Compact course notes

Combinatorics and Optimization 331, Winter 2011

Coding Theory

Professor: D.Jao
transcribed by: J. Lazovskis
University of Waterloo
April 6, 2011

Contents

1 Introduction

It is always assumed that the source and the receiver are separated by space and/or time.

1.1 Fundamentals

Definition 1.1.1. An alphabet is a finite set of symbols.

Definition 1.1.2. A word is a finite sequence of symbols from a given alphabet.

Definition 1.1.3. The length of a word is the number of symbols in the word.

Definition 1.1.4. A code is a subset of the set of words in a given alphabet.

Definition 1.1.5. A code word is a word in a particular code.

Definition 1.1.6. A block code is a code where every code word has the same length.

Definition 1.1.7. The length of a block code is the length of any code word in the block code.

Definition 1.1.8. An $[n, M]$-code is a block code C of length n with $|C| = M$.

1.2 Channels

Definition 1.2.1. A channel is a medium over which a symbol is sent.

Definition 1.2.2. A symmetric channel is a channel satisfying the following properties:

1. Only symbols from a set alphabet A are received.
2. No symbols are deleted, inserted, or translated.
3. Random independent probability p of error for each symbol.

Definition 1.2.3. Given an alphabet $A = \{a_1, a_2, \ldots, a_q\}$, let X_i be the ith symbol sent, and let Y_i be the ith symbol received. Then a q-symmetric channel with symbol error probability p has the property that

$$\text{for all } 1 \leqslant j, k \leqslant q, \quad P(Y_i = a_k | X_i = a_j) = \begin{cases} 1-p & j = k \\ \frac{p}{q-1} & j \neq k \end{cases}$$

Definition 1.2.4. A binary symmetric channel is a symmetric channel using only the binary alphabet.

Definition 1.2.5. The information rate of an $[n, M]$-code defined over an alphabet A of size q is $r = \frac{\log_q(M)}{n}$

Definition 1.2.6. Let A be an alphabet with words $x, y \in A^n$. Then the Hamming distance of x and y is defined to be the number of positions in which x and y differ in symbols. It is denoted by $d(x, y)$.

Theorem 1.2.7. [PROPERTIES OF HAMMING DISTANCE]

1. $d(x, y) \geqslant 0$ and $d(x, y) = 0 \iff x = y$
2. $d(x, y) = d(y, x)$
3. $d(x, y) + d(y, z) \geqslant d(x, z)$

Remark 1.2.8. The main goals of coding theory are:

1. High error correction capability
2. High information rate
3. Efficient encoding and decoding algorithms

1.3 Decoding

Algorithm 1.3.1. [Incomplete maximum likelihood decoding (IMLD)]
Suppose $r \in A^n$ is received.
If $r \in C$, accept r.
If $r \notin C$, then:
If there exists a unique $c_o \in C$ such that $d(r, c_o) < d(r, c)$ for all $c \in C, c \neq c_o$, return c_o.
Else reject r.

Algorithm 1.3.2. [Complete maximum likelihood decoding (CMLD)]
Identical to IMLD, except in last step choose a c_o arbitrarily from $\{c_o \in C | d(r, c_o) \leqslant d(r, c) \forall c \in C, c \neq c_o\}$

Theorem 1.3.3. For $r \in A^n$, IMLD outputs the code word $c \in C$ with the property that it maximizes $P(r|c) := P(r \text{ is received } | c \text{ is sent})$.

Algorithm 1.3.4. [Minimum error decoding (MED)]
Suppose $r \in A^n$ is received.
Return $c \in C$ such that $P(c|r) = P(r|c)\frac{P(c)}{P(r)}$ is maximized.

1.4 Error detection & correction

Definition 1.4.1. A code C can correct e errors if the decoder always returns the correct code word whenever e or fewer errors occur per received code word.

Theorem 1.4.2. If $d(C) = d_o$, then C can detect at most $d_o - 1$ errors per word.

Theorem 1.4.3. If $d(C) = d_o$, then C can correct at most $\left\lfloor \dfrac{d-1}{2} \right\rfloor$ errors.

Definition 1.4.4. The error probability of a code is the probability that an incorrect code word is output by IMLD for a received word.

Lemma 1.4.5. Suppose C is an $[n, M]$-code and each code word is sent with equal probability. Write for $c \in C$, $w(c) = P(\text{CMLD is wrong } | c \text{ is sent})$. Then the error probability of C is given by $P(C) = \dfrac{1}{M}\sum_{c \in C} w(c)$.

Definition 1.4.6. Define $P^*(n, M, p) = \max\{P(C) | C \text{ is an } [n, M]\text{-code}\}$.

Definition 1.4.7. The channel capacity of a binary symmetric channel, for p the symbol error probability, is given by $c(p) = 1 + p\log(p) + (1-p)\log(1-p)$.

Theorem 1.4.8. Set $R = \dfrac{\log(M)}{n}$. Then for fixed $R < c(p)$, $\lim_{n\to\infty} [P^*(n, M, p)] = 0$.

2 Finite fields

2.1 Basics

Definition 2.1.1. A field is a set $\mathbb{F}$ closed under the operations $+ : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ and $\cdot : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$.

1. $(a + b) + c = a + (b + c)$
2. $a + b = b + 1$
3. $\exists\ 0 \in \mathbb{F}$ such that $a + 0 = a$ for all $a \in \mathbb{F}$
4. $\exists\ -a \in \mathbb{F}\ \forall\ a \in \mathbb{F}$ such that $a + (-a) = 0$
5. $a \cdot (b + c) = a \cdot b + a \cdot c$
6. $(a \cdot b) \cdot c = a \cdot (b \cdot c)$
7. $a \cdot b = b \cdot a$
8. $\exists\ 1 \in \mathbb{F}$ such that $a \cdot 1 = a$ for all $a \in \mathbb{F}$
9. $\exists\ a^{-1} \in \mathbb{F}\ \forall\ a \in \mathbb{F}$ such that $a \cdot a^{-1} = 1$

Definition 2.1.2. The order of a field is defined to be its cardinality: $\text{ord}(\mathbb{F}) = |\mathbb{F}|$

Definition 2.1.3. A field is finite is its order is finite. Else it is infinite.

Definition 2.1.4. $\mathbb{Z}_n$ is a field $\iff$ n is prime.

Remark 2.1.5. A field can also be defined as a commutative ring with inverses and the identity element.

Definition 2.1.6. The characteristic of a finite field $\mathbb{F}$ is defined to be the smallest positive integer n such that $\underbrace{1+1+\cdots+1}_{n \text{ times}} = 0$ for 1 the multiplicative identity of $\mathbb{F}$. If no such n exists, then the characteristic of $\mathbb{F}$ is defined to be 0. It is denoted $\operatorname{char}(\mathbb{F})$.

Definition 2.1.7. For a field $\mathbb{F}$, $\operatorname{char}(\mathbb{F}) = 0 \iff \mathbb{F}$ is not finite.

Definition 2.1.8. For a field $\mathbb{F}$, a subfield of $\mathbb{F}$ is a subset of $\mathbb{F}$ that is a field itself.

Definition 2.1.9. For $\mathbb{F}$ a field with $\operatorname{char}(\mathbb{F}) = p$ prime, the set $\{0, 1+1, 1+1+1, \dots\}$ is termed the prime subfield of $\mathbb{F}$.

Remark 2.1.10. The prime subfield is the smallest subfield of any field.

2.2 Polynomial rings

Definition 2.2.1. For any field $\mathbb{F}$, the polynomial ring $\mathbb{F}[x]$ is the set of all polynomials:

$$\mathbb{F}[x] = \left\{ \sum_{k=0}^{n} a + kx^k \;\middle|\; a_k \in \mathbb{F}, n \in \mathbb{N} \right\}$$

Theorem 2.2.2. For any polynomials $a(x), b(x) \neq 0 \in \mathbb{F}[x]$, there exist unique polynomials $q(x), r(x)$ such that $a(x) = q(x)b(x) + r(x)$ such that $\deg(r(x)) < \deg(b(x))$.

Note that $\deg(0) = -\infty$ by definition.

Definition 2.2.3. Fix $f(x) \in \mathbb{F}[x]$. The equivalence class of $a(x)$ modulo $f(x)$ is denoted $[a(x)]$.

Definition 2.2.4. For any $f(x) \in \mathbb{F}[x]$, the set $\mathbb{F}[x]/f(x)$ is the set of all equivalence classes of polynomials in $\mathbb{F}[x]$ modulo $f(x)$.

Remark 2.2.5. This set may be defined as $\mathbb{F}[x]/f(x) = \{r(x) | \deg(r) < \deg(f)\}$, with $\mathbb{F}[x]/f(x)$ is a field $\iff$ $f(x)$ is irreducible in $\mathbb{F}[x]$.

Definition 2.2.6. The polynomial $f(x)$ is termed irreducible over a field $\mathbb{F}[x]$ is there exists no factorization $f(x) = p(x)q(x)$ with $\deg(p(x)) < \deg(f(x))$ and $\deg(q(x)) < \deg(f(x))$.

Corollary 2.2.7. If $f(x) \in \mathbb{Z}[x]/f(x)$ is irreducible, then $\mathbb{Z}[x]/f(x)$ is a field of order p^n for $p = \deg(f(x))$.

Theorem 2.2.8. For every prime p and every positive integer n, there exists an irreducible polynomial in $\mathbb{Z}_p[x]$ of degree n.

Corollary 2.2.9. For every prime p and every positive integer n, there exists a finite field of order p^n with $p \geqslant 2$ and $n > 0$.

Theorem 2.2.10. Any two fields of the same order are isomorphic to each other.

Definition 2.2.11. Denote by $GF(q)$ or $\mathbb{F}_q$ the unique (up to isomorphism) finite field of order q

Lemma 2.2.12. [Anti-calculus lemma]
In a field $\mathbb{F}$ with $\operatorname{char}(\mathbb{F}) = p$ prime, $(x+y)^{p^k} = x^{p^k} + y^{p^k}$ for all $x, y \in \mathbb{F}$.

Theorem 2.2.13. [Fermat]
In a finite field $GF(q)$ for q prime, $\alpha^{q-1} = 1$ for all $\alpha \in \mathbb{F}$.

Corollary 2.2.14. In $GF(q)$, $\alpha^q = \alpha$ for all $\alpha \in \mathbb{F}$.

Definition 2.2.15. For any $\alpha \in GF(q)^*$, the order of α is the smallest positive integer $t = \text{ord}(\alpha)$ such that $\alpha^t = 1$, where $GF(q)^* = GF(q) \setminus \{0\}$.

Theorem 2.2.16. Let $\alpha \in GF(q)^*$ with $\text{ord}(\alpha) = t$. Then $\alpha^s = 1 \iff s|t$.

Definition 2.2.17. An element α of $GF(q)^*$ is termed a generator (or primitive element or primitive root) of $GF(q)^*$ if $\text{ord}(\alpha) = q - 1$.

In this case, $GF(q)^* = \{\alpha^1, \alpha^2, \ldots, \alpha^{q-1}\}$.

Theorem 2.2.18. Every $GF(q)^*$ contains a generator.

Theorem 2.2.19. Let $u_1, u_2 \in GF(q)$ with $\text{ord}(u_1) = t_1$ and $\text{ord}(u_2) = t_2$ and $\gcd(t_1, t_2) = 1$. Then $\text{ord}(u_1 u_2) = t_1 t_2$.

Theorem 2.2.20. Let $u_1, u_2 \in GF(q)$ with $\text{ord}(u_1) = t_1$ and $\text{ord}(u_2) = t_2$. Then there exists $u \in GF(q)^*$ with $\text{ord}(u_1 u_2) = \text{lcm}(t_1 t_2)$.

3 Linear codes

3.1 Fundamentals

Remark 3.1.1. For $\mathbb{F} = GF(q)$, denote the set $\mathbb{F}^n = \underbrace{\mathbb{F} \times \mathbb{F} \times \cdots \times \mathbb{F}}_{n \text{ times}}$ by $V_n(\mathbb{F})$.

Definition 3.1.2. A linear (n, k)-code C is defined to be a subspace $C \subset V_n(\mathbb{F})$ of dimension k.

Remark 3.1.3. A linear (n, k)-code C has the following properties:

i. The number of code words in C is $|C| = q^k$.

ii. The distance of C is $d(C) = \min\{d(x, y)|x, y \in C, x \neq y\}$.

iii. The information rate of C is $R = \frac{\log_q(M)}{n} = \frac{k}{n}$

Definition 3.1.4. The Hamming weight of a code C is $w(C) = \min\{w(x)|x \in C, x \neq 0\}$.

Theorem 3.1.5. For a linear code C, $w(C) = d(C)$.

Definition 3.1.6. Let C be an (n, k)-code. A generator matrix for C isa $k \times n$ matrix G with coefficients in $\mathbb{F}$ whose rows are a basis of C.

A generator matrix need not be unique.

Definition 3.1.7. A generator matrix of the form $G = [I_k|A]$ is said to be in standard form.

Definition 3.1.8. If C has at least 1 generator matrix in standard form, then C is a systematic code.

Remark 3.1.9. For any linear (n, k)-code C with generator matrix G, the encoding function is:

$$\mathbb{F}^k \to C, \text{ given by } m = (m_1 \ \ m_2 \ \ \cdots \ \ m_k) \in \mathbb{F}^k \mapsto m_1 c_1 + m_2 c_2 + \cdots + m_k c_k = mG \in C$$

Definition 3.1.10. Two linear codes C and C' are equivalent if there exists a permutation of the coordinates of $V_n(\mathbb{F})$ mapping C to C'.

Theorem 3.1.11. Every linear code is equivalent to a systematic code.

3.2 Dual codes and parity-check matrices

Definition 3.2.1. Let C be an (n, k)-code over $\mathbb{F}$. The dual code (or orthogonal code) of C is given by

$$C^\perp = \{x \in V_n(\mathbb{F}) \mid x \cdot y = 0 \ \forall \ y \in \mathbb{C}\}$$

Note that $(C^\perp)^\perp = C$.

Theorem 3.2.2. If C is an (n, k)-code, then $C^\perp$ is an $(n, n-k)$-code.

Theorem 3.2.3. If C is a systematic code with generator matrix $G = [I_k|A] \in M_{k\times n}$, then a generator matrix for $C^\perp$ is $H = [-A^T|I_{n-k}]$.

Definition 3.2.4. A parity-check matrix for C is a generator matrix for $C^\perp$.

Proposition 3.2.5. $x \in C \iff Hx^T = 0$ for H a parity-check matrix of C.

Theorem 3.2.6. For any $s \in \mathbb{N}$, $d(C) \geqslant s \iff$ every subset of $s-1$ columns of the parity-check matrix H for C is linearly independent.

Remark 3.2.7. The following hold for a linear code C with parity-check matrix H:

1. $d(C) \geqslant 2 \iff$ no column of H is the zero vector

2. $d(C) \geqslant 3 \iff$ no column of H is a multiple of another column of H

Definition 3.2.8. A Hamming code of order r over $GF(q)$ is an (n, k)-code where $n = \frac{q^r-1}{q-1}$ and $k = n - r$ with a parity-check matrix $H \in M_{k\times n}(GF(q))$ such that no column of H is a zero column and no column is a multiple of another column.

Hamming codes are 1-error correcting codes.

Definition 3.2.9. Given a transmitted code word c and received code word r, the error vector is defined to be $e = r - c$.

Definition 3.2.10. For every $r \in V_n(\mathbb{F})$, the syndrome of r is Hr^T.

Proposition 3.2.11. For a Hamming code C, every $r \in V_n(\mathbb{F})$ is within distance 1 of a code word.

Definition 3.2.12. Let C be an $[n, m]$-code of distance d with $e = \left\lfloor \frac{d-1}{2} \right\rfloor$. Then C is termed a perfect code if every $r \in A^n$ is within distance e of some $r \in C$.

Definition 3.2.13. Let C be an (n, k)-code. An element c of a coset of C (in $V_n(\mathbb{F})$) is termed a coset leader if no other element of the coset has lesser weight than c.

Definition 3.2.14. Let C be an (n, k)-code over $GF(q)$. A standard array of C is a $q^{n-k} \times q^k$ matrix with entries vectors over $V_n(GF(q))$ with:

1. Each code word appears exactly once in the first row

2. Each row is a coset of C

3. Every element of $V_n(GF(q))$ appears exactly once in the array

4. In each row, the entry in the first column is a coset leader for its respective row

5. $A_{ij} = A_{i1} + A_{1j}$ for all i, j

Theorem 3.2.15. If c is a coset leader for C with $d(C) = d$ and $w(c) = \left\lfloor \frac{d-1}{2} \right\rfloor$, then c is the unique coset leader in its coset.

Remark 3.2.16. Some properties of C_{24}, the code presented below:

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

2. C_{24} can correct 3 errors

Definition 3.2.17. This is the generator matrix for C_{24}, the Extended binary Golay code:

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

4 Cyclic codes

4.1 Fundamentals

Definition 4.1.1. A subspace S of $V_n(\mathbb{F})$ is termed a cyclic subspace if $(a_0, a_1, \ldots, a_{n-1}) \in S$ implies $(a_{n-1}, a_0, a_1, \ldots, a_{n-2}) \in S$.

Definition 4.1.2. A linear code is a cyclic code if it is a cyclic subspace of $V_n(\mathbb{F})$.
Equivalently, it is a code that is invariant under cyclic shifts.

Definition 4.1.3. A commutative ring is a set R closed under the two operations $+ : R \times R \to R$ and $\cdot : R \times R \to R$ with the following properties:
i. Addition is associative and commutative
ii. The identity and inverses exist for addition
iii. Multiplication is associative and commutative
iv. The identity exists for multiplication
v. The distributive law holds

Definition 4.1.4. Let R be a commutative ring. An ideal of R is a non-empty set $I \subset R$ with
i. if $a, b \in I$, then $a + b \in I$
ii. if $a \in I$, then $-a \in I$
iii. for all $a \in I$ and $r \in R$, $ar \in I$

Theorem 4.1.5. A subset $S \subset V_n(\mathbb{F})$ is a cyclic subspace $\iff S \subset \mathbb{F}[x]/x^n - 1$ is an ideal of $\mathbb{F}[x]/x^n - 1$.

Theorem 4.1.6. Every ideal in $\mathbb{F}[x]$ is of the form $I = \langle g(x) \rangle$ for $g(x) \in \mathbb{F}[x]$.

Definition 4.1.7. An ideal I of R is termed a principal ideal if $I = \langle g \rangle$ for $g \in R$ and $\langle g \rangle = \{gr | r \in R\}$. Then I is said to be generated by g.

Definition 4.1.8. For any ideal $I \subset R = \mathbb{F}[x]/x^n - 1$, the generator polynomial of I is the unique monic polynomial $g(x)$ such that $I = \langle g(x) \rangle$.

Remark 4.1.9. Let $\mathbb{F}$ be a field. Then the only ideals of $\mathbb{F}$ are $\langle 0 \rangle$ and $\langle 1 \rangle$.

4.2 Encoding with cyclic codes

Proposition 4.2.1. Given a cyclic (n,k)-code C over $\mathbb{F}$ with generator polynomial $g(x)|x^n-1$, a generator matrix for C is

$$G = \begin{bmatrix} g(x) & xg(x) & \cdots & x^{k-1}g(x) \end{bmatrix}^T \in M_{k\times n}$$

where the entries are the coefficients of unique powers of x.

Remark 4.2.2. For any message $m \in V_k(\mathbb{F})$, $mG = m(x)g(x)$.

Proposition 4.2.3. For a cyclic (n,k)-code C over $\mathbb{F}$ with generator polynomial $g(x)|x^n-1$, the parity check matrix of C is the generator matrix for $C^\perp$ with generator polynomial $h(x) = x^n - 1/g(x)$.

Note that $C^\perp$ is also cyclic.

Definition 4.2.4. For any polynomial $h(x) = h_0 + h_1x + h_2x^2 + \cdots + h_kx^k$, define $h_R(x) := \frac{x^k}{h_0}h\left(\frac{1}{x}\right)$. Then $h_R(x) = (h_k, h_{k-1}, \ldots, h_1, h_0, h_{n-1}, h_{n-2}, \ldots, h_{k+1})$ and a generator matrix for $C^\perp$ is

$$H = \begin{bmatrix} h_R(x) & xh_R(x) & \cdots & x^{n-k-1}h_R(x) \end{bmatrix}^T \in M_{(n-k)\times n}$$

Proposition 4.2.5. Let $C \subset V_n(\mathbb{F}) = \mathbb{F}[x]/x^n - 1$ be a cyclic code with generator polynomial $g(x) = g_0 + g_1x + \cdots + g_{n-k}x^{n-k}$ and $g(x)|x^n-1$. Then

$$G = \begin{bmatrix} g_0 & g_1 & g_2 & \cdots & g_{n-k} & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & g_2 & \cdots & g_{n-k} & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & & \ddots & \\ 0 & 0 & \cdots & g_0 & g_1 & g_2 & \cdots & g_{n-k} \end{bmatrix}$$

$$H = \begin{bmatrix} h_k & h_{k-1} & \cdots & h_0 & 0 & \cdots & 0 \\ 0 & h_k & h_{k-1} & \cdots & h_0 & 0 & 0 \\ \vdots & & \ddots & \ddots & & \ddots & \\ 0 & 0 & \cdots & h_k & h_{k-1} & \cdots & h_0 \end{bmatrix}$$

are the generator and parity-check matrices for C.

Proposition 4.2.6. Consider a re-expression of the generator matrix G as described above by:

$$\begin{aligned} x^{n-k} &= q_0g(x) + r_0(x) \\ x^{n-k+1} &= q_1(x)g(x) + r_1(x) \\ x^{n-k+2} &= q_2(x)g(x) + r_2(x) \\ \vdots \quad & \quad\quad \vdots \\ x^{n-1} &= q_{k-1}(x)g(x) + r_{k-1}(x) \end{aligned} \qquad G = \begin{bmatrix} x^{n-k} - r_0(x) \\ x^{n-k+1} - r_1(x) \\ x^{n-k+2} - r_2(x) \\ \vdots \\ x^{n+1} - r_{k-1}(x) \end{bmatrix}$$

Then the parity-check matrix can be re-written as:

$$H = \begin{bmatrix} I_{n-k} & | & r_0(x) & r_1(x) & \cdots & r_k(x) \end{bmatrix} = \begin{bmatrix} x^0 \mod g & x^1 \mod g & \cdots & x^{n-1} \mod g \end{bmatrix} \in M_{(n-k)\times n}$$

4.3 Burst errors

Definition 4.3.1. A vector $e \in V_n(\mathbb{F})$ has <u>cyclic burst length</u> $\leqslant t$ if there exists a cyclic block of (at most) t contiguous positions such that every non-zero coordinate of e lies within the t positions.

Theorem 4.3.2. [Rieger bound]
For C an (n,k)-code, C has burst error correction capability $t \leqslant \frac{n-k}{2}$.

Theorem 4.3.3. Given a cyclic (n, k)-code C with burst error correction capability t and a received vector $r(x)$, the syndrome of the sent code word is

$$s(x) = x^i r(x) \quad \text{such that} \quad i = \min\{i \mid x^i r(x) \pmod{g(x)} \text{ has every error in the first } t \text{ positions}\}$$

This is termed the error trapping method. Further, the error vector is

$$e(x) = x^{n-i} s(x) \pmod{x^n - 1}$$

Definition 4.3.4. A cyclotomic coset of q modulo n is a set of unique elements,

$$C_k = \{k \pmod{n}, qk \pmod{n}, q^2 k \pmod{n}, \ldots, q^{m_k} k \pmod{n}\} \quad \text{with} \quad 0 \leqslant k \leqslant q - 1$$

Proposition 4.3.5. With respect to the above, the irreducible monic polynomials of $x^n - 1$ over $\mathrm{GF}(q)$ are

4.4 BCH codes

Definition 4.4.1. Suppose $\mathbb{F} = \mathrm{GF}(q)$ and $K = \mathrm{GF}(q^m)$, with $\mathbb{F} \subset K$. For any $\alpha \in K$, the minimal polynomial of α over $\mathbb{F}$ is the non-zero monic polynomial of smallest degree in $\mathbb{F}[x]$ such that it has α as a root.

The minimal polynomial is denoted $m_\alpha(x)$.

Theorem 4.4.2.

1. $m_\alpha(x)$ is unique
2. $m_\alpha(x)$ is irreducible
3. $\deg(m_\alpha(x)) \leqslant m$
4. if $f(x) \in \mathbb{F}[x]$ and $f(\alpha) = 0$, then $m_\alpha | f$

Corollary 4.4.3. If $f(x)$ is monic irreducible and $f(\alpha) = 0$, then $f(x) = m_\alpha(x)$.

Lemma 4.4.4. For any $\alpha \in \mathrm{GF}(q^m)$, $\alpha \in \mathrm{GF}(q) \iff \alpha^q = \alpha$.

Definition 4.4.5. For $\alpha \in \mathrm{GF}(q^m)$, the set of conjugates of α over $\mathrm{GF}(q)$ is $C(\alpha) = \{\alpha^{q^n} | n \in \mathbb{N} \cup \{0\}\}$.

Remark 4.4.6. Suppose t is the smallest positive integer such that $\alpha^{q^t} = \alpha$. Note $t \leqslant m$. Then the elements of $C(\alpha) = \{\alpha, \alpha^{q^1}, \ldots, \alpha^{q^t}\}$ are all pairwise distinct.

Theorem 4.4.7. Let $\alpha \in \mathrm{GF}(q^m)$. Then the minimal polynomial of α over $\mathrm{GF}(q)$ is

$$\prod_{\beta \in C(\alpha)} (x - \beta) = (x - \alpha)(x - \alpha^{q^1}) \cdots (x - \alpha^{q^t})$$

Theorem 4.4.8. Given the following conditions:

$$\left.\begin{array}{l} p \text{ prime}, k \in \mathbb{N}, q = p^k \\ n \text{ block length with } \gcd(n, q) = 1 \\ m \text{ such that } o(q) = m \text{ in } \mathbb{Z}_n \\ \alpha \in \mathrm{GF}(q^m)^* \text{ a generator} \\ \beta \in \mathrm{GF}(q^m) \text{ of order } n \text{ with } \beta = \alpha^{(q^m-1)/n} \\ a, \delta \in \mathbb{Z}_n \end{array}\right\}$$

Then the code C generated by $g(x)$ is a BCH code over $\mathrm{GF}(q)$ of block length n and designed distance δ.

$$g(x) = \operatorname{lcm}\{m_{\beta^i}(x) \in \mathrm{GF}(q) \,\big|\, a \leqslant i \leqslant a + \delta - 2\}$$

Theorem 4.4.9. A BCH code of designed distance δ has distance at least δ.

References

Vanstone, Scott A. and Paul C. van Oorschot. *An introduction to Error Correcting Codes with Applications.* Kluwer Academic Publishers: 1989