MATH-315301

This question paper consists of 5 printed pages, each of which is identified by the reference MATH-3153

Only approved basic scientific calculators may be used.

Examination for the Module MATH-3153

(May/June 2011)

Coding Theory

Time allowed: 2.5 hours

Attempt no more than **four** questions. All questions carry equal marks.

1. Let Σ_q denote a set of symbols (an 'alphabet') of size q. That is, $|\Sigma_q| = q$. We shall assume that there is a 'zero' element $0 \in \Sigma_q$. Let Σ_q^n denote the n-th Cartesian power of Σ_q.

   (a) (i). A q-ary code of length n is a subset of Σ_q^n. How many of these are there (as a function of q and n).

   (ii). Write down $P(\Sigma_2^2)$, the complete set of codes from Σ_2^2 (you may take $\Sigma_2 = \{0, 1\}$).

   (b) (i). Define the Hamming distance d on Σ_q^n.

   (ii). Note that $00...0 \in \Sigma_q^n$. How many elements of Σ_q^n have Hamming distance 2 or less from the element 00...0?

   (iii). Give the definition of the notion of *equivalence* of codes. How many equivalence classes of codes are there in $P(\Sigma_2^2)$ as defined above.

   (iv). Define the minimum distance $d(C)$ of a code $C \subset \Sigma_q^n$.

   (v). Given that code C is 5 error correcting, what is the smallest that $d(C)$ could be.

   (vi). State the singleton bound on the size M of a q-ary (n, M, d)-code C.

   (c) For each of the following triples (n, M, d) construct, if possible, a binary (n, M, d)-code:

$$(5, 2, 5) \qquad (3, 5, 1) \qquad (4, 8, 2) \qquad (7, 90, 3)$$

   If no such code exists, then prove it, stating any theorems used.

2. Write $\mathcal{M}_{n,m}(F)$ for the set of $n \times m$ matrices with entries in field F. For example

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix} \in \mathcal{M}_{2,5}(\mathbb{Z}_2)$$

(a) If F is a field of order p^e, what is $|\mathcal{M}_{n,m}(F)|$?

(b) Let $r > 1$. Define H_r to be the matrix whose columns are the non-zero vectors in $\mathbb{Z}_2^r$ (where $\mathbb{Z}_2$ is the field of order 2), arranged in any order of your choice, so long as the leading square submatrix is the identity matrix. The Hamming $[n = 2^r - 1,\ k = 2^r - r - 1,\ 3]$-code C may be defined as the linear code whose parity check matrix is H_r.

Write out H_2 and H_3. Write out the corresponding generator matrices.

(c) Write down the binary linear code C_1 with generator matrix

$$G_1 = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

(d) Consider the following sets: $S_1 = \{000000, 000100, 100000, 110000, 111000\} \subset \mathbb{Z}_2^6$;

$$S_2 = \{0000, 0010, 1000, 1010\} \subset \mathbb{Z}_2^4;$$

$$S_3 = \{00000, 01111, 01000, 00111\} \subset \mathbb{Z}_2^5;$$

$$S_4 = \{0000, 0001, 1000, 1001\} \subset \mathbb{Z}_5^4$$

$$S_5 = \{0000, 1000, 1001\} \subset \mathbb{Z}_2^4$$

Determine which of these are linear codes (giving the reasons for your answers).

(e) Define the minimum weight $w(C)$ for a code. Prove that, for a linear code, the minimum distance $d(C)$ is equal to $w(C)$.

(f) Define $C^\perp$, the dual code to a linear code C. Prove that $C^\perp$ is also a linear code.

(g) A linear code is self-dual if $C^\perp = C$. Compute the dual of C_1 above, and hence or otherwise determine if it is self-dual.

Continued ...

3. (a) Consider the ring $\mathbb{Z}_2[x]$ of polynomials with coefficients in $\mathbb{Z}_2$. Explain why $(x^2+1) = (x+1)(x+1)$ is an identity in this ring. Show that the polynomial $x^3+x+1 \in \mathbb{Z}_2[x]$ does not have a root in $\mathbb{Z}_2$.

Suppose that we extend the field $\mathbb{Z}_2$ by an element x obeying $x^3+x+1=0$. Write down the set F of elements of the resultant field.

Check that this is a field by computing the multiplicative inverses of all nonzero elements in F.

(b) Let $C \subset \mathbb{Z}_7^5$ be the linear code with generator matrix

$$G' = \begin{pmatrix} 1 & 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 5 & 6 \\ 0 & 1 & 0 & 3 & 4 \end{pmatrix}$$

(i). By a suitable row permutation, bring the matrix G' above into a standard form G. Hence write down a parity check matrix H for C.

(ii). Explain the following statement: Permuting rows of a generator matrix for a code does not change the code; permuting columns does not change the equivalence class of code.

(iii). Compute the matrix $G.H^t$ (where H^t is the transpose of H, with G and H as above). Interpret your result.

(iv). Show that $d(C) = 3$.

(v). How many of the coset leaders of C have weight 1? Explain your answer.

(vi). Codeword x is transmitted down a noisy channel, so that $y = 11244$ is received, with exactly one error having occured. What was the transmitted codeword x?

Continued ...

4. Let C, C' be the binary linear codes with generator matrices

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

respectively.

(a) Construct a standard array for each of C and C'.

(b) Decode the received message 1101 using your array for C.

(c) Code C is transmitted down a binary symmetric channel with symbol error probability $p = 0.01$, with the received vectors being decoded by the coset decoding method. Let x denote the sent word, and y the received word, so that $e = y - x$ is the transmission error vector. Then $P(e = 0000)$ denotes the probability of a codeword being transmitted without error. What is $P(e = 0000)$? (Explain your answer.)
Explain why the probability of a transmitted word being decoded correctly can be written as

$$P_{corr}(C) \;=\; P(e = 0000) + P(e = 1000) + P(e = 0100) + P(e = 1100)$$

Calculate $P_{err}(C)$, the word error probability of the code; and $P_{undetec}(C)$, the probability of there being an undetected error in a transmitted word.

If code C' is transmitted, and $p = 0.02$, what is $P(e = 00001)$ in this case?

(d) Code C is again transmitted down a binary symmetric channel with symbol error probability $p = 0.01$, but is now used only for error detection. If an error is detected in a received vector, the receiving device requests retransmission of the codeword. Calculate $P_{retrans}(C)$, the probability that a single codeword transmission will result in a request to retransmit.

Continued ...

5. (a) A cyclic shift of a codeword removes the first symbol from the word and places it at the end. A linear code C_c is said to be cyclic if any cyclic shift of a codeword in C_c is also a codeword in C_c. Write down a subset of $\mathbb{Z}_2^3$ of order 4 that is a cyclic code.

(b) (i). The 26 letters of the alphabet may be represented in $\mathbb{Z}_3^3$ by $A \mapsto 001$, $B \mapsto 002$, $C \mapsto 010$, ..., $Z \mapsto 222$. Let us also represent ‘space’ by 000.

We are given the parity check matrix

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

of a 3-ary $[6, 3, d]$-code C.

Compute $H(112000)^t$ and hence determine whether $112000 \in C$.

(ii). Note that H as above is not in standard form. Confirm that

$$G = \begin{pmatrix} 2 & 2 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 & 2 & 1 \end{pmatrix}$$

is a generator matrix for C.

(iii). Recall that G as above may be used to encode elements $u = (u_1, u_2, u_3)$ of $\mathbb{Z}_3^3$ by $u \mapsto x = uG$. Thus it may be used to encode letters of the alphabet, via our representation above. Compute the encoded form of the letter R.

(iv). What is $d(C)$ here? How many coset leaders lie within distance 1 of 000000? Compute their syndromes.

(v). Decode as much as possible of the following received message. You may assume that the transmitted message was encoded using C with generator matrix G as above, and use nearest neighbour decoding. (Marks are available for partial decodings, but all working must be shown.)

Message:

000000 012212 220112 112100 220112 000001
200021 112000 220112 000000 022021 221000
022200 002000 212012 202100 111112 220112
012022

End.