STEP Support Programme

2023 STEP 2 Worked Paper

General comments

These solutions have a lot more words in them than you would expect to see in an exam script and in places I have tried to explain some of my thought processes as I was attempting the questions. What you will not find in these solutions is my crossed out mistakes and wrong turns, but please be assured that they did happen!

You can find the examiners report and mark schemes for this paper from the OCR website. These are the general comments for the STEP 2023 exam from the Examiner's report:

Many candidates were able to express their reasoning clearly and presented good solutions to the questions that they attempted. There were excellent solutions seen for all of the questions.

An area where candidates struggled in several questions was in the direction of the logic that was required in a solution. Some candidates failed to appreciate that separate arguments may be needed for the "if" and "only if" parts of a question and, in some cases, candidates produced correct arguments, but for the wrong direction.

In several questions it was clear that candidates who used sketches or diagrams generally performed much better that those who did not. Sketches often also helped to make the solution clearer and easier to understand.

Several questions on the STEP papers ask candidates to show a given result. Candidates should be aware that there is a need to present sufficient detail in their solutions so that it is clear that the reasoning is well understood.

The three main points to note are:

1. The difference between "if" and "only if" is not always understood. Some resources that might help are:
   - Assignment 10 from the STEP Support Programme Foundation Modules.
   - This collection of NRICH problems.
2. Sketches are often useful! Definitely for mechanics, but also for many other questions as well. They can also be used to support an argument.
3. If a question asks you "show that", then you do need to fully support your argument. An examiner cannot tell the difference between a candidate who did a load of algebra in their head and one that just writes down the result they are aiming for without knowing how to show it.

Please send any corrections, comments or suggestions to step@maths.org.

Question 1

1 **(i)** Show that making the substitution $x = \dfrac{1}{t}$ in the integral

$$\int_a^b \frac{1}{(1+x^2)^{\frac{3}{2}}}\,\mathrm{d}x\,,$$

where $b > a > 0$, gives the integral

$$\int_{a^{-1}}^{b^{-1}} \frac{-t}{(1+t^2)^{\frac{3}{2}}}\,\mathrm{d}t\,.$$

(ii) Evaluate:

(a)
$$\int_{\frac{1}{2}}^{2} \frac{1}{(1+x^2)^{\frac{3}{2}}}\,\mathrm{d}x\,;$$

(b)
$$\int_{-2}^{2} \frac{1}{(1+x^2)^{\frac{3}{2}}}\,\mathrm{d}x\,.$$

(iii) **(a)** Show that

$$\int_{\frac{1}{2}}^{2} \frac{1}{(1+x^2)^2}\,\mathrm{d}x = \int_{\frac{1}{2}}^{2} \frac{x^2}{(1+x^2)^2}\,\mathrm{d}x = \tfrac{1}{2}\int_{\frac{1}{2}}^{2} \frac{1}{1+x^2}\,\mathrm{d}x\,,$$

and hence evaluate

$$\int_{\frac{1}{2}}^{2} \frac{1}{(1+x^2)^2}\,\mathrm{d}x\,.$$

(b) Evaluate

$$\int_{\frac{1}{2}}^{2} \frac{1-x}{x(1+x^2)^{\frac{1}{2}}}\,\mathrm{d}x\,.$$

Examiner's report

The first part of this question was often completed well, although candidates should note that in questions where the result is given it is important to show enough detail in the solution. Weaker candidates failed to change the limits or did not differentiate $\frac{1}{x}$ correctly when completing the substitution.

Most candidates realised that part **(ii)(a)** could be completed by applying the result from part **(i)** and were able to select the correct values for a and b. However, many did not realise that the result from part **(i)** was not directly applicable to part **(ii)(b)** and so did not gain any marks for that part, although some candidates did realise that the answer of zero could not be correct and received some credit for recognising that the function was even and so could identify the start of a correct solution. Solutions that applied the result from part **(i)** successfully often achieved full marks, although in some cases the way in which limits were dealt with was not sufficient. A significant number of candidates recognised that part **(ii)(b)** could be solved with a tan substitution and while this approach was successful, in some cases the final answer was not written in its simplest form.

In part **(iii)(a)** many candidates recognised that the same substitution would produce the required results, but as in part **(i)** several cases did not produce clear enough solutions to earn all of the marks. Most candidates were able to successfully calculate the value of the integral. Many candidates did not choose a suitable substitution for part **(iii)(b)**, but those who did generally managed to reach an appropriate form of the integral that could be compared to the original. Many then deduced the correct answer from this, but several did not recognise the significance of the new integral and then attempted other substitutions with little success.

Solution

(i) This is a "show that" question (as discussed earlier), so make sure that you show sufficient working to justify the result! Don't skip steps, even if they are "obvious", and don't run too many steps together.

We have $\dfrac{\mathrm{d}x}{\mathrm{d}t} = -t^{-2}$, and so we have:

$$\begin{aligned}\int_a^b \frac{1}{(1+x^2)^{\frac{3}{2}}}\,\mathrm{d}x &= \int_{a^{-1}}^{b^{-1}} \frac{1}{\left(1+\left(\frac{1}{t}\right)^2\right)^{\frac{3}{2}}} \times -t^{-2}\,\mathrm{d}t \\ &= \int_{a^{-1}}^{b^{-1}} \frac{-1}{t^2\left(1+\left(\frac{1}{t}\right)^2\right)^{\frac{3}{2}}}\,\mathrm{d}t \\ &= \int_{a^{-1}}^{b^{-1}} \frac{-t}{t^3\left(1+\left(\frac{1}{t}\right)^2\right)^{\frac{3}{2}}}\,\mathrm{d}t \\ &= \int_{a^{-1}}^{b^{-1}} \frac{-t}{(t^2+1)^{\frac{3}{2}}}\,\mathrm{d}t\end{aligned}$$

(ii) **(a)** Using the result from part **(i)** gives:

$$\int_{\frac{1}{2}}^{2} \frac{1}{(1+x^2)^{\frac{3}{2}}}\,\mathrm{d}x = \int_{2}^{\frac{1}{2}} \frac{-t}{(1+t^2)^{\frac{3}{2}}}\,\mathrm{d}t$$

Differentiating $(1+t^2)^{-\frac{1}{2}}$ gives:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(1+t^2)^{-\frac{1}{2}} &= -\tfrac{1}{2} \times 2t \times (1+t^2)^{-\frac{3}{2}} \\ &= -t \times (1+t^2)^{-\frac{3}{2}}\end{aligned}$$

Therefore we have:

$$\begin{aligned}\int_{2}^{\frac{1}{2}} \frac{-t}{(1+t^2)^{\frac{3}{2}}}\,\mathrm{d}t &= \left[\frac{1}{(1+t^2)^{\frac{1}{2}}}\right]_2^{\frac{1}{2}}\\ &= \frac{1}{\left(\frac{5}{4}\right)^{\frac{1}{2}}} - \frac{1}{(5)^{\frac{1}{2}}}\\ &= \frac{2}{\sqrt{5}} - \frac{1}{\sqrt{5}}\\ &= \frac{1}{\sqrt{5}}\end{aligned}$$

(b) Since the lower limit is negative, we cannot use the result from part **(i)**. However, since the integrand is an even function we have:

$$\int_{-2}^{2} \frac{1}{(1+x^2)^{\frac{3}{2}}}\,\mathrm{d}x = 2\int_{0}^{2} \frac{1}{(1+x^2)^{\frac{3}{2}}}\,\mathrm{d}x$$

If we let the bottom limit instead be a very small number (traditionally represented by ε), we can calculate the integral in terms of ε.

$$\begin{aligned}2\int_{\varepsilon}^{2} \frac{1}{(1+x^2)^{\frac{3}{2}}}\,\mathrm{d}x &= 2\int_{\frac{1}{\varepsilon}}^{\frac{1}{2}} \frac{-t}{(1+t^2)^{\frac{3}{2}}}\,\mathrm{d}t\\ &= 2\left[\frac{1}{(1+t^2)^{\frac{1}{2}}}\right]_{\frac{1}{\varepsilon}}^{\frac{1}{2}}\end{aligned}$$

The contribution from the bottom limit is given by:

$$\frac{2}{\left(1+\left(\frac{1}{\varepsilon}\right)^2\right)^{\frac{1}{2}}} = \frac{2\varepsilon}{\sqrt{1+\varepsilon^2}}$$

Which tends to zero as $\varepsilon \to 0$. We then have:

$$\int_{-2}^{2} \frac{1}{(1+x^2)^{\frac{3}{2}}}\,\mathrm{d}x = 2 \times \frac{1}{\left(1+\left(\frac{1}{2}\right)^2\right)^{\frac{1}{2}}} = \frac{4}{\sqrt{5}}$$

(iii) **(a)** Using the $x = \frac{1}{t}$ substitution (as in part **(i)**) we have:

$$\begin{aligned}\int_{\frac{1}{2}}^{2} \frac{1}{(1+x^2)^2}\,\mathrm{d}x &= \int_{2}^{\frac{1}{2}} \frac{1}{\left(1+\left(\frac{1}{t}\right)^2\right)^2} \times -t^{-2}\,\mathrm{d}t\\ &= -\int_{2}^{\frac{1}{2}} \frac{1}{t^2\left(1+\left(\frac{1}{t}\right)^2\right)^2}\,\mathrm{d}t\\ &= \int_{\frac{1}{2}}^{2} \frac{t^2}{(t^2+1)^2}\,\mathrm{d}t\\ &= \int_{\frac{1}{2}}^{2} \frac{x^2}{(x^2+1)^2}\,\mathrm{d}x\end{aligned}$$

where the last step uses a $t = x$ substitution. Since both of these are the same, let them both be equal to I. We then have:

$$\begin{aligned} 2I &= \int_{\frac{1}{2}}^{2} \frac{1}{(1+x^2)^2}\,\mathrm{d}x + \int_{\frac{1}{2}}^{2} \frac{x^2}{(x^2+1)^2}\,\mathrm{d}x \\ &= \int_{\frac{1}{2}}^{2} \frac{1+x^2}{(x^2+1)^2}\,\mathrm{d}x \\ &= \int_{\frac{1}{2}}^{2} \frac{1}{1+x^2}\,\mathrm{d}x \end{aligned}$$

Therefore we have:

$$\int_{\frac{1}{2}}^{2} \frac{1}{(1+x^2)^2}\,\mathrm{d}x = \int_{\frac{1}{2}}^{2} \frac{x^2}{(x^2+1)^2}\,\mathrm{d}x = I = \tfrac{1}{2}\int_{\frac{1}{2}}^{2} \frac{1}{1+x^2}\,\mathrm{d}x$$

as required.

Using the substitution $x = \tan\theta$ we have:

$$\begin{aligned} \tfrac{1}{2}\int_{\frac{1}{2}}^{2} \frac{1}{1+x^2}\,\mathrm{d}x &= \tfrac{1}{2}\int_{\tan^{-1}\frac{1}{2}}^{\tan^{-1}2} \frac{1}{1+\tan^2\theta} \times \sec^2\theta\,\mathrm{d}\theta \\ &= \tfrac{1}{2}\left(\tan^{-1}2 - \tan^{-1}\frac{1}{2}\right) \end{aligned}$$

(b) Using the $x = \frac{1}{t}$ substitution again we have:

$$\begin{aligned} \int_{\frac{1}{2}}^{2} \frac{1-x}{x(1+x^2)^{\frac{1}{2}}}\,\mathrm{d}x &= \int_{2}^{\frac{1}{2}} \frac{1-\frac{1}{t}}{\frac{1}{t}\left(1+\left(\frac{1}{t}\right)^2\right)^{\frac{1}{2}}} \times -t^{-2}\mathrm{d}t \\ &= \int_{\frac{1}{2}}^{2} \frac{t-1}{\left(1+\left(\frac{1}{t}\right)^2\right)^{\frac{1}{2}}} \times \frac{1}{t^2}\mathrm{d}t \\ &= \int_{\frac{1}{2}}^{2} \frac{t-1}{t\,(t^2+1)^{\frac{1}{2}}}\mathrm{d}t \\ &= -\int_{\frac{1}{2}}^{2} \frac{1-t}{t\,(1+t^2)^{\frac{1}{2}}}\mathrm{d}t \end{aligned}$$

Therefore we have $\displaystyle\int_{\frac{1}{2}}^{2} \frac{1-x}{x(1+x^2)^{\frac{1}{2}}}\,\mathrm{d}x = -\int_{\frac{1}{2}}^{2} \frac{1-x}{x(1+x^2)^{\frac{1}{2}}}\,\mathrm{d}x$, and so we have $\displaystyle\int_{\frac{1}{2}}^{2} \frac{1-x}{x(1+x^2)^{\frac{1}{2}}}\,\mathrm{d}x = 0$.

Question 2

2 **(i)** The real numbers x, y and z satisfy the equations

$$y = \frac{2x}{1-x^2}\,,$$
$$z = \frac{2y}{1-y^2}\,,$$
$$x = \frac{2z}{1-z^2}\,.$$

Let $x = \tan\alpha$. Deduce that $y = \tan 2\alpha$ and show that $\tan\alpha = \tan 8\alpha$.

Find all solutions of the equations, giving each value of x, y and z in the form $\tan\theta$ where $-\frac{1}{2}\pi < \theta < \frac{1}{2}\pi$.

(ii) Determine the number of real solutions of the simultaneous equations

$$y = \frac{3x - x^3}{1-3x^2}\,,$$
$$z = \frac{3y - y^3}{1-3y^2}\,,$$
$$x = \frac{3z - z^3}{1-3z^2}\,.$$

(iii) Consider the simultaneous equations

$$y = 2x^2 - 1\,,$$
$$z = 2y^2 - 1\,,$$
$$x = 2z^2 - 1\,.$$

(a) Determine the number of real solutions of these simultaneous equations with $|x| \leqslant 1$, $|y| \leqslant 1$, $|z| \leqslant 1$.

(b) By finding the degree of a single polynomial equation which is satisfied by x, show that all solutions of these simultaneous equations have $|x| \leqslant 1$, $|y| \leqslant 1$, $|z| \leqslant 1$.

Examiner's report

In part **(i)** most attempts to show $\tan\alpha = \tan 8\alpha$ were successful and included sufficient detail to earn the marks. Some candidates attempted to use the half-angle formula instead of the double-angle formula. This does not work, as the logic goes in the wrong direction, and leads to a quadratic with two solutions following which candidates simply asserted that the half angle is the correct solution. A large proportion of students made no further progress on this question.

Of the students that did progress further on this part, many became confused by the restrictions on range. They realised solutions were of the form $(\tan\alpha, \tan 2\alpha, \tan 4\alpha)$, but then tried to simultaneously have $\alpha, 2\alpha, 4\alpha$ between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$. These solutions erroneously discarded (or did not even find) solutions other than $(0, 0, 0)$. However, some students did realise that they could subtract or add multiples of π to some of the arguments until all were in the required range. These attempts often obtained full marks for this part. Some attempts used an alternative method, rather than using periodicity of tan to solve $\tan\alpha = \tan 8\alpha$, they rewrote in terms of sin and cos and used addition formulae to obtain $\sin 7\alpha = 0$. However, these attempts often only checked the logic in one direction and did not comment that $\cos\alpha$ and $\cos 8\alpha$ were non-zero in these cases.

In part **(ii)** a good number of students attempted the substitution $x = \tan\alpha$, and many of these either quoted or proved the triple angle formula for tan. Again, some made no further progress (often scripts that attempted both **(i)** and **(ii)** made similar amounts of progress on both parts). Since this part asked for the number of solutions, rather than finding all solutions, many students only found solutions for x. This lost credit unless it was accompanied by a check that each value of x led to a value of y and z. Several candidates failed to discard $x = \pm\frac{\pi}{2}$, for which $\tan x$ is undefined, leading to an answer of 27, rather than 25.

In part **(iii)** **(a)** many students attempted a useful trigonometric substitution in this part (either cos or sin). Most scripts that attempted a useful substitution made correct use of the double angle formula to arrive at $\cos\alpha = \cos 8\alpha$ or a similar equation using a sin substitution. Solving the equation $\cos\alpha = \cos 8\alpha$ gave many students significant difficulty. Many attempts used only the periodicity of cos, and not the evenness, thus only obtaining half of the solutions. Others failed to restrict to a range where cos is single-valued, thus finding the same solution for x for different values of α and erroneously counting these as different solutions. Those who chose to draw a sketch of the graph to aid their thinking generally produced better solutions in this part. In part **(iii)** **(b)** most attempts found the correct octic [*Edit: an octic is a polynomial of degree 8*]. Attempts that had found fewer than 8 solutions in **(a)** often made no further progress. Candidates who had found 8 solutions in **(a)** often obtained full marks in this part.

Solution

(i) Substituting $x = \tan\alpha$ into the expression for y gives:

$$\begin{aligned} y &= \frac{2\tan\alpha}{1-\tan^2\alpha} \\ &= \tan 2\alpha \end{aligned}$$

Using the same method we have $z = \tan 4\alpha$, and using this in the expression for x in terms of z gives $x = \tan 8\alpha$. Therefore we have $\tan\alpha = \tan 8\alpha$.

One solution is given by $\alpha = 8\alpha \implies \alpha = 0$, which gives $x = y = z = 0$. Using the symmetry of $\tan\theta$ we also have solutions when $8\alpha = \alpha + n\pi \implies \alpha = \frac{1}{7}n\pi$ (you can sketch a graph of

$\tan\theta$ to help with this — see the comments in part **(iii)(a)** for a similar situation). For θ to be in the required range we need $-3 \leqslant n \leqslant 3$.

If we take $x = \tan\frac{1}{7}\pi$, then this gives $y = \tan\frac{2}{7}\pi$ and $z = \tan\frac{4}{7}\pi = \tan\frac{-3}{7}\pi$ (noting that we need $\frac{1}{2}\pi < \theta < \frac{1}{2}\pi$, and so we cannot have $\frac{4}{7}\pi$).

The non-zero solutions are:

$$\begin{aligned}(x,\, y,\, z) = &\text{ Cyclic permutations of } \left(\tan\tfrac{1}{7}\pi,\, \tan\tfrac{2}{7}\pi,\, \tan\tfrac{-3}{7}\pi\right)\\ \text{and } (x,\, y,\, z) = &\text{ Cyclic permutations of } \left(\tan\tfrac{-1}{7}\pi,\, \tan\tfrac{-2}{7}\pi,\, \tan\tfrac{3}{7}\pi\right)\end{aligned}$$

A *Cyclic permutation* means keeping the relative order of the object the same, so the cyclic permutations of $(A,\, B,\, C)$ are $(A,\, B,\, C)$, $(B,\, C,\, A)$ and $(C,\, A,\, B)$.
You can also just write out all the possible solutions — there are only 7 of them including the $(0,\, 0,\, 0)$ one, so it shouldn't take too long.

(ii) In the previous part, we used an identity for $\tan 2\alpha$. Here there are some 3's appearing, so let's start by considering $\tan 3\alpha$

$$\begin{aligned}\tan 3\alpha &= \tan(\alpha + 2\alpha)\\ &= \frac{\tan\alpha + \tan 2\alpha}{1 - \tan\alpha\tan 2\alpha}\\ &= \frac{t + \frac{2t}{1-t^2}}{1 - t \times \frac{2t}{1-t^2}} \qquad \text{where } t = \tan\alpha\\ &= \frac{t - t^3 + 2t}{1 - t^2 - 2t^2}\\ &= \frac{3t - t^3}{1 - 3t^2}\end{aligned}$$

This is the same form as the expressions for x, y and z in this part. Therefore if $x = \tan\alpha$, then $y = \tan 3\alpha$ and $z = \tan 9\alpha$. This means that we have $\tan 27\alpha = \tan\alpha$, and in a similar way to before we have $26\alpha = n\pi$.

There are 25 values of n which give distinct values of α, which are $\alpha = -12, -11, \cdots, 11, 12$. $n = -13$ and $n = 13$ do not give finite values for $\tan\alpha$. We therefore have 25 possible values for x, but we need to check that these give finite values for y and z as well.

If $\alpha = \frac{n}{26}\pi$, where $n \neq -13$ and $n \neq 13$, then $3\alpha = \frac{3n}{26}\pi$ and $9\alpha = \frac{9n}{26}\pi$ are not multiples of $\frac{1}{2}\pi$, as n and 3 are co-prime to 13. Therefore for each possible value of x there are corresponding possible values of y and z, and there are 25 solutions altogether.

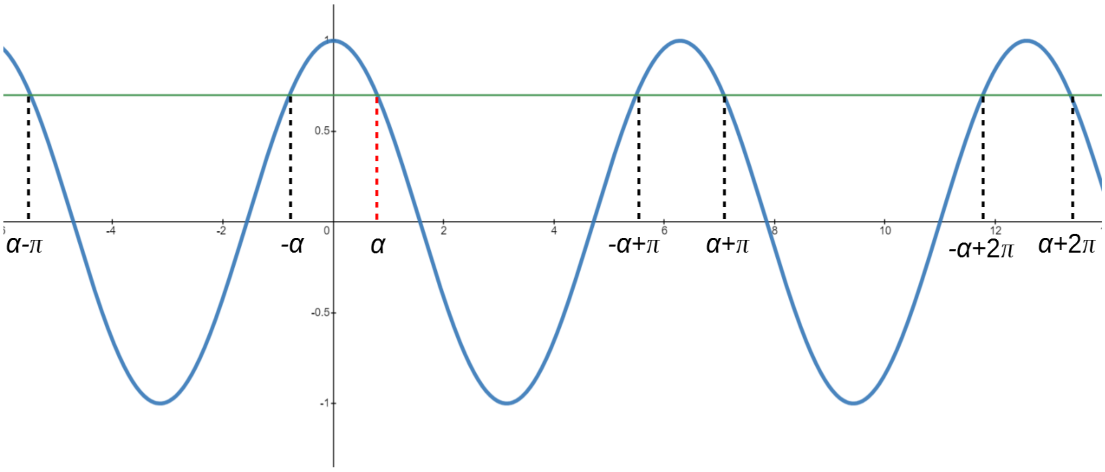
You do need to justify that each of the 25 values of α give finite values for x, y and z, but this explanation was only worth one mark, so don't worry too much if you missed it out!

(iii) These expressions look suspiciously like $\cos 2\alpha$.

(a) We are told that $|x| \leqslant 1$, so we can use the substitution $x = \cos\alpha$, with $0 \leqslant \alpha \leqslant \pi$ (we need to make sure we cover the range of values $-1 \leqslant x \leqslant 1$, which this range of α satisfies). If $x = \cos\alpha$, then $y = \cos 2\alpha$ and $z = \cos 4\alpha$, and we have $\cos 8\alpha = \cos\alpha$.

This means that we need $8\alpha = \alpha + 2n\pi$ or $8\alpha = -\alpha + 2m\pi$.

When trying to find general solutions of trig equations, I rarely remember the general forms, I tend to sketch a graph so that I can work out what the general solution might look like. Below is the sort of thing I would (roughly!) sketch to check that I had the correct general form:



Using $8\alpha = \alpha + 2n\pi$ we have $\alpha = \frac{2}{7}n\pi$, so with $0 \leqslant \alpha \leqslant \pi$ there are four solutions, $n = 0, 1, 2, 3$.

If instead we take $8\alpha = -\alpha + 2m\pi$ we have $\alpha = \frac{2}{9}m\pi$, so with $0 \leqslant \alpha \leqslant \pi$ there are five solutions, $m = 0, 1, 2, 3, 4$.

The solutions with $n = 0$ and $m = 0$ are the same (both give $\alpha = 0$), but the rest are distinct so we have 8 distinct solutions.

(b) We have:

$$\begin{aligned} x &= 2z^2 - 1 \\ &= 2\left(2y^2 - 1\right)^2 - 1 \\ &= 2\left[2\left(2x^2 - 1\right)^2 - 1\right]^2 - 1 \end{aligned}$$

The highest power of x on the right-hand side is 8, so the polynomial in x has degree 8, meaning that there are at most 8 distinct roots of the polynomial, so at most 8 distinct values of x. Since we found 8 distinct values of x in the previous part, these must be all of the possible solutions of the equations, and therefore all of the solutions to the simultaneous equations have $|x| \leqslant 1$, $|y| \leqslant 1$ and $|z| \leqslant 1$.

Question 3

3 Let $\mathrm{p}(x)$ be a polynomial of degree n with $\mathrm{p}(x) > 0$ for all x and let

$$\mathrm{q}(x) = \sum_{k=0}^{n} \mathrm{p}^{(k)}(x)\,,$$

where $\mathrm{p}^{(k)}(x) \equiv \dfrac{\mathrm{d}^k\mathrm{p}(x)}{\mathrm{d}x^k}$ for $k \geqslant 1$ and $\mathrm{p}^{(0)}(x) \equiv \mathrm{p}(x)$.

(i) **(a)** Explain why n must be even and show that $\mathrm{q}(x)$ takes positive values for some values of x.

(b) Show that $\mathrm{q}'(x) = \mathrm{q}(x) - \mathrm{p}(x)$.

(ii) In this part you will be asked to show the same result in three different ways.

(a) Show that the curves $y = \mathrm{p}(x)$ and $y = \mathrm{q}(x)$ meet at every stationary point of $y = \mathrm{q}(x)$.

Hence show that $\mathrm{q}(x) > 0$ for all x.

(b) Show that $\mathrm{e}^{-x}\mathrm{q}(x)$ is a decreasing function.

Hence show that $\mathrm{q}(x) > 0$ for all x.

(c) Show that

$$\int_0^\infty \mathrm{p}(x+t)\mathrm{e}^{-t}\,\mathrm{d}t = \mathrm{p}(x) + \int_0^\infty \mathrm{p}^{(1)}(x+t)\mathrm{e}^{-t}\,\mathrm{d}t\,.$$

Show further that

$$\int_0^\infty \mathrm{p}(x+t)\mathrm{e}^{-t}\,\mathrm{d}t = \mathrm{q}(x)\,.$$

Hence show that $\mathrm{q}(x) > 0$ for all x.

Examiner's report

In part **(i)** **(a)** when assuming that the degree of p is odd for a contradiction, many also assumed that the lead coefficient of $\mathrm{p}(x)$ was positive and so made the statement that $\mathrm{p}(x)$ tends to minus infinity as x tends to minus infinity which is not necessarily correct (unless an argument that the lead coefficient is positive was provided). Many candidates did not provide sufficient detail and so were not awarded full marks for this part.

In part **(i)(b)** candidates generally produced good answers, but a number lost marks for not stating that the $(n+1)^{\text{th}}$ derivative of p is zero sufficiently clearly. Some used +... at the end of the sum of polynomials that define $\mathrm{q}(x)$ or just didn't discuss the final term of $\mathrm{q}'(x)$ and again in these cases it was not sufficiently clear that the key idea had been understood.

In part **(ii) (a)** candidates generally completed the first part well, but a significant number of candidates lost a mark because their argument was the wrong way round, arguing that B implies A rather than A implies B. A significant number of candidates realised that all the stationary points of q must have a positive y-coordinate but they didn't link this to $\mathrm{q}(x)$ being positive for large $|x|$ to get all the marks.

In part **(ii) (b)** the first part was again usually very well done. In a similar way to part **(a)** there were a good number of impressive answers to '$\mathrm{q}(x) > 0$ for all x' but many lost marks by not providing sufficient detail or not including all aspects of the argument (particularly that $\mathrm{q}(x) > 0$ for large x).

Part **(ii) (c)** was generally very well done. Virtually all candidates used the right method for the first part but some lost a mark for not providing sufficient detail in the substitution in integration by parts. Most did the rest of this part well but quite a few candidates lost marks for not dealing with the end term of the summation correctly - in the main line of the solution it is an integral which candidates should explain is zero. Some neglected to include this term without comment or used +... at the end of the sum and, in these cases, it was not clear that the idea had been understood.

Solution

(i) **(a)** If we think about any polynomial with an odd degree, then as $x \to \infty$ in different directions, the polynomial will tend to $+\infty$ in one direction and $-\infty$ in the other, which means that there will be some values where $\mathrm{p}(x) < 0$. Hence we must have n even.

Note that whilst it is *necessary* for n to be even, this is not *sufficient* for $\mathrm{p}(x) > 0$ to hold for all x. Another way of saying this is that $\mathrm{p}(x) > 0$ *only if* n is even — but this is not an "if and only if" situation.

If $\mathrm{p}(x) = a_n x^n + a_{n-1}x^{n-1} \cdots$, then we must have $a_n > 0$ if we are going to have $\mathrm{p}(x) > 0$ for all x.

We have:

$$\begin{aligned}
\mathrm{q}(x) &= \mathrm{p}(x) + \mathrm{p}'(x) + \mathrm{p}''(x) + \cdots \\
&= \left[a_n x^n + a_{n-1}x^{n-1} + \cdots\right] + \left[na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots\right] \\
&\qquad\qquad + \left[n(n-1)a_n x^{n-2} + (n-1)(n-2)a_{n-1}x^{n-3} + \cdots\right] \\
&= a_n x^n + \left[a_{n-1} + na_n\right] x^{n-1} + \cdots
\end{aligned}$$

Therefore the degree of $\mathrm{q}(x)$ is n, where n is even, and the coefficient of x^n is postive, so $\mathrm{q}(x)$ will tend to $+\infty$ as $x \to \pm\infty$. Therefore there must be some values of x where $\mathrm{q}(x)$ is positive.

(b) Since $\mathrm{p}(x)$ is a polynomial of degree n, the $(n+1)^{\text{th}}$ derivative and higher will be zero. For example, if $n = 2$ and $\mathrm{p}(x) = x^2 + 2x + 3$, then when we differentiate three times we get zero.

Differentiating the expression for $\mathrm{q}(x)$ gives:

$$\begin{aligned}
\mathrm{q}(x) &= \sum_{k=0}^{n} \mathrm{p}^{(k)}(x) \\
\mathrm{q}'(x) &= \sum_{k=0}^{n} \mathrm{p}^{(k+1)}(x) \\
&= \mathrm{p}^{(1)}(x) + \mathrm{p}^{(2)}(x) + \cdots + \mathrm{p}^{(n)}(x) + \mathrm{p}^{(n+1)}(x) \\
&= \mathrm{p}^{(1)}(x) + \mathrm{p}^{(2)}(x) + \cdots + \mathrm{p}^{(n)}(x) \qquad \text{as } \ \mathrm{p}^{(n+1)}(x) = 0 \\
&= \sum_{k=0}^{n} \mathrm{p}^{(k)}(x) - \mathrm{p}(x) \\
&= \mathrm{q}(x) - \mathrm{p}(x)
\end{aligned}$$

(ii) **(a)** At a stationary point we have $\mathrm{q}'(x) = 0$, and so we have $\mathrm{p}(x) = \mathrm{q}(x)$, and so the two curves meet at the stationary points of $\mathrm{q}(x)$.

Care is needed here to get the implication in the correct direction. The question is asking you to show that "stationary point of $\mathrm{q}(x) \implies$ the curves meet", and not the other way around. This means that arguments that started with $\mathrm{p}(x) = \mathrm{q}(x)$ were not valid.

We know that $\mathrm{q}(x)$ takes positive values as $x \to \pm\infty$, so at some point it must have at least one minimum value. At every stationary point of $\mathrm{q}(x)$ we have $\mathrm{q}(x) = \mathrm{p}(x)$, so at every (local) minimum of the curve $y = \mathrm{q}(x)$ we have $\mathrm{q}(x) = \mathrm{p}(x) > 0$. Therefore $\mathrm{q}(x) > 0$ for all values of x.

(b) Differentiating $\mathrm{e}^{-x}\mathrm{q}(x)$ gives:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{-x}\mathrm{q}(x)\right) &= \mathrm{e}^{-x}\mathrm{q}'(x) - \mathrm{e}^{-x}\mathrm{q}(x) \\
&= \mathrm{e}^{-x}\big[\mathrm{q}(x) - \mathrm{p}(x)\big] - \mathrm{e}^{-x}\mathrm{q}(x) \\
&= -\mathrm{p}(x)\mathrm{e}^{-x}
\end{aligned}$$

Since we are give $\mathrm{p}(x) > 0$, and we know that $\mathrm{e}^{-x} > 0$, we have $\dfrac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{e}^{-x}\mathrm{q}(x)\right) < 0$ and the function is decreasing.

In part **(i)** we showed that $\mathrm{q}(x) \to +\infty$ as $x \to \infty$. We also have $\mathrm{e}^{-x} > 0$ for all x, and so as $x \to +\infty$ we have $\mathrm{e}^{-x}\mathrm{q}(x) > 0$. Since $\mathrm{e}^{-x}\mathrm{q}(x)$ is a decreasing function, we therefore have $\mathrm{e}^{-x}\mathrm{q}(x) > 0$ for all x, and as e^{-x} is always positive we have $\mathrm{q}(x) > 0$ for all x.

(c) We have:

$$\begin{aligned}\int_0^\infty \mathrm{p}(x+t)\mathrm{e}^{-t}\,\mathrm{d}t &= \Big[-\mathrm{e}^{-t}\mathrm{p}(x+t)\Big]_0^\infty - \int_0^\infty -\mathrm{e}^{-t}\mathrm{p}'(x+t)\,\mathrm{d}t \\ &= \Big[-0 - -\mathrm{e}^0 \times \mathrm{p}(x+0)\Big] + \int_0^\infty \mathrm{p}^{(1)}(x+t)\mathrm{e}^{-t}\,\mathrm{d}t \\ &= \mathrm{p}(x) + \int_0^\infty \mathrm{p}^{(1)}(x+t)\mathrm{e}^{-t}\,\mathrm{d}t\end{aligned}$$

This process can then be repeated until we reach:

$$\int_0^\infty \mathrm{p}(x+t)\mathrm{e}^{-t}\,\mathrm{d}t = \mathrm{p}(x) + \mathrm{p}^{(1)}(x) + \mathrm{p}^{(2)}(x) + \cdots + \mathrm{p}^{(n)}(x) + \int_0^\infty \mathrm{p}^{(n+1)}(x+t)\mathrm{e}^{-t}\,\mathrm{d}t$$

But we have $\mathrm{p}^{(n+1)}(x) = 0$ and so

$$\int_0^\infty \mathrm{p}(x+t)\mathrm{e}^{-t}\,\mathrm{d}t = \mathrm{p}(x) + \mathrm{p}^{(1)}(x) + \mathrm{p}^{(2)}(x) + \cdots + \mathrm{p}^{(n)}(x) = \mathrm{q}(x)$$

We are told that $\mathrm{p}(x) > 0$ for all x, and we know that $\mathrm{e}^{-x} > 0$ for all x, hence $\mathrm{p}(x+t)\mathrm{e}^{-t}$ is positive for all t.

Therefore the integral $\displaystyle\int_0^\infty \mathrm{p}(x+t)\mathrm{e}^{-t}\,\mathrm{d}t = \mathrm{q}(x)$ represents a positive area above the x axis. Hence we have $\mathrm{q}(x) > 0$ for all x.

Question 4

4 **(i)** Show that, if $\left(x - \sqrt{2}\right)^2 = 3$, then $x^4 - 10x^2 + 1 = 0$.

Deduce that, if $\mathrm{f}(x) = x^4 - 10x^2 + 1$, then $\mathrm{f}\left(\sqrt{2} + \sqrt{3}\right) = 0$.

(ii) Find a polynomial g of degree 8 with integer coefficients such that $\mathrm{g}\left(\sqrt{2} + \sqrt{3} + \sqrt{5}\right) = 0$. Write your answer in a form without brackets.

(iii) Let a, b and c be the three roots of $t^3 - 3t + 1 = 0$.

Find a polynomial h of degree 6 with integer coefficients such that $\mathrm{h}\left(a+\sqrt{2}\right) = 0$, $\mathrm{h}\left(b+\sqrt{2}\right) = 0$ and $\mathrm{h}\left(c+\sqrt{2}\right) = 0$. Write your answer in a form without brackets.

(iv) Find a polynomial k with integer coefficients such that $\mathrm{k}\left(\sqrt[3]{2} + \sqrt[3]{3}\right) = 0$. Write your answer in a form without brackets.

Examiner's report

There were a wide variety of different approaches to part **(i)**, including some which identified what the four roots of a quartic with integer coefficients would have to be in order for the required condition to be met.

In part **(ii)** many candidates were able to identify a valid approach to the question although some algebraic errors meant that some did not reach the correct final polynomial. As with part **(i)** there were a number of different approaches that were taken.

In part **(iii)** most candidates recognised that a translation of the graph would provide a cubic with the correct roots. Many were then able to apply similar methods to the earlier parts of the question to obtain the required polynomial with integer coefficients.

Many candidates did not attempt the final part of the question, but those who did were generally able to adapt the methods from the previous parts successfully to make good progress.

Solution

(i) We have:

$$\begin{aligned}\left(x-\sqrt{2}\right)^2 &= 3\\ x^2-2\sqrt{2}x+2 &= 3\\ x^2-1 &= 2\sqrt{2}x\\ x^4-2x^2+1 &= 8x^2\\ x^4-10x^2+1 &= 0\end{aligned}$$

One root of $\left(x-\sqrt{2}\right)^2=3$ is $x=\sqrt{2}+\sqrt{3}$, and so this is also a root of $x^4-10x^2+1=0$, so we have $\mathrm{f}\left(\sqrt{2}+\sqrt{3}\right)=0$.

(ii) $\sqrt{2}+\sqrt{3}+\sqrt{5}$ is a root of $(x-\sqrt{2})^2=\left(\sqrt{3}+\sqrt{5}\right)^2$. Simplifying gives:

$$\begin{aligned}(x-\sqrt{2})^2 &= \left(\sqrt{3}+\sqrt{5}\right)^2\\ x^2-2\sqrt{2}x+2 &= 3+2\sqrt{15}+5\\ x^2-6 &= 2\sqrt{2}x+2\sqrt{15}\\ x^4-12x^2+36 &= 8x^2+8\sqrt{30}x+60\\ x^4-20x^2-24 &= 8\sqrt{30}x\\ x^8-40x^6+352x^4+960x^2+576 &= 1920x^2\\ x^8-40x^6+352x^4-960x^2+576 &= 0\end{aligned}$$

(iii) Using a substitution of $t=x-\sqrt{2}$ gives:

$$\begin{aligned}(x-\sqrt{2})^3-3(x-\sqrt{2})+1 &= 0\\ x^3-3\sqrt{2}x^2+6x-2\sqrt{2}-3x+3\sqrt{2}+1 &= 0\\ x^3-3\sqrt{2}x^2+3x+\sqrt{2}+1 &= 0\\ x^3+3x+1 &= \sqrt{2}\big(3x^2-1\big)\\ x^6+6x^4+2x^3+9x^2+6x+1 &= 2\big(9x^4-6x^2+1\big)\\ x^6-12x^4+2x^3+21x^2+6x-1 &= 0\end{aligned}$$

(iv) Let $x=\sqrt[3]{2}+\sqrt[3]{3}$. We then have:

$$\begin{aligned}x^3 &= \left(\sqrt[3]{2}+\sqrt[3]{3}\right)^3\\ x^3 &= 2+3\times\sqrt[3]{2^2}\times\sqrt[3]{3}+3\times\sqrt[3]{2}\times\sqrt[3]{3^2}+3\\ x^3 &= 5+3\sqrt[3]{12}+3\sqrt[3]{18}\\ x^3-5 &= 3\sqrt[3]{6}\left(\sqrt[3]{2}+\sqrt[3]{3}\right)\\ x^3-5 &= 3\sqrt[3]{6}\,(x) \qquad \textit{Spotting this is the trickiest step!}\\ \left(x^3-5\right)^3 &= 27\times 6\times x^3\\ x^9-15x^6+75x^3-125 &= 162x^3\\ x^9-15x^6-87x^3-125 &= 0\end{aligned}$$

Question 5

5 **(i)** The sequence x_n for $n = 0, 1, 2, \ldots$ is defined by $x_0 = 1$ and by

$$x_{n+1} = \frac{x_n + 2}{x_n + 1}$$

for $n \geqslant 0$.

(a) Explain briefly why $x_n \geqslant 1$ for all n.

(b) Show that $x_{n+1}^2 - 2$ and $x_n^2 - 2$ have opposite sign, and that

$$\left|x_{n+1}^2 - 2\right| \leqslant \tfrac{1}{4}\left|x_n^2 - 2\right| .$$

(c) Show that

$$2 - 10^{-6} \leqslant x_{10}^2 \leqslant 2 .$$

(ii) The sequence y_n for $n = 0, 1, 2, \ldots$ is defined by $y_0 = 1$ and by

$$y_{n+1} = \frac{y_n^2 + 2}{2y_n}$$

for $n \geqslant 0$.

(a) Show that, for $n \geqslant 0$,

$$y_{n+1} - \sqrt{2} = \frac{\left(y_n - \sqrt{2}\right)^2}{2y_n}$$

and deduce that $y_n \geqslant 1$ for $n \geqslant 0$.

(b) Show that

$$y_n - \sqrt{2} \leqslant 2\left(\frac{\sqrt{2} - 1}{2}\right)^{2^n}$$

for $n \geqslant 1$.

(c) Using the fact that

$$\sqrt{2} - 1 < \tfrac{1}{2} ,$$

or otherwise, show that

$$\sqrt{2} \leqslant y_{10} \leqslant \sqrt{2} + 10^{-600} .$$

Examiner's report

In part **(i) (a)** most candidates realised that induction was necessary. Although "explain briefly" was written in the question, some candidates omitted necessary components of an inductive argument here. Some candidates incorrectly stated that the sequence always increased. A popular alternative method was stating $x_n + 2 > x_n + 1$. In this case it is necessary to observe that the denominator is positive to secure full marks.

In part **(i) (b)** many candidates were successful here in rewriting $x_{n+1}^2 - 2$ in terms of x_n but some failed to assert (and very briefly justify) the strict positivity of $(x_n + 1)^2$ in order to show that $x_{n+1}^2 - 2$ and $x_n^2 - 2$ have opposite signs. When showing $|x_{n+1}^2 - 2| \leqslant \frac{1}{4}|x_n^2 - 2|$ the most common mistake was to not use absolute value signs, and write false assertions like $x_{n+1}^2 - 2 \leqslant \frac{1}{4}(x_n^2 - 2)$, which is false for odd n.

In part **(i) (c)** many students used the inequality in the previous part repeatedly to write $|x_{10}^2 - 2| \leqslant \frac{1}{4^{10}}|x_0^2 - 2|$ but did not give a justification that $4^{10} > 10^6$. A small number of candidates were able to calculate x_{10}, and x_{10}^2 successfully and numerically compare these to 2 and $2 - 10^{-6}$, however almost all attempts at this were unsuccessful.

Almost all candidates who attempted part **(ii) (a)** earned at least one mark. In several cases candidates did not formulate a standard inductive argument, either missing the base case or not using an inductive hypothesis.

In part **(ii) (b)** many candidates used $n = 0$ as a base case, but this is not valid here. Of those who opted for an alternative method of using recursion to write $y_n - \sqrt{2}$ in terms of $y_0 - \sqrt{2}$, few were able to justify the exponent for powers of 2. Candidates who attempted a full inductive proof often earned at least 2 of the 4 marks for this part.

Candidates attempting part **(ii)(c)** often earned some marks for showing the correct method, but errors in the accuracy of the work meant that few were able to achieve full marks here.

Solution

(i) **(a)** Rearranging gives:

$$x_{n+1} - 1 = \frac{x_n + 2}{x_n + 1} - 1$$
$$x_{n+1} - 1 = \frac{(x_n + 2) - (x_n + 1)}{x_n + 1}$$
$$x_{n+1} - 1 = \frac{1}{x_n + 1}$$

If $x_n \geqslant 1$, then we have $x_{n+1} - 1 \geqslant 0 \implies x_{n+1} \geqslant 1$. We also have $x_0 = 1$, and so $x_n \geqslant 1$ for $n \geqslant 0$ by induction.

When trying to prove an inequality, it's often a good idea to try and rearrange to compare to zero instead.

(b) We have:

$$\begin{aligned} x_{n+1}^2 - 2 &= \left(\frac{x_n+2}{x_n+1}\right)^2 - 2 \\ &= \frac{(x_n+2)^2 - 2(x_n+1)^2}{(x_n+1)^2} \\ &= \frac{x_n^2 + 4x_n + 4 - 2x_n^2 - 4x_n - 2}{(x_n+1)^2} \\ &= \frac{2 - x_n^2}{(x_n+1)^2} \\ &= \frac{-(x_n^2-2)}{(x_n+1)^2} \end{aligned}$$

Then as $(x_n+1)^2 > 0$, $x_{n+1}^2 - 2$ and $x_n^2 - 2$ have different signs.

Since we have $x_n \geqslant 1$, $(x_n+1)^2 \geqslant 2^2$, and so since $x_{n+1}^2 - 2 = \dfrac{-(x_n^2-2)}{(x_n+1)^2}$ we have $|x_{n+1}^2 - 2| \leqslant \frac{1}{4}|x_n^2 - 2|$.

(c) $x_{10}^2 - 2$ and $x_0^2 - 2$ have the same sign. Since $x_0 = 1$, we know $x_0^2 - 2 < 0$, so $x_{10}^2 - 2 < 0 \implies x_{10}^2 < 2$.

We also have $|x_{n+1}^2 - 2| \leqslant \frac{1}{4}|x_n^2 - 2|$, and so $|x_{10}^2 - 2| \leqslant \frac{1}{4^{10}}|x_0^2 - 2| = \frac{1}{4^{10}}$ (as $x_0 = 1$). Therefore we have $|x_{10}^2 - 2| \leqslant \frac{1}{2^{20}}$. Noting that $2^{10} = 1024 > 1000 = 10^3$, so we have $2^{20} > 10^6$, therefore $2^{-20} < 10^{-6}$ and we have:

$$\begin{aligned} |x_{10}^2 - 2| &\leqslant 10^{-6} \\ 2 - 10^{-6} \leqslant x_{10}^2 &\leqslant 2 + 10^{-6} \\ 2 - 10^{-6} \leqslant x_{10}^2 &\leqslant 2 \qquad \text{as we have previously shown } x_{10}^2 < 2 \end{aligned}$$

(ii) **(a)** We have:

$$\begin{aligned} y_{n+1} - \sqrt{2} &= \frac{y_n^2+2}{2y_n} - \sqrt{2} \\ y_{n+1} - \sqrt{2} &= \frac{y_n^2 + 2 - 2\sqrt{2}y_n}{2y_n} \\ y_{n+1} - \sqrt{2} &= \frac{(y_n - \sqrt{2})^2}{2y_n} \end{aligned}$$

We know that $(y_n - \sqrt{2})^2 \geqslant 0$, so as long as $y_n > 0$, we have $y_{n+1} \geqslant \sqrt{2}$. Since $y_0 = 1$ we have $y_{n+1} \geqslant \sqrt{2}$ for $n \geqslant 0$, so $y_n \geqslant 1$ for $n \geqslant 0$.

(b) Be a little bit careful here - in this case we are looking at $n \geqslant 1$.

Considering the case $n = 1$ we have $y_1 = \dfrac{1^2+2}{2} = \dfrac{3}{2}$, and so $y_1 - \sqrt{2} = \dfrac{3 - 2\sqrt{2}}{2}$.

Looking at the right hand side of the inequality we have

$$2\left(\frac{\sqrt{2}-1}{2}\right)^{2^1} = 2\left(\frac{2 + 1 - 2\sqrt{2}}{4}\right) = \frac{3 - 2\sqrt{2}}{2}$$

Therefore the inequality is true when $n = 1$ (when it is actually an equality).

Assuming the statement is true when $n = k$ we have:

$$y_k - \sqrt{2} \leqslant 2\left(\frac{\sqrt{2}-1}{2}\right)^{2^k}$$

and since $k \geqslant 1$ we know from before that $y_k - \sqrt{2} \geqslant 0$. This means that we have:

$$\begin{aligned} y_k - \sqrt{2} &\leqslant 2\left(\frac{\sqrt{2}-1}{2}\right)^{2^k} \\ \implies \left(y_k - \sqrt{2}\right)^2 &\leqslant 2\left[\left(\frac{\sqrt{2}-1}{2}\right)^{2^k}\right]^2 \end{aligned}$$

Note that if $y_k - \sqrt{2}$ was negative then we couldn't square the inequality and be sure that it would still be true. For example $-3 < 2$ but if we square both sides then the inequality would have to be reversed.

Then considering $n = k + 1$:

$$\begin{aligned} y_{k+1} - \sqrt{2} &= \frac{(y_k - \sqrt{2})^2}{2y_k} && \text{from part } \textbf{(ii)(a)} \\ y_{k+1} - \sqrt{2} &\leqslant \frac{1}{2y_k}\left[2\left(\frac{\sqrt{2}-1}{2}\right)^{2^k}\right]^2 && \text{using inductive hypothesis} \\ y_{k+1} - \sqrt{2} &\leqslant \frac{4}{2y_k}\left(\frac{\sqrt{2}-1}{2}\right)^{2\times 2^k} && \text{squaring} \\ y_{k+1} - \sqrt{2} &\leqslant \frac{2}{y_k}\left(\frac{\sqrt{2}-1}{2}\right)^{2^{k+1}} && \text{simplifying} \\ y_{k+1} - \sqrt{2} &\leqslant 2\left(\frac{\sqrt{2}-1}{2}\right)^{2^{k+1}} && \text{as } y_k \geqslant \sqrt{2} > 1 \text{ for } n \geqslant 1 \end{aligned}$$

Therefore the statement is true when $n = 1$, and if it is true for $n = k$ then it is true for $n = k + 1$, hence it is true for all integers $n \geqslant 1$.

(c) From the work done in **(ii)(a)** we have $y_{n+1} \geqslant \sqrt{2}$ for $n \geqslant 0$, and so we have $y_{10} \geqslant \sqrt{2}$.

Using the result from part **(ii)(b)** we have

$$y_{10} - \sqrt{2} \leqslant 2\left(\frac{\sqrt{2}-1}{2}\right)^{2^{10}}$$

Using the given fact that $\sqrt{2} - 1 < \frac{1}{2}$ gives

$$y_{10} - \sqrt{2} \leqslant 2\left(\frac{1}{4}\right)^{2^{10}}$$

We want to convert this limit into powers of 10, rather than powers of 4. We have:

$$\frac{1}{2^{10}} < 10^{-3}$$

using the fact that $2^{10} = 1024 > 1000$.

Manipulating the inequality:

$$y_{10} - \sqrt{2} \leqslant 2\left(\frac{1}{4}\right)^{2^{10}} = 2 \times \frac{1}{2^{2048}}$$
$$y_{10} - \sqrt{2} \leqslant \frac{2}{2^8} \times \frac{1}{2^{2040}} < \frac{2}{2^8} \times 10^{-612}$$

Where the last line uses the fact that $2^{-10} < 10^{-3}$ so $2^{-10\times 204} < 10^{-3\times 204}$.

Therefore we have $y_{10} - \sqrt{2} \leqslant 10^{-600}$ and putting the two inequalities for y_{10} together gives:

$$\sqrt{2} \leqslant y_{10} \leqslant \sqrt{2} + 10^{-600}$$

as required.

Question 6

6 The sequence F_n, for $n = 0, 1, 2, \ldots$, is defined by $F_0 = 0$, $F_1 = 1$ and by $F_{n+2} = F_{n+1} + F_n$ for $n \geqslant 0$.

Prove by induction that, for all positive integers n,

$$\begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix} = \mathbf{Q}^n,$$

where the matrix $\mathbf{Q}$ is given by

$$\mathbf{Q} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

(i) By considering the matrix $\mathbf{Q}^n$, show that $F_{n+1}F_{n-1} - F_n^2 = (-1)^n$ for all positive integers n.

(ii) By considering the matrix $\mathbf{Q}^{m+n}$, show that $F_{m+n} = F_{m+1}F_n + F_mF_{n-1}$ for all positive integers m and n.

(iii) Show that $\mathbf{Q}^2 = \mathbf{I} + \mathbf{Q}$.

In the following parts, you may use without proof the Binomial Theorem for matrices:

$$(\mathbf{I} + \mathbf{A})^n = \sum_{k=0}^{n} \binom{n}{k} \mathbf{A}^k.$$

(a) Show that, for all positive integers n,

$$F_{2n} = \sum_{k=0}^{n} \binom{n}{k} F_k\,.$$

(b) Show that, for all positive integers n,

$$F_{3n} = \sum_{k=0}^{n} \binom{n}{k} 2^k F_k$$

and also that

$$F_{3n} = \sum_{k=0}^{n} \binom{n}{k} F_{n+k}\,.$$

(c) Show that, for all positive integers n,

$$\sum_{k=0}^{n} (-1)^{n+k} \binom{n}{k} F_{n+k} = 0\,.$$

Examiner's report

Most candidates were able to complete the proof by induction on which the other parts of the question are based. In some cases, the matrix multiplication was not completed correctly (such as calculating the product AB rather than BA). Throughout the question some candidates also got confused about the different variables involved although in some cases where this was clearly simply a mislabelling, they were given the benefit of the doubt.

Most candidates were able to see how the relevant matrices could be used to obtain answers for both part **(i)** and part **(ii)**, but in a small number of cases there was insufficient justification to show that the way in which the result was deduced had been understood.

In part **(iii)** most candidates were able to show that $\mathbf{Q}^2 = \mathbf{I} + \mathbf{Q}$, but many candidates were unable to make much more progress from this point. There were a small number of excellent solutions, carefully checking all of the relevant cases in each part and providing very clear explanations of the reasoning.

Solution

Be careful not to miss the request in the stem of the question!
As is often the case, it turns out that F_n are the Fibonacci numbers.

We are being asked to prove the result for all positive integers, so the smallest value of n is $n = 1$. When $n = 1$ we have:

$$\mathbf{Q}^1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} F_2 & F_1 \\ F_1 & F_0 \end{pmatrix}$$

and so the result is true when $n = 1$.

Assume the result is true when $n = k$, and so we have:

$$\mathbf{Q}^k = \begin{pmatrix} F_{k+1} & F_k \\ F_k & F_{k-1} \end{pmatrix}$$

Then considering $n = k + 1$ we have:

$$\begin{aligned} \mathbf{Q}^{k+1} = \mathbf{Q}\mathbf{Q}^k &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_{k+1} & F_k \\ F_k & F_{k-1} \end{pmatrix} \\ &= \begin{pmatrix} F_{k+1} + F_k & F_k + F_{k-1} \\ F_{k+1} & F_k \end{pmatrix} \\ &= \begin{pmatrix} F_{k+2} & F_{k+1} \\ F_{k+1} & F_k \end{pmatrix} \end{aligned}$$

which is the required form of the matrix when $n = k + 1$, and so the result is true for all positive integers n.

(i) We have $\det \mathbf{Q} = 0 - 1 = -1$. We also have $\det(\mathbf{Q}^n) = [\det \mathbf{Q}]^n$. This gives:

$$F_{n+1}F_{n-1} - F_n^2 = (-1)^n$$

as required.

(ii) Considering $\mathbf{Q}^{m+n} = \mathbf{Q}^m\mathbf{Q}^n$ gives:

$$\begin{pmatrix} F_{m+n+1} & F_{m+n} \\ F_{m+n} & F_{m+n-1} \end{pmatrix} = \begin{pmatrix} F_{m+1} & F_m \\ F_m & F_{m-1} \end{pmatrix} \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix}$$

By considering the top right element of $\mathbf{Q}^{m+n}$ we have:

$$F_{m+n} = F_{m+1}F_n + F_mF_{n-1}$$

as required.

You could have also considered the bottom left element, but would have needed to do a m/n swap in order to get the exact result requested.

(iii) We have:

$$\begin{aligned} \mathbf{Q}^2 &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \mathbf{I} + \mathbf{Q} \end{aligned}$$

(a) Using the result just shown we have:

$$\begin{aligned} \mathbf{Q}^{2n} &= (\mathbf{I} + \mathbf{Q})^n \\ \begin{pmatrix} F_{2n+1} & F_{2n} \\ F_{2n} & F_{2n-1} \end{pmatrix} &= \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} \mathbf{Q}^k \end{aligned}$$

Then considering the top right (or bottom left) element we have:

$$F_{2n} = \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} F_k$$

(b) We have:

$$\begin{aligned} \mathbf{Q}^3 &= \mathbf{Q}\mathbf{Q}^2 \\ &= \mathbf{Q}\left[\mathbf{I} + \mathbf{Q}\right] \\ &= \mathbf{Q} + \mathbf{Q}^2 \\ &= \mathbf{Q} + \mathbf{I} + \mathbf{Q} \\ &= \mathbf{I} + 2\mathbf{Q} \end{aligned}$$

Using this result we have:

$$\begin{aligned} \mathbf{Q}^{3n} &= (\mathbf{I} + 2\mathbf{Q})^n \\ \begin{pmatrix} F_{3n+1} & F_{3n} \\ F_{3n} & F_{3n-1} \end{pmatrix} &= \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} 2^k\mathbf{Q}^k \end{aligned}$$

Taking the top right element gives:

$$F_{3n} = \sum_{k=0}^{n} \binom{n}{k} 2^k F_k$$

Alternatively, using $\mathbf{Q}^3 = \mathbf{Q}\big[\mathbf{I} + \mathbf{Q}\big]$ gives:

$$\begin{aligned}
\mathbf{Q}^{3n} &= \mathbf{Q}^n \big[\mathbf{I} + \mathbf{Q}\big]^n \\
\begin{pmatrix} F_{3n+1} & F_{3n} \\ F_{3n} & F_{3n-1} \end{pmatrix} &= \mathbf{Q}^n \sum_{k=0}^{n} \binom{n}{k} \mathbf{Q}^k \\
&= \sum_{k=0}^{n} \binom{n}{k} \mathbf{Q}^{n+k}
\end{aligned}$$

Taking the top right element gives:

$$F_{3n} = \sum_{k=0}^{n} \binom{n}{k} F_{n+k}$$

An alternative approach for this last result:

From **(ii)** we have

$$F_{3n} = F_{2n+n} = F_{2n+1}F_n + F_{2n}F_{n-1}$$

Using the matrices in **(iii)(a)** and considering the top left element we have:

$$F_{2n+1} = \sum_{k=0}^{n} \binom{n}{k} F_{k+1}$$

Using this and the result from **(iiii)(a)** we have:

$$\begin{aligned}
F_{3n} &= F_{2n+1}F_n + F_{2n}F_{n-1} \\
&= \sum_{k=0}^{n} \binom{n}{k} F_{k+1}F_n + \sum_{k=0}^{n} \binom{n}{k} F_k F_{n-1} \\
&= \sum_{k=0}^{n} \binom{n}{k} \Big[F_{k+1}F_n + F_k F_{n-1}\Big] \\
&= \sum_{k=0}^{n} \binom{n}{k} F_{n+k}
\end{aligned}$$

(c) For this part we want a negative sign to appear in the Binomial Theorem. We have:

$$\mathbf{Q}^2 = \mathbf{I} + \mathbf{Q} \implies \mathbf{I} = \mathbf{Q}\big[\mathbf{Q} - \mathbf{I}\big]$$

We then have:

$$\begin{aligned}
\mathbf{I}^n &= (-1)^n \mathbf{Q}^n \big[\mathbf{I} - \mathbf{Q}\big]^n \\
\mathbf{I}^n &= (-1)^n \mathbf{Q}^n \sum_{k=0}^{n} \binom{n}{k} (-\mathbf{Q})^k \\
\mathbf{I}^n &= \sum_{k=0}^{n} \binom{n}{k} (-1)^{n+k}\, \mathbf{Q}^{n+k}
\end{aligned}$$

Then considering the top right element gives:

$$0 = \sum_{k=0}^{n} (-1)^{n+k} \left(\begin{array}{c} n \\ k \end{array} \right) F_{n+k}$$

Question 7

7 **(i)** The complex numbers z and w have real and imaginary parts given by $z = a + \mathrm{i}b$ and $w = c + \mathrm{i}d$. Prove that $|zw| = |z||w|$.

(ii) By considering the complex numbers $2 + \mathrm{i}$ and $10 + 11\mathrm{i}$, find positive integers h and k such that $h^2 + k^2 = 5 \times 221$.

(iii) Find positive integers m and n such that $m^2 + n^2 = 8045$.

(iv) You are given that $102^2 + 201^2 = 50805$.

Find positive integers p and q such that $p^2 + q^2 = 36 \times 50805$.

(v) Find three distinct pairs of positive integers r and s such that $r^2 + s^2 = 25 \times 1002082$ and $r < s$.

(vi) You are given that $109 \times 9193 = 1002037$.

Find positive integers t and u such that $t^2 + u^2 = 9193$.

Examiner's report

Most candidates were successful in the first two parts, with marks being lost mostly due to the small inaccuracy of forgetting the square root in the expression for the modulus of a complex number.

Part **(iii)** was also typically done well, with most candidates picking up the idea of dividing by 5, however with mixed accuracy on the other factor. The candidates who picked up that the other factor can be written as a sum of squares were mostly successful in this part, as were almost all the candidates who attempted part **(iv)**.

Parts **(v)** and **(vi)** discriminated between candidates, with many successfully getting through **(i)-(iv)** with full marks but unfortunately making little to no progress on these two. Many failed to spot the decompositions $1001^2 + 9^2$ in **(v)** and $1001^2 + 6^2$ in **(vi)**. The candidates who found these got access to the marks, though many didn't manage to find three solutions in part **(v)**. This was from either overlooking the Pythagorean triple of $(3, 4, 5)$ or the simpler solution obtained by noting that 25 is the square of 5. In part **(vi)**, many candidates either chose the wrong complex number and did not try another one or by failing to notice that 10028 or 2943 are divisible by 109.

Solution

(i) We have:

$$\begin{aligned}
|zw|^2 &= |(a+\mathrm{i}b)(c+\mathrm{i}d)|^2 \\
&= |(ac-bd)+\mathrm{i}(ad+bc)|^2 \\
&= (ac-bd)^2+(ad+bc)^2 \\
&= a^2c^2-2abcd+b^2d^2+a^2d^2+2abcd+b^2c^2 \\
&= a^2c^2+b^2d^2+a^2d^2+b^2c^2
\end{aligned}$$

Also:

$$\begin{aligned}
(|z||w|)^2 = |z|^2|w|^2 &= (a^2+b^2)(c^2+d^2) \\
&= a^2c^2+a^2d^2+b^2c^2+b^2d^2
\end{aligned}$$

Therefore we have $|zw|^2 = (|z||w|)^2$ and so $|zw| = |z||w|$ as both sides have to be positive.

(ii) Let $z = 2+\mathrm{i}$ and $w = 10+11\mathrm{i}$, which gives $|z|^2 = 5$ and $|w|^2 = 221$. Using $|zw|^2 = |z|^2|w|^2$, we know that $|zw|^2 = 5\times 221$. We have:

$$\begin{aligned}
zw &= (2+\mathrm{i})(10+11\mathrm{i}) \\
&= 20-11+10\mathrm{i}+22\mathrm{i} \\
&= 9+32\mathrm{i}
\end{aligned}$$

So we have $h = 9$ and $k = 32$.

(iii) Using the same idea as before, start by factorising 8045, and we can see that we have a factor of 5. We have $8045 = 5\times 1609$, and we know that $|2+\mathrm{i}|^2 = 5$. We can write $1609 = 40^2+3^2$, and so we can take $w = 40+3\mathrm{i}$. Expanding dives:

$$\begin{aligned}
(2+\mathrm{i})(40+3\mathrm{i}) &= 80-3+40\mathrm{i}+6\mathrm{i} \\
&= 77+46\mathrm{i}
\end{aligned}$$

and so we have $77^2+46^2 = 8045$.

You could instead consider $(2+\mathrm{i})(3+40\mathrm{i})$ which leads to the answer 34^2+83^2.

(iv) The first thought is to try to write 36 as a sum of two squares, but this is not possible. Instead, consider $|kz|^2 = k^2|z|^2$, where k is a real constant.

We have $36\times 50805 = 6^2\times 50805$. Therefore since $|102+201\mathrm{i}|^2 = 50805$, we have $612^2+1206^2 = 6^2\times 102^2+6^2\times 201^2 = 36\times 50805$.

(v) First note that $1002082 = 1002001+81 = 1001^2+9^2$.

Using the same idea as in part **(iv)** we have one pair of values given by 5005^2+45^2.

In this case, we **can** write 25 as the sum as two squares, i.e. $25 = 3^2+4^2$. Therefore we can also have:

$$\begin{aligned}
(3+4\mathrm{i})(1001+9\mathrm{i}) &= 3003-36+4004\mathrm{i}+27\mathrm{i} \\
&= 2967+4031\mathrm{i}
\end{aligned}$$

So another solution is $2967^2 + 4031^2$.

Alternatively we have:

$$\begin{aligned}(4+3\mathrm{i})(1001+9\mathrm{i}) &= 4004 - 27 + 3003\mathrm{i} + 36\mathrm{i} \\ &= 3977 + 3039\mathrm{i}\end{aligned}$$

So a third solution is $3977^2 + 3039^2$

(vi) We have $10^2 + 3^2 = 109$ and $1001^2 + 6^2 = 1002037$. Let $z = 10 + 3\mathrm{i}$, $w = x + \mathrm{i}y$ and $zw = 1001 + 6\mathrm{i}$, which then gives:

$$\begin{aligned}x + \mathrm{i}y &= \frac{1001+6\mathrm{i}}{10+3\mathrm{i}} \\ &= \frac{(1001+6\mathrm{i})(10-3\mathrm{i})}{109} \\ &= \frac{10028-2943\mathrm{i}}{109} \\ &= 92 - 27\mathrm{i}\end{aligned}$$

Therefore we have $92^2 + 27^2 = 9193$.

The arithmetic here has numbers which are a little too large to handle very comfortably without a calculator (though it is still possible to do so!). A couple of alternative methods are shown below, which you may, or may not, find to be less work.

Neither of the following methods use the statement given at the start of part **(iv)**, but since the question did not say "hence" that's fine!

Method 1

We can write 9193 as a product of two prime numbers, $9193 = 29 \times 317$. We have $29 = 5^2 + 2^2$ and $317 = 11^2 + 14^2$. Considering $(2+5\mathrm{i})(11+14\mathrm{i})$ gives:

$$\begin{aligned}(2+5\mathrm{i})(11+14\mathrm{i}) &= 22 - 70 + (55+28)\mathrm{i} \\ &= -48 + 83\mathrm{i}\end{aligned}$$

Therefore we have $48^2 + 83^2 = 9193$.

Method 2

Square numbers have to end in 0, 1, 4, 5, 6 or 9. The only way to get a sum of two squares which ends in 3 is by picking two square numbers which end in 4 and 9, so WLOG let u^2 end in 4 and t^2 end in 9. This means that u ends in 2 or 8, and t end in 3 or 7. Working systematically through possible values of u or t will (eventually) lead to a solution. For example:

$$\begin{aligned}t = 7 &\implies u^2 = 9193 - 49 = 9144 \quad \text{not a square number} \\ t = 17 &\implies u^2 = 9193 - 289 = 8904 \quad \text{not a square number} \\ t = 27 &\implies u^2 = 9193 - 729 = 8464 = 92^2\end{aligned}$$

Question 8

8 A tetrahedron is called isosceles if each pair of edges which do not share a vertex have equal length.

(i) Prove that a tetrahedron is isosceles if and only if all four faces have the same perimeter.

Let $OABC$ be an isosceles tetrahedron and let $\overrightarrow{OA} = \mathbf{a}$, $\overrightarrow{OB} = \mathbf{b}$ and $\overrightarrow{OC} = \mathbf{c}$.

(ii) By considering the lengths of OA and BC, show that

$$2\mathbf{b}.\mathbf{c} = |\mathbf{b}|^2 + |\mathbf{c}|^2 - |\mathbf{a}|^2.$$

Show that

$$\mathbf{a}.(\mathbf{b} + \mathbf{c}) = |\mathbf{a}|^2.$$

(iii) Let G be the *centroid* of the tetrahedron, defined by $\overrightarrow{OG} = \frac{1}{4}(\mathbf{a} + \mathbf{b} + \mathbf{c})$.

Show that G is equidistant from all four vertices of the tetrahedron.

(iv) By considering the length of the vector $\mathbf{a} - \mathbf{b} - \mathbf{c}$, or otherwise, show that, in an isosceles tetrahedron, none of the angles between pairs of edges which share a vertex can be obtuse. Can any of them be right angles?

Examiner's report

In part **(i)**, most candidates answered the "only if" direction of the argument successfully, often with a diagram. Many candidates did not realise they needed to prove "if" separately, but those that did usually answered this well. Most students wrote a list of simultaneous equations in the edge lengths to solve. Appeals to symmetry were accepted without much detail needed. Many candidates did not attempt the later parts of the question.

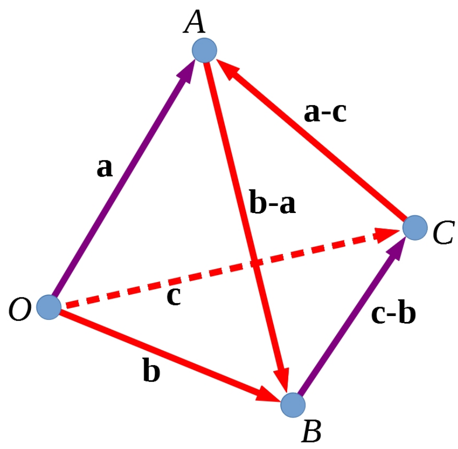
Part **(ii)** was generally answered well. Many students attempted the cosine rule, which required some detail relating it to the given problem. A surprising number of candidates would set the direction vectors of each edge equal, rather than just their lengths. Some ended up confusing scalars and vectors due to poor notation.

In part **(iii)** many candidates found $\mathbf{a}.\mathbf{g}$ etc. rather than $|\overrightarrow{AG}|$. This could be made to work but needed to be made relevant to the question to earn marks. Some candidates thought $\mathbf{a}, \mathbf{b}, \mathbf{c}$ were the components of a vector and attempted to use Pythagoras, which got no credit. Several candidates found $|\overrightarrow{AG}|^2$ etc. in a non-symmetric form and attempted to appeal to symmetry, which was not accepted.

Part **(iv)** proved to be quite tricky for most candidates. Many ignored the questions prompting entirely [*Edit: i.e. they did not consider the length of the vector* $\mathbf{a}-\mathbf{b}-\mathbf{c}$ *as was suggested!*] or failed to relate it to a relevant geometrical idea. Few candidates used cosine successfully and attained the final two marks. Several candidates stated that right angles were possible at the end.

Solution

(i) The first thing to do is to come up with a labelling system. There is a suggested labelling system after part **(i)**, so lets try using that. A clearly labelled diagram will be helpful.



The pairs of edges which do not share a vertex are:

$$\mathbf{a} \text{ and } \mathbf{c}-\mathbf{b}$$
$$\mathbf{b} \text{ and } \mathbf{a}-\mathbf{c}$$
$$\mathbf{c} \text{ and } \mathbf{b}-\mathbf{a}$$

If the tetrahedron is isosceles then we have:

$$|\mathbf{a}| = |\mathbf{c}-\mathbf{b}|$$
$$|\mathbf{b}| = |\mathbf{a}-\mathbf{c}|$$
$$|\mathbf{c}| = |\mathbf{b}-\mathbf{a}|$$

Consider the perimeter of triangle OAB. This has perimeter $|\mathbf{a}|+|\mathbf{b}|+|\mathbf{b}-\mathbf{a}|$, which using the equal lengths above can also be written as $|\mathbf{a}|+|\mathbf{b}|+|\mathbf{c}|$. In a similar way the perimeter of all the other triangular faces is also equal to $|\mathbf{a}|+|\mathbf{b}|+|\mathbf{c}|$. Therefore if the tetrahedron is isosceles then the perimeter of each face is the same.

Working in the opposite direction, assume that the perimeters of all the faces are equal, and then try to show that $|\mathbf{b} - \mathbf{a}| = |\mathbf{c}|$ (and then by symmetry the other cases will also follow).

Considering faces OAB and OAC we have:

$$\begin{aligned} |\mathbf{a}| + |\mathbf{b}| + |\mathbf{b} - \mathbf{a}| &= |\mathbf{a}| + |\mathbf{c}| + |\mathbf{a} - \mathbf{c}| \\ \implies |\mathbf{b}| + |\mathbf{b} - \mathbf{a}| &= |\mathbf{c}| + |\mathbf{a} - \mathbf{c}| \qquad (1) \end{aligned}$$

and considering faces OCB and ABC:

$$\begin{aligned} |\mathbf{c}| + |\mathbf{b}| + |\mathbf{c} - \mathbf{b}| &= |\mathbf{b} - \mathbf{a}| + |\mathbf{a} - \mathbf{c}| + |\mathbf{c} - \mathbf{b}| \\ \implies |\mathbf{c}| + |\mathbf{b}| &= |\mathbf{b} - \mathbf{a}| + |\mathbf{a} - \mathbf{c}| \qquad (2) \\ (1) - (2) \implies & \\ |\mathbf{b} - \mathbf{a}| - |\mathbf{c}| &= |\mathbf{c}| - |\mathbf{b} - \mathbf{a}| \\ 2|\mathbf{b} - \mathbf{a}| &= 2|\mathbf{c}| \end{aligned}$$

and so we have $|\mathbf{b} - \mathbf{a}| = |\mathbf{c}|$. By symmetry the equivalent results hold for the other pairs of edges which do not share a vertex.

Therefore we have that the tetrahedron is isosceles if and only if each pair of edges which do not share a vertex are equal in length.

(ii) Since the tetrahedron is isosceles we have $|\mathbf{a}| = |\mathbf{c} - \mathbf{b}|$ and so:

$$\begin{aligned} |\mathbf{a}|^2 &= |\mathbf{c} - \mathbf{b}|^2 \\ |\mathbf{a}|^2 &= (\mathbf{c} - \mathbf{b}) \cdot (\mathbf{c} - \mathbf{b}) \\ |\mathbf{a}|^2 &= |\mathbf{c}|^2 - 2\mathbf{b} \cdot \mathbf{c} + |\mathbf{b}|^2 \end{aligned}$$

Therefore we have $2\mathbf{b} \cdot \mathbf{c} = |\mathbf{b}|^2 + |\mathbf{c}|^2 - |\mathbf{a}|^2$ as required.

By symmetry we also have $2\mathbf{a} \cdot \mathbf{c} = |\mathbf{a}|^2 + |\mathbf{c}|^2 - |\mathbf{b}|^2$ and $2\mathbf{b} \cdot \mathbf{a} = |\mathbf{b}|^2 + |\mathbf{a}|^2 - |\mathbf{c}|^2$. Adding these two gives:

$$\begin{aligned} 2\mathbf{a} \cdot \mathbf{b} + 2\mathbf{a} \cdot \mathbf{c} &= |\mathbf{b}|^2 + |\mathbf{a}|^2 - |\mathbf{c}|^2 + |\mathbf{a}|^2 + |\mathbf{c}|^2 - |\mathbf{b}|^2 \\ 2\mathbf{a} \cdot \big(\mathbf{b} + \mathbf{c}\big) &= 2|\mathbf{a}|^2 \\ \mathbf{a} \cdot \big(\mathbf{b} + \mathbf{c}\big) &= |\mathbf{a}|^2 \end{aligned}$$

(iii) Consider the distance between A and G given by $|\overrightarrow{AG}| = |\mathbf{g} - \mathbf{a}|$. We have:

$$\begin{aligned}
|\mathbf{g} - \mathbf{a}|^2 &= \frac{1}{16}|4\mathbf{g} - 4\mathbf{a}|^2 \\
&= \frac{1}{16}|\mathbf{a} + \mathbf{b} + \mathbf{c} - 4\mathbf{a}|^2 \\
&= \frac{1}{16}|\mathbf{b} + \mathbf{c} - 3\mathbf{a}|^2 \\
&= \frac{1}{16}\left(9|\mathbf{a}|^2 + |\mathbf{b}|^2 + |\mathbf{c}|^2 + 2\mathbf{b}\cdot\mathbf{c} - 6\mathbf{a}\cdot\mathbf{c} - 6\mathbf{a}\cdot\mathbf{b}\right) \\
&= \frac{1}{16}\left(9|\mathbf{a}|^2 + |\mathbf{b}|^2 + |\mathbf{c}|^2 + 2\mathbf{b}\cdot\mathbf{c} - 6\mathbf{a}\cdot(\mathbf{b} + \mathbf{c})\right) \\
&= \frac{1}{16}\left(9|\mathbf{a}|^2 + |\mathbf{b}|^2 + |\mathbf{c}|^2 + |\mathbf{b}|^2 + |\mathbf{c}|^2 - |\mathbf{a}|^2 - 6|\mathbf{a}|^2\right) \\
&= \frac{1}{8}\left(|\mathbf{a}|^2 + |\mathbf{b}|^2 + |\mathbf{c}|^2\right)
\end{aligned}$$

This is symmetric in $\mathbf{a}, \mathbf{b}, \mathbf{c}$, therefore G is equidistant from A, B and C. Considering the distance OG gives:

$$\begin{aligned}
|\mathbf{g}|^2 &= \frac{1}{16}|\mathbf{a} + \mathbf{b} + \mathbf{c}|^2 \\
&= \frac{1}{16}\left(|\mathbf{a}|^2 + |\mathbf{b}|^2 + |\mathbf{c}|^2 + 2\mathbf{a}\cdot\mathbf{b} + 2\mathbf{b}\cdot\mathbf{c} + 2\mathbf{c}\cdot\mathbf{a}\right) \\
&= \frac{1}{16}\left(|\mathbf{a}|^2 + |\mathbf{b}|^2 + |\mathbf{c}|^2 + (\mathbf{a}\cdot\mathbf{b} + \mathbf{a}\cdot\mathbf{c}) + (\mathbf{b}\cdot\mathbf{a} + \mathbf{b}\cdot\mathbf{c}) + (\mathbf{c}\cdot\mathbf{a} + \mathbf{c}\cdot\mathbf{b})\right) \\
&= \frac{1}{16}\left(|\mathbf{a}|^2 + |\mathbf{b}|^2 + |\mathbf{c}|^2 + |\mathbf{a}|^2 + |\mathbf{b}|^2 + |\mathbf{c}|^2\right) \\
&= \frac{1}{8}\left(|\mathbf{a}|^2 + |\mathbf{b}|^2 + |\mathbf{c}|^2\right)
\end{aligned}$$

Therefore G is equidistant from all four vertices.

(iv) The "or otherwise" statement in this part is so that those who took a different route can still potentially get the marks, but it is almost always easier to use the method suggested by the question even if you cannot immediately see how it will help.

We have:

$$\begin{aligned}
|\mathbf{a} - \mathbf{b} - \mathbf{c}|^2 &= |\mathbf{a}|^2 + |\mathbf{b}|^2 + |\mathbf{c}|^2 - 2\mathbf{a}\cdot(\mathbf{b} + \mathbf{c}) + 2\mathbf{b}\cdot\mathbf{c} \\
&= |\mathbf{a}|^2 + |\mathbf{b}|^2 + |\mathbf{c}|^2 - 2|\mathbf{a}|^2 + (|\mathbf{b}|^2 + |\mathbf{c}|^2 - |\mathbf{a}|^2) \\
&= 2(|\mathbf{b}|^2 + |\mathbf{c}|^2 - |\mathbf{a}|^2)
\end{aligned}$$

Since we have $|\mathbf{a} - \mathbf{b} - \mathbf{c}|^2 \geqslant 0$ we must have $|\mathbf{b}|^2 + |\mathbf{c}|^2 - |\mathbf{a}|^2 \geqslant 0$. Since the tetrahedron is isosceles we have $|BC| = |\mathbf{a}|$ etc., and so we have $\cos BAC \geqslant 0$, hence the angle BAC cannot be obtuse. By symmetry, none of the angles can be obtuse.

If BAC is a right angle then we have $|\mathbf{b}|^2 + |\mathbf{c}|^2 = |\mathbf{a}|^2$, and so we must have $\mathbf{a} - \mathbf{b} - \mathbf{c} = 0 \implies \mathbf{a} = \mathbf{b} + \mathbf{c}$. This means that $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are in the same plane and the shape cannot be a tetrahedron.

Question 9

9 A truck of mass M is connected by a light, rigid tow-bar, which is parallel to the ground, to a trailer of mass kM. A constant driving force D which is parallel to the ground acts on the truck, and the only resistance to motion is a frictional force acting on the trailer, with coefficient of friction μ.

- When the truck pulls the trailer up a slope which makes an angle α to the horizontal, the acceleration is a_1 and there is a tension T_1 in the tow-bar.
- When the truck pulls the trailer on horizontal ground, the acceleration is a_2 and there is a tension T_2 in the tow-bar.
- When the truck pulls the trailer down a slope which makes an angle α to the horizontal, the acceleration is a_3 and there is a tension T_3 in the tow-bar.

All accelerations are taken to be positive when in the direction of motion of the truck.

(i) Show that $T_1 = T_3$ and that $M(a_1 + a_3 - 2a_2) = 2(T_2 - T_1)$.

(ii) It is given that $\mu < 1$.

(a) Show that

$$a_2 < \tfrac{1}{2}(a_1 + a_3) < a_3\,.$$

(b) Show further that

$$a_1 < a_2\,.$$

Examiner's report

Less than half of the candidates produced an accurate diagram for this question, with many leaving off some forces, or making errors with the gravitational force by not including g. This had an impact on their ability to proceed with the question, and often those with poorly presented diagrams had sign errors in their force balance equations (for example, with tension in the wrong direction). Most seemed to understand how to calculate frictional force. The inclusion of friction on the trailer and not the truck clearly confused some candidates, causing many of the question parts to be inaccessible.

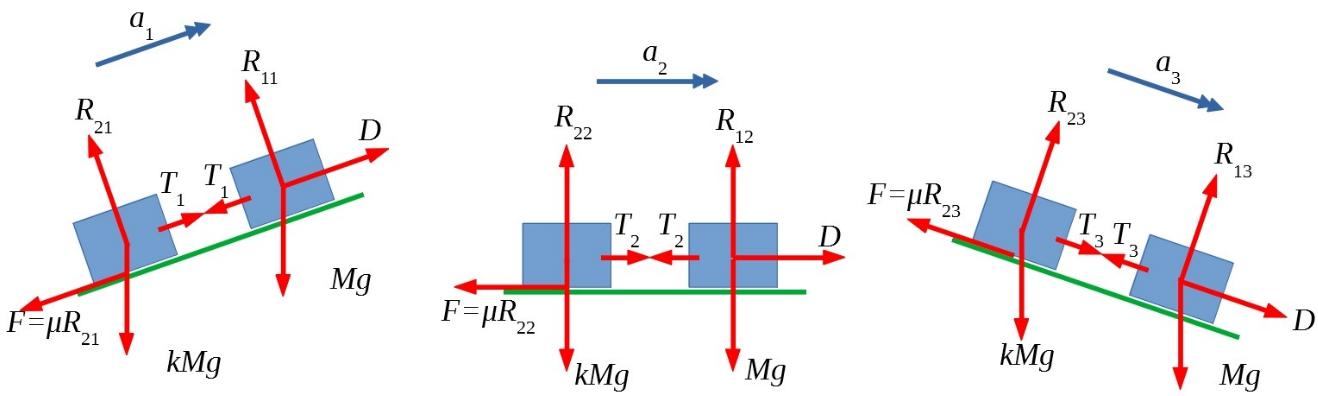
Many candidates seemed to struggle with the fact that 6 equations had to be dealt with, and so struggled to identify which variables to eliminate and how to eliminate them. Additionally, some candidates did not realise that some of the forces would take different values in the different cases being considered.

Part **(i)** was done quite well overall, although for the second part, a fair number of candidates showed each side was equal to some expression involving the other variables, which is valid but took much more time than the direction approach using the equation of motion for the truck.

Part **(ii)** **(a)** was done well in some cases, although less well than the previous part. Most candidates who attempted this part were able to show the upper inequality, but the lower one proved to be more difficult. Most attempts to part **(ii)** **(b)** only achieved two of the marks available. Many candidates did not recognise that the half angle formula was useful here and so struggled to make progress on the question.

Solution

As is often the case, it's a good idea to start with a clear diagram! I have used different labels for the 6 different reaction forces, with the convention that R_{13} is the reaction force on the truck in the third diagram (when acceleration is a_3). Your diagrams would probably be a little larger than the ones shown below!



(i) Resolving forces when the truck and trailer are going uphill we have:

$$D - T_1 - Mg\sin\alpha = Ma_1 \tag{1}$$
$$R_{11} = Mg\cos\alpha \quad \text{(This turns out not to be useful!)}$$
$$T_1 - \mu R_{21} - kMg\sin\alpha = kMa_1$$
$$R_{21} = kMg\cos\alpha$$
$$\Longrightarrow \; T_1 - \mu kMg\cos\alpha - kMg\sin\alpha = kMa_1 \tag{2}$$

Resolving for the downhill case gives:

$$D - T_3 + Mg\sin\alpha = Ma_3 \tag{3}$$
$$T_3 - \mu kMg\cos\alpha + kMg\sin\alpha = kMa_3 \tag{4}$$

The horizontal case gives:

$$D - T_2 = Ma_2 \tag{5}$$
$$T_2 - \mu kMg = kMa_2 \tag{6}$$

There are lots of equations here, so it is a good idea to number the ones that we think we will use later! It doesn't matter if you label ones you don't end up using.

The first thing we are asked to show is that $T_1 = T_3$, so we probably don't need the horizontal equations for this part. Using (1) and (2) to eliminate a_1 gives:

$$\begin{aligned} T_1 - \mu k M g \cos\alpha - kMg\sin\alpha &= k(D - T_1 - Mg\sin\alpha) \\ \implies (1+k)T_1 &= kD + \mu k M g \cos\alpha \end{aligned}$$

Similarly we can use (3) and (4) to eliminate a_3:

$$\begin{aligned} T_3 - \mu k M g \cos\alpha + kMg\sin\alpha &= k(D - T_3 + Mg\sin\alpha) \\ \implies (1+k)T_3 &= kD + \mu k M g \cos\alpha \end{aligned}$$

Therefore the expressions for T_1 and T_3 are the same, so we have $T_1 = T_3$.

For the second result consider $(1) + (3) - 2(5)$ to give:

$$\begin{aligned} \big[D - T_1 - Mg\sin\alpha\big] + \big[D - T_3 + Mg\sin\alpha\big] - 2\big[D - T_2\big] &= Ma_1 + Ma_3 - 2Ma_2 \\ 2T_2 - T_1 - T_3 &= M\big(a_1 + a_3 - 2a_2\big) \\ 2\big(T_2 - T_1\big) &= M\big(a_1 + a_3 - 2a_2\big) \end{aligned}$$

where the last step uses the fact that $T_1 = T_3$.

(ii)(a) If we can show that $T_2 > T_1$ then using the second result from part **(i)** will give $a_1 + a_3 > 2a_2$.

From earlier we have:

$$T_1 = \frac{kD + \mu k M g \cos\alpha}{1+k}$$

Eliminating a_2 from equations (5) and (6) gives:

$$\begin{aligned} T_2 - \mu k M g &= k\big(D - T_2\big) \\ \implies T_2 &= \frac{kD + \mu k M g}{1+k} \end{aligned}$$

Then since $\cos\alpha < 1$ (as $\alpha < 90°$; we are not trying to drive up a vertical cliff face!) we have $T_2 > T_1$, and so we have $a_1 + a_3 > 2a_2 \implies a_2 < \frac{1}{2}(a_1 + a_3)$.

Rearranging (2) and (4) to find a_1 and a_3 gives:

$$\begin{aligned} a_1 &= \frac{1}{kM}\big(T_1 - \mu k M g \cos\alpha - kMg\sin\alpha\big) \\ a_3 &= \frac{1}{kM}\big(T_3 - \mu k M g \cos\alpha + kMg\sin\alpha\big) \end{aligned}$$

and since $T_1 = T_3$ and $\sin\alpha > 0$ we have $a_3 > a_1 \implies a_3 > \frac{1}{2}(a_3 + a_1)$. Putting these together gives:

$$a_2 < \frac{1}{2}(a_1 + a_3) < a_3$$

as required.

Note that it seems very sensible that $a_3 > a_1$ — you would expect your acceleration going down a slope to be greater than your acceleration going up!

(ii)(b) This also seems like a sensible result!

It would probably be useful to have an expression of the form $a_2 - a_1$, and then try to show that $a_2 - a_1 > 0$. It is usually easier to show that an inequality is positive or negative than to compare non-zero values.

Adding equations (1) and (2) gives:

$$D - Mg\sin\alpha - \mu kMg\cos\alpha - kMg\sin\alpha = M(k+1)a_1 \tag{7}$$

Adding (5) and (6) gives:

$$D - \mu kMg = M(k+1)a_2 \tag{8}$$

Then using (8) − (7) gives:

$$\begin{aligned}
M(k+1)\big(a_2 - a_1\big) &= \big[D - \mu kMg\big] - \big[D - Mg\sin\alpha - \mu kMg\cos\alpha - kMg\sin\alpha\big]\\
M(k+1)\big(a_2 - a_1\big) &= M\big[g\sin\alpha + \mu kg\cos\alpha + kg\sin\alpha - \mu kg\big]\\
(k+1)\big(a_2 - a_1\big) &= g\sin\alpha + \mu kg\cos\alpha + kg\sin\alpha - \mu kg\\
(k+1)\big(a_2 - a_1\big) &= (k+1)g\sin\alpha + \mu kg\big(\cos\alpha - 1\big)\\
(k+1)\big(a_2 - a_1\big) &= (k+1)g\sin\alpha - \mu kg\big(1 - \cos\alpha\big) \qquad (*)
\end{aligned}$$

At this stage it's not obvious what to try next. We know that $\mu < 1$, so $\mu k < k$. We also know that $\alpha < 90°$, but we really need a way to compare $\sin\alpha$ and $1 - \cos\alpha$. Using the double angle formulae we have:

$$\begin{aligned}
\sin\alpha &= 2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}\\
1 - \cos\alpha &= 2\sin^2\frac{\alpha}{2}
\end{aligned}$$

Substituting these into $(*)$ gives:

$$\begin{aligned}
(k+1)\big(a_2 - a_1\big) &= (k+1)g\sin\alpha - \mu kg\big(1 - \cos\alpha\big)\\
(k+1)\big(a_2 - a_1\big) &= 2(k+1)g\sin\frac{\alpha}{2}\cos\frac{\alpha}{2} - 2\mu kg\sin^2\frac{\alpha}{2}\\
(k+1)\big(a_2 - a_1\big) &= 2g\sin\frac{\alpha}{2}\left[(1+k)\cos\frac{\alpha}{2} - \mu k\sin\frac{\alpha}{2}\right] \qquad (\dagger)
\end{aligned}$$

We have $1 + k > k > \mu k$ and we have $\cos\frac{\alpha}{2} > \sin\frac{\alpha}{2}$ because $\frac{\alpha}{2} < 45°$. Hence the right hand side of $(\dagger)$ is positive, and so we have $a_2 > a_1$.

Question 10

10 In this question, the x- and y-axes are horizontal and the z-axis is vertically upwards.

(i) A particle P_α is projected from the origin with speed u at an acute angle α above the positive x-axis.

The curve E is given by $z = A - Bx^2$ and $y = 0$. If E and the trajectory of P_α touch exactly once, show that

$$u^2 - 2gA = u^2\big(1 - 4AB\big)\cos^2\alpha\,.$$

E and the trajectory of P_α touch exactly once for all α with $0 < \alpha < \frac{1}{2}\pi$. Write down the values of A and B in terms of u and g.

An explosion takes place at the origin and results in a large number of particles being simultaneously projected with speed u in different directions. You may assume that all the particles move freely under gravity for $t \geqslant 0$.

(ii) Describe the set of points which can be hit by particles from the explosion, explaining your answer.

(iii) Show that, at a time t after the explosion, the particles lie on a sphere whose centre and radius you should find.

(iv) Another particle Q is projected horizontally from the point $(0, 0, A)$ with speed u in the positive x direction.

Show that, at all times, Q lies on the curve E.

(v) Show that for particles Q and P_α to collide, Q must be projected a time $\dfrac{u(1-\cos\alpha)}{g\sin\alpha}$ after the explosion.

Examiner's report

Part **(i)** was answered well. Almost all candidates used the discriminant condition correctly, even if their quadratic contained an error. The final two marks in this part were trickier to achieve. Many candidates substituted two values for α and then solved for A, B. Of these, it was fairly common for them to use $\alpha = 0, \frac{1}{2}\pi$ which were excluded from the range being considered. Many missed that one can simply set the coefficients of each side to zero. Some treated α as a variable to be solved for rather than varied.

Part **(ii)** was found very tricky. Some candidates were able to guess or intuit that the safe zone should be the parabola considered in part **(i)**, but almost no candidate gave a proper explanation as to why it was only these points that could be reached. A few considered the 2D parabola correctly but did not consider 3D. Some wondered about the presence of a "floor" at $z = 0$. This only caused an issue if the candidate thought we were only interested in points of intersection with this plane, possibly caused by imagining the problem in the context of artillery as it is often presented in schools. In this case they answered with a circle and received no credit.

Part **(iii)** was fairly tough for most candidates. Many candidates who could not find the equation of a circle/sphere would attempt an explanation in words, which was almost never sufficient to earn credit. It was possible for candidates to guess the correct centre and radius with no mathematical justification, but this earned no credit.

Many candidates who attempted part **(iv)** were able to complete it successfully.

Candidates who attempted part **(v)** were generally able to pick up at least one mark by appealing to earlier calculations. Most appreciated the need to introduce separate times for Q and P, or a time delay between the two.

Solution

(i) Initially we are only interested in the $x-z$ plane, and there is no movement in the y direction.

The particle follows the trajectory $x = ut\cos\alpha$ and $z = ut\sin\alpha - \frac{1}{2}gt^2$. If it is to touch the curve E then there has to be a repeated root of

$$ut\sin\alpha - \tfrac{1}{2}gt^2 = A - B\left(ut\cos\alpha\right)^2$$

Rearranging gives:

$$\left(Bu^2\cos^2\alpha - \tfrac{1}{2}g\right)t^2 + (u\sin\alpha)t - A = 0$$

For a repeated root we need the discriminant to be equal to zero.

$$\begin{aligned}
u^2\sin^2\alpha + 4A\left(Bu^2\cos^2\alpha - \tfrac{1}{2}g\right) &= 0\\
u^2(1-\cos^2\alpha) + 4A\left(Bu^2\cos^2\alpha - \tfrac{1}{2}g\right) &= 0\\
u^2 - u^2\cos^2\alpha + 4ABu^2\cos^2\alpha - 2Ag &= 0\\
\implies \quad u^2 - 2Ag &= u^2(1-4AB)\cos^2\alpha
\end{aligned}$$

If the above is going to be true for all values of α then both the LHS and the RHS have to equal to 0.

This means we have:

$$\begin{aligned}
u^2 - 2gA = 0 &\implies A = \frac{u^2}{2g}\\
1 - 4AB = 0 &\implies B = \frac{1}{4A} = \frac{g}{2u^2}
\end{aligned}$$

A different method is to consider the fact that the path of the particle and curve E touch if where they meet the gradients are the same.

The path of the particle is given by $x = ut\cos\alpha$ and $z = ut\sin\alpha - \frac{1}{2}gt^2$. Using $t = \dfrac{x}{u\cos\alpha}$ to eliminate t in the z equation gives:

$$\begin{aligned}
z &= x\tan\alpha - \frac{gx^2\sec^2\alpha}{2u^2}\\
\implies \frac{\mathrm{d}z}{\mathrm{d}x} &= \tan\alpha - \frac{gx\sec^2\alpha}{u^2}
\end{aligned}$$

Curve E has equation $z = A - Bx^2$ and gradient $\dfrac{\mathrm{d}z}{\mathrm{d}x} = -2Bx$. Equating the gradients gives:

$$\begin{aligned} -2Bx &= \tan\alpha - \frac{gx\sec^2\alpha}{u^2} \\ \implies \frac{gx\sec^2\alpha}{u^2} &= \tan\alpha + Bx \qquad (*) \\ \text{and} \quad x &= \frac{u^2\tan\alpha}{g\sec^2\alpha - 2Bu^2} \qquad (\dagger) \end{aligned}$$

Equating expressions for z gives:

$$\begin{aligned} A - Bx^2 &= x\tan\alpha - \frac{gx^2\sec^2\alpha}{2u^2} \\ A - Bx^2 &= x\tan\alpha - \frac{x}{2}\big[\tan\alpha + 2Bx\big] \quad \text{using } (*) \\ A - Bx^2 &= x\tan\alpha - \frac{1}{2}x\tan\alpha - Bx^2 \\ \implies A &= \frac{1}{2}x\tan\alpha \\ A &= \frac{u^2\tan^2\alpha}{2(g\sec^2\alpha - 2Bu^2)} \quad \text{using } (\dagger) \\ \implies 2gA\sec^2\alpha - 4ABu^2 &= u^2\tan^2\alpha \\ \implies 2gA - 4ABu^2\cos^2\alpha &= u^2\sin^2\alpha \\ 2gA - 4ABu^2\cos^2\alpha &= u^2(1-\cos^2\alpha) \\ \implies u^2 - 2gA &= u^2(1-4AB)\cos^2\alpha \end{aligned}$$

(ii) If S is the surface formed when you rotate $z = A - Bx^2$ around the z-axis, then the particle cannot cross the boundary of S, so everything inside and on the boundary S could be hit.

(iii) Considering the particle path $x = ut\cos\alpha$, $z = ut\sin\alpha - \frac{1}{2}gt^2$ we can eliminate $\cos\alpha$ to get:

$$x^2 + \left(z + \tfrac{1}{2}gt^2\right)^2 = (ut)^2$$

This is the equation of a circle in the $x - z$ plane, with centre $\left(0, -\frac{1}{2}gt^2\right)$ and radius ut. If you rotate this around the z axis you get a sphere radius ut and centre $\left(0, 0, -\frac{1}{2}gt^2\right)$.

(iv) Particle Q follows the trajectory $(ut, 0, A - \frac{1}{2}gt^2)$. Eliminating t gives:

$$\begin{aligned} z &= A - \frac{1}{2}g\left(\frac{x}{u}\right)^2 \\ z &= A - \frac{g}{2u^2}x^2 \\ z &= A - Bx^2 \end{aligned}$$

therefore Q lies on the curve E.

(v) First work out when P meets E. From part **(i)** this is when:

$$\begin{aligned}
ut\sin\alpha - \tfrac{1}{2}gt^2 &= A - B\left(ut\cos\alpha\right)^2\\
ut\sin\alpha - \tfrac{1}{2}gt^2 &= \frac{u^2}{2g} - \frac{g}{2u^2}\left(ut\cos\alpha\right)^2\\
ut\sin\alpha - \tfrac{1}{2}gt^2 &= \frac{u^2}{2g} - \frac{g}{2}\left(t\cos\alpha\right)^2\\
ut\sin\alpha - \tfrac{1}{2}gt^2 &= \frac{u^2}{2g} - \frac{g}{2}t^2\left(1-\sin^2\alpha\right)\\
ut\sin\alpha &= \frac{u^2}{2g} + \frac{g}{2}t^2\sin^2\alpha
\end{aligned}$$

This can be rearranged to give:

$$\begin{aligned}
\implies g^2t^2\sin^2\alpha - 2gut\sin\alpha + u^2 &= 0\\
(gt\sin\alpha - u)^2 &= 0\\
\implies t &= \frac{u}{g\sin\alpha}
\end{aligned}$$

At this time the x component of the position of P is given by $x = ut\cos\alpha = \dfrac{u^2}{g\tan\alpha}$. The time it takes Q to reach this x value if given by ut', so we have:

$$t' = \frac{u}{g\tan\alpha}$$

The difference between the times it takes P and Q to reach this point is give:

$$\begin{aligned}
t - t' &= \frac{u}{g\sin\alpha} - \frac{u}{g\tan\alpha}\\
&= \frac{u(1-\cos\alpha)}{g\sin\alpha}
\end{aligned}$$

Therefore Q has to be projected a time $\dfrac{u(1-\cos\alpha)}{g\sin\alpha}$ after the explosion.

Question 11

11 **(i)** X_1 and X_2 are both random variables which take values $x_1, x_2, \ldots, x_n$, with probabilities $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ respectively.

The value of random variable Y is defined to be that of X_1 with probability p and that of X_2 with probability $q = 1 - p$.

If X_1 has mean μ_1 and variance σ_1^2, and X_2 has mean μ_2 and variance σ_2^2, find the mean of Y and show that the variance of Y is $p\sigma_1^2 + q\sigma_2^2 + pq(\mu_1 - \mu_2)^2$.

(ii) To find the value of random variable B, a fair coin is tossed and a fair six-sided die is rolled. If the coin shows heads, then $B = 1$ if the die shows a six and $B = 0$ otherwise; if the coin shows tails, then $B = 1$ if the die does **not** show a six and $B = 0$ if it does. The value of Z_1 is the sum of n independent values of B, where n is large.

Show that Z_1 is a Binomial random variable with probability of success $\frac{1}{2}$.

Using a Normal approximation, show that the probability that Z_1 is within 10% of its mean tends to 1 as $n \longrightarrow \infty$.

(iii) To find the value of random variable Z_2, a fair coin is tossed and n fair six-sided dice are rolled, where n is large. If the coin shows heads, then the value of Z_2 is the number of dice showing a six; if the coin shows tails, then the value of Z_2 is the number of dice **not** showing a six.

Use part **(i)** to write down the mean and variance of Z_2.

Explain why a Normal distribution with this mean and variance will not be a good approximation to the distribution of Z_2.

Show that the probability that Z_2 is within 10% of its mean tends to 0 as $n \longrightarrow \infty$.

Examiner's report

There were very few substantial attempts at this question overall.

In part **(i)** a large number of candidates incorrectly stated that $Y = pX_1 + qX_2$. However, there were several good responses to this question with many candidates obtaining the correct value for at least one of the mean and variance of Y.

Similarly, in part **(ii)** many candidates were able to compute the mean and variance of Z_1 correctly. However, several candidates only computed $\mathrm{P}(B = 1)$ when asked to justify that Z_1 is a binomial variable.

Candidates generally struggled with part **(iii)**, often comparing the variance and the mean incorrectly for the two facts that were required to be shown.

Solution

(i) The probability that Y takes the value x_i is given by:

$$\mathrm{P}(Y = x_i) = p\mathrm{P}(X_1 = x_i) + q\mathrm{P}(X_2 = x_i) = pa_i + qb_i$$

We have:

$$\begin{aligned}
\mathrm{E}(Y) &= \sum_{i=1}^{n} x_i \mathrm{P}(Y = x_i) \\
&= \sum_{i=1}^{n} x_i \big(pa_i + qb_i\big) \\
&= p\sum_{i=1}^{n} a_i x_i + q\sum_{i=1}^{n} b_1 x_i \\
&= p\mu_1 + q\mu_2
\end{aligned}$$

Using $\mathrm{Var}(Y) = \mathrm{E}(Y^2) - \big[\mathrm{E}(Y)\big]^2$:

$$\begin{aligned}
\mathrm{Var}(Y) &= \sum_{i=1}^{n} x_i^2 \big(pa_i + qb_i\big) - \big[p\mu_1 + q\mu_2\big]^2 \\
&= p\mathrm{E}(X_1^2) + q\mathrm{E}(X_2^2) - \big[p\mu_1 + q\mu_2\big]^2 \\
&= p\big(\sigma_1^2 + \mu_1^2\big) + q\big(\sigma_2^2 + \mu_2^2\big) - \big[p\mu_1 + q\mu_2\big]^2 \\
&= p\sigma_1^2 + q\sigma_2^2 + (p - p^2)\mu_1^2 + (q - q^2)\mu_2^2 - 2pq\mu_1\mu_2 \\
&= p\sigma_1^2 + q\sigma_2^2 + p(1 - p)\mu_1^2 + q(1 - q)\mu_2^2 - 2pq\mu_1\mu_2 \\
&= p\sigma_1^2 + q\sigma_2^2 + pq\mu_1^2 + pq\mu_2^2 - 2pq\mu_1\mu_2 \\
&= p\sigma_1^2 + q\sigma_2^2 + pq(\mu_1 - \mu_2)^2
\end{aligned}$$

(ii) We have $\mathrm{P}(B = 1) = \frac{1}{2} \times \frac{1}{6} + \frac{1}{2} \times \frac{5}{6} = \frac{1}{2} = \mathrm{P}(B = 0)$. The values of Z_1 is the same as the number of B values which equal 1, which is the same as the number of "successes" in n trials.

Z_1 has mean $\frac{1}{2}n$ and variance $\frac{1}{2}\times\frac{1}{2}n = \frac{1}{4}n$. If n is "large enough" then Z_1 can be approximated with a normal distribution with the same mean and variance. The probability that Z_1 is within 10% of its mean is given by:

$$\begin{aligned}
&\mathrm{P}\left(\tfrac{1}{2}n - \tfrac{1}{20}n \leqslant Z_1 \leqslant \tfrac{1}{2}n + \tfrac{1}{20}n\right) \\
\approx &\mathrm{P}\left(\frac{\frac{1}{2}n - \frac{1}{20}n - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}} \leqslant \Phi \leqslant \frac{\frac{1}{2}n + \frac{1}{20}n - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}}\right) \qquad \text{where } \Phi \sim \mathrm{N}(0,1) \\
= &\mathrm{P}\left(-\frac{1}{10}\sqrt{n} \leqslant \Phi \leqslant \frac{1}{10}\sqrt{n}\right) \\
= &1 - 2\mathrm{P}\left(\Phi \geqslant \frac{1}{10}\sqrt{n}\right)
\end{aligned}$$

Since we have $\mathrm{P}\left(\Phi \geqslant \frac{1}{10}\sqrt{n}\right) \to 0$ as $n \to \infty$, the probability that Z_1 lies within 10% of its mean tends to 1 as $n \to \infty$.

(iii) The difference between this part and the previous part is that in part **(ii)** we are deciding whether to look at if a single dice shows a six or not each time for n repeats. In part **(iii)** we decide whether to count the number of sixes or the number of non-sixes for all n dice simultaneously.

Using the results from part **(i)** gives:

$$\begin{aligned}
\mathrm{E}(Z_2) &= p\mu_1 + q\mu_2 \\
&= \tfrac{1}{2} \times \tfrac{1}{6}n + \tfrac{1}{2} \times \tfrac{5}{6}n = \tfrac{1}{2}n \\
\mathrm{Var} &= p\sigma_1^2 + q\sigma_2^2 + pq(\mu_1 - \mu_2)^2 \\
&= \tfrac{1}{2} \times \tfrac{5}{36}n + \tfrac{1}{2} \times \tfrac{5}{36}n + \tfrac{1}{4}\left(\tfrac{1}{6}n - \tfrac{5}{6}n\right)^2 \\
&= \tfrac{5}{36}n + \tfrac{1}{4} \times \left(\tfrac{2}{3}n\right)^2 \\
&= \tfrac{5}{36}n + \tfrac{1}{9}n^2
\end{aligned}$$

Z_2 will take values either close to $\frac{1}{6}n$, or close to $\frac{5}{6}n$ so it will be bi-modal, and hence the normal distribution will not be a good approximation.

Let B_1 be the number of sixes that are shown when the n dice are rolled, and let B_2 be the number of non-sixes, so $\mathrm{E}(B_1) = \frac{1}{6}n$ and $\mathrm{E}(B_2) = \frac{5}{6}n$. In a similar way to part **(ii)** we can see that as $n \to \infty$, the probability that B_1 lies within 10% of its mean tends to 1, and similarly for B_2. This means that we have :

$$\begin{aligned}
&\mathrm{P}\left(\tfrac{9}{10} \times \tfrac{1}{6}n \leqslant B_1 \leqslant \tfrac{11}{10} \times \tfrac{1}{6}n\right) \to 1 \\
&\mathrm{P}\left(\tfrac{9}{60} \leqslant B_1 \leqslant \tfrac{11}{60}\right) \to 1
\end{aligned}$$

and

$$\begin{aligned}
&\mathrm{P}\left(\tfrac{9}{10} \times \tfrac{5}{6}n \leqslant B_2 \leqslant \tfrac{11}{10} \times \tfrac{5}{6}n\right) \to 1 \\
&\mathrm{P}\left(\tfrac{45}{60} \leqslant B_2 \leqslant \tfrac{55}{60}\right) \to 1
\end{aligned}$$

Then as Z_2 with either take the value of B_1 or B_2 then as $n \to \infty$, Z_2 will lie either in the range $\left[\frac{9}{60}n, \frac{11}{60}n\right]$ or $\left[\frac{45}{60}n, \frac{55}{60}n\right]$.

If Z_2 is to be within 10% of its mean then we need:

$$\begin{aligned}
&Z_2 \in [\tfrac{9}{10} \times \tfrac{1}{2}n, \tfrac{11}{10} \times \tfrac{1}{2}n] \\
&Z_2 \in [\tfrac{27}{60}n, \tfrac{33}{60}n]
\end{aligned}$$

There is no overlap between the range we want Z_2 to lie in and the ranges that B_1 and B_2 tend to as $n \to \infty$, hence the probability that Z_2 lies within 10% of its mean tends to 0 as $n \to \infty$.

It is perhaps surprising that two processes which on the surface seem very similar lead to very different results!

Question 12

12 Each of the independent random variables $X_1, X_2, \ldots, X_n$ has the probability density function $\mathrm{f}(x) = \frac{1}{2}\sin x$ for $0 \leqslant x \leqslant \pi$ (and zero otherwise). Let Y be the random variable whose value is the maximum of the values of $X_1, X_2, \ldots, X_n$.

(i) Explain why $\mathrm{P}(Y \leqslant t) = \big[\mathrm{P}(X_1 \leqslant t)\big]^n$ and hence, or otherwise, find the probability density function of Y.

Let $m(n)$ be the median of Y and $\mu(n)$ be the mean of Y.

(ii) Find an expression for $m(n)$ in terms of n. How does $m(n)$ change as n increases?

(iii) Show that

$$\mu(n) = \pi - \frac{1}{2^n}\int_0^\pi (1 - \cos x)^n \, \mathrm{d}x \,.$$

(a) Show that $\mu(n)$ increases with n.

(b) Show that $\mu(2) < m(2)$.

Examiner's report

This was the more popular of the two "Probability and Statistics" questions and a larger number of substantial attempts was seen.

Part **(i)** was generally completed well although in some cases there was insufficient explanation that "$Y \leqslant t$" is equivalent to "$X_i \leqslant t$ for all i".

Many candidates successfully calculated the value of $m(n)$ for part **(ii)**, but some only stated that $m(n)$ increases, rather than considering the value of the limit.

In part **(iii)** many candidates successfully showed the formula for $\mu(n)$. A number of candidates attempted to prove that $\mu(n)$ is increasing by differentiating with respect to n and showing that this is a positive quantity. However, none of these candidates were able to produce a fully correct version of this approach.

In part **(iii)** **(b)** most candidates were able to calculate $\mu(2)$ correctly, but then a number of errors were seen in the subsequent argument. Common errors were to fail to consider which choice of square root is appropriate and to omit to consider the effect of squaring on an inequality.

Solution

(i) If we are going to have $Y \leqslant t$, then we need all of the $X_i \leqslant t$. The X_i are independent and all have the same probability density function, hence we have:

$$\mathrm{P}(Y \leqslant t) = \big[\mathrm{P}(X_1 \leqslant t)\big]^n$$

Looking at X_1 we have:

$$\begin{aligned}\mathrm{P}(X_1 \leqslant t) &= \int_0^t \tfrac{1}{2}\sin x \,\mathrm{d}x \\ &= \big[-\tfrac{1}{2}\cos x\big]_0^t \\ &= \tfrac{1}{2}(1-\cos t)\end{aligned}$$

and so we have:

$$\mathrm{P}(Y \leqslant t) = \tfrac{1}{2^n}\big[1-\cos t\big]^n$$

Differentiating to find the pdf of Y gives:

$$f_Y(t) = \frac{n\sin t}{2^n}(1-\cos t)^{n-1}$$

(ii) To find the median we need:

$$\begin{aligned}\mathrm{P}(Y \leqslant m) &= \tfrac{1}{2} \\ \tfrac{1}{2^n}\big[1-\cos m\big]^n &= \tfrac{1}{2} \\ \big[1-\cos m\big]^n &= 2^{n-1} \\ 1-\cos m &= 2^{\frac{n-1}{n}} \\ \implies m(n) &= \cos^{-1}\left(1-2^{\frac{n-1}{n}}\right)\end{aligned}$$

As $n \to \infty$ we have $2^{\frac{n-1}{n}} \to 2$, and so $m(n) \to \cos^{-1}(-1) = \pi$. As n increases, $2^{1-\frac{1}{n}}$ increases and so $\cos^{-1}\left(1-2^{1-\frac{1}{n}}\right)$ increases towards π.

This Desmos Graph shows how the pdf of Y changes as n increases. This might help you visualise what happens to the median as n increases.

(iii) We have:

$$\mu(n) = \int_0^\pi x \times \frac{n\sin x}{2^n}(1-\cos x)^{n-1}\,\mathrm{d}x$$

Noting that $\dfrac{\mathrm{d}}{\mathrm{d}x}(1-\cos x)^n = n\sin x(1-\cos x)^{n-1}$ we have:

$$\begin{aligned}\mu(n) &= \left[\frac{x}{2^n}(1-\cos x)^n\right]_0^\pi - \int_0^\pi \frac{1}{2^n}(1-\cos x)^n\,\mathrm{d}x \\ &= \frac{\pi}{2^n}\times 2^n - 0 - \int_0^\pi \frac{1}{2^n}(1-\cos x)^n\,\mathrm{d}x \\ &= \pi - \int_0^\pi \frac{1}{2^n}(1-\cos x)^n\,\mathrm{d}x\end{aligned}$$

as required.

(a) We have:

$$\mu(n) = \pi - \int_0^\pi \left(\frac{1-\cos x}{2}\right)^n \, \mathrm{d}x$$

We have $0 < \dfrac{1-\cos x}{2} < 1$, and so as n increases we have $\left(\dfrac{1-\cos x}{2}\right)^n \to 0$ as $n \to \infty$. So $\mu(n) \to \pi$ as $n \to \infty$ (i.e. $\mu(n)$ increases as n increases).

(b) We have:

$$\begin{aligned}
\mu(2) &= \pi - \int_0^\pi \left(\frac{1-\cos x}{2}\right)^2 \, \mathrm{d}x \\
&= \pi - \tfrac{1}{4}\int_0^\pi 1 - 2\cos x + \cos^2 x \, \mathrm{d}x \\
&= \pi - \tfrac{1}{4}\int_0^\pi 1 - 2\cos x + \tfrac{1}{2}(1+\cos 2x) \, \mathrm{d}x \\
&= \pi - \tfrac{1}{4}\int_0^\pi \tfrac{3}{2} - 2\cos x + \tfrac{1}{2}\cos 2x \, \mathrm{d}x \\
&= \pi - \tfrac{1}{4}\left[\tfrac{3}{2}x - 2\sin x + \tfrac{1}{4}\sin 2x\right]_0^\pi \\
&= \pi - \tfrac{3}{8}\pi \\
&= \tfrac{5}{8}\pi
\end{aligned}$$

If $\mu(2) < m(2)$, then we must have $\mathrm{P}\big[Y \leqslant \mu(2)\big] < 0.5$ (i.e. μ is less than the median). We have:

$$\begin{aligned}
\mathrm{P}\big[Y \leqslant \mu(2)\big] &= \mathrm{P}\big[Y \leqslant \tfrac{5}{8}\pi\big] \\
&= \frac{1}{2^2}\big(1 - \cos\tfrac{5}{8}\pi\big)^2
\end{aligned}$$

Evaluating $\cos\frac{5}{8}\pi$:

$$\begin{aligned}
\cos\tfrac{5}{8}\pi &= \cos\left(\frac{\pi}{2} + \frac{\pi}{8}\right) \\
&= \cos\frac{\pi}{2}\cos\frac{\pi}{8} - \sin\frac{\pi}{2}\sin\frac{\pi}{8} \\
&= -\sin\frac{\pi}{8}
\end{aligned}$$

Then using $\cos 2A = 1 - 2\sin^2 A$ gives:

$$\begin{aligned}
\sin^2\tfrac{\pi}{8} &= \tfrac{1}{2}\big(1 - \cos\tfrac{\pi}{4}\big) \\
\sin^2\tfrac{\pi}{8} &= \tfrac{1}{2}\big(1 - \tfrac{\sqrt{2}}{2}\big) \\
\sin^2\tfrac{\pi}{8} &= \tfrac{2-\sqrt{2}}{4} \\
\implies \sin\tfrac{\pi}{8} &= \frac{\sqrt{2-\sqrt{2}}}{2}
\end{aligned}$$

Noting that since $\frac{\pi}{8} < \pi$, $\sin\frac{\pi}{8} > 0$.

This means that we have $\cos\frac{5\pi}{8} = -\sin\frac{\pi}{8} = -\frac{\sqrt{2-\sqrt{2}}}{2}$, and so:

$$\begin{aligned}
\mathrm{P}\big[Y \leqslant \mu(2)\big] &= \frac{1}{2^2}\big(1 - \cos\tfrac{5}{8}\pi\big)^2 \\
&= \frac{1}{4}\left(1 + \frac{\sqrt{2-\sqrt{2}}}{2}\right)^2 \\
&= \frac{1}{16}\left(2 + \sqrt{2-\sqrt{2}}\right)^2
\end{aligned}$$

We want $\mathrm{P}\big[Y \leqslant \mu(2)\big] < 0.5$:

$$\begin{aligned}
& & \frac{1}{16}\left(2 + \sqrt{2-\sqrt{2}}\right)^2 &< \frac{1}{2} \\
&\iff & \left(2 + \sqrt{2-\sqrt{2}}\right)^2 &< 8 \\
&\iff & 2 + \sqrt{2-\sqrt{2}} &< 2\sqrt{2} \\
&\iff & \sqrt{2-\sqrt{2}} &< 2\sqrt{2} - 2 \\
&\iff & 2 - \sqrt{2} &< (2\sqrt{2} - 2)^2 \\
&\iff & 2 - \sqrt{2} &< (\sqrt{2})^2(2 - \sqrt{2})^2 \\
&\iff & 1 &< 2(2 - \sqrt{2}) \\
&\iff & 2\sqrt{2} &< 3 \\
&\iff & 8 &< 9
\end{aligned}$$

Note that all of these steps are valid and reversible as we have $2 - \sqrt{2} > 0$.

Therefore we have $\mathrm{P}\big[Y \leqslant \mu(2)\big] < 0.5$ and so $\mu(2) < m(2)$.