Coding Theory 2011 Answers

Questions are SEEN or similar to seen unless otherwise stated.

1. (a) $|P(\Sigma_q^n)| = 2^{(q^n)}$
(also accept count excluding empty set). (1 marks)

(b) $P(\{0,1\}^2) = P(\{00, 01, 10, 11\}) =$
$\{\emptyset, \{00\}, \{01\}, \{10\}, \{11\}, \{00, 01\} ... \{00, 01, 10, 11\}\}$
(answer excluding $\emptyset$ is also ok) (3 marks)

(c) i. Hamming distance $d(x, y) = \#\{i | x_i \neq y_i\}$ (1 marks)

ii. Weight 0/1: Vectors of form $(0, 0, ..., 0, X, 0, ..., 0)$. There are $n \times (q - 1) + 1$ of these.
Weight 2: Vectors of form $(0, 0, ..., 0, X, 0, .., Y, 0, ..., 0)$ with $X, Y \neq 0$. There are $\frac{n(n-1)}{2} \times (q - 1)^2$ of these. (2 marks)

iii. Two codes are equivalent if there is a set bijection f between them such that $d(f(x), f(y)) = d(x, y)$ for all x, y.

Note that our example is binary. It follows that the map f_i on $\mathbb{Z}_2^n$ defined by flipping the i-th symbol in each string is a bijection (indeed an involution) and reduces to an equivalence on any code. Let us write $f_{-}C$ for the orbit of equivalent codes under the action of all the f_is.
Discard $\emptyset$. All codes of order 1 are trivially equivalent. Codes of order 2 are equivalent iff the unique nontrivial distance is the same. The codes of order 3 are equivalent, since their complements are order 1 and f_{-} acts transitively. There is only one code of order 4 here.
(2 marks)

iv. minimum distance $d(C) = min\{d(x, y) | x, y \in C, \ x \neq y\}$
(2 marks)

v. For C to be t error correcting requires $d(C) \geq 2t+1$. (1 marks)

vi. ball-packing bound on the size M of a q-ary (n, M, d)-code C:

$$M \sum_{r=0}^{t} \binom{n}{r} (q-1)^r \leq q^n$$

where t such that $d \geq 2t+1$.

singleton bound: $M \leq q^{n-(d-1)}$ (2 marks)

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

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

(for given values of X, M). If no such code exists, then prove it, stating any theorems used.
ANSWER: (X,2,X): $\{000000...0, 111111...1\}$.
(3,8,1): $\{000, 001, 010, 011, 100, 101, 110, 111\}$
Throw any three away to get (3,5,1).
(4,8,2): $\{0000, 0011, 0101, 0110, 1001, 1010, 1100, 1111\}$
(8,M,3): fails the BP bound if:

$$M(1+8) = 9 * M \not\leq 2^8 = 256$$

so fails for $M > 256/9$ (e.g. $M > 28$).
(7,90,3) also fails the singleton bound.

(12 marks)

(/25 marks)

2. (a) Set $q = p^e$, then $|\mathcal{M}_{n,m}(F)| = q^{nm}$ (2 marks)

(b) $H_2 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$

$$H_3 = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$

$G_2 = (1, 1, 1)$

$$G_3 = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$

(4 marks)

(c)

$$C_1 = \{00000, 10100, 01100, 11000\}$$

(2 marks)

(d) S_1 not closed under $+$.
S_2 is closed under linear combinations, so linear code.
S_3 is closed under linear combinations, so linear code.
S_4 is not closed under $+$.
S_5 is not closed under $+$. (4 marks)

(e) minimum weight $w(C) = min\{w(x) | x \in C \setminus \{0\}\}$ (where $w(x)$ is weight of x (define it!), and 0 denotes the zero vector).
Prove that, for a linear code, the minimum distance $d(C)$ is equal to $w(C)$.:
$d(x, y) = d(x - y, 0) = w(x - y)$ $\square$ (3 marks)

(f) Define $C^\perp$, the dual code to a linear code C.
$C \subset F_q^n$, $C^\perp = \{v \in F_q^n | v.x = 0 \forall x \in C\}$ where $v.x = \sum_i v_i x_i$ (over F).

Prove that $C^\perp$ is also a linear code:
$v.x = 0$ is a linear constraint on $\{v_i\}$ for any given x. (5 marks)

(g) Compute the dual of C_1 above, and hence or otherwise determine if it is self-dual.
Ignoring the last two digits (which are always zero in C_1) for now, we have
$(x, y, z).(1, 0, 1) = x + z = 0$
$(x, y, z).(0, 1, 1) = y + z = 0$
$(x, y, z).(1, 1, 0) = x + y = 0$
These imply $x = y = z$. The last two digits in the dual are not constrained, so $C^\perp = \{000, 111\} \times \mathbb{Z}_2^2$ (in the obvious notation) (any equivalent, such as giving a PCM, is acceptable).
So $C_1^\perp \neq C_1$. So C_1 is not self-dual. (5 marks)

3. (a) $(x+1)(x+1) = x^2 + 2x + 1 = x^2 + 0x + 1$

$p(1) = p(0) = 1$

Since the polynomial is cubic is it enough to evaluate at all points to show irreducibility.

(3 marks)

$F = \{0, 1, x, 1+x, x^2, 1+x^2, x+x^2, 1+x+x^2\}$

$x(x^2+1) = 1$

$(1+x)(x^2+x) = 1$

$x^2(1+x+x^2) = 1$

(5 marks)

(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.

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

(2 marks)

ii. The code is generated by the rows, and the order of writing the generators does not matter. The column perm changes the roles of the components of the vectors — i.e. perms the 'axes'. The Hamming distance is invariant under this perm.

(2 marks)

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

$GH^t = 0$ (show calculations)

Columns of H^t are rows of H so GH^t assembles the various inner product calculations for C and $^{\perp}$, which must all be zero by definition. (2 marks)

iv. Show that $d(C) = 3$.
H has no zero or parallel columns, but $w(G_3) = 3$ so $d(C) \leq 3$.
So $d(C) = 3$. (3 marks)

v. How many of the coset leaders of C have weight 1?
There are 7^2 coset leaders. There are $5 \times 6 = 30$ weight 1 vectors, none of which lie in C, and no distint pair of which have $x - y \in C$. So number $= 30$.
(full marks for any legitimate argument with right final answer) (3 marks)

vi. Codeword x is transmitted down a noisy channel, so that $y = 112\tau4$ is received (some τ), with exactly one error having occured. What was the transmitted codeword x?

$Hy^t = \begin{pmatrix} \tau \\ 0 \end{pmatrix}$. Now find coset leader: $H \begin{pmatrix} 0 \\ 0 \\ 0 \\ \tau \\ 0 \end{pmatrix} = \begin{pmatrix} \tau \\ 0 \end{pmatrix}$

(3 marks)

so $x = 112\tau4 - 000\tau0 = 11204$. (2 marks)

4. (a) a standard array for $C = \{0000, 1010, 0101, 1111\}$:

0000 1010 0101 1111

1000 0010 1101 0111

0100 1110 0001 1011

1100 0110 1001 0011

(any correctly formed standard array is acceptable)

For C' we get the first four rows by appending 0 to each vector in the array above; and a further four rows formed by appending 1 to each vector in the array above.

(8 marks)

(b) Decode the received message 1101 using your array:
(IF the coset leaders are as above then)
the coset leader is 1000, so 1101-1000=0101 is the decoding.

(3 marks)

(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. ...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.
ANSWER:
$P(e = 0000) = (1 - p)^4$

(2 marks)

$P_{corr}(C)$ takes the given form since an error (including the null error) is corrected if it takes any of these forms, and not corrected otherwise. Thus

$$P_{corr}(C) =$$

$$P(e = 0000) + P(e = 1000) + P(e = 0100) + P(e = 1100) =$$

$$(1 - p)^4 + 2p(1 - p)^3 + p^2(1 - p)^2 = 0.9801$$

$$P_{err}(C) = 1 - P_{corr}(C) = 0.0199$$

(3 marks)

For there to be an undetected error in the transmitted word the received word would have to be in C, but in error. That means both transmitted word x and received word y are in C (and are different), so the error $e = x - y$ is also in C (and of course is not the zero word). Thus

$$P_{undetec}(C)|_{p=0.01} = P(e = 1010) + P(e = 0101) + P(e = 1111)$$
$$= 2 \times (0.01)^2(0.99)^2 + (0.01)^4 = 0.00019603$$

(3 marks)

For C', $P(e = 00001) = (1-p)^4 p$

(1 marks)

(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.

$$P_{retrans} = 1 - P(\textit{no error detected})$$
$$= 1 - P(\textit{no error}) - P(\textit{undetected error})$$
$$P_{undetec} = 2p^2(1-p)^2 + p^4$$

(2 marks)

so

$$P_{retrans}(C) = 1 - (1-p)^4 - 2p^2(1-p)^2 - p^4 \ = \ 4p + O(p^2)$$

so

$$P_{retrans}(C)|_{p=0.01} = etc. \sim 0.04$$

(3 marks)

(UNSEEN)

5. (a) $C_c = \{000, 101, 011, 110\}$

(3 marks)

(b) $H(221000)^t = 0$ so $221000 \in C$. Similarly 112000. (1 marks)

(c) Confirm that G is a generator matrix for C:
1. rows linearly independent
2. $GH^t = ...calculation... = 0$
3. # rows = 6-3

All ok. (2 marks)

(d) Compute the encoded form of the letter U:
(another example: E is 5th letter, so rep is 012 and encoding is 220112)
U is represented by 210 and its encoding is 202100
Similarly, R (200) encodes as 112000. (3 marks)

(e) What is $d(C)$?
Clearly $d(C) \leq 3$, but no column of H is "parallel" to another, so $d(C) = 3$. (3 marks)

This implies no y with $w(y) = 1$ or 2 lies in C.

Now suppose x, y of wt 1 lie in $C + x$. Then $y - x$ lies in C. But $w(y - x) \leq 2$, so wt.1 vectors lie in distinct cosets.
There are $6 \times 2 = 12$ of them. Syndromes:

$S(000000) = 000$

$S(100000) = 100$

$S(200000) = 200$

etc

$S(000001) = 101$

$S(000002) = 202$

(4 marks)

(f) Message:

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

Now:

012212.$H^t = 220$ so we have 012212-002000=010212, which gives 022, which gives H;

we already encoded E to get the next word;

212012.$H^t = 000$ so we have 202, which gives T;

...and so on, until

$$(022021)\begin{pmatrix} 100 \\ 010 \\ 110 \\ 200 \\ 001 \\ 101 \end{pmatrix} = (010) = S(010000)$$

giving 201, and hence S;

and so on, until the last vector:

$$(012022)\begin{pmatrix} 100 \\ 010 \\ 110 \\ 200 \\ 001 \\ 101 \end{pmatrix} = (101) = S(000001)$$

which thus corrects to 012022-000001=012021, giving 201, and hence S again.

Altogether we get:

[space]HERE ARE SIX TUNES

(9 marks)