EE4-07
EE9-SO11

IMPERIAL COLLEGE LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2013

MSc and EEE/EIE PART IV: MEng and ACGI

CODING THEORY

Wednesday, 22 May 10:00 am

Time allowed: 3:00 hours

There are FIVE questions on this paper.

Answer ALL questions.

All the questions carry equal marks.

Any special instructions for invigilators and information for candidates are on page 1.

Examiners responsible	First Marker(s) :	W. Dai
	Second Marker(s) :	C. Ling

EE4-07 Coding Theory

Instructions for Candidates

Answer all five questions. The star notation * right after the sub-question numbering means that the particular sub-question may be difficult to solve.

1. (Linear Codes and Error Probability) Consider a linear code C over $\mathbb{F}_3$ defined by the parity check matrix

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

(a) Find the minimum distance of C. Prove your answer. [4]

(b) Find the generator matrix of C in systematic form. [4]

(c) A codeword in C is transmitted through a ternary symmetric channel and the word $\boldsymbol{y} = [1\ 1\ 2\ 2]$ is received.

i). Find the syndrome vector of $\boldsymbol{y}$. [2]

ii). Find the codeword in C that is produced by applying the minimum distance decoder. [2]

(d) A codeword in C is transmitted through an erasure channel. Let $\boldsymbol{y} = [?\ ?\ 2\ 2]$ be the received word, where the question marks denote that the corresponding symbols have been erased. Find the most plausible correction of $\boldsymbol{y}$. [4]

(e) Consider the case where a code $C \subset \mathbb{F}_q^n$ is used for the memoryless q-ary symmetric channel with crossover probability p:

$$\Pr\left(y_i \text{ received} | c_i \text{ transmitted}\right) = \begin{cases} 1-p & \text{if } y_i = c_i, \\ p/\left(q-1\right) & \text{otherwise.} \end{cases}$$

Find the the relationship between p and q such that the minimum distance decoder (MDD) is equivalent to the maximum likelihood decoder (MLD). [4]

2. (Linear Codes) Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be two linear codes over $\mathbb{F}_q$ of the same length n. Let $\boldsymbol{G}_i$, $\boldsymbol{H}_i$, k_i, and d_i be the generator matrix, the parity-check matrix, the dimension, and the minimum distance of $\mathcal{C}_i$, $i = 1, 2$, respectively. Define

- $\mathcal{C}_3 = \{[\boldsymbol{c}_1, \boldsymbol{c}_2] : \boldsymbol{c}_1 \in \mathcal{C}_1, \ \boldsymbol{c}_2 \in \mathcal{C}_2\}$.
- $\mathcal{C}_4 = \mathcal{C}_1 \bigcup \mathcal{C}_2$.

(a) Show that $\mathcal{C}_3$ is a linear code. [5]

(b) Write the generator matrix $\boldsymbol{G}_3$ and the parity-check matrix $\boldsymbol{H}_3$ for the code $\mathcal{C}_3$ in terms of $\boldsymbol{G}_1$, $\boldsymbol{G}_2$, $\boldsymbol{H}_1$, and $\boldsymbol{H}_2$. [5]

(c) Determine the dimension k_3 and the minimum distance d_3 of the code $\mathcal{C}_3$ in terms of k_1, k_2, d_1, and d_2. Prove your results. [5]

(d) *Suppose that $k_1 < k_2$. Show that $\mathcal{C}_4$ is a linear code if and only if $\mathcal{C}_1 \subset \mathcal{C}_2$. [5]

3. (Finite Field)

(a) Consider finite fields $\mathcal{F}_1 = \mathbb{F}_2[x] / (x^3 + x^2 + 1)$ and $\mathcal{F}_2 = \mathbb{F}_2[y] / (y^3 + y + 1)$. Consider the mapping φ from $\mathcal{F}_1$ to $\mathcal{F}_2$ defined by $\varphi(x) = y + 1$. Consider two polynomials $f_1(x) = x^2 + 1$ and $f_2(x) = x^2 + x + 1$ taken from $\mathcal{F}_1$. Compute the following: [10]

i). $f_1(x) \cdot f_2(x)$,

ii). $\varphi(f_1(x) \cdot f_2(x))$,

iii). $\varphi(f_1(x))$,

iv). $\varphi(f_2(x))$, and

v). $\varphi(f_1(x)) \cdot \varphi(f_2(x))$.

(b)

i). Find all the cyclotomic cosets of 2 mod 7. [3]

ii). Let α be a primitive element of $\mathbb{F}_8$. Find the irreducible polynomials in $\mathbb{F}_2[x]$ that factor $x^7 - 1$. [3]

iii). *Count the number of distinct cyclic codes in $\mathbb{F}_2$ of length 7. [4]

4. (Reed-Solomon, cyclic, and BCH codes)

Let α be a primitive element of $\mathbb{F}_9$. Define $\boldsymbol{A} \in \mathbb{F}_9^{4\times 8}$ by

$$\boldsymbol{A} = \begin{bmatrix} 1 & \alpha & \alpha^2 & \cdots & \alpha^7 \\ 1 & \alpha^2 & \alpha^4 & \cdots & \alpha^{14} \\ 1 & \alpha^3 & \alpha^6 & \cdots & \alpha^{21} \\ 1 & \alpha^4 & \alpha^8 & \cdots & \alpha^{28} \end{bmatrix}.$$

(a) Let C be a code over $\mathbb{F}_9$ and its parity-check matrix be the matrix $\boldsymbol{A}$. What are the parameters, $[n, k, d]$, of C? Find the number of errors that this code can correct. [4]

(b) Suppose that a codeword $\boldsymbol{c} \in C$ is transmitted and a word $\boldsymbol{y} = \boldsymbol{c} + \boldsymbol{e}$ is received where $\boldsymbol{e} \in \mathbb{F}_9^8$ is the error vector. Define the set of error positions $\mathcal{I} = \{0 \le i \le 7 : \ e_i \ne 0\}$. Define the error locator polynomial as

$$L(z) = \prod_{i\in\mathcal{I}} \left(1 - \alpha^i z\right).$$

Prove that

$$\begin{cases} L\left(\alpha^{-k}\right) = 0 & \text{if } 0 \le k \le 7 \text{ and } k \in \mathcal{I}, \\ L\left(\alpha^{-k}\right) \ne 0 & \text{if } 0 \le k \le 7 \text{ and } k \notin \mathcal{I}. \end{cases}$$

[4]

(c) Suppose that $\mathcal{I} \ne \phi$. Fix an $i \in \mathcal{I}$. Find $\frac{d}{dz} L(z)$ when $z = \alpha^{-i}$. [4]

(d) The syndrome vector is defined via $\boldsymbol{s} = [s_0, \cdots, s_3] = \boldsymbol{e}\boldsymbol{A}^T$. The syndrome polynomial is defined as $S(z) = \sum_{j=0}^{3} s_j z^j$. Prove that

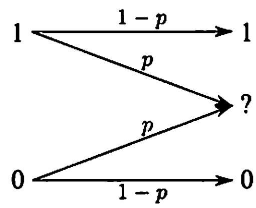
$$S(z) = \sum_{i\in\mathcal{I}} \frac{e_i \alpha^i}{1 - \alpha^i z} \bmod z^4.$$

[4]

(e) Define a cyclic code $C' = \left\{\boldsymbol{c} \in \mathbb{F}_3^8 : \ \boldsymbol{c}\boldsymbol{A}^T = \boldsymbol{0}\right\} \subset \mathbb{F}_3^8$. Find its generator polynomial $g(x)$. [4]

5. (Decoding on graphs with the application of erasure correction)

Consider the following binary erasure channel:



(a) Assume that $\Pr(x=0) = \Pr(x=1) = 1/2$. Find $\Pr(x|y)$ when x varies in $\{0,1\}$ and y varies in $\{0,1,?\}$. [5]

(b) Draw the Tanner graph that corresponds to the parity-check matrix

$$\boldsymbol{H} = \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ \left[\begin{array}{ccccc} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{array} \right] \end{array} \begin{array}{c} \\ c_1 \\ c_2 \end{array} .$$

(Notations in the Tanner graph are required to be consistent with the above ones.) [5]

(c) Assume conditional independence, i.e., $\Pr(\boldsymbol{x}|\boldsymbol{y}) = \prod_i \Pr(x_i|y_i)$. Assume that $\boldsymbol{y}_{3:5} = [y_3, y_4, y_5] = [?, 1, 1]$. Decode x_3 by computing $\Pr(x_3|\boldsymbol{y}_{3:5})$. [5]

(d) * Consider a code C which may or may not be linear. Assume that the minimum distance of C is d. Prove that it can correct $d-1$ erasures. (Note that the "up to" part is not required in your solution, in order to simplify the problem.) [5]

Solution of Question 1.

(a) Since every pair of columns of $\boldsymbol{H}$ are linearly independent, one has $d \geq 3$. On the other hand, there exist three columns in $\boldsymbol{H}$, for example, columns 1, 2 and 3, linearly dependent, $d \leq 3$. Hence, $d(\mathcal{C}) = 3$. [4]

(b) The generator matrix in the systematic form is given by

$$\boldsymbol{G} = \begin{bmatrix} 1 & 1 & -1 & 0 \\ 2 & 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix}.$$

[4]

(c)

i). $\boldsymbol{s} = \boldsymbol{y}\boldsymbol{H}^T = [1\ 2]$. [2]

ii). Since $\boldsymbol{s} = (\boldsymbol{c} + \boldsymbol{e})\, \boldsymbol{H}^T = \boldsymbol{e}\boldsymbol{H}^T$, it is clear that $\hat{\boldsymbol{e}} = [0\ 0\ 0\ 2]$. As a result, $\hat{\boldsymbol{c}} = \boldsymbol{y} - \hat{\boldsymbol{e}} = [1\ 1\ 2\ 0]$. [2]

(d) The only codeword of which the last two symbols are both 2 is given by $[2\ 2] \cdot \boldsymbol{G} = [0\ 2\ 2\ 2]$. Hence the decoding result is $\hat{\boldsymbol{c}} = [0\ 2\ 2\ 2]$. [4]

(e) For any given $\boldsymbol{c} \in \mathcal{C} \subset \mathbb{F}_q^n$ and $\boldsymbol{y} \in \mathbb{F}_q^n$, let $d = d_H(\boldsymbol{y}, \boldsymbol{c})$. Then

$$\Pr(\boldsymbol{y}|\boldsymbol{c}) = \left(\frac{p}{q-1}\right)^d (1-p)^{n-d} = (1-p)^n \left(\frac{p}{1-p} \cdot \frac{1}{q-1}\right)^d.$$

The sufficient and necessary condition for the equivalence is that

$$\frac{p}{1-p} \cdot \frac{1}{q-1} < 1.$$

Simple algebra shows that this is equivalent to $q > \frac{1}{1-p}$ or $p < 1 - \frac{1}{q}$. [4]

Solution of Question 2.

(a) For any $c, c' \in C_3$ and $\lambda, \lambda' \in \mathbb{F}_q$, one has

$$\begin{aligned} \lambda c + \lambda' c' &= \lambda [c_1, c_2] + \lambda' [c_1', c_2'] \\ &= [\lambda c_1 + \lambda' c_1', \ \lambda c_2 + \lambda' c_2'] \\ &= [c_1'', c_2''] \in C_3, \end{aligned}$$

for some $c_1'' \in C_1$ and $c_2'' \in C_2$, where the last equality follows from the linearity of C_1 and C_2. [5]

(b) It holds that

$$G_3 = \begin{bmatrix} G_1 & 0 \\ 0 & G_2 \end{bmatrix}, \text{ and } H_3 = \begin{bmatrix} H_1 & 0 \\ 0 & H_2 \end{bmatrix}$$

[5]

(c) Note that the size of C_3 is $q^{k_1} \cdot q^{k_2}$. It is clear that $k_3 = k_1 + k_2$. The same conclusion can be derived from the form of the generator matrix G as well.

The computation of the distance of C_3 relies on the fact that C_3 is linear. It holds that

$$\begin{aligned} d(C_3) &= \min_{c \in C_3} \mathrm{wt}(c) \\ &= \min_{c_1 \in C_1, \, c_2 \in C_2} \mathrm{wt}([c_1, c_2]) \\ &= \min(d_1, d_2). \end{aligned}$$

[5]

(d) The "if" part: $C_1 \subset C_2 \Rightarrow C_4 = C_1 \bigcup C_2 = C_2 \Rightarrow C_4$ is linear.

The "only if" part: Suppose that $C_1 \not\subseteq C_2$. It follows that $C_1 \backslash C_2 \neq \phi$ and $C_2 \backslash C_1 \neq \phi$ where the second inequality holds as $|C_1| < |C_2|$. Hence, there exist $c_1 \in C_1 \backslash C_2 \subset C_4$ and $c_2 \in C_2 \backslash C_1 \subset C_4$. We claim that $c_1 + c_2 \notin C_1$ because otherwise $c_2 = (c_1 + c_2) - c_1$ will be in C_1, which contradicts the choice of c_2. Similarly, $c_1 + c_2 \notin C_2$. Hence, $c_1 + c_2 \notin C_1 \bigcup C_2 = C_4$. Therefore, C_4 is not linear. The above arguments show that if C_4 is linear, then $C_1 \subset C_2$. The "only if" part is therefore proved. [5]

Solutions of Question 3.

(a)

$$
\begin{aligned}
f_1(x)\cdot f_2(x) &= \left(x^2+1\right)\left(x^2+x+1\right)\\
&= x^4+x^3+x^2+x^2+x+1\\
&= x^4+x^3+x+1\\
&= x\left(x^2+1\right)+x^3+x+1\\
&= 1,
\end{aligned}
$$

$$\varphi\left(f_1(x)\cdot f_2(x)\right)=\varphi(1)=1,$$

$$\varphi\left(f_1(x)\right)=(y+1)^2+1=y^2,$$

$$
\begin{aligned}
\varphi\left(f_2(x)\right) &= (y+1)^2+(y+1)+1\\
&= y^2+y+1,
\end{aligned}
$$

and

$$
\begin{aligned}
\varphi\left(f_1(x)\right)\cdot\varphi\left(f_2(x)\right) &= y^2\left(y^2+y+1\right)\\
&= y^4+y^3+y^2\\
&= y(y+1)+y+1+y^2\\
&= 1.
\end{aligned}
$$

[10]

(b)

i). $C_0=\{0\}$, $C_1=\{1,2,4\}$, $C_3=\{3,6,5\}$. [3]

ii). x^7-1 can be factored into three irreducible polynomials: $x-1$, $(x-\alpha)\cdot(x-\alpha^2)\cdot(x-\alpha^4)$ and $(x-\alpha^3)\cdot(x-\alpha^5)\cdot(x-\alpha^6)$. [3]

iii). Note that the generator polynomial $g(x)$ of a cyclic code of length 7 has to satisfy that $g(x)\in\mathbb{F}_2[x]$ and $g(x)\,|x^7-1$. From the factorization of x^7-1, the number of distinct cyclic codes are given by $\binom{3}{0}+\binom{3}{1}+\binom{3}{2}+\binom{3}{3}=2^3=8$, where $\binom{3}{0}$ and $\binom{3}{3}$ are included as the constant polynomial 1 (corresponding to the code $C = \mathbb{F}_2^7$) and the polynomial $x^7 - 1$ (corresponding to the trial code $C = \{0\}$) generate cyclic does as well. [4]

Solutions of Question 4.

(a) $[n, k, d] = [8, 4, 5]$: It is clear that $n = 8$ and $k = n - (n - k) = 8 - 4 = 4$. Note that every four columns of $\boldsymbol{A}$ forms a Vandemonde matrix, every four columns of $\boldsymbol{A}$ are linearly independent, i.e., $d \geq 5$. Furthermore, every five columns of $\boldsymbol{A}$ must be linearly dependent and hence $d = 5$.

The number of errors that this code can correct is $\lfloor \frac{d-1}{2} \rfloor = 2$. [4]

(b) Given $0 \leq k \leq 7$, it holds that $\alpha^i \alpha^{-k} = 1$ if and only if $k = i$.

When $k \notin \mathcal{I}$, $L(z)$ is a product of nonzero elements and hence nonzero.

When $k \in \mathcal{I}$, there exists an $i \in \mathcal{I}$ such that $\left(1 - z^i z^{-k}\right) = 0$ and therefore $L(z) = 0$. [4]

(c) Elementary algebra shows that

$$\frac{d}{dz} L(z) = \sum_{j \in \mathcal{I}} \left(-\alpha^j\right) \prod_{\ell \in \mathcal{I},\, \ell \neq j} \left(1 - \alpha^\ell z\right).$$

Let $z = \alpha^{-i}$. Then the term

$$\prod_{\ell \in \mathcal{I},\, \ell \neq j} \left(1 - \alpha^\ell z\right) = \begin{cases} 0 & \text{if } j \neq i \\ \prod_{\ell \in \mathcal{I},\, \ell \neq i} \left(1 - \alpha^{\ell - i}\right) & \text{if } j = i \end{cases}.$$

Hence,

$$\left.\frac{d}{dz} L(z)\right|_{z = \alpha^{-i}} = -\alpha^i \prod_{\ell \in \mathcal{I},\, \ell \neq i} \left(1 - \alpha^{\ell - i}\right).$$

[4]

(d) It is straightforward to see that $s_j = \sum_{i \in \mathcal{I}} e_i \alpha^{i(j+1)}$ and

$$\begin{aligned} S(z) &= \sum_{j=0}^{3} z^j \sum_{i \in \mathcal{I}} e_i \alpha^{i(j+1)} \\ &= \sum_{i \in \mathcal{I}} e_i \alpha^i \sum_{j=0}^{3} \alpha^{ij} z^j \\ &\equiv \sum_{i \in \mathcal{I}} e_i \alpha^i \sum_{j=0}^{\infty} \left(\alpha^i z\right)^j \bmod z^4 \\ &= \sum_{i \in \mathcal{I}} e_i \alpha^i \frac{1}{1 - \alpha^i z}. \end{aligned}$$

[4]

(e) Let $g(x) = \sum_{i=0}^{\ell} g_i x^i$ where $g_i \in \mathbb{F}_3$. Since $\boldsymbol{g} = [g_0, \cdots, g_7] \in \mathcal{C}$, it holds $\boldsymbol{g}\boldsymbol{A}^T = \boldsymbol{0}$. Hence, α, α^2, α^3, and α^4 are roots of $g(x)$. The corresponding cyclotomic cosets (of 3 mod 8) are given by $C_1 = \{1, 3\}$, $C_2 = \{2, 6\}$, and $C_4 = \{4\}$. Therefore,

$$
\begin{aligned}
g(x) &= \operatorname{lcm}\left(M^{(1)}(x), M^{(2)}(x), M^{(3)}(x), M^{(4)}(x)\right) \\
&= \operatorname{lcm}\left(M^{(1)}(x), M^{(2)}(x), M^{(4)}(x)\right) \\
&= M^{(1)}(x) \cdot M^{(2)}(x) \cdot M^{(4)}(x),
\end{aligned}
$$

where

$$
\begin{aligned}
M^{(1)}(x) &= (x - \alpha)\left(x - \alpha^3\right), \\
M^{(2)}(x) &= \left(x - \alpha^2\right)\left(x - \alpha^6\right), \\
M^{(4)}(x) &= \left(x - \alpha^4\right).
\end{aligned}
$$

[4]

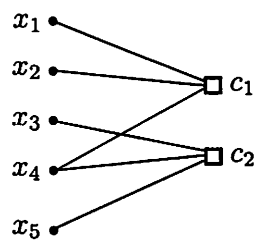
Solutions of Question 5.

(a) It is straightforward to compute that

$$\Pr(x=0|y=0)=1,\ \Pr(x=1|y=0)=0,$$
$$\Pr(x=0|y=1)=0,\ \Pr(x=1|y=1)=1,$$
$$\Pr(x=0|y=?)=\frac{1}{2},\ \Pr(x=1|y=?)=\frac{1}{2},$$

by using the Bayes' theorem $\Pr(x|y)=\Pr(y|x)\Pr(x)/\Pr(y)$ and the fact that $\Pr(y)=\sum_x \Pr(y|x)\Pr(x)$. [5]

(b) The Tanner graph is given by



[5]

(c) Note that

$$\begin{aligned}\Pr(x_3=0|\mathbf{y}_{3:5}) &= \Pr(x_3=0,\ x_4+x_5=0|\mathbf{y}_{3:5})\\ &= \Pr(x_3=0|y_3)\Pr(x_4+x_5=0|y_4y_5)\\ &= \frac{1}{2}(\Pr(x_4=1|y_4)\Pr(x_5=1|y_5)\\ &\quad + \Pr(x_4=0|y_4)\Pr(x_5=0|y_5))\\ &= \frac{1}{2}(1+0)=\frac{1}{2},\end{aligned}$$

and

$$\begin{aligned}\Pr(x_3=1|\mathbf{y}_{3:5}) &= \Pr(x_3=1,\ x_4+x_5=1|\mathbf{y}_{3:5})\\ &= \Pr(x_3=1|y_3)\Pr(x_4+x_5=1|y_4y_5)\\ &= \frac{1}{2}(0+0)=0.\end{aligned}$$

We decode x_3 as 0. [5]

(d) Let $\mathcal{I}$ denote the index set containing the locations of the erased bits and $\mathcal{I}^c$ be its complement. Assume that $|\mathcal{I}|\le d-1$. For any two different codewords $\boldsymbol{c}$ and $\boldsymbol{c}'$ from the codebook $\mathcal{C}$, it holds

$$\begin{aligned} d &\leq d_H(\boldsymbol{c}, \boldsymbol{c}') = d_H(\boldsymbol{c}_{\mathcal{I}}, \boldsymbol{c}'_{\mathcal{I}}) + d_H(\boldsymbol{c}_{\mathcal{I}^c}, \boldsymbol{c}'_{\mathcal{I}^c}) \\ &\leq d - 1 + d_H(\boldsymbol{c}_{\mathcal{I}^c}, \boldsymbol{c}'_{\mathcal{I}^c}), \end{aligned}$$

which implies that $d_H(\boldsymbol{c}_{\mathcal{I}^c}, \boldsymbol{c}'_{\mathcal{I}^c}) \geq 1$. In other words, the mapping that maps a codeword $\boldsymbol{c}$ to $\boldsymbol{c}_{\mathcal{I}^c}$ is one to one. The inverse mapping is well defined. The erasures can be corrected. [5]