Electronic calculators are permitted, provided they cannot store text.
A1.

(a) Let $C$ be a linear code in $\mathbb{F}_q^n$. Explain what is meant by:

1. the weight $w(\mathbf{x})$ of a vector $\mathbf{x}$;
2. the weight $w(C)$ of $C$;
3. a generator matrix of $C$;
4. the inner product $\mathbf{x} \cdot \mathbf{y}$ of vectors $\mathbf{x}$, $\mathbf{y}$;
5. the dual code $C^\perp$.

**Answer.**

1. $w(\mathbf{x})$ is equal to the number of non-zero coordinates in $\mathbf{x}$;
2. $w(C) = \min \{ w(\mathbf{x}) : \mathbf{x} \in C, \ \mathbf{x} \neq \mathbf{0} \}$;
3. a matrix whose rows form a basis of $C$;
4. if $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n)$, then $\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n x_i y_i$;
5. $C^\perp = \{ \mathbf{x} \in \mathbb{F}_q^n | \mathbf{x} \cdot \mathbf{c} = 0 \ \forall \mathbf{c} \in C \}$.

**Feedback:** Mostly done well. Some students forgot to check that $\mathbf{0}$ is in $C$. One has to check this, otherwise one can have an empty set which is trivially closed under taking a linear combination, but is not a vector space and not a code.

(b) If $H$ is a matrix with $n$ columns over $\mathbb{F}_q$, prove that $C = \{ \mathbf{x} \in \mathbb{F}_q^n | \mathbf{x} H^T = \mathbf{0} \}$ is a linear code.

**Answer.**

$0 \in C$ because $0 H^T = 0$;
if $\mathbf{u}, \mathbf{v} \in C$, $\lambda \in \mathbb{F}_q$, then $\mathbf{(u + \lambda v)} H^T = \mathbf{u} H^T + \lambda (\mathbf{v} H^T) = \mathbf{0} + \lambda \mathbf{0} = \mathbf{0}$, so $\mathbf{u} + \lambda \mathbf{v}$ is in $C$.

**Feedback:** Mostly done well. Some students forgot to check that $\mathbf{0}$ is in $C$. One has to check this, otherwise one can have an empty set which is trivially closed under taking a linear combination, but is not a vector space and not a code.

(c) Let $E_5$ denote the binary even weight code of length 5. Write down a generator matrix of $E_5$.

**Answer.**

$$
\begin{bmatrix}
1 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 1 & 1 \\
\end{bmatrix}
$$

For example, Other answers possible.

**Feedback:** Mostly done well. Mark given if the matrix was indeed a generator matrix of $E_5$.

(d) Write down a generator matrix of $(E_5)^\perp$. Identify the code $(E_5)^\perp$ by its well-known name.

**Answer.**

$$
\begin{bmatrix}
1 & 1 & 1 & 1 & 1 \\
\end{bmatrix}
$$

binary repetition code of length 5.

**Feedback:** Done well. One mark was for the correct matrix — the answer is unique here. The other mark was for the correct name of the code.
A2.

(a) State the Distance Theorem for linear codes.

Answer.
Let $C$ be a linear code with parity check matrix $H$. Then $d(C) = d$ if all sets of $d - 1$ columns of $H$ are linearly independent, but some set of $d$ columns of $H$ is linearly dependent.

Feedback: Mistakes were made by a number of students, e.g., putting $d$ and $d + 1$ instead of $d - 1$ and $d$.

Let $C$ be the binary linear code with parity check matrix $H = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$.

(b) State the length and the dimension of $C$.

Answer.
The length is $n = 6$ (the number of columns in $H$) and the dimension is $k = 4$ (the length minus the number of rows in $H$).

Feedback: A mark was given if both numbers were correct.

(c) Using the Distance Theorem, find the minimum distance $d(C)$ of $C$.

Answer.
Columns of $H$ are non-zero so $d(C) > 1$; there are two columns of $H$ that sum to 0 so $d(C) = 2$.

Feedback: A common mistake was not to point out that there are no zero columns. This is necessary: if there were a zero column in $H$, then $d(C)$ would be 1, even though there are two columns that sum to 0.

(d) How many errors can the code detect?

Answer.
$C$ can detect $d(C) - 1 = 1$ error.

Feedback: A mark was given if it was clear that the formula $d(C) - 1$ was used.

(e) How many errors can the code correct?

Answer.
$C$ can correct $\left\lfloor \frac{d(C) - 1}{2} \right\rfloor = 0$ errors.

Feedback: A mark was given if it was clear that the correct formula was used. Note that the formula $\frac{d(C) - 1}{2}$ is NOT correct.

(f) Find three vectors of weight 1 with pairwise distinct syndromes.

Answer.
For example, $x = 100000$, $S(x) = 11$; $y = 010000$, $S(y) = 10$; $z = 001000$, $S(z) = 01$.

Feedback: The answer is not unique.
(g) Using (f), write down a set of coset leaders. Calculate the table of syndromes, and decode the vectors 000101 and 111000 using syndrome decoding.

**Answer.**
The vectors shown in (f) are three coset leaders of weight 1. The vector 000000 is a coset leader of weight 0 with syndrome 00. These are all cosets, as the number of cosets is $2^{n-k} = 4$.

Now $S(000101) = 10 = S(010000)$ (a coset leader found earlier), so 000101 is decoded as 000101−010000 = 010101. Also, $S(111000) = 00$ so the vector 111000 is a codeword and is decoded as 111000.

**Feedback:** This part was mostly done well.

(h) If the code $C$ is transmitted down a binary symmetric channel with bit error rate $r$, write down a formula for $P_{\text{corr}}(C)$, the probability that a received vector is decoded correctly.

**Answer.**

$P_{\text{corr}} = (1-r)^6 + 3r(1-r)^5$.

**Feedback:** For some reason, this question was not done very well — students apparently confused formulas for $P_{\text{corr}}$ and $P_{\text{undetect}}$.

[15 marks]

A3.

(a) Define projective $(n-1)$-space, $\mathbb{P}_{n-1}(\mathbb{F}_q)$.

**Answer.**

A line in $\mathbb{F}_q^n$ is a one-dimensional vector subspace of $\mathbb{F}_q^n$ (or an equivalent definition); $\mathbb{P}_{n-1}(\mathbb{F}_q)$ is the set of all lines in $\mathbb{F}_q^n$.

**Feedback:** Various mistakes were made in this definition: “set of all lines in $\mathbb{F}_q$” was popular.

(b) Define a Hamming code $\text{Ham}(s, q)$.

**Answer.**

$\text{Ham}(s, q)$ is a linear $q$-ary code for which the check matrix has as columns representative vectors of each element of projective space $\mathbb{P}_{s-1}(\mathbb{F}_q)$.

**Feedback:** Again, various mistakes were made. For example, it was claimed that the check matrix of $\text{Ham}(s, q)$ has columns which are representatives of $\mathbb{P}_{n-1}(\mathbb{F}_q)$, without saying what $n$ is. Or it was said that the check matrix has as columns elements of $\mathbb{P}_{s-1}(\mathbb{F}_q)$ — this is wrong because elements of $\mathbb{P}_{s-1}(\mathbb{F}_q)$ are sets not vectors.

(c) Write down, without proof, formulas for the length $n$, dimension $k$ and minimum distance $d$ of $\text{Ham}(s, q)$.

**Answer.**

$$n = \frac{q^s - 1}{q - 1}, \quad k = \frac{q^s - 1}{q - 1} - s, \quad d = 3.$$ 

**Feedback:** No problem with this question — done almost universally.
(d) Let $C$ be a Ham$(2, 3)$ code. In the following, you may quote any result from the course without proof:

1. Write down a generator matrix for $C$.
   
   **Answer.**
   
   For example, $G = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 2 & 1 \end{bmatrix}$.
   
   **Feedback:** Done mostly correctly except in the cases where students got the dimensions of the generating matrix wrong.

2. Show that $C$ is a self-dual code, that is, $C^\perp = C$.
   
   **Answer.**
   
   By a result from the course, it is enough to observe that $\dim C = \frac{1}{2}n$ and that $GG^T$ is a zero matrix. Indeed, $\dim C = 2 = \frac{1}{2}4$, and a calculation gives $GG^T = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.
   
   **Feedback:** A common mistake was to check that $GG^T = 0$ but not to state that $\dim C = \frac{1}{2}$ of length of $C$. The fact that $GG^T = 0$ only implies that $C \subseteq C^\perp$, not necessarily $C = C^\perp$.

3. Show that Ham($s$, $q$) is not a self-dual code if $s > 2$.
   
   **Answer.**
   
   Assume that $s > 2$ and Ham($s$, $q$) is self-dual. Then $k = n - k$, so $\frac{q^s - 1}{q - 1} - s = s$. Rewrite this as $q^{s-1} + \ldots + q + 1 = 2s$. Out of the $2s$ summands on the left-hand side, the first, $q^{s-1}$, is at least $2^{s-1} = 4$; the other powers of $q$ are at least 2, and the last term is 1. Hence the left-hand side is at least $4 + 2(s - 2) + 1 = 2s + 1$, and cannot equal $2s$. This contradiction shows that the assumption about self-duality was wrong.
   
   **Feedback:** This turned out to be one of the hardest questions of the exam, done by very few people; many arrived at the equation $\frac{q^{s-1}}{q - 1} - s = s$ but did not solve it.

[15 marks]
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) Define the Hamming sphere $S_t(u)$ with centre $u$ and radius $t$ in the vector space $\mathbb{F}_q^n$.

**Answer.**

$S_t(u) = \{ y \in \mathbb{F}_q^n : d(u, y) \leq t \}$.

**Feedback:** Done very well.

(b) Write down a formula for $|S_t(u)|$, the number of elements in $S_t(u)$.

**Answer.**

$|S_t(u)| = \sum_{i=0}^{t} \binom{n}{i} (q-1)^i$.

**Feedback:** Done well. In an answer, it was wrong to state or assume that $t = \left[ \frac{d-1}{2} \right]$, the formula should have been given for general $t$.

(c) Let $C$ be a code in $\mathbb{F}_q^n$. Show that if $t < \frac{d(C)}{2}$, $u, v \in C$, $u \neq v$, then $S_t(u) \cap S_t(v) = \emptyset$.

**Answer.**

Assume for contradiction that there exits an element $x \in S_t(u) \cap S_t(v)$. Then by the triangle inequality $d(u, v) \leq d(u, x) + d(x, v) \leq t + t < d$. This contradicts $d$ being the minimum distance of $C$.

**Feedback:** Done well. Some people argued that $t < \frac{d(C)}{2}$ necessarily implies $t = \left[ \frac{d-1}{2} \right]$, but this was not necessary and not true.

(d) State and prove the Hamming bound for the number $M$ of elements of a code in $\mathbb{F}_q^n$ of minimum distance $d$.

**Answer.**

$M|S_t(0)| \leq q^n$ where $t = \left[ \frac{d-1}{2} \right]$ (or any of the equivalent statements).

Proof: observe that $t < \frac{d}{2}$, so by (c), the $M$ spheres $S_t(u), u \in C$, are disjoint. Then by (b), the union of these spheres contains a total of $M|S_t(0)|$ elements; it is a subset of the set $\mathbb{F}_q^n$ with $q^n$ elements, hence the bound.

**Feedback:** A bookwork question, done mostly ok.

(e) Show that a $k$-dimensional linear code in $\mathbb{F}_2^n$ of minimum distance 3 satisfies $k \leq n - \log_2(n+1)$.

**Answer.**

$t = 1$, so the Hamming bound reads $2^k(1 + n) \leq 2^n$. Taking $\log_2$ of both sides gives $k + \log_2(1 + n) \leq n$ as required.

**Feedback:** Not an insignificant number of students tried to deduce this inequality from the Singleton bound. This is not possible. Moreover, for binary codes the Singleton bound is generally weaker than the Hamming bound, and the required inequality is literally the Hamming bound for $d = 3$. 

P.T.O.
(f) Define what is meant by a perfect code.

**Answer.**
A perfect code is a code which attains the Hamming bound.

**Feedback:** Answered very well.

(g) State without proof for which pairs \((q, d)\) where \(q\) is a prime there exists a perfect \(q\)-ary code of minimum distance \(d\). Name a perfect code corresponding to each pair.

**Answer.**
Pairs \((q, 1)\), resp. \((q, 3)\), where \(q\) is any prime, are realised by trivial codes, resp. Hamming codes. Pairs \((2, d)\) where \(d\) is odd are realised by binary repetition codes of odd length. \([Also, (2, 7) is the binary Golay code — not necessary for a complete answer.]\) The pair \((3, 5)\) is realised by the ternary Golay code.

**Feedback:** Not an easy question — many people forgot that trivial codes or binary repetition codes of odd length are perfect, or forgot about the ternary Golay code. A small minority got full marks here by recalling the classification theorem for perfect codes.

[20 marks]

B5. Let \(p\) be a prime.

(a) Let a cyclic code \(C \subseteq \mathbb{F}_p^n\) be considered as an ideal of \(R_n = \mathbb{F}_p[x]/(x^n - 1)\) in the usual way. Show that there is a unique monic polynomial \(g\) of minimum degree such that \(C = gR_n\).

**Answer.**
Among non-zero polynomials with images in \(C\), take one with minimum degree. Multiply it by a scalar (if necessary) to make it monic, and call it \(g\). If \(f\) is another non-zero polynomial such that \(\overline{f} \in C\), then \(\deg f \geq \deg g\). Dividing \(f\) by \(g\), we obtain \(f = qg + r\) where \(\deg r < \deg g\). Note that \(\overline{r} = \overline{f} - \overline{qg}\) where the right-hand side is in the ideal \(C\). So \(\overline{r} \in C\) and \(\deg r < \deg g\).

By the choice of \(g\) it follows that \(r = 0\), so \(f = qg\), \(\overline{f} = \overline{qg}\). This shows that \(C \subseteq \overline{gR}_n\); \(\supseteq\) is also true because \(C\) is an ideal. The polynomial \(g\) is unique: another \(h\) must be of the same degree as \(g\) and be divisible by \(g\) by the above, so be equal to \(g\) as they are both monic.

**Feedback:** A straight bookwork question, but many answers contained mistakes or omissions which led to loss of marks.

(b) For \(C\) and \(g\) as above, write down a generator matrix for \(C\) in terms of the coefficients of \(g\).

**Answer.**
If \(g(x) = g_0 + g_1 x + \ldots + g_r x^r\), \(G = \begin{bmatrix} g_0 & g_1 & \cdots & g_r \\ 0 & g_0 & \cdots & g_r \\ \vdots & \ddots & \ddots & \vdots \\ g_0 & \cdots & \cdots & g_r \end{bmatrix}\), with \(n\) columns.

**Feedback:** One of the common mistakes was to give a matrix with the number of columns equal to the power of \(g\), or a matrix which evidently had \(n\) rows.
You are given that, in $\mathbb{F}_3[x]$,
\[ x^8 - 1 = (x^5 - x^4 + x - 1)(x^3 + x^2 + x + 1). \]

(c) Write down a generator polynomial and a check polynomial for a ternary cyclic code $D$ of length 8 and dimension 5.

**Answer.**
Generator $g(x) = x^3 + x^2 + x + 1$. Check polynomial $h(x) = x^5 - x^4 + x - 1$.

**Feedback:** Generally done well.

In the rest of the question, $D$ refers to the code that you obtained in part (c).

(d) Let $a, b, c \in \mathbb{F}_3$ be such that the vector $ab000000c$ is a codeword of $D$. Show that $a = b = c = 0$.

**Answer.**
$ab000000c$ is a codeword iff its cyclic shift $cab00000$ is a codeword, iff the polynomial $(c + ax + bx^2)h(x)$ is divisible by $x^8 - 1$. However, a non-zero polynomial of this form has degree at most 7 hence cannot be divisible by $x^8 - 1$. Thus, the only solution is $a = b = c = 0$.

**Feedback:** Most students, instead of going down the above route, preferred to write out the check matrix of the code and to show that if the syndrome of $ab000000c$ is 0 then $a = b = c = 0$. Some mars were sometimes lost due to mistakes in the check matrix.

(e) Does $D$ contain codewords of weight 2? If so, write down a codeword of $D$ of weight 2.

**Answer.**
Yes, for example, 20001000 is a codeword: indeed, 11110000 is the codeword corresponding to the generator polynomial, 01111000 is its cyclic shift, and one has 20001000 = 01111000 - 11110000.

[In fact, every weight 2 codeword in the code $D$ generated by $g(x) = 1 + x + x^2 + x^3$ is a cyclic shift of this one, because vectors of the form $ab0000000$, $a0b000000$, $a00b0000$ are not codewords, for reasons similar to those given in the solution to (d)]

**Feedback:** Done in a minority of scripts. A common mistake was to write 11110000+01111000 and to claim that this gives 10001000 — unfortunately this is not true as the code is not binary!

B6.

**Feedback:** This question was less popular than the previous two.

(a) Let $C_1$, $C_2$ be linear codes in $\mathbb{F}^q^n$.

1. Define the code $|C_1|C_2|$. 

**Answer.**
$|C_1|C_2| = \{ [u | u + v] : u \in C_1, v \in C_2 \}$.

**Feedback:** Done well.
2. Prove that \( d(|C_1|C_2|) \geq \min\{2d(C_1), d(C_2)\} \).

**Answer.**

Let \([u \mid u + v]\) be a non-zero codeword in \(|C_1|C_2|\), where \(u \in C_1\) and \(v \in C_2\).

Its weight is \(w(u) + w(u + v)\). If \(v = 0\), this is \(2w(u) \geq 2d(C_1) \geq \min\{2d(C_1), d(C_2)\}\). If \(v \neq 0\), we use the (in)equalities \(w(a) = w(-a)\) and \(w(a + b) \leq w(a) + w(b)\) to argue that \(w(u) + w(u + v) = w(-u) + w(u + v) \geq w(-u + u + v) = w(v) \geq d(C_2) \geq \min\{2d(C_1), d(C_2)\}\).

The code \(|C_1|C_2|\) is a Reed-Muller code.

**Feedback:** Done well — by those who got up to here.

(b) Define the \(r\)th order Reed-Muller code \(R(r, m)\) in terms of Boolean functions.

**Answer.**

The Reed-Muller code \(R(r, m)\) is the space of value vectors of all Boolean functions in \(m\) variables of degree at most \(r\).

**Feedback:** Done well.

(c) Explain why \(R(r + 1, m + 1) = |R(r + 1, m)|R(r, m)|\).

**Answer.**

A codeword in \(R(r + 1, m + 1)\) corresponds to linear combination of square-free monomials in \(v_1, \ldots, v_{m+1}\), of degree at most \(r + 1\). Each monomial contains a copy of \(v_{m+1}\) or none. Hence \(f(v_1, \ldots, v_{m+1}) = g(v_1, \ldots, v_m) + v_{m+1}h(v_1, \ldots, v_m)\) where \(g \in R(r + 1, m)\) and \(h \in R(r, m)\).

If the truth table is arranged so that \(v_{m+1} = 0\) \(0 \ldots 0 \mid 1 \ldots 1\), then \(g\) corresponds to \([g \mid g]\) and \(v_{m+1}h\) corresponds to \([0 \mid h]\). Thus, \(f\) corresponds to \([g \mid g + h]\), and any \(g, h\) as above yield \(f = g + v_{m+1}h \in R(r + 1, m + 1)\).

**Feedback:** Done well.

(d) Give an example of a Reed-Muller code \(R(r, m)\) which has relative distance \(\delta = 0.5\).

**Answer.**

The minimum distance of \(R(r, m)\) is \(2^{m-r}\), the length is \(2^m\), hence the relative distance is \(\delta = \frac{2^{m-r}}{2^m} = 2^{-r}\). Any code \(R(1, m)\) is suitable.

**Feedback:** Done well — by those who got up to here.

(e) You are given that \(C\) is some linear code in \(F_2^n\) which has relative distance \(\delta > 0.5\) and contains the codeword \(11 \ldots 1\) (where all \(n\) bits are 1). Prove that \(\dim C = 1\).

**Answer.**

Assume that \(C\) contains a non-zero codeword \(v \neq 11 \ldots 1\). Either \(w(v) \leq \frac{n}{2}\), or the weight of the codeword \(11 \ldots 1 - v\), which is \(n - w(v)\), is less than \(\frac{n}{2}\). In both cases the minimum weight of \(C\) is \(w(C) \leq \frac{n}{2}\) which means that the relative distance of \(C\) is \(\delta(C) = \frac{d(C)}{n} = \frac{w(C)}{n} \leq \frac{1}{2}\). This contradiction shows that \(C = \{00 \ldots 0, 11 \ldots 1\}\), a repetition code of dimension 1.

**Feedback:** Done by only a few students. Importantly, in (e) one should not have assumed that \(C\) is a Reed-Muller code.