MATH-315201

This question paper consists of 6 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

(May 2009)

Coding Theory

Time allowed: 2 hours

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

1. In this question S will be our alphabet set, with

$$|S| = q$$

(a) Explain precisely what is meant by the Cartesian product $S \times S$, also denoted S^2.

Explain the formal sense in which

$$(S \times S) \times S \neq S \times (S \times S)$$

Construct a natural bijection between these sets, and hence explain how the n-fold Cartesian product S^n is understood.

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 S^n. How many of these are there (as a function of q and n)?

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

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

(iii). Suppose that $0 \in S$, so that $00...0 \in S^n$. How many elements of S^n have Hamming distance 2 or less from the element 00...0?

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

(v). Given that code C is 5 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:

$$(4, 2, 4) \qquad (3, 8, 1) \qquad (4, 8, 2) \qquad (8, 41, 3)$$

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

Continued ...

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

   (a) Recall that the rank of a matrix is the number of linearly independent rows. Write out all the elements of $\mathcal{M}_{2,2}(\mathbb{Z}_2)$ of rank 2.

   (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

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

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

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

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

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

$$S_5 = \{0000, 0001, 0002\} \subset \mathbb{Z}_3^4.$$

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

   (e) In each of the cases S_i above that *are* linear codes, write down a generator matrix for this code.

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

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

3. (a) Let $D \subset F^n$ be a linear code over field F. For $x \in F^n$, give the formal definition of the coset $x + D$.

(b) 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 1110 using your array.

(iii). By reference to your array, or otherwise, construct an example which shows that C is not single error correcting.

Under what circumstances *does* C correct a single error?

(iv). Code C is transmitted down a binary symmetric channel. Given that a single error has occurred, what is the probability that this single error will be corrected?

(v). 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.

Continued ...

4. (a) Write down all degree 2 polynomials over $\mathbb{Z}_2$. Show that only one of these is irreducible.

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

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

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

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

   (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?

5. 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.)

(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) The 26 letters of the Roman 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.

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 Y.

(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:

222221 101200 202100 000000 200021 112000

220112 000001 212012 012212 220112 000000

011021 212012 200021 112100 022022 000000

End.