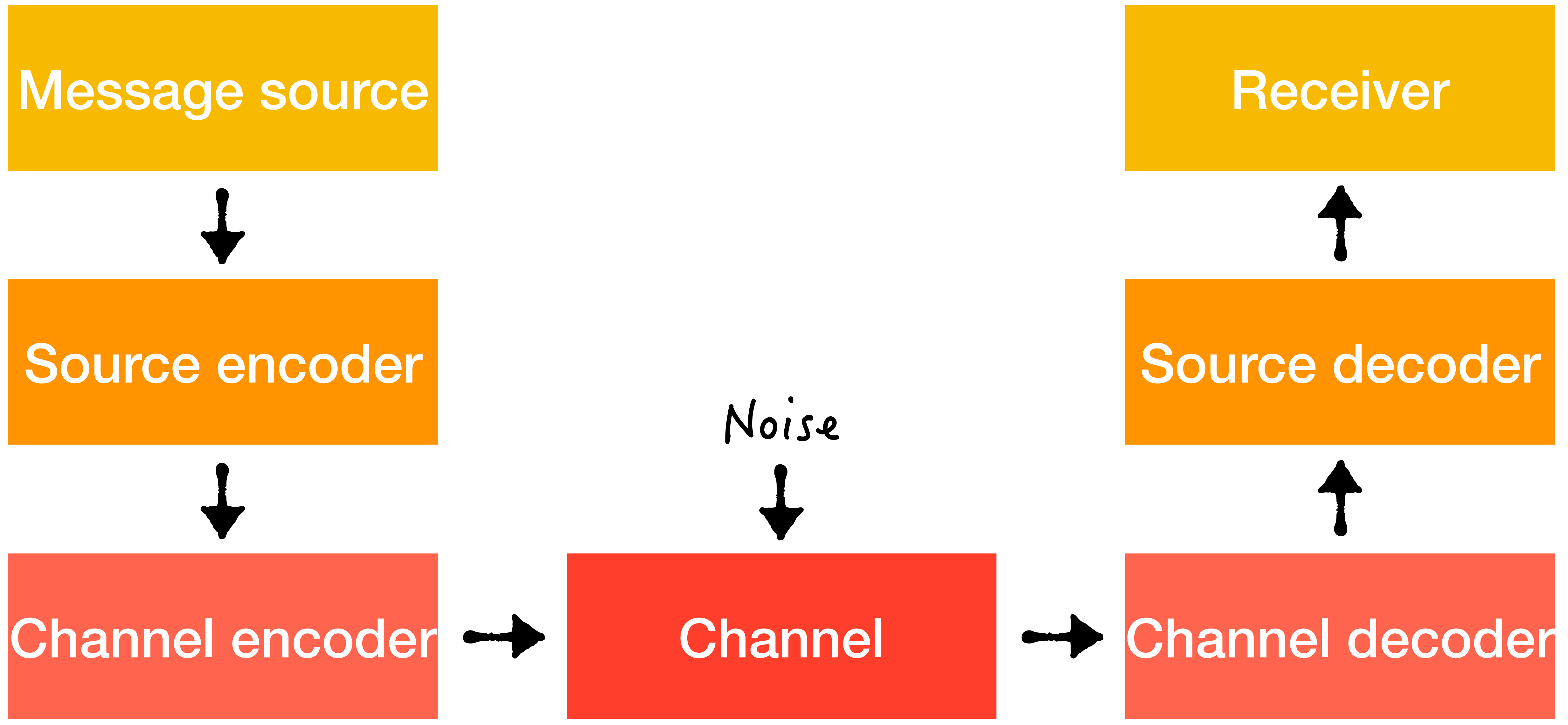
Contents

- Introduction
- Error detection, correction and decoding
- Linear codes



T. M. Thompson- From Error Correcting Codes Through Sphere Packings, to Simple Groups:
1.The origin of error correcting codes

Introduction



Source Coding:
Changing the message source to a suitable code

Channel Coding:
Encoding the message by introducing some form of redundancy

Introduction

Example: repetition code -> r+1 times for detecting r errors

Errors → Detect → Correct

The goal of channel coding:
(1) fast encoding of messages;
(2) easy transmission of encoded messages;
(3) fast decoding of received messages;
(4) **maximum transfer of information per unit time;**
(5) **maximal detection or correction capability.**

Error detection, correction and decoding

Definitions:

(1) **Code Alphabet**: $A = \{a_1, \ldots, a_q\}$ -> finite field F_q of order q

(2) **Code Symbols**: $a_1, \ldots, a_q$

(3) **q-ary word**: $W = w_1 \ldots w_n : w_i \in A$

(4) **q-ary block code**: non empty set C of q-ary words having the same length n.

(5) **Code word**: $c \in C$

(6) **Size of C**: $|C|$

(7) **Information rate** of a code C of length n: $(log_q|C|)/n$

(8) **(n, M)-code**: A code of length n and size M

+	0	1
0	0	1
1	1	0

x	0	1
0	0	0
1	0	1

Addition and multiplication tables for Z_2

Error detection, correction and decoding

Communication channel:

consists of a finite channel alphabet $A = \{a_1, \ldots, a_q\}$ as well as a set of forward channel probabilities $P(a_j \, received \mid a_i \, sent)$, satisfying:

$$\sum_{j=1}^{q} P(a_j \, received \mid a_i \, sent) = 1 \text{ for all } i.$$

Memoryless:

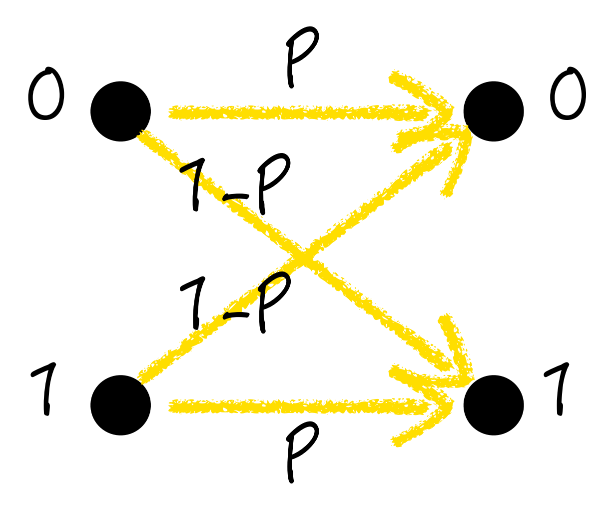
if $c = c_1 c_2 \ldots c_n$ and $x = x_1 x_2 \ldots x_n$ are words of length n, then $P(x \, received \mid c \, sent) = \Pi_{i=1}^{n} P(x_i \, received \mid c_i \, sent)$

q-ary symmetric channel:

a memoryless channel which has a channel alphabet of size q such that

(i) each symbol transmitted has the same probability p (<1/2) of being received in error;

(ii) if a symbol is received in error, then each of the q − 1 possible errors is equally likely.



Binary symmetric channel (BSC)

Error detection, correction and decoding

Maximum Likelihood decoding (MLD):

$$P(x \; received \,|\, c_x \; sent) = \max_{c \in C} P(x \; received \,|\, c \; sent)$$

(i) **Complete maximum likelihood decoding (CMLD):**

If a word **x** is received, find the most likely codeword transmitted.

If there are more than one such codewords, select one of them arbitrarily.

(ii) **Incomplete maximum likelihood decoding (IMLD):**

If a word **x** is received, find the most likely codeword transmitted.

If there are more than one such codewords, request a retransmission.

Error detection, correction and decoding

Hamming distance:

If $x = x_1x_2\ldots x_n$ and $y = y_1y_2\ldots y_n$ then $d(x, y) = d(x_1, y_1) + \ldots + d(x_n, y_n)$

$$\begin{cases} d(x_i, y_i) = 1 & \text{if } x_i \neq y_i \\ d(x_i, y_i) = 0 & \text{if } x_i = y_i \end{cases}$$

Nearest neighbour/minimum distance decoding (NND-MDD):

$$d(x, c_x) = \min_{c \in C} d(x, c)$$

(i) Complete nearest neighbour decoding (CNND)

(ii) Incomplete nearest neighbour decoding (INND)

Theorem 2.4.1 *For a BSC with crossover probability $p < 1/2$, the maximum likelihood decoding rule is the same as the nearest neighbour decoding rule.*

Proof. Let C denote the code in use and let $\mathbf{x}$ denote the received word (of length n). For any vector $\mathbf{c}$ of length n, and for any $0 \leq i \leq n$,

$$d(\mathbf{x}, \mathbf{c}) = i \Leftrightarrow \mathcal{P}(\mathbf{x} \text{ received} \mid \mathbf{c} \text{ sent}) = p^i(1-p)^{n-i}.$$

Since $p < 1/2$, it follows that

$$p^0(1-p)^n > p^1(1-p)^{n-1} > p^2(1-p)^{n-2} > \cdots > p^n(1-p)^0.$$

By definition, the maximum likelihood decoding rule decodes $\mathbf{x}$ to $\mathbf{c} \in C$ such that $\mathcal{P}(\mathbf{x} \text{ received} \mid \mathbf{c} \text{ sent})$ is the largest, i.e., such that $d(\mathbf{x}, \mathbf{c})$ is the smallest (or seeks retransmission if incomplete decoding is in use and $\mathbf{c}$ is not unique). Hence, it is the same as the nearest neighbour decoding rule.

Error detection, correction and decoding

Definitions:

(1) **Distdnce of C**: $d(C) = \min\{d(x, y) : x, y \in C, x \neq y\}$

(2) **(n, M, d)-code**: A code of length n and size M and distance d

(3) **u-error-detecting**: codeword incurs at least one but at most u errors <-> $d(C) \geq u + 1$

(4) **v-error-correcting**: minimum distance (incomplete) decoding is able to correct v or fewer errors <-> $d(C) \geq 2v + 1$

Linear Codes

Let F_q be the finite field of order q . A nonempty set V , together with some (vector) addition + and scalar multiplication by elements of F_q, is a vector space (or linear space) over F_q if it satisfies all of the following conditions. For all **u**, **v**, **w** ∈ V and for all λ, $\mu \in F_q$:

- (i) $u+v \in V$;
- (ii) $(u+v)+w = u+(v+w)$;
- (iii) there is an element $0 \in V$ with the property $0+v = v = v+0$ for all $v \in V$;
- (iv) for each $u \in V$ there is an element of V, called $-u$, such that $u+(-u)= 0 = (-u)+u$;
- (v) $u+v=v+u$;
- (vi) $\lambda v \in V$;
- (vii) $\lambda(u+v)=\lambda u+\lambda v, (\lambda+\mu)u=\lambda u+\mu u$;
- (viii) $(\lambda\mu)u = \lambda(\mu u)$;
- (ix) if 1 is the multiplicative identity of F_q, then $1u=u$.

Linear Codes

$C^{\perp} = \{x \in F_q \mid < x, y > = 0 \ \textit{for all}\ y \in C\}$

A linear code C of length n over F_q is a subspace of F_q^n

(1) **Dual code** of C: $C^{\perp}$, the orthogonal complement of the subspace C of F_q^n.
(2) **Dimension** of C: the dimension of C as a vector space over F_q . -> $\dim(C) = \log_q |C|$. ($|C| = q^{\dim(C)}$)
-> $\dim(C) + \dim(C^{\perp}) = n$

Example (1):
$C = \{0000,1010,0101,1111\}$
$\dim(C) = \log_2 |C| = \log_2 4 = 2$
$C^{\perp} = \{0000,1010,0101,1111\}$
$\dim(C^{\perp}) = \log_2 |C| = \log_2 4 = 2$

Example (2):
$C = \{000,001,002,010,020,011,012,021,022\}$
$\dim(C) = \log_3 |C| = \log_3 9 = 2$
$C^{\perp} = \{000,100,200\}$
$\dim(C^{\perp}) = \log_3 |C| = \log_3 3 = 1$

Linear Codes

Hamming weight: $wt(x) = d(x,0) = \sum_{i=1}^{n} wt(x_i)$

$$\begin{cases} wt(x_i) = 1 \;\; if\, x \neq 0 \\ wt(x_i) = 0 \;\; if\, x = 0 \end{cases}$$

-> Lemma: $d(x, y) = wt(x - y)$

The minimum (Hamming) weight: $wt(C) = \min_{0 \neq x \in C} wt(x)$

Theorem 4.3.8 *Let C be a linear code over $\mathbf{F}_q$. Then $d(C) = \mathrm{wt}(C)$.*

Proof. Recall that for any words $\mathbf{x}, \mathbf{y}$ we have $d(\mathbf{x}, \mathbf{y}) = \mathrm{wt}(\mathbf{x} - \mathbf{y})$.
By definition, there exist $\mathbf{x}', \mathbf{y}' \in C$ such that $d(\mathbf{x}', \mathbf{y}') = d(C)$, so

$$d(C) = d(\mathbf{x}', \mathbf{y}') = \mathrm{wt}(\mathbf{x}' - \mathbf{y}') \geq \mathrm{wt}(C),$$

since $\mathbf{x}' - \mathbf{y}' \in C$.
Conversely, there is a $\mathbf{z} \in C \backslash \{\mathbf{0}\}$ such that $\mathrm{wt}(C) = \mathrm{wt}(\mathbf{z})$, so

$$\mathrm{wt}(C) = \mathrm{wt}(\mathbf{z}) = d(\mathbf{z}, \mathbf{0}) \geq d(C).$$

Linear Codes

Bases for linear codes:

Input: A nonempty subset S of F_q^n.

Output: A basis for C = < S >, the linear code generated by S.

Algorithm 1:

Description: Form the matrix A whose rows are the words in S. Use elementary row operations to find an REF of A. Then the nonzero rows of the REF form a basis for C.

$$S = \{12101, 20110, 01122, 11010\}.$$

$$A = \begin{pmatrix} 12101 \\ 20110 \\ 01122 \\ 11010 \end{pmatrix} \to \begin{pmatrix} 12101 \\ 02211 \\ 01122 \\ 02212 \end{pmatrix} \to \begin{pmatrix} 12101 \\ 01122 \\ 00001 \\ 00000 \end{pmatrix}.$$

Linear Codes

Algorithm 2:

Description: Form the matrix A whose columns are the words in S. Use elementary row operations to put A in REF and locate the leading columns in the REF. Then the original columns of A corresponding to these leading columns form a basis for C.

$$S = \{11101, 10110, 01011, 11010\}.$$

$$A = \begin{pmatrix} 1101 \\ 1011 \\ 1100 \\ 0111 \\ 1010 \end{pmatrix} \to \begin{pmatrix} 1101 \\ 0110 \\ 0001 \\ 0111 \\ 0111 \end{pmatrix} \to \begin{array}{c} \begin{array}{cccc} 1 & 2 & & 4 \end{array} \\ \begin{pmatrix} 1&1&0&1 \\ 0&1&1&0 \\ 0&0&0&1 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix} \end{array}.$$

Linear Codes

Algorithm 3:

Description: Form the matrix A whose rows are the words in S. Use elementary row operations to place A in **RREF**. Let G be the k x n matrix consisting of all the nonzero rows of the **RREF**: $A \to \begin{pmatrix} G \\ 0 \end{pmatrix}$

The matrix G contains k leading columns. Permute the columns of G to form $G' = (I_k | X)$.

Form a matrix H' as follows: $H' = (-X^T | I_{n-k})$.

Apply the inverse of the permutation applied to the columns of G to the columns of H' to form H. Then the rows of H **form a basis for $C^{\perp}$**.

$$G = \begin{array}{c} \begin{array}{cccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{array} \\ \begin{pmatrix} 1 & 0 & 2 & 0 & 0 & 2 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 2 \end{pmatrix} \end{array}. \quad G' = (I_5|X) = \begin{array}{c} \begin{array}{cccccccccc} 1 & 4 & 5 & 7 & 9 & 2 & 3 & 6 & 8 & 10 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 2 & 2 & 1 & 2 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 2 \end{pmatrix} \end{array}.$$

$$H' = \begin{array}{c} \begin{array}{cccccccccc} 1 & 4 & 5 & 7 & 9 & 2 & 3 & 6 & 8 & 10 \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 2 & 0 & 2 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}, \quad H = \begin{array}{c} \begin{array}{cccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 2 & 0 & 0 & 2 & 0 & 1 & 1 \end{pmatrix} \end{array}.$$

Linear Codes

Generator matrix: Matrix G whose rows form a basis for C. -> Standard: $G = (I_k | X)$
Parity-check matrix: Matrix H is a generator matrix for the dual code $C^{\perp}$. -> Standard: $H = (Y | I_{n-k})$

Let C be a linear code and let H be a parity-check matrix for C. Then the following statements are equivalent:
(i) C has distance d;
(ii) any $d-1$ columns of H are linearly independent and H has d columns that are linearly dependent.

-> If $G = (I_k | X)$ is the standard form generator matrix of an [n,k]-code C, then a parity-check matrix for C is $H = (-X^T | I_{n-k})$.

Linear Codes

Two (n, M)-codes over F_q are **equivalent** if one can be obtained from the other by a combination of operations of the following types:
(i) permutation of the n digits of the codewords;
(ii) multiplication of the symbols appearing in a fixed position by a nonzero scalar.

-> Any linear code C is equivalent to a linear code C' with a generator matrix in standard form.

Example (1):
Let $q = 2$ and $n = 4$. Choosing to rearrange the bits in the order 2, 4, 1, 3, we see that the code
$C = \{0000, 0101, 0010, 0111\}$ is equivalent to the code
$C' = \{0000, 1100, 0001, 1101\}$.

Example (2):
Let $q = 3$ and $n = 3$. Consider the ternary code
$C = \{000, 011, 022\}$.
Permuting the first and second positions, followed by multiplying the third position by 2, we obtain the equivalent code
$C' = \{000, 102, 201\}$.

Linear Codes

Encoding with a linear code:

Let C be an [n, k, d]-**linear code** over the finite field F_q. Each codeword of C can represent one piece of information, so C can represent q^k distinct pieces of information. Once a basis $\{r_1, \ldots, r_k\}$ is fixed for C, each codeword v, or, equivalently, each of the q^k pieces of information, can be uniquely written as a linear combination, $v = u_1 r_1 + \ldots + u_k r_k = uG$. (G is the generator matrix of C whose ith row is the vector r_i in the chosen basis.)

-> The process of representing the elements u of F_q^k as codewords v = uG in C is called **encoding**.

Example (1):

Let C be the binary [5,3]-linear code with the generator matrix

$G = \begin{pmatrix} 10110 \\ 01011 \\ 00101 \end{pmatrix}$ -> for $u = 101$: $v = uG = (101)\begin{pmatrix} 10110 \\ 01011 \\ 00101 \end{pmatrix} = 10011$

Linear Codes

If an [n,k,d]-linear code C has a generator matrix G in standard form, $G = (I_k|X)$, then it is trivial to recover the message u from the codeword v = uG -> $v = uG = u(I|X) = (u, uX)$;
Message digits: the first k digits in the codeword **v** = **u**G give the message **u**.
Check digits: The remaining $n - k$ digits!

The check digits represent the **redundancy** which has been added to the message for protection against noise.

Linear Codes

Coset: (of C) determined by u to be the set: $u + C = C + u = \{v + u : v \in C\}$.

Coset leader: A word of the least (Hamming) weight in a coset.

Example (1):

Let $q = 2$ and $C = \{000, 101, 010, 111\}$. Then

$C + 000 = \{000, 101, 010, 111\}$,

$C + 001 = \{001, 100, 011, 110\}$,

$C + 010 = \{010, 111, 000, 101\}$,

$C + 011 = \{011, 110, 001, 100\}$,

$C + 100 = \{100, 001, 110, 011\}$,

$C + 101 = \{101, 000, 111, 010\}$,

$C + 110 = \{110, 011, 100, 001\}$,

$C + 111 = \{111, 010, 101, 000\}$.

Theorem 4.8.4 *Let C be an $[n, k, d]$-linear code over the finite field $\mathbf{F}_q$. Then,*

(i) *every vector of $\mathbf{F}_q^n$ is contained in some coset of C;*
(ii) *for all $\mathbf{u} \in \mathbf{F}_q^n$, $|C + \mathbf{u}| = |C| = q^k$;*
(iii) *for all $\mathbf{u}, \mathbf{v} \in \mathbf{F}_q^n$, $\mathbf{u} \in C + \mathbf{v}$ implies that $C + \mathbf{u} = C + \mathbf{v}$;*
(iv) *two cosets are either identical or they have empty intersection;*
(v) *there are q^{n-k} different cosets of C;*
(vi) *for all $\mathbf{u}, \mathbf{v} \in \mathbf{F}_q^n$, $\mathbf{u} - \mathbf{v} \in C$ if and only if $\mathbf{u}$ and $\mathbf{v}$ are in the same coset.*

Linear Codes

Assume the codeword v is transmitted and the word w is received,

Error pattern (Error string): $e = w - v \in w + C$.

Then $w - e = v \in C$, so the error pattern e and the received word w are in the same coset.

Nearest neighbour decoding:

Since error patterns of **small weight** are the most likely to occur, nearest neighbour decoding works for a linear code C in the following manner. Upon receiving the word w, we choose a word e of least weight in the coset $w + C$ and conclude that $v = w - e$ was the codeword transmitted.

Example (1):

Let $q = 2$ and $C = \{0000, 1011, 0101, 1110\}$.

Decode the following received words:

(i) $w = 1101$; (ii) $w = 1111$.

First, we write down the standard array of C

$C + 0000 = \{0000, 1011, 0101, 1110\}$,

$C + 0001 = \{0001, 1010, 0100, 1111\}$ -> (ii) $e = 0001$ or 0100 -> complete or incomplete ?!

$C + 0010 = \{0010, 1001, 0111, 1100\}$,

$C + 1000 = \{1000, 0011, 1101, 0110\}$ -> (i) $e = 1000$ -> $w-e=0101$

Linear Codes

Syndrome decoding: $\forall w \in F_q^n : S(w) = wH^T \in F_q^{n-k}$

(i) $S(u+v)=S(u)+S(v)$;

(ii) $S(u)=0$ if and only if u is a codeword in C;

(iii) $S(u)=S(v)$ if and only if u and v are in the same coset of C.

Syndrome look-up table (Standard Decoding Array (SDA).): table which matches each coset leader with its syndrome.

-> Steps to construct a syndrome look-up table assuming complete nearest neighbour decoding

Step 1: List all the cosets for the code, choose from each coset a word of least weight as coset leader **u**.

Step 2: Find a parity-check matrix H for the code and, for each coset leader **u**, calculate its syndrome $S(\mathbf{u}) = \mathbf{u}H^T$.

Coset leader **u**	Syndrome $S(\mathbf{u})$
0000	00
0001	01
0010	10
1000	11

$C = \{0000, 1011, 0101, 1110\}$.

Linear Codes

Decoding procedure for syndrome decoding:

Step 1: For the received word w, compute the syndrome S(w).

Step 2: Find the coset leader u next to the syndrome S(w) = S(u) in the syndrome look-up table.

Step3: Decode w as v=w−u.

Example (1): Let $q = 2$ and let $C = \{0000, 1011, 0101, 1110\}$. Use the syndrome look-up table below to decode (i) $w = 1101$; (ii) $w = 1111$.

(i) $w = 1101$. The syndrome is $S(w) = wH^T = 11$. From Table, we see that the coset leader is 1000. Hence, $1101 + 1000 = 0101$ was a most likely codeword sent.

(ii) $w = 1111$. The syndrome is $S(w) = wH^T = 01$. From Table, we see that the coset leader is 0001. Hence, $1111 + 0001 = 1110$ was a most likely codeword sent.

Coset leader $\mathbf{u}$	Syndrome $S(\mathbf{u})$
0000	00
0001	01
0010	10
1000	11

Thank You

For your Attention