MATH-315201

This question paper consists of 5 printed pages, each of which is identified by the reference MATH-3152

Only approved basic scientific calculators may be used.

Examination for the Module MATH-3152

(June 2008)

Coding Theory

Time allowed: 2 hours

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

1. (a) Let Σ_q be an alphabet of size q. Explain what is meant by the n-th Cartesian product of Σ_q, denoted Σ_q^n. Illustrate your answer by writing out all elements of $\{0,1\}^3$ explicitly, carefully explaining any notation you use.

   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).

   Roughly determine an n and a q such that this entire document is a codeword in some $C \subset \Sigma_q^n$. (You do not need to know how many Greek letters there are. A rough estimate of the number of symbols will do.)

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

   (ii). Show that Hamming distance satisfies the triangle inequality.

   (iii). Suppose that $0 \in \Sigma_q$, so 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?

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

   (v). Given that code C is 7 error correcting, state a lower bound on $d(C)$.

   (vi). State the ball-packing 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:

   $$(6,2,6) \qquad (3,8,1) \qquad (4,8,2) \qquad (8,40,3)$$

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

   (d) Suppose that the probability of error in transmission of a single digit is $p < 1/2$. Show that, given a particular message w received, and a codeword v such that $d(w, v)$ is minimal, then there is no better guess than v for the transmitted codeword.

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

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \in \mathcal{M}_{2,3}(\mathbb{Z}_2)$$

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

(b) Associated to each $M \in \mathcal{M}_{n,m}(F)$ is the row space $R(M)$ of M. This is the vector space over F spanned by the rows of M (regarded as vectors). Under what conditions is M a generator matrix for a linear code; and what kind of code does it generate (that is, what are the parameters of the code it generates)?

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

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

(d) Consider the following sets: $S_1 = \{0000, 0001, 1000\} \subset \mathbb{Z}_2^4$;

$$S_2 = \{0000, 0001, 1000, 1001\} \subset \mathbb{Z}_2^4;$$

$$S_3 = \{00000, 01110, 01000, 00110\} \subset \mathbb{Z}_2^5;$$

$$S_4 = \{0000, 0001, 1000, 1001\} \subset \mathbb{Z}_3^4.$$

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.

(h) Give an example of an error correcting linear code used by humans in everyday life.

3. (a) Explain a way to construct a field of order 4. Write down the addition and multiplication tables for this field.

Construct the table of multiplicative inverses for the field $\mathbb{Z}_7$.

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

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

(i). Write down a parity check matrix H for C.

(ii). Compute the matrix $G.H^t$ (where H^t is the transpose of H). Interpret your result.

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

(iv). How many of the coset leaders of C have weight 1?

(v). Codeword x is transmitted down a noisy channel, so that $y = 11254$ is received, with exactly one error having occured. What was the transmitted codeword x?

Continued ...

4. (a) Let C be the binary linear code with generator matrix

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

(i). Construct a standard array for C.

(ii). Decode the received message 1101 using your array.

(iii). 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. 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.

(iv). 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.

(b) Give the definition of the syndrome of a received word. Prove that two words have the same syndrome iff they lie in the same coset of the code C.

Continued ...

5. 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. That is, $w \in C$ iff $Hw^t = 0$. (As usual we write simply 0 for the zero vector, where no ambiguity can arise.)

For example $H(100012)^t = 0$, so $100012 \in C$.

(a) Note that H 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.

(b) Recall that G 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 E.

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

(d) Decode as much as possible of the following received message, given that the transmitted message was encoded using C with generator matrix G, assuming nearest neighbour decoding. (Marks are available for partial decodings, but all working must be shown.)

Message:

212012 012212 220112 112100 220112 000000

200021 112000 220112 000000 022021 221000

022200 002000 022021 221000 111112 022022

Hints:

(i). The message digits in 212012 are 202 (why?)

(ii). 202 is the representation of the 20-th letter: T.

(iii). The message digits in 012212 are 222. What is going on here?

End.