SECTION A

Answer **ALL** questions in this section (40 marks in total)

Feedback: ***General feedback on the 2018 exam***

For most students, MATH32032 was their very first encounter with error-correcting codes and the associated theory. As well as introducing those new objects, the course offered the students a good opportunity to apply their mathematical skills to problems at a level of difficulty consistent with the final semester of a three-year Honours degree in Mathematics.

Having marked the exam, I can say that the students who demonstrated excellent knowledge of Coding Theory language, setup, methods and algorithms, and were able to correctly apply the methods and algorithms in a practical situation, generally obtained at least a 2:1 class result.

Out of those, students who were comfortable with the most difficult proofs given in the course and confident in tricky manipulations with matrices and polynomials typically achieved a 1st class result.

Finally, those who obtained high first-class marks were typically able to make progress in some of the more challenging questions which required top problem-solving skills and an ability to use non-trivial ideas from basic combinatorics (A3b), linear algebra (B4d), calculus (B5d) or elementary number theory (B6d).

Below are the questions, solutions and post-examination feedback. The paper tested the Intended Learning Outcomes (ILO) of the course; ILO numbers next to each question refer to the following list.

ILO1 Define and illustrate main concepts and prove fundamental theorems concerning error-correcting codes, given in the course.

ILO2 Calculate the parameters of given codes and their dual codes using standard matrix and polynomial operations.

ILO3 Encode and decode information by applying algorithms associated with well-known families of codes.

ILO4 Compare the error-detecting/correcting facilities of given codes for a given binary symmetric channel.

ILO5 Design simple linear or cyclic codes with required properties.

ILO6 Solve mathematical problems involving error-correcting codes by linking them to concepts from elementary number theory, combinatorics, linear algebra and elementary calculus.

A1. (a) $\boxed{\text{ILO1}}$ Define what is meant by:

— the *weight* $w(\underline{x})$ *of a vector* $\underline{x} \in \mathbb{F}_q^n$;

[**Answer:** the number of non-zero coordinates (*components*, *symbols*) in $\underline{x}$]

— a *linear code* C of length n over the field $\mathbb{F}_q$;

[**Answer:** a subspace of the vector space $\mathbb{F}_q^n$]

— the *weight* $w(C)$ *of a linear code* C;

[**Answer:** $w(C) = \min\{w(\underline{c}) : \underline{c} \in C \setminus \{\underline{0}\}\}$]

— the *inner product* $\underline{x} \cdot \underline{y}$ of vectors $\underline{x}, \underline{y} \in \mathbb{F}_q^n$;

[**Answer:** if $\underline{x} = (x_1, ..., x_n)$, $\underline{y} = (y_1, ..., y_n)$, then $\underline{x} \cdot \underline{y} = \sum\limits_{i=1}^{n} x_i y_i$, *or* $\underline{x} \cdot \underline{y} = \underline{x}\underline{y}^T$]

— the *dual code*.

[**Answer:** if $C \subseteq \mathbb{F}_q^n$ is a linear code, $C^\perp = \{\underline{y} \in \mathbb{F}_q^n : \underline{y} \cdot \underline{c} = 0$ for all $\underline{c} \in C\}$]

Feedback: Done well but typical mistakes included using just a subset instead of a subspace in the definition of a linear code and not leaving out $\underline{0}$ in the definition of $w(C)$. Marks were deducted for either mistake.

(b) $\boxed{\text{ILO2}}$ Consider the binary code $C = \{\underline{x} \in \mathbb{F}_2^n : \underline{x} \cdot \underline{x} = 0\}$ of length n where $n \geqslant 2$. Explain why C is a linear code. State without proof the cardinality, dimension and weight of C. Identify the code C by its well-known name.

Answer. Recall that $\underline{x} \cdot \underline{x} = \sum\limits_{i=1}^{n} x_i^2$ where $x_i \in \mathbb{F}_2 = \{0, 1\}$ for all i so $x_i^2 = x_i$. Hence $C = \{\underline{x} \in \mathbb{F}_2^n : x_1 + \ldots + x_n = 0\}$. Being the set of solutions to a homogeneous linear equation, it is a vector space so a linear code.

Alternatively, the fact that the code is linear can be checked directly: $(\underline{x} + \lambda\underline{y}) \cdot (\underline{x} + \lambda\underline{y}) = \underline{x} \cdot \underline{x} + 2\lambda\underline{x} \cdot \underline{y} + \lambda^2\underline{y} \cdot \underline{y}$ where the first and last term are 0 whenever $\underline{x}, \underline{y} \in C$ and the middle term is 0 because the field is $\mathbb{F}_2$ where $2 = 0$. Hence $\underline{x}, \underline{y} \in C$, $\lambda \in \mathbb{F}_2$ implies $\underline{x} + \lambda\underline{y} \in C$.

The well-known name of this code is the (binary) even weight code of length n.

The code has cardinality 2^{n-1}, dimension $n - 1$ and weight 2.

Feedback: This part was done well, although some students mistakenly stated that C was a self-dual code, which led them to believe that $\dim C = n/2$. The condition $\underline{x} \cdot \underline{y} = 0$ for all $\underline{x}, \underline{y} \in D$ implies that D is a *self-orthogonal* code — not necessarily self-dual — but note that C was defined by a weaker condition. The even weight code E_n is not self-orthogonal if $n > 2$.

[10 marks]

A2.

(a) $\boxed{\text{ILO1 ILO3}}$ Let an $(n - k) \times n$ check matrix H for a q-ary linear code C be given. What is meant by a *table of syndromes* for C? How many rows does such a table have? Explain how to decode a received vector using the table of syndromes.

Answer. A table of syndromes has a row for every coset of $C \subseteq \mathbb{F}_q^n$; the row contains a coset leader $\underline{a}$ and its syndrome $S(\underline{a}) = \underline{a}H^T$. The table has q^{n-k} rows.

To decode a received vector $\underline{y}$, calculate its syndrome $S(\underline{y}) = \underline{y}H^T$ and look it up in the table, locating a row $(\underline{a},\ S(\underline{a}))$ where $S(\underline{a}) = S(\underline{y})$. Then $\underline{a}$ is the coset leader of the coset $\underline{y} + C$, so by the general nearest neighbour decoding for linear codes, $\underline{y}$ is decoded to $\underline{y} - \underline{a}$.

Feedback: A number of students stated that a table of syndromes has a row for each coset leader. This is technically incorrect because if a coset has multiple coset leaders, only one of them appears in the table of syndromes.

From now on, let D be the binary linear code with parity check matrix $H = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$.

(b) $\boxed{\text{ILO3}}$ Construct a table of syndromes for D and use it to decode the vector 11100.

Answer. *Detailed steps (not required)*: the top row of the table of syndromes is 00000, $S(00000) = 00000 \cdot H^T = 00$. Next, we consider vectors of weight 1. Note that the syndrome of such a vector is a (transposed) column of H. Hence we can pick two coset leaders of weight 1 as H has two different columns: say, 10000 with syndrome 01 and 01000 with syndrome 11. We are now forced to look for a coset leader of weight 2 with syndrome 10; we pick 11000.

A possible table of syndromes:
(*answers may vary depending on coset leaders chosen*)

$\underline{a}$	$S(\underline{a})$
00000	00
10000	01
01000	11
11000	10

$S(11100) = 11100 \cdot H^T = 11$ which is the syndrome of 01000 in our table. Hence we decode 11100 to $11100 - 01000 = 10100$ (*answers may vary depending on the table of syndromes*).

Feedback: Done well, showing that the students learned to apply this algoithm well.

(c) $\boxed{\text{ILO2 ILO4}}$ Use the table of syndromes constructed in (b) to show that if the code D is transmitted via $BSC(p)$, then the probability $P_{\text{corr}}(D)$ that the received vector is decoded correctly is $(1-p)^3$.

Answer. If the code $D \subseteq \mathbb{F}_2^n$ is transmitted via the binary symmetric channel with bit error rate p, then by a fact from the course, $P_{\text{corr}}(D) = \sum_{i=0}^{n} \alpha_i (1-p)^{n-i} p^i$ where α_i is the number of cosets with coset leader of weight i.
From the table, $\alpha_0 = 1$, $\alpha_1 = 2$, $\alpha_2 = 1$, $\alpha_3 = \alpha_4 = \alpha_5 = 0$ so that

$$\begin{aligned} P_{\text{corr}}(D) &= (1-p)^5 + 2(1-p)^4 p + (1-p)^3 p^2 \\ &= (1-p)^3((1-p)^2 + 2(1-p)p + p^2) = (1-p)^3(1-p+p)^2 = (1-p)^3. \end{aligned}$$

Feedback: Most students used the correct formula for P_{corr} but unfortunately some had problems with the binomial expansion — it is a basic skill which needs to stay with you throughout your mathematical career (and life!)

(d) $\boxed{\text{ILO4}}$ Write down the probability that an unencoded three-bit message, sent via $BSC(p)$, is received without errors. Compare this probability with $P_{\text{corr}}(D)$ from part (c) and determine whether encoding three-bit messages using the code D improves *error correction* if transmitting via $BSC(p)$. Suggest one possible advantage and one disadvantage of using the code D versus transmitting unencoded messages of length 3.

Answer. The probability that no errors occur in a three-bit message sent via $BSC(p)$ is $(1-p)^3$. This is the same as $P_{\text{corr}}(D)$, hence the code D does not offer any improvement in error correction.

A possible advantage of using D is improved error detection compared to the trivial code (*one can observe that $w(D) = 2$ so, unlike the trivial code, D detects a single bit error*).

However, as with any non-trivial code, a disadvantage of D is that its rate is less than 1 hence using D will increase the amount of data transferred (*other valid reasons include: a decoder for D requires storage/computation*).

Feedback: Some students used wrong probability of receiving three bits without errors via $BSC(p)$. Furthermore, marks were deducted for stating that D improves error correction (as an advantage of D) — the P_{corr} clearly shows that it doesn't — and for suggesting that D somehow protects information from eavesdroppers. Remember, coding theory is not cryptography!

[20 marks]

A3.

(a) $\boxed{\text{ILO1}}$ Consider the Reed-Muller code $R(r,m)$ where $0 \leqslant r < m$. Write down the parameters $[n,k,d]_q$ of this code. State without proof which Reed-Muller code coincides with $R(r,m)^\perp$. Hence write down the condition on r and m equivalent to $R(r,m)$ being self-dual. Explain briefly why $R(r,m)$ is self-orthogonal, if and only if $r < m/2$.

Answer. $R(r,m)$ is a $[2^m, k, 2^{m-r}]_2$-code, with $k = \dim R(r,m) = \binom{m}{0} + \binom{m}{1} + \cdots + \binom{m}{r}$.

$R(r,m)^\perp = R(m-1-r, m)$.

Hence $R(r,m)$ is self-dual, if and only if $r = m-1-r$, equivalently $r = (m-1)/2$.

"Self-orthogonal" means that $R(r,m) \subseteq R(r,m)^\perp = R(m-1-r,m)$. Since $R(r,m)$ is a subspace of $R(r',m)$ whenever $r \leqslant r'$, self-orthogonality is equivalent to $r \leqslant m-1-r$, or, the same, $2r \leqslant m-1$, equivalently $2r < m$ and $r < m/2$.

Feedback: The parameters of $R(r,m)$ were correctly recalled by many students, but the questions about orthogonality were not well-attempted. Reed-Muller codes were the last part of the course which was not reinforced by coursework.

(b) $\boxed{\text{ILO6 combinatorics}}$ Use the result of (a) to prove that all Reed-Muller codes of dimension 2018 are self-orthogonal codes.

Answer. Assume for contradiction that $R(r,m)$ of dimension 2018 is not self-orthogonal. Then by part (a), $r \geqslant m/2$ hence the dimension, $2018 = \binom{m}{0} + \binom{m}{1} + \cdots + \binom{m}{r}$, is at least one half of $2^m = \sum_{i=0}^{m} \binom{m}{i}$. It follows that $2^m = 2^{11} = 2048$.

But then $\binom{11}{r+1} + \cdots + \binom{11}{11}$ must be $2^{11} - 2018 = 30$ which is impossible as $\binom{11}{11} = 1$, $\binom{11}{10} + \binom{11}{11} = 12$ and $\binom{11}{9} + \binom{11}{10} + \binom{11}{11} > 30$. The contradiction finishes the proof.

Remark: in fact, the only Reed-Muller code of dimension 2018 is $R(1, 2017)$, but the proof of that is much longer and was not required in this exam paper.

Feedback: A challenging question where only a few made progress.

[10 marks]

SECTION B

Answer **TWO** of the three questions in this section (40 marks in total).
If more than TWO questions from this section are attempted, then credit will be given for the best TWO answers.

B4. (a) $\boxed{\text{ILO1 ILO5}}$ State without proof the *Hamming bound* for codes in $\mathbb{F}_q^n$ of minimum distance d. Define what is meant by a *perfect code*. Name a perfect code of minimum distance 9.

Answer. Hamming bound: $M \times \#S_t(\underline{0}) \leqslant q^n$ where $t = \left[\frac{d-1}{2}\right]$ (*or any of the equivalent statements*).
A code for which the Hamming bound is attained (*holds with an equality*) is a perfect code.
The binary repetition code $\mathrm{Rep}(9,2) = \{000000000, 111111111\}$ is a perfect code with $d = 9$.

Feedback: Done very well.

(b) $\boxed{\text{ILO1}}$ Prove that if the minimum distance of a code is even, then the code is not perfect. You can use any other facts from the course without proof, but you should state the facts you use.

Answer. Assume for contradiction that a perfect code $C \subseteq \mathbb{F}_q^n$ has *even* minimum distance d. Take any codeword $\underline{w}$ of C and change the first $d/2$ symbols in $\underline{w}$ to obtain a word $\underline{z} \in \mathbb{F}_q^n$ with $d(\underline{z}, \underline{w}) = d/2$.

Since C is perfect, by a fact from the course the Hamming spheres $S_t(\underline{c})$ of radius $t = [(d-1)/2]$ centred at codewords of C cover $\mathbb{F}_q^n$, so there must be a codeword $\underline{v}$ such that $d(\underline{v}, \underline{z}) \leqslant t$.
Then $\underline{v} \neq \underline{w}$ as $d(\underline{w}, \underline{z}) > t$, and by the triangle inequality

$$d(\underline{v}, \underline{w}) \leqslant d(\underline{v}, \underline{z}) + d(\underline{z}, \underline{w}) \leqslant t + \frac{d}{2} < \frac{d}{2} + \frac{d}{2} = d,$$

contradicting d being the minimum distance of C.

Feedback: Done by a considerable number of students, although in many of the attempts it was not pointed out that $\underline{v} \neq \underline{w}$ — e.g., because, by construction, $d(\underline{v}, \underline{z}) < d(\underline{w}, \underline{z})$. Two codewords $\underline{v}, \underline{w}$ with $d(\underline{v}, \underline{w}) < d$ lead to a contradiction only if it is shown that $\underline{v} \neq \underline{w}$, hence marks were deducted for this omission.

(c) $\boxed{\text{ILO1 ILO5}}$ Let $C \subseteq \mathbb{F}_q^n$ be a linear code. Define what is meant by a *generator matrix* of C. Assuming that C has a generator matrix G such that all rows of G have even weight: (i) Show that if $q = 2$, then C is not a perfect code. (ii) If $q = 3$, can such a code C be perfect? Justify your answer.

Answer. A generator matrix G of a linear code C is a matrix whose rows form a basis of the vector space C.

(i) If $q = 2$, vectors of even weight in $\mathbb{F}_2^n$ form a vector space — the even weight code E_n as in question A1, hence C, spanned by the rows of G, is a subspace of E_n. This means that all codevectors of C have even weight; this applies to vectors of weight $w(C)$, so $w(C)$ is even. For linear codes, $d(C) = w(C)$ which must then be even, so by part (b) the code is not perfect.

(ii) Yes, C can be perfect if $q = 3$: consider the matrix $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$. Its rows have even weight 2 and are linearly independent over $\mathbb{F}_3$ so it is an invertible square matrix hence generator matrix for the trivial code $\mathbb{F}_3^2$ which is perfect.

Feedback: In (i), marks were deducted for stating that the weight of all rows of G is even implies that $w(C)$ is even without an explanation (for example, using E_n). And indeed, some students went on to say that the same argument works when $q = 3$ in (ii). This probably stemmed from assuming that $w(C)$ equals the minimum row weight in G for all generator matrices G — a grave mistake! (ii) was not done as well as (i), although a variety of examples were given.

(d) ILO6 linear algebra A ternary linear code C has a generator matrix G with the following property: if the last row of G is removed, the remaining rows form a generator matrix of a ternary Golay code. Find all the possible values of the parameters $[n, k, d]_3$ of such codes C and justify your answer. Any results from the course can be used without particular comment.

Answer. A ternary Golay code is an $[11, 6, 5]_3$-code, hence its generator matrix has 6 rows and 11 columns. Therefore, G must be a 7×11 matrix so $n = 11$ and $k = 7$.

It remains to find the possible values of $d = d(C) = w(C)$. We claim that $d \leqslant 2$. Indeed, let Γ be the span of the first six rows of G and let $\underline{r}_7 \in C$ be the last row of G. Being a Golay code, Γ is perfect, hence by the fact about Hamming spheres recalled in the solution to (b), any vector in $\mathbb{F}_3^{11}$ is at distance of at most $[(5-1)/2] = 2$ from a codevector of Γ. Therefore, $d(\underline{r}_7, \underline{g}) \leqslant 2$ for some $\underline{g} \in \Gamma$. Note that $\underline{r}_7 \notin \Gamma$ because G is a generator matrix, hence has linearly independent rows, so $\underline{r}_7 \neq \underline{g}$. Also, $\underline{g} \in \Gamma \subset C$ and $\underline{r}_7 \in C$, which shows that the minimum distance of C is at most 2 as claimed.

We now show that both $d = 1$ and $d = 2$ are possible. Let $\underline{r}_7 \in \mathbb{F}_3^{11}$ be any vector of weight 1 or 2. Append $\underline{r}_7$ as a row to any generator matrix of a ternary linear Golay code Γ. Since $\underline{r}_7 \notin \Gamma$ (codevectors of Γ have weight at least 5), the resulting matrix G has linearly independent rows, hence is a generator matrix for a linear code C.

If $w(\underline{r}_7) = 1$, $\underline{r}_7 \in C$ automatically implies that $w(C) = 1$.

Finally, if $w(\underline{r}_7) = 2$, then automatically $w(C) \leqslant 2$; moreover, if $\underline{v} \in C$, then $\underline{v} = \lambda \underline{r}_7 + \underline{g}$ for some $\lambda \in \mathbb{F}_3$ and $\underline{g} \in \Gamma$. If $\underline{g} \neq \underline{0}$, then $w(\underline{v}) \geq w(\underline{g}) - w(\lambda \underline{r}_7) \geq 5 - 2 = 3$, and if $\underline{g} = \underline{0}$ and $\lambda \neq 0$, then $w(\underline{v}) = 2$. Thus, $w(C) = 2$.

Answer: $[11, 7, 1]_3$ or $[11, 7, 2]_3$.

Feedback: A difficult question only done by few. Having written down the answer, students did not try to show that there indeed exists a code with each of the given combinations of parameters — but this is necessary in order to justify the answer!

[20 marks]

B5. In this question, Z is the ternary linear code with generator matrix $\begin{bmatrix} 1 & 0 & 1 & 2 \\ 1 & 2 & 0 & 1 \end{bmatrix}$.

(a) $\boxed{\text{ILO1}}$ What is meant by saying that a generator matrix G of a linear code $C \subseteq \mathbb{F}_q^n$ is in *standard form*? Given G in standard form, explain how one can find a generator matrix for the dual code $C^\perp$. Define what is meant by the *weight enumerator* $W_C(x, y)$.

Answer. A $k \times n$ generator matrix of G is in standard form if its first k columns form an identity matrix: $G = \big[I_k|A\big]$ for some $k \times (n-k)$ matrix A.

In this case, a check matrix H (a generator matrix for $C^\perp$) can be $\big[-A^T|I_{n-k}\big]$.

The weight enumerator $W_C(x, y)$ is defined as $\sum_{i=0}^n A_i x^{n-i} y^i$ where A_i is the number of codevectors of C of weight i, *or, equivalently,* $W_C(x, y) = \sum_{\underline{c} \in C} x^{n-w(\underline{c})} y^{w(\underline{c})}$.

Feedback: Done well. Marks were deducted for mistakes such as writing A^T instead of $-A^T$.

(b) $\boxed{\text{ILO3 ILO2}}$ List the codevectors of Z and find the weight enumerator $W_Z(x, y)$ of Z.

Answer. The codevectors of Z are obtained by encoding all possible two-symbol vectors: $[00]G = 0000$, $[01]G = 1201$, $[02]G = 2102$, $[10]G = 1012$, $[20]G = 2021$, $[11]G = 2210$, $[22]G = 1120$, $[12]G = 0111$, $[21]G = 0222$.

By inspection, $W_Z(x, y) = x^4 + 8xy^3$.

Feedback: Done well. Some students forgot that 2 is an element of the field $\mathbb{F}_3$ and encoded only vectors made up of 0 and 1. This gave only four codevectors and led to an incorrect $W_z(x, y)$.

(c) $\boxed{\text{ILO1 ILO2}}$ State the *MacWilliams identity* for q-ary linear codes. Using the MacWilliams identity, or otherwise, for each $i = 0, 1, 2, 3, 4$ calculate the number of codevectors of weight i in the dual code $Z^\perp$.

Answer. The MacWilliams identity: if $C \subseteq \mathbb{F}_q^n$ is a linear code, then

$$W_{C^\perp}(x, y) = \frac{1}{\#C} W_C(x + (q-1)y,\, x - y).$$

By the MacWilliams identity with $q = 3$, $9W_{Z^\perp}(x, y) = (x+2y)^4 + 8(x+2y)(x-y)^3$. Expand this: e.g.,

$$\begin{aligned}(x+2y)^4 + 8(x-y)^4 + 8 \times 3y(x-y)^3 = \;& (x^4 + 8x^3y + 24x^2y^2 + 32xy^3 + 16y^4)\\ &+ 8(x^4 - 4x^3y + 6x^2y^2 - 4xy^3 + y^4)\\ &+ 24(\qquad x^3y - 3x^2y^2 + 3xy^3 - y^4)\\ =\;& 9x^4 + 72xy^3 = 9(x^4 + 8xy^3).\end{aligned}$$

We can now see that $W_{Z^\perp}(x, y) = W_Z(x, y)$, meaning that $Z^\perp$ contains one codevector of weight 0, eight codevectors of weight 3, and no vectors of any other weight.

Alternatively, the expansion can be done differently:

$$W_Z(x, y) = x^4 + 8xy^3 = x(x^3 + (2y)^3) = x(x+2y)(x^2 - 2xy + 4y^2) = x(x+2y)((x-y)^2 + 3y^2)$$

so $W_Z(x+2y, x-y) = (x+2y)(3x)((3y)^2 + 3(x-y)^2) = 9W_Z(x,y)$. Hence $W_{Z^\perp}(x,y) = W_Z(x,y)$.

Alternatively, a check matrix H for Z can be found by bringing G to standard form and applying (a); H can then be seen to generate a code with the same weights (in fact, the same code) as Z.

Yet another alternative way is to notice that $GG^T = 0$ and $\dim Z$ is one half of the length of Z, so Z is self-dual; in fact, Z is a Ham(2,3) Hamming code which is self-dual.

Feedback: Most stated the MacWilliams identity correctly, although some stated the form for $q = 2$. A number of the attempts to calculate $W_{Z^\perp}$ by direct expansion, though, ended in failure due to arithmetic mistakes or not using the binomial theorem correctly (!). A mistake revealing a serious misunderstanding was seen in a couple of scripts: the coefficients of $W_Z(x,y)$ were reduced mod 3. This does not make sense because the coefficients are integers showing the number of codevectors of certain weight, not elements of $\mathbb{F}_3$.

Those who chose to work out the generator matrix of $Z^\perp$ were fewer in number but more successful computationally. A small number of students did notice that $Z^\perp$ was self-dual straight away.

(d) $\boxed{\text{ILO6 calculus}}$ Let $D \subseteq \mathbb{F}_q^{2q}$ be a linear code which consists of the zero vector and $(q-1)^3$ vectors of weight q. Show that $w(D^\perp) = 1$. Find all q for which a code D with these properties exists.

Answer. By the Average Weight Equation, the average weight of a codevector of D is $(n-z)(1-q^{-1})$ where $n = 2q$ is the length of the code and z is the number of zero columns in a generator matrix G of D. Since $q \geqslant 2$, this is at least $(2q-z)/2 = q - z/2$, but the weights of codevectors are 0 and q hence the average weight is less than q. Therefore, $z > 0$ and G has a zero column, so $w(D^\perp) = 1$ (*this follows by the Distance Theorem for linear codes*).

If $k = \dim D$, the number of non-zero codevectors is $q^k - 1$ which must be equal to $(q-1)^3$. Note that $q^3 - 1 > (q-1)^3$ (seen e.g. from the binomial formula) so there are two cases:
$k = 1$, then $q - 1 = (q-1)^3$ which gives $\boldsymbol{q = 2}$;
$k = 2$, when $q^2 - 1 = (q-1)^3$, divide by $q-1$ to get $q+1 = (q-1)^2$, $0 = q^2 - 3q$ and $\boldsymbol{q = 3}$.

A full solution must include an example of a code D for each value of q. The two examples are omitted here and are left as an exercise to the reader.

Feedback: Challenging: the only real progress was made in proving that $w(D^\perp) = 1$. Some achieved this without the use of AWE directly via MacWilliams, starting from $W_D(x,y) = x^{2q} + (q-1)^3 x^q y^q$, finding $W_{D^\perp}(x,y)$, differentiating with respect to y then putting $y = 0$ to find the coefficient of $x^{2q-1}y$ and prove that it is positive. This was difficult and was in fact a repetition of the proof of AWE.

[20 marks]

B6. (a) $\boxed{\text{ILO1}}$ What is a *cyclic code*? What are the properties required of a polynomial $g(x) \in \mathbb{F}_q[x]$ to be a generator polynomial of some cyclic code of length n over $\mathbb{F}_q$?

Answer. A cyclic code is a linear code $C \subseteq \mathbb{F}_q^n$ which is closed under the cyclic shift, i.e., for any codevector $(a_0, a_1, \dots, a_{n-1}) \in C$, the vector $(a_1, \dots, a_{n-1}, a_0)$ also lies in C.

A polynomial $g(x) \in \mathbb{F}_q[x]$ is a generator polynomial of some cyclic code of length n if $g(x)$ is monic and $g(x)$ divides the polynomial $x^n - 1$ in $\mathbb{F}_q[x]$.

Feedback: Done well. An absolute majority remembered that cyclic codes are linear!

(b) ILO5 ILO2 Factorise $x^3 - 1$ into irreducible polynomials in $\mathbb{F}_3[x]$. Hence list all the cyclic codes in $\mathbb{F}_3^3$, stating the cardinality, the minimum distance, a generator polynomial and a generator matrix for each code.

Answer. One has $x^3 - 1 = (x-1)^3$ where $x - 1$ is irreducible.

The cyclic codes in $\mathbb{F}_3^3$ correspond to generator polynomials, i.e., to all the possible monic factors of $x^3 - 1$:

- $g(x) = 1$, $G = I_3$ (the identity matrix) generates the trivial code $\mathbb{F}_3^3$ of cardinality 27, $d = 1$;
- $g(x) = x - 1$, $G = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$ generates the "zero sum code" of cardinality 9, $d = 2$;
- $g(x) = (x-1)^2 = x^2 + x + 1$, $G = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$ generates the ternary repetition code of length 3, of cardinality 3 and minimum distance 3;
- $g(x) = x^3 - 1$ generates the code $\{000\}$, with empty generator matrix (0 rows), cardinality 1 and undefined d.

Feedback: A mistake often made was to factorise $(x-1)^3 = (x-1)(x^2+x+1)$ and stop. Over $\mathbb{F}_3$, $x^2 + x + 1$ is not irreducible and equals $(x-1)^2$. Also, recall the *Freshman's Dream*, a lemma which says that $(a+b)^p = a^p + b^p$ in $\mathbb{F}_p$, and apply this to $(x-1)^3$. This mistake, however, does not affect the rest of the solution.

Some students forgot the polynomials $g(x) = 1$ and $g(x) = x^3 - 1$ which *are* generator polynomials and must be included in the classification.

(c) ILO1 Let C be a linear code. Prove that C is cyclic, if and only if $C^\perp$ is cyclic.

Answer. We will prove that C is cyclic, only if $C^\perp$ is cyclic. The "if" part will also follow because $(C^\perp)^\perp = C$.

Let C be a cyclic code in $\mathbb{F}_q^n$. As $C^\perp$ is always a linear code, we only need to prove that $C^\perp$ is closed under the cyclic shift.

Write $s(\underline{b})$ for the cyclic shift of a vector $\underline{b}$ and observe that the inner product does not cnange under the cyclic shift, that is, $s(\underline{b}) \cdot s(\underline{c}) = \underline{b} \cdot \underline{c}$. Now, if $\underline{b} \in C^\perp$, then $\underline{b} \cdot C = \{0\}$, then $s(\underline{b}) \cdot s(C) = \{0\}$, but $s(C) = C$ because C is a cyclic code, hence $s(\underline{b}) \cdot C = \{0\}$ meaning that $s(\underline{b}) \in C^\perp$. Thus, $C^\perp$ is closed under the cyclic shift.

Feedback: Done reasonably well, although some students tried to take a completely different route and claimed that $C^\perp$ is a cyclic code with generator polynomial $h(x)$ – WRONG – where $h(x)$ is the check polynomial of C. In fact, $C^\perp$ has generator polynomial given by $h(x)$ written in reverse order... but one still needs to prove that $C^\perp$ is cyclic!

(d) $\boxed{\text{ILO6 number theory}}$ For a prime p, let $I_p = \{(a_1, \dots, a_{p-1}) \in \mathbb{F}_p^{p-1} \mid \sum_{i=1}^{p-1} i a_i = 0 \text{ in } \mathbb{F}_p\}$.

i. What are the odd primes p for which I_p is an MDS code?

ii. What are the odd primes p for which I_p is a cyclic code?

iii. What are the odd primes p for which $I_p^\perp \subseteq I_p$?

Justify your answer in each case. Any facts from the course can be freely used.

Answer. i. The length of I_p is $n = p - 1$ by definition. To determine if I_p is MDS, we need to know the dimension k and the minimum distance d of I_p. Note that I_p is a linear code whose dual code $I_p^\perp$ is the one-dimensional code in $\mathbb{F}_p^{p-1}$ generated by the matrix

$$H = \begin{bmatrix} 1 & 2 & \dots & p-1 \end{bmatrix}.$$

It follows that $k = n - \dim I_p^\perp = n - 1$.

Similarly to the ISBN-10 code (which is seen to be I_{11}), observe that I_p contains the vector $100\dots001$ of weight 2 but no vectors of weight 1 (e.g., because H has no zero entries), so $d = w(I_p) = 2$.

We have $k = n - d + 1$ as $n - 1 = n - 2 + 1$ so I_p is an MDS code *for all odd primes* p.

ii. If I_p is cyclic, then together with $100\dots001$ the code must contain its cyclic shift $\underline{v} = 1100\dots00$. Yet the checksum of $\underline{v}$ is $1 + 2 = 3$ which must be zero in $\mathbb{F}_p$. Hence $p = 3$. We proved that if I_p is cyclic, then $p = 3$. Now, if $p = 3$, then $I_p = I_3 = \{00, 11, 22\}$ is clearly cyclic. *Answer*: I_p is cyclic iff $p = 3$.

iii. The condition says that $I_p^\perp$ is self-orthogonal. This is equivalent to $HH^T = 0$, equivalently $1^2 + 2^2 + \dots + (p-1)^2$ is 0 mod p.

Recall that $1^2 + 2^2 + \dots + (p-1)^2 = \dfrac{(p-1)p(2p-1)}{6}$. If $p > 3$, this integer is divisible by p. If $p = 3$, one has $1^2 + 2^2 = 2 \neq 0$ in $\mathbb{F}_3$. *Answer*: $I_p^\perp \subseteq I_p$ iff $p > 3$.

Feedback: Challenging: progress made by few. iii was the most difficult part, where partial marks were given for stating the condition $1^2 + 2^2 + \dots + (p-1)^2 = 0$ in $\mathbb{F}_p$. Writing the left-hand side as $(p-1)p(2p-1)/6$ turned out to be a step very few thought of.

[20 marks]

END OF EXAMINATION PAPER