This question paper consists of
4 printed pages, each of which
is identified by the reference MATH315301.

All calculators must carry
an approval sticker issued
by the School of Mathematics.

School of Mathematics

January 2017

MATH315301

Coding Theory

Time Allowed: $2\frac{1}{2}$ hours

Answer no more than 4 questions.
If you attempt 5, only the best 4 will be counted.
All questions carry equal marks.

Turn Over

1. (a) Let Σ_q be an alphabet of size q and $C \subseteq \Sigma_q^n$ be a q-ary block code of length n such that $|C| = M$. (You are reminded that, for $x, y \in \Sigma_q^n$, the Hamming distance $d(x, y)$ is the number of positions at which x and y differ.)

(i) Define the *minimum distance* $d(C)$ of the code C.

(ii) Define $A_q(n, d)$ (in terms of q-ary (n, M, d) codes).

(iii) Prove that $A_2(4, 3) = 2$.

(b) (i) State the *ball packing bound* for any q-ary (n, M, d)-code in the form "$A_q(n, d) \leq$ (a suitable expression)".

(ii) Define what it means for a q-ary (n, M, d)-code C to be perfect.

(iii) For each of the following triples determine whether a perfect binary (n, M, d)-code exists. (Marks are only awarded if your reasons are clearly stated.)

$$(24, 2^{12}, 8) \quad (63, 2^{57}, 3)$$

(c) For each of the following triples construct a binary (n, M, d)-code if one such exists.

$$(4, 8, 2) \quad (7, 5, 5)$$

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

(d) Codewords from the binary Hamming code $C = \mathrm{Ham}(\mathbb{Z}_2^5)$ are transmitted via a binary symmetric channel with the probability of error in transmission of a single digit (i.e. the symbol error probability) being p. Calculate in terms of p the probability that, when a codeword u is transmitted, a different codeword $v \neq u$ is recovered (using nearest neighbour decoding).

2. You are reminded that $\mathcal{M}_{m,n}(F)$ denotes the set of $m \times n$ matrices with entries in a field F, and that F_q denotes the finite field of q elements.

(a) Define what it means to say that C is a *linear* $[n, k]$*-code* over F_q. Determine the number of codewords of such a code C and the *rate of information* of C, stating clearly your reasons.

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

(c) Consider the matrices $M_1, M_2, M_3 \in \mathcal{M}_{2,4}(\mathbb{Z}_7)$ such that

$$M_1 = \begin{pmatrix} 1 & 1 & 5 & 1 \\ 2 & 1 & 1 & 6 \end{pmatrix} \qquad M_2 = \begin{pmatrix} 0 & 1 & 2 & 4 \\ 1 & 0 & 5 & 3 \end{pmatrix} \qquad M_3 = \begin{pmatrix} 1 & 4 & 5 & 6 \\ 2 & 1 & 3 & 5 \end{pmatrix}$$

(i) Explain which of these matrices are generator matrices for codes in $\mathbb{Z}_7^4$.

(ii) For the generator matrices in part (i) find generator matrices in standard form and hence find parity check matrices for the associated codes.

(iii) Determine $d(C)$ for the codes found in part (ii) and in each case exhibit two codewords $x, y \in C$ such that $d(x, y) = d(C)$. Also decide whether any of these codes is self dual, stating clearly your reasoning.

3. (a) Let C be the binary linear code with generator matrix $G = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$.

(i) Construct a standard array for C.

(ii) Code C is transmitted down a binary symmetric channel with symbol error probability $p = 0.03$ with the received words being decoded using the coset decoding method. Calculate $P_{\mathsf{err}}(C)$, the word error probability of the code, and $P_{\mathsf{undetec}}(C)$, the probability of there being an undetected error in a transmitted word.

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

(c) Suppose that n is some positive integer.

(i) Let x and y be words in $\mathbb{Z}_2^n$. Show that, if x and y are either both of even weight, or both of odd weight, then the word $x + y$ has even weight.

(ii) Let x and y be words in $\mathbb{Z}_2^n$. Show that, if exactly one of x, y has odd weight, then the word $x + y$ has odd weight.

(iii) Using parts (i) and (ii) or otherwise prove that, for a binary linear code C either all the codewords have even weight or exactly half of the codewords have even weight.

4. (a) (i) Explain why $\mathbb{Z}_9$ is *not* a field under its natural operations of addition and multiplication.

(ii) Let $f(x) = 2 + x + x^2 \in \mathbb{Z}_3[x]$. Consider the quotient ring $\mathbb{Z}_3[x]/f(x)$ equipped with its natural operations of addition and multiplication. Write down the set of elements belonging to $\mathbb{Z}_3[x]/f(x)$ and write down the row of its multiplication table corresponding to the element x. Explain briefly whether or not $\mathbb{Z}_3[x]/f(x)$ is a field.

(b) We can construct the field F_4 as the set of elements $\{0, 1, a, b\}$ equipped with addition and multiplication satisfying the following tables.

$+$	0	1	a	b
0	0	1	a	b
1	1	0	b	a
a	a	b	0	1
b	b	a	1	0

$+$	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

Consider the linear code $C \subseteq F_4^5$ with generator matrix $G' = \begin{pmatrix} 1 & a & 1 & a & 1 \\ 1 & 0 & 0 & b & b \\ 1 & a & 0 & 1 & 0 \end{pmatrix}$.

(i) Derive a generator matrix G for C in standard form and write down a parity check matrix H for C. Also determine $d(C)$.

(ii) Your generator matrix G is used to encode messagewords. Compute the codeword corresponding to the messageword $ba1$.

(iii) You receive words $v = abb0b$ and $w = ba11a$ after transmission via a noisy channel. For $y \in \{v, w\}$ compute the syndrome $S(y)$. Given that at most one error occurs during transmission find the two codewords that were sent explaining precisely why the decoding method that you use is correct.

5. You are given that the following matrix

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

is a *parity check matrix* of a 3-ary $[6, 3, d]$ (linear) code C. That is $w \in C$ if and only if $wH^t = 0$. (As usual we write 0 for the zero vector.) Note that H is not in standard form.

(a) Verify that

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

is a generator matrix for C.

(b) The 26 letters of the Roman alphabet may be represented in $\mathbb{Z}_3^3$ by mapping $A \mapsto 001$, $B \mapsto 002$, $C \mapsto 010$, $D \mapsto 011 \ldots H \mapsto 022 \ldots Z \mapsto 222$. 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 S.

(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 codings, but all work must be shown.)

The Message:

$$111210 \quad 111122 \quad 021112 \quad 000000 \quad 2002202 \quad 020020 \quad 220211$$

(e) A single codeword $x \in C$ is transmitted twice down a noisy channel so that you receive two words $y_0, y_1 \in \mathbb{Z}_3^6$. You are given that $n_0 + n_1 \leq 3$ where $n_i = d(x, y_i)$ for each $0 \leq i \leq 1$. Using the syndromes that you computed and the technique that you applied in part (d) you recover from y_0, y_1 two codewords $z_0 \neq z_1$.

Describe, in terms of the syndromes $S(y_0), S(y_1)$ and the number of errors n_0, n_1, the case(s) in which you are ARE NOT able to decide which of z_0, z_1 is the original codeword x and the case(s) in which you ARE able to decide this.

End.