UNIVERSITY OF CAMBRIDGE

maths.org/step

STEP Support Programme

2019 STEP 2

General comments - In the examiners' report for this paper it was noted that the pure questions were much more popular than the applied ones (7 out of the 8 pure questions were attempted by more than half of candidates, whereas only 1 of the 4 applied questions was attempted by more than a quarter). It was also noted that in some cases solutions reached the correct results, but without sufficient justification of all of the steps. Take particular care if the question says "show that" or "prove that". Another aspect that came up in several questions is candidates being confused with the direction of implications (so showing that $P \implies Q$ rather than $Q \implies P$ or not showing that an implication works in both directions for an "if and only if"). If a question says "show that P is true if Q is true" then it is asking you to show that $Q \implies P$.

There are often many ways to approach a STEP question. Your methods may be different to the ones shown here but correct maths done correctly (and explained fully, especially in the case of a "show that") always gets the marks.

The full examiners report and mark-schemes for this paper can be found on the Cambridge Assessment Admissions Testing website

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

Question 1

1 Let $f(x) = (x - p)g(x)$, where g is a polynomial. Show that the tangent to the curve $y = f(x)$ at the point with $x = a$, where $a \neq p$, passes through the point $(p, 0)$ if and only if $g'(a) = 0$.

The curve C has equation

$$y = A(x - p)(x - q)(x - r),$$

where p, q and r are constants with $p < q < r$, and A is a non-zero constant.

(i) The tangent to C at the point with $x = a$, where $a \neq p$, passes through the point $(p, 0)$. Show that $2a = q + r$ and find an expression for the gradient of this tangent in terms of A, q and r.

(ii) The tangent to C at the point with $x = c$, where $c \neq r$, passes through the point $(r, 0)$. Show that this tangent is parallel to the tangent in part (i) if and only if the tangent to C at the point with $x = q$ does not meet the curve again.

Examiner's report

This was the question answered by the largest proportion of candidates and many good solutions were seen. However, many candidates did not appreciate the importance of the phrase *if and only if* in parts of this question. As a result a large number of attempts failed to achieve full marks as it was not made clear that the reasoning presented also worked in the opposite direction.

Having shown the first result, many candidates were able to identify the appropriate choice of $g(x)$ when attempting part **(i)** and successfully showed that $2a = q + r$. Many were also able to find a correct expression for the gradient, although some did not find this expression in terms of the variables requested.

In part **(ii)** a pleasing number of candidates were able to recognise that the results from part **(i)** would be relevant here as well. Again, some of the solutions to this part failed to recognise that the question required the result to be shown in both directions.

Solution

Notice that this question has a "stem" followed by two parts. I would expect to use the result we are asked to show in the stem in one (or both) of the following parts. Curve C is defined in the stem so when C is referred to in parts **(i)** and **(ii)** it is this definition we need to look back to.

Most of this question concerns a cubic graph with three distinct roots, and some of the tangents that can be drawn to that curve. You might like to sketch a cubic to help you see what is going on.

Stem - 6 marks

There were quite a few conditions given in the stem, such as $p < q < r$ and $a \neq p$. Its a good idea to keep these in mind as you work through the question — STEP questions do not usually give you unnecessary conditions.

Let's start by finding the equation of the tangent to $\mathrm{f}(x)$. Differentiating gives:

$$\mathrm{f}'(x) = \mathrm{g}(x) + (x - p)\mathrm{g}'(x) \implies \mathrm{f}'(a) = \mathrm{g}(a) + (a - p)\mathrm{g}'(a)$$

The equation of the tangent is therefore:

$$\begin{aligned} y - \mathrm{f}(a) &= \Big[\mathrm{g}(a) + (a-p)\mathrm{g}'(a)\Big](x - a) \\ y - (a-p)\mathrm{g}(a) &= \Big[\mathrm{g}(a) + (a-p)\mathrm{g}'(a)\Big](x - a) \end{aligned}$$

Now consider whether $(p, 0) \implies \mathrm{g}'(a) = 0$ (the *only if* part of the statement). Start by substituting $x = p, y = 0$ into the equation of the tangent.

$$\begin{aligned} 0 - (a-p)\mathrm{g}(a) &= \Big[\mathrm{g}(a) + (a-p)\mathrm{g}'(a)\Big](p - a) \\ \cancel{(p-a)\mathrm{g}(a)} &= \cancel{\mathrm{g}(a)(p-a)} + (a-p)(p-a)\mathrm{g}'(a) \\ (a-p)^2\mathrm{g}'(a) &= 0 \\ \implies \mathrm{g}'(a) &= 0 \qquad \text{(since } a \neq p\text{)} \end{aligned}$$

Note that we had to use the condition $a \neq p$ here in order to divide by $(a-p)^2$ and conclude that $\mathrm{g}'(a) = 0$.

Now we need to consider whether $\mathrm{g}'(a) = 0 \implies$ tangent passes through $(p, 0)$. The equation of the tangent is:

$$\begin{aligned} & y - (a-p)\mathrm{g}(a) = \Big[\mathrm{g}(a) + (a-p)\mathrm{g}'(a)\Big](x - a) \\ \mathrm{g}'(a) = 0 \implies & y - (a-p)\mathrm{g}(a) = \mathrm{g}(a)(x - a) \\ \text{substitute } x = p \implies & y - (a-p)\mathrm{g}(a) = \mathrm{g}(a)(p - a) \\ & y = \mathrm{g}(a)(p-a) + (a-p)\mathrm{g}(a) \\ & y = 0 \end{aligned}$$

Therefore we have that the tangent passes through the point $(p, 0) \iff \mathrm{g}'(a) = 0$.

Note that I was careful to show that the implication works in both directions.

Part (i) - 4 marks

Start by looking at the links to the stem. We have a curve C which is $y = A(x-p)(x-q)(x-r)$ which can be written as $y = (x-p)\mathrm{g}(x)$ where $\mathrm{g}(x) = A(x-q)(x-r)$. We are told that the tangent to the point with $x = a$, where $a \neq p$, passes through the point $(p, 0)$ and so by the stem result we know that we have $\mathrm{g}'(a) = 0$. Therefore:

$$\begin{aligned} \mathrm{g}(x) &= A(x-q)(x-r) \\ \mathrm{g}'(x) &= A(x-r) + A(x-q) \\ \mathrm{g}'(a) &= A(a-r) + A(a-q) = 0 \end{aligned}$$

So we have $A[(a-r)+(a-q)] = 0 \implies 2a = q + r$ (note that $A \neq 0$).

Be careful with A and a, they represent different quantities and in general $A \neq a$.

We are then asked to find the expression for the gradient of the tangent in terms of A, q and r.

The gradient of the tangent is:

$$\begin{aligned} \mathrm{g}(a) + (a-p)\mathrm{g}'(a) &= A(a-q)(a-r) + 0 \\ &= A\left(\tfrac{1}{2}(q+r) - q\right)\left(\tfrac{1}{2}(q+r) - r\right) \\ &= \tfrac{1}{4}A(r-q)(q-r) \\ &= -\tfrac{1}{4}A(r-q)^2 \end{aligned}$$

Note that I had to substitute for a, using $2a = p + r$, as the question asked for the gradient to be in terms of A, r and q.

Part (ii) - 10 marks

In this part we are asked to consider another tangent to $y = A(x-p)(x-q)(x-r)$, but this time it is the tangent to $x = c$ (rather than $x = a$) and it passes through $(r, 0)$ (rather than $(p, 0)$). By substituting c for a, and swapping p and r we can see that $2c = p + q$ and the gradient of this tangent is $-\frac{1}{4}A(p-q)^2$.

We are being asked to show that the tangents in part **(i)** and **(ii)** are parallel if and only if a third tangent through $x = q$ does not intersect the curve again.

The tangents are parallel if and only if:

$$\begin{aligned} -\tfrac{1}{4}A(p-q)^2 &= -\frac{1}{4}A(r-q)^2 \\ (p-q)^2 &= (r-q)^2 \\ \iff q-p &= r-q \qquad \text{since we are given } p < q < r \end{aligned}$$

Curve C has equation $y = A(x-p)(x-q)(x-r)$, and so passes through the point $(q, 0)$. The gradient of the line is $\dfrac{\mathrm{d}y}{\mathrm{d}x} = A\big((x-q)(x-r) + (x-p)(x-r) + (x-p)(x-q)\big)$. This can be found via the product rule. When $x = q$ the gradient of the curve is $A(q-p)(q-r)$.

The equation of the tangent when $x = q$ is therefore:

$$y - 0 = A(q-p)(q-r)(x-q)$$

Where this tangent meets the curve again we must have:

$$\begin{aligned} A(x-p)(x-q)(x-r) &= A(q-p)(q-r)(x-q) && (*) \\ (x-p)(x-r) &= (q-p)(q-r) && \text{since we must have } x \neq q \end{aligned}$$

We have $x \neq q$ as we are looking for another point where the tangent meets to curve, i.e. a point where x isn't q.

Expanding brackets gives:

$$\begin{aligned} x^2 - px - rx + \cancel{pr} &= q^2 - pq - rq + \cancel{pr} \\ x^2 - px - rx + pq + rq - q^2 &= 0 \\ (x-q)(x-p-r+q) &= 0 \\ \implies x = q &\text{ or } x = p + r - q \end{aligned}$$

It is reasonably obvious that $x = q$ is a root of $x^2 - px - rx + pq + rq - q^2 = 0$ (if you substitute $x = q$ in you can see that this will give 0). Alternatively, you would expect $x = q$ to be a double root of equation $(*)$ as the line is a tangent to the curve at $x = q$. Either way, if $x = q$ is a root then $(x - q)$ will be a factor, and you can use inspection to find the other factor.

Since $x \neq q$ then we must have $x = p + r - q$ for the other point where the tangent meets the curve C. The tangent only meets the curve at $x = q$ if and only if $p + r - q = q$.

So we now have that:

- the tangents are parallel if and only if $q - p = r - q$
- the third tangent only meets the curve at $x = q$ if and only if $p + r - q = q \iff r - q = q - p$

These are the same condition so we have that the first two tangents are parallel if and only the third tangent only meets the curve at $x = q$.

Question 2

2 The function f satisfies $\mathrm{f}(0) = 0$ and $\mathrm{f}'(t) > 0$ for $t > 0$. Show by means of a sketch that, for $x > 0$,

$$\int_0^x \mathrm{f}(t)\,\mathrm{d}t + \int_0^{\mathrm{f}(x)} \mathrm{f}^{-1}(y)\,\mathrm{d}y = x\mathrm{f}(x).$$

(i) The (real) function g is defined, for all t, by

$$\big(\mathrm{g}(t)\big)^3 + \mathrm{g}(t) = t.$$

Prove that $\mathrm{g}(0) = 0$, and that $\mathrm{g}'(t) > 0$ for all t.

Evaluate $\displaystyle\int_0^2 \mathrm{g}(t)\,\mathrm{d}t$.

(ii) The (real) function h is defined, for all t, by

$$\big(\mathrm{h}(t)\big)^3 + \mathrm{h}(t) = t + 2.$$

Evaluate $\displaystyle\int_0^8 \mathrm{h}(t)\,\mathrm{d}t$.

Examiner's report

This question was another popular question that was generally well answered, achieving the second best average mark of all of the questions and was also the question for which the largest number of solutions received full marks. Most candidates drew a convincing sketch to demonstrate that the two integrals make a rectangle. Arguments from sketches showing the inverse function and reflective symmetry were less successful and often candidates' diagrams assumed x to be a fixed point of $\mathrm{f}(t)$.

By far the most common mistake in the first part was to notice the solution $\mathrm{g}(2) = 1$ but not to factorise and use the quadratic discriminant to show that no other solutions were possible. The conceptually difficult part was to use $\mathrm{g}^{-1}(y) = y^3 + y$, and many candidates stopped just before this point.

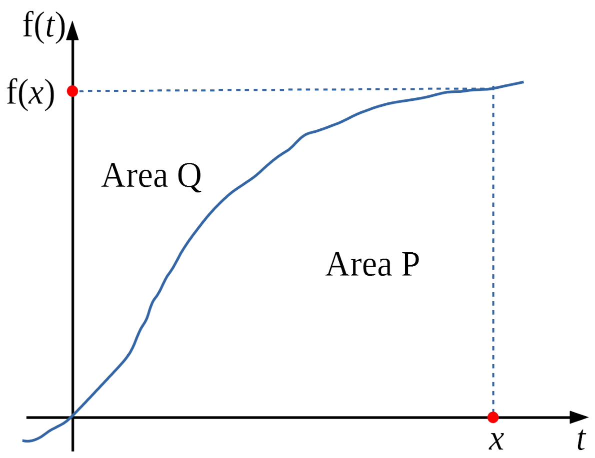
In the final part, many candidates tried to apply the stem identity in its original form, without noticing that $\mathrm{h}(0) \neq 0$. This was the most difficult part, and those who modified it correctly generally did well. Candidates sometimes failed to check that $\mathrm{h}'(t) > 0$, but this was not necessary for those who used $\mathrm{h}(t) = \mathrm{g}(t + 2)$.

Solution

This is another "stem" question. Remember that even if you cannot show the "stem" result you can still answer the other parts of the question. There are not many marks allocated to the stem in this case, so don't write off the whole question if you are unsure how to approach the stem.

Stem - 2 marks

The question tells you to use a sketch, so you need to draw a sketch!



Note that t is the parameter on the horizontal axis (rather than x), and that $\mathrm{f}'(t) > 0$ so the graph is always heading "upwards". Area P is equal to $\displaystyle\int_0^x \mathrm{f}(t)\,\mathrm{d}t$. Area Q is the area to the y-axis[1]. This is equal to $\displaystyle\int_0^{\mathrm{f}(x)} t\,\mathrm{d}y = \int_0^{\mathrm{f}(x)} \mathrm{f}^{-1}(y)\,\mathrm{d}y$, remembering that $y = \mathrm{f}(t) \implies t = \mathrm{f}^{-1}(y)$.
Since Area P + Area Q $= x \times \mathrm{f}(x)$ we have:

$$\int_0^x \mathrm{f}(t)\,\mathrm{d}t + \int_0^{\mathrm{f}(x)} \mathrm{f}^{-1}(y)\,\mathrm{d}y = x\mathrm{f}(x)$$

as required.

Part (i) - 9 marks

Substituting in $t = 0$ gives:

$$\begin{aligned}
\big(\mathrm{g}(0)\big)^3 + \mathrm{g}(0) &= 0 \\
\mathrm{g}(0)\big((\mathrm{g}(0))^2 + 1\big) &= 0 \\
&\implies \mathrm{g}(0) = 0 \text{ or } (\mathrm{g}(0))^2 + 1 = 0 \\
&\implies \mathrm{g}(0) = 0 \text{ as } (\mathrm{g}(0))^2 + 1 > 0
\end{aligned}$$

[1]Area to the y axis is similar to the area to the x axis. You can consider the limit of lots of small rectangles of height δy and length $x (= \mathrm{f}^{-1}(y))$ to get $\int x\,\mathrm{d}y$.

Differentiating with respect to t gives:

$$
\begin{aligned}
3\big(\mathrm{g}(t)\big)^2 \times \mathrm{g}'(t) + \mathrm{g}'(t) &= 1 \\
\mathrm{g}'(t)\left[3\left(\mathrm{g}(t)\right)^2 + 1\right] &= 1 \\
\mathrm{g}'(t) &= \frac{1}{3\big(\mathrm{g}(t)\big)^2 + 1}
\end{aligned}
$$

As $3\big(\mathrm{g}(t)\big)^2 + 1 > 0$, we have $\mathrm{g}'(t) > 0$ for all t.

Function g now satisfies the requirements of the stem and so we have:

$$\int_0^x \mathrm{g}(t)\,\mathrm{d}t + \int_0^{\mathrm{g}(x)} \mathrm{g}^{-1}(y)\,\mathrm{d}y = x\mathrm{g}(x)$$

and so by substituting $x = 2$ we have:

$$\int_0^2 \mathrm{g}(t)\,\mathrm{d}t + \int_0^{\mathrm{g}(2)} \mathrm{g}^{-1}(y)\,\mathrm{d}y = 2\mathrm{g}(2)$$

There are two things we don't know here, what $\mathrm{g}(2)$ is and what $\mathrm{g}^{-1}(t)$ is. Starting with the definition $\big(\mathrm{g}(t)\big)^3 + \mathrm{g}(t) = t$ we have:

$$
\begin{aligned}
\big(\mathrm{g}(2)\big)^3 + \mathrm{g}(2) &= 2 \\
\big(\mathrm{g}(2)\big)^3 + \mathrm{g}(2) - 2 &= 0 \\
\big(\mathrm{g}(2) - 1\big)\left[\big(\mathrm{g}(2)\big)^2 + \mathrm{g}(2) + 2\right] &= 0
\end{aligned}
$$

and so $\mathrm{g}(2) = 1$ or $\big(\mathrm{g}(2)\big)^2 + \mathrm{g}(2) + 2 = 0$, which has no real solutions as the discriminant is $1 - 4 \times 2 < 0$ and so we must have $\mathrm{g}(2) = 1$.

For the inverse function let $\mathrm{g}^{-1}(y) = t$, and so we have $y = \mathrm{g}(t)$. Substituting these into the definition of $\mathrm{g}(t)$ gives:

$$y^3 + y = \mathrm{g}^{-1}(y)$$

The stem relationship now becomes:

$$
\begin{aligned}
\int_0^2 \mathrm{g}(t)\,\mathrm{d}t + \int_0^1 y^3 + y\,\mathrm{d}y &= 2 \times 1 \\
\int_0^2 \mathrm{g}(t)\,\mathrm{d}t &= 2 - \left[\tfrac{1}{4}y^4 + \tfrac{1}{2}y^2\right]_0^1 \\
&= 2 - \frac{3}{4} \\
&= \frac{5}{4}
\end{aligned}
$$

Part (ii) - 9 marks

The first thing to notice is that the function $\mathrm{h}(x)$ does not satisfy $\mathrm{h}(0) = 0$ so we cannot apply the stem result to $\mathrm{h}(t)$. However there is a relationship between $\mathrm{h}(t)$ and $\mathrm{g}(t)$ i.e. $\mathrm{h}(t) = \mathrm{g}(t+2)$. Since we have $\mathrm{g}'(t) > 0$ we also have $\mathrm{h}'(t) > 0$.

Drawing a sketch might also be a good idea. In this case $\mathrm{h}(0) \neq 0$. We have:

$$\begin{aligned}\big(\mathrm{h}(0)\big)^3 + \mathrm{h}(0) &= 2 \\ \mathrm{h}^3 + \mathrm{h} - 2 &= 0 \qquad \text{using } \mathrm{h}(0) = \mathrm{h} \\ (\mathrm{h}-1)(\mathrm{h}^2 + \mathrm{h} + 2) &= 0\end{aligned}$$

and so $\mathrm{h}(0) = 1$ ($\mathrm{h}^2 + \mathrm{h} + 2 = 0$ has no real solutions).

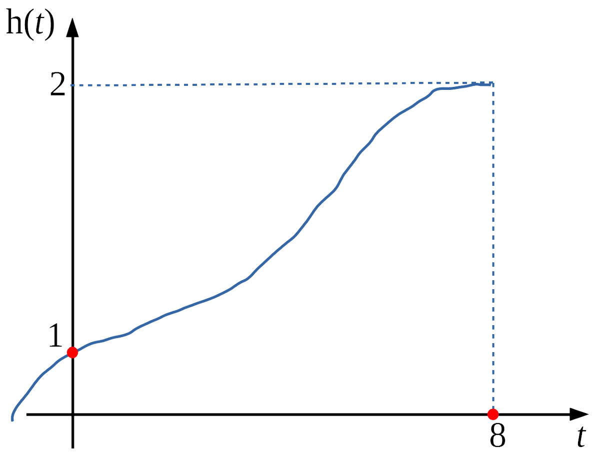
Having read that back, I realise that I could have just written down $\mathrm{h}(0) = 1$, as we have $\mathrm{h}(0) = \mathrm{g}(2)$, which we showed was equal to 1 in the previous part.

It would also be helpful to know what $\mathrm{h}(8)$ is. We have:

$$\begin{aligned}\big(\mathrm{h}(8)\big)^3 + \mathrm{h}(8) &= 10 \\ \mathrm{h}^3 + \mathrm{h} - 10 &= 0 \qquad \text{using } \mathrm{h}(8) = \mathrm{h} \\ (\mathrm{h}-2)(\mathrm{h}^2 + 2\mathrm{h} + 5) &= 0\end{aligned}$$

and so $\mathrm{h}(8) = 2$.

A sketch might look like:



Note that $\mathrm{h}'(t) > 0$ so this curve also keeps heading "upwards".

We now have:

$$\int_0^8 \mathrm{h}(t)\,\mathrm{d}t + \int_1^2 \mathrm{h}^{-1}(y)\,\mathrm{d}y = 16$$

Using a substitution of $y = \mathrm{h}(t) \implies t = \mathrm{h}^{-1}(y)$ in $\big(\mathrm{h}(t)\big)^3 + \mathrm{h}(t) = t + 2$ gives $\mathrm{h}^{-1}(y) = y^3 + y - 2$.

The required integral is now given by:

$$\begin{aligned}\int_0^8 \mathrm{h}(t)\,\mathrm{d}t &= 16 - \int_1^2 \left(y^3 + y - 2\right)\,\mathrm{d}y \\ &= 16 - \left[\tfrac{1}{4}y^4 + \tfrac{1}{2}y^2 - 2y\right]_1^2 \\ &= 16 - \left(\tfrac{1}{4}\times 16 + \tfrac{1}{2}\times 4 - 4\right) + \left(\tfrac{1}{4} + \tfrac{1}{2} - 2\right) \\ &= 16 - 2 - 1\tfrac{1}{4} \\ &= 12\tfrac{3}{4}\end{aligned}$$

Question 3

3 For any two real numbers x_1 and x_2, show that

$$|x_1 + x_2| \leqslant |x_1| + |x_2|.$$

Show further that, for any real numbers x_1, x_2, $\ldots$, x_n,

$$|x_1 + x_2 + \cdots + x_n| \leqslant |x_1| + |x_2| + \cdots + |x_n|.$$

(i) The polynomial f is defined by

$$\mathrm{f}(x) = 1 + a_1 x + a_2 x^2 + \cdots + a_{n-1}x^{n-1} + x^n$$

where the coefficients are real and satisfy $|a_i| \leqslant A$ for $i = 1, 2, \ldots, n-1$, where $A \geqslant 1$.

(a) If $|x| < 1$, show that

$$|\mathrm{f}(x) - 1| \leqslant \frac{A|x|}{1 - |x|}.$$

(b) Let ω be a real root of f, so that $\mathrm{f}(\omega) = 0$. In the case $|\omega| < 1$, show that

$$\frac{1}{1+A} \leqslant |\omega| \leqslant 1 + A. \qquad (*)$$

(c) Show further that the inequalities $(*)$ also hold if $|\omega| \geqslant 1$.

(ii) Find the integer root or roots of the quintic equation

$$135x^5 - 135x^4 - 100x^3 - 91x^2 - 126x + 135 = 0.$$

Examiner's report

While this was a popular question it was also the one where the average mark achieved by candidates was the lowest. In this question many of the results to be reached were given in the question. Students therefore need to recognise that it is necessary for solutions to be presented very clearly, and it is for this reason that many solutions in the first parts did not achieve full marks. For example, justifications of the generalised result for a set of n real numbers expressed in the form of an inductive proof were the most successful. For most candidates the majority of marks were scored in the sections up to and including part (i)(b). Many candidates were then unable to see how to work in the cases where $|x| \geqslant 1$ for part (i)(c). In the final part, candidates were often unable to put the equation into the form that had been used in the earlier parts of the questions and therefore did not manage to reduce the possible values of the integer roots to a sufficiently small set.

Solution

As the Examiner's report points out, lots of the required results are given in the question so you must show enough reasoning to fully justify your results. You can include words (and sometimes diagrams) to support your reasoning!

The stem - 2 marks

If x_1 and x_2 have the same sign (i.e. are both positive or are both negative) then $|x_1+x_2|$ will reach its maximum and we have $|x_1+x_2| = |x_1|+|x_2|$. If x_1 and x_2 have different signs then we will have $|x_1+x_2| = |x_1| - |x_2|$ if $|x_1| > |x_2|$ and $|x_2| - |x_1|$ otherwise. Hence we have $|x_1+x_2| \leqslant |x_1|+|x_2|$.

For the second part, use proof by induction. Trivially we have $|x_1| = |x_1|$, so the statement is true when $n=1$. We know that $|x_1+x_2| \leqslant |x_1|+|x_2|$ and so the statement is true when $n=2$. Assume that the statement hold when $n=k$, so we have:

$$|x_1 + x_2 + \cdots + x_k| \leqslant |x_1| + |x_2| + \cdots + |x_k|$$

Now consider the $n = k+1$ case:

$$\begin{aligned} |x_1 + x_2 + \cdots + x_{k+1}| &\leqslant |x_1 + x_2 + \cdots + x_k| + |x_{k+1}| && \text{Using the 2 case result} \\ &\leqslant |x_1| + |x_2| + \cdots + |x_k| + |x_{k+1}| && \text{Using the } n=k \text{ result} \end{aligned}$$

Hence we have $|x_1 + x_2 + \cdots + x_n| \leqslant |x_1| + |x_2| + \cdots + |x_n|$ for all $n \geqslant 1$.

Part (i)(a) - 6 marks We have:

$$\begin{aligned} |\mathrm{f}(x) - 1| &= |a_1 x + a_2 x^2 + \cdots + a_{n-1}x^{n-1} + x^n| \\ &\leqslant |a_1 x| + |a_2 x^2| + \cdots + |a_{n-1}x^{n-1}| + |x^n| \\ &= |a_1||x| + |a_2||x|^2 + \cdots + |a_{n-1}||x|^{n-1} + |x|^n \\ &\leqslant A|x| + A|x|^2 + \cdots + A|x|^{n-1} + |x|^n \qquad \text{as } |a_i| \leqslant A \\ &\leqslant A(|x| + |x|^2 + \cdots + |x|^{n-1} + |x|^n) \qquad \text{as } 1 \leqslant A \\ &= A|x|\Big(1 + |x| + \cdots + |x|^{n-1}\Big) \\ &= A|x| \times \frac{1-|x|^n}{1-|x|} \\ &\leqslant \frac{A|x|}{1-|x|} \qquad \text{as } 1 - |x|^n \leqslant 1 \end{aligned}$$

Part (i)(b) - 3 marks Since $\mathrm{f}(\omega) = 0$ then result from part (i) (a) gives us:

$$\begin{aligned} |-1| &\leqslant \frac{A|\omega|}{1-|\omega|} \\ \implies 1 &\leqslant \frac{A|\omega|}{1-|\omega|} \end{aligned}$$

Since $|\omega| < 1$ we have $1 - |\omega| > 0$, and so we can multiply throughout by $1 - |\omega|$[2] to get:

$$\begin{aligned} 1 - |\omega| &\leqslant A|\omega| \\ 1 &\leqslant (A+1)|\omega| \\ \frac{1}{A+1} &\leqslant |\omega| \end{aligned}$$

We also have $|\omega| < 1$ and $A \geqslant 1$, and so $|\omega| \leqslant 1 + A$ giving:

$$\frac{1}{1+A} \leqslant |\omega| \leqslant 1 + A$$

as required.

Part (i)(c) - 3 marks

The phrase "show further" suggests some linking to the previous part. Note that if $|\omega| > 1$ then $\left|\frac{1}{\omega}\right| < 1$ which might be useful.

We have:

$$\begin{aligned} \mathrm{f}(\omega) &= 1 + a_1\omega + a_2\omega^2 + \cdots + a_{n-1}\omega^{n-1} + \omega^n = 0 \\ \implies \frac{\mathrm{f}(\omega)}{\omega^n} &= \left(\frac{1}{\omega}\right)^n + a_1\left(\frac{1}{\omega}\right)^{n-1} + \cdots + a_{n-1}\left(\frac{1}{\omega}\right) + 1 = 0 \end{aligned}$$

We therefore have a polynomial $\mathrm{g}\,(x)$ with $\mathrm{g}\left(\frac{1}{\omega}\right) = 0$. The coefficients are all less than or equal to A, and if $|\omega| > 1$ then we have $\left|\dfrac{1}{\omega}\right| < 1$. we therefore have:

$$\begin{aligned} \frac{1}{1+A} &\leqslant \left|\frac{1}{\omega}\right| \leqslant 1 + A \\ \frac{1}{1+A} &\leqslant \frac{1}{|\omega|} \leqslant 1 + A \\ 1 + A &\geqslant |\omega| \geqslant \frac{1}{1+A} \end{aligned}$$

For the last line note that all the terms are positive and if two positive numbers a and b satisfy $a < b$ then $\frac{1}{a} > \frac{1}{b}$.

We also need to consider the case when $|\omega| = 1$. In this case we must have $\dfrac{1}{1+A} \leqslant |\omega| \leqslant 1 + A$ as the left hand side is less than 1 and the right hand side is greater than 1 (remember that $A \geqslant 1$).

[2]Always check that terms cannot be negative before you multiply or divide both sides of an inequality by them.

Part (ii) - 6 marks

If would be good if the polynomial here was similar to $\mathrm{f}(x)$ considered in part (i). If we divide throughout by 135 we get:

$$\mathrm{f}(x) = x^5 - x^4 - \frac{100}{135}x^3 - \frac{91}{135}x^2 - \frac{126}{135}x + 1$$

This now has the same form as $\mathrm{f}(x)$ and all of the coefficients are less than or equal to 1.

Take $A = 1$, then the roots satisfy:

$$\frac{1}{2} \leqslant |\omega| \leqslant 2$$

Therefore any integer roots of $\mathrm{f}(x)$ must satisfy $\omega = \pm 1, \pm 2$.

If we take $\omega = \pm 2$, and consider the x^3, x^2 and x terms then these will have the form:

$$\frac{\text{even}}{135} + \frac{\text{even}}{135} + \frac{\text{even}}{135} = \frac{\text{even}}{135}$$

and so we will have a non-integer fraction for these three terms, and hence $\mathrm{f}(\pm 2) \neq 0$.

$\mathrm{f}(1) = 1 - 1 - \dfrac{317}{135} + 1 \neq 0$ so 1 is not a root.

$\mathrm{f}(-1) = -1 - 1 + \frac{100}{135} - \frac{91}{135} + \frac{126}{135} + 1 = -2 + \frac{135}{135} + 1 = 0$, and so -1 is the only integer root of $\mathrm{f}(x)$.

Question 4

4 You are not required to consider issues of convergence in this question.

For any sequence of numbers $a_1, a_2, \ldots, a_m, \ldots, a_n$, the notation $\prod_{i=m}^{n} a_i$ denotes the product $a_m a_{m+1} \cdots a_n$.

(i) Use the identity $2\cos x \sin x = \sin(2x)$ to evaluate the product $\cos(\frac{\pi}{9})\cos(\frac{2\pi}{9})\cos(\frac{4\pi}{9})$.

(ii) Simplify the expression

$$\prod_{k=0}^{n} \cos\left(\frac{x}{2^k}\right) \qquad (0 < x < \tfrac{1}{2}\pi).$$

Using differentiation, or otherwise, show that, for $0 < x < \frac{1}{2}\pi$,

$$\sum_{k=0}^{n} \frac{1}{2^k} \tan\left(\frac{x}{2^k}\right) = \frac{1}{2^n} \cot\left(\frac{x}{2^n}\right) - 2\cot(2x).$$

(iii) Using the results $\lim_{\theta \to 0} \frac{\sin\theta}{\theta} = 1$ and $\lim_{\theta \to 0} \frac{\tan\theta}{\theta} = 1$, show that

$$\prod_{k=1}^{\infty} \cos\left(\frac{x}{2^k}\right) = \frac{\sin x}{x}$$

and evaluate

$$\sum_{j=2}^{\infty} \frac{1}{2^{j-2}} \tan\left(\frac{\pi}{2^j}\right).$$

Examiner's report

This was a well-answered question, but also one in which a fairly large number of solutions scored very low marks. The majority of candidates were able to evaluate the first product using the identify provided and most were then able to apply the same technique to simplify the first expression in part (ii). Many students then differentiated, but some then struggled to manage the notation correctly to reach the second result requested in part (ii).

Part (iii) required some care to ensure that the sums and products were over the correct range, but those who managed to adjust correctly for this were then able to reach the required results.

Solution

The question starts with a note: "You are not required to consider issues of convergence in this question". This means that you don't have to formally show that the products or sums converge, it is intended to be helpful rather than make you start worrying about what it means! This is the first question on this paper for which there is no request in the stem, just some information for the question.

Part (i) - 4 marks

We are told to use the given identity, so it might be a good idea to start by multiplying the given expression by $\sin\left(\frac{\pi}{9}\right)$ so that we have something of the form $\cos x \sin x$ on the left hand side.

$$\begin{aligned}
\sin\left(\tfrac{\pi}{9}\right)\cos\left(\tfrac{\pi}{9}\right)\cos\left(\tfrac{2\pi}{9}\right)\cos\left(\tfrac{4\pi}{9}\right) &= \tfrac{1}{2}\sin\left(\tfrac{2\pi}{9}\right)\cos\left(\tfrac{2\pi}{9}\right)\cos\left(\tfrac{4\pi}{9}\right)\\
&= \tfrac{1}{2}\times\tfrac{1}{2}\sin\left(\tfrac{4\pi}{9}\right)\cos\left(\tfrac{4\pi}{9}\right)\\
&= \tfrac{1}{2}\times\tfrac{1}{2}\times\tfrac{1}{2}\sin\left(\tfrac{8\pi}{9}\right)\\
&= \tfrac{1}{8}\sin\left(\pi-\tfrac{8\pi}{9}\right)\ \text{ using }\ \sin\alpha=\sin(\pi-\alpha)
\end{aligned}$$

Hence we have $\sin\left(\frac{\pi}{9}\right)\cos\left(\frac{\pi}{9}\right)\cos\left(\frac{2\pi}{9}\right)\cos\left(\frac{4\pi}{9}\right) = \frac{1}{8}\sin\left(\frac{\pi}{9}\right)$ and so $\cos\left(\frac{\pi}{9}\right)\cos\left(\frac{2\pi}{9}\right)\cos\left(\frac{4\pi}{9}\right) = \frac{1}{8}$. ($\sin\left(\frac{\pi}{9}\right)\neq 0$ so we can divide by it.)

Part (ii) - 7 marks

It might be helpful to expand the product to see more clearly what is being considered.[3]

$$\begin{aligned}
\prod_{k=0}^{n}\cos\left(\frac{x}{2^k}\right) &= \cos\left(\frac{x}{2^0}\right)\times\cos\left(\frac{x}{2^1}\right)\times\cos\left(\frac{x}{2^2}\right)\times\cdots\times\cos\left(\frac{x}{2^{n-1}}\right)\times\cos\left(\frac{x}{2^n}\right)\\
&= \cos\left(\frac{x}{1}\right)\times\cos\left(\frac{x}{2}\right)\times\cos\left(\frac{x}{4}\right)\times\cdots\times\cos\left(\frac{x}{2^{n-1}}\right)\times\cos\left(\frac{x}{2^n}\right)
\end{aligned}$$

By comparing to part (i), it looks like multiplying by $\sin\left(\dfrac{x}{2^n}\right)$ might be useful:

$$\begin{aligned}
\sin\left(\frac{x}{2^n}\right)\times\prod_{k=0}^{n}\cos\left(\frac{x}{2^k}\right) &= \cos\left(\frac{x}{1}\right)\times\cos\left(\frac{x}{2}\right)\times\cos\left(\frac{x}{4}\right)\times\cdots\times\cos\left(\frac{x}{2^{n-1}}\right)\times\cos\left(\frac{x}{2^n}\right)\times\sin\left(\frac{x}{2^n}\right)\\
&= \cos\left(\frac{x}{1}\right)\times\cos\left(\frac{x}{2}\right)\times\cos\left(\frac{x}{4}\right)\times\cdots\times\cos\left(\frac{x}{2^{n-1}}\right)\times\frac{1}{2}\sin\left(\frac{x}{2^{n-1}}\right)\\
&= \cos\left(\frac{x}{1}\right)\times\cos\left(\frac{x}{2}\right)\times\cos\left(\frac{x}{4}\right)\times\cdots\times\frac{1}{4}\sin\left(\frac{x}{2^{n-2}}\right)\\
&= \frac{1}{2^{n+1}}\sin(2x)
\end{aligned}$$

Therefore we have $\displaystyle\prod_{k=0}^{n}\cos\left(\frac{x}{2^k}\right) = \frac{\sin(2x)}{2^{n+1}\sin\left(\frac{x}{2^n}\right)}$.

[3] I do this quite often with sums and products as I can find it easier to manipulate them in this form.

The question then tells us to use differentiation. Note that the product has become a sum — this suggests that taking logs[4] might be a good idea:

$$\log\left(\prod_{k=0}^{n}\cos\left(\frac{x}{2^k}\right)\right) = \log\left(\frac{\sin(2x)}{2^{n+1}\sin\left(\frac{x}{2^n}\right)}\right)$$
$$\sum_{k=0}^{n}\log\left(\cos\left(\frac{x}{2^k}\right)\right) = \log\left(\sin(2x)\right) - \log\left(2^{n+1}\right) - \log\left(\sin\left(\frac{x}{2^n}\right)\right)$$

Now we can use the facts that $\dfrac{\mathrm{d}}{\mathrm{d}x}\log\left(\cos x\right) = \dfrac{1}{\cos x} \times -\sin x = -\tan x$ and $\dfrac{\mathrm{d}}{\mathrm{d}x}\log\left(\sin x\right) = \dfrac{1}{\sin x} \times \cos x = \cot x$ to differentiate the above expression to get:

$$-\sum_{k=0}^{n}\frac{1}{2^k}\tan\left(\frac{x}{2^k}\right) = 2\cot(2x) - \frac{1}{2^n}\cot\left(\frac{x}{2^n}\right)$$

and then multiplying throughout by -1 gives us the required result.

Since this is a "show that" question is is probably best to state or derive $\dfrac{\mathrm{d}}{\mathrm{d}x}\log\left(\cos x\right) = -\tan x$ rather than jumping straight to the required result.

Part (iii) - 9 marks

We have:

$$\prod_{k=0}^{n}\cos\left(\frac{x}{2^k}\right) = \frac{\sin(2x)}{2^{n+1}\sin\left(\frac{x}{2^n}\right)}$$

but this has a $k = 0$ term — $\cos\left(\frac{x}{2^0}\right)$— which isn't in the required result. Dividing through by $\cos x$ gives:

$$\begin{aligned}
\prod_{k=1}^{n}\cos\left(\frac{x}{2^k}\right) &= \frac{\sin(2x)}{2^{n+1}\sin\left(\frac{x}{2^n}\right)\cos x} \\
&= \frac{2\sin x}{2^{n+1}\sin\left(\frac{x}{2^n}\right)} && \text{as } \sin 2x = 2\sin x\cos x \\
&= \frac{\sin x}{2^{n}\sin\left(\frac{x}{2^n}\right)} \\
&\approx \frac{\sin x}{2^{n}\times\left(\frac{x}{2^n}\right)} && \text{as } \sin\theta \approx \theta \text{ as } \theta \to 0
\end{aligned}$$

Then letting $n \to \infty$ gives:

$$\prod_{k=1}^{n}\cos\left(\frac{x}{2^k}\right) = \frac{\sin x}{x}$$

[4]Remember that $\log(AB) = \log A + \log B$.

For the second result it looks as if we need to use the second result from part (ii), but it would be nice if the power of 2 outside the tan was the same as the power of 2 inside the tan. Start by manipulating the sum until it looks more like the one in part (ii).

$$\begin{aligned}\sum_{j=2}^{n} \frac{1}{2^{j-2}} \tan\left(\frac{\pi}{2^j}\right) &= \sum_{j=2}^{n} \frac{1}{2^{j-2}} \tan\left(\frac{\pi/4}{2^{j-2}}\right) \\ &= \sum_{k=0}^{n-2} \frac{1}{2^k} \tan\left(\frac{\pi/4}{2^k}\right) \qquad \text{using } k = j-2\end{aligned}$$

So we have:

$$\sum_{j=2}^{n} \frac{1}{2^{j-2}} \tan\left(\frac{\pi}{2^j}\right) = \sum_{k=0}^{n-2} \frac{1}{2^k} \tan\left(\frac{\pi/4}{2^k}\right) = \frac{1}{2^{n-2}} \cot\left(\frac{\pi/4}{2^{n-2}}\right) - 2\cot\left(2 \times \frac{\pi}{4}\right)$$

We have $\lim\limits_{\theta\to 0} \dfrac{\tan\theta}{\theta} = 1$ and so $\tan\theta \approx \theta$ for "small" θ. Taking the limit as $n \to \infty$ we have:

$$\begin{aligned}\sum_{k=0}^{\infty} \frac{1}{2^k} \tan\left(\frac{\pi/4}{2^k}\right) &= \lim_{n\to\infty} \left[\frac{1}{2^{n-2}} \cot\left(\frac{\pi/4}{2^{n-2}}\right) - 2\cot\left(\frac{\pi}{2}\right)\right] \\ &= \lim_{n\to\infty} \left[\frac{1}{2^{n-2}} \times \frac{2^{n-2}}{\pi/4}\right] - 0 \\ &= \frac{4}{\pi}\end{aligned}$$

Question 5

5 The sequence u_0, u_1, ... is said to be a *constant sequence* if $u_n = u_{n+1}$ for $n = 0, 1, 2, \ldots$. The sequence is said to be a *sequence of period* 2 if $u_n = u_{n+2}$ for $n = 0, 1, 2, \ldots$ and the sequence is not constant.

(i) A sequence of real numbers is defined by $u_0 = a$ and $u_{n+1} = \mathrm{f}(u_n)$ for $n = 0, 1, 2, \ldots$, where

$$\mathrm{f}(x) = p + (x - p)x,$$

and p is a given real number.

Find the values of a for which the sequence is constant.

Show that the sequence has period 2 for some value of a if and only if $p > 3$ or $p < -1$.

(ii) A sequence of real numbers is defined by $u_0 = a$ and $u_{n+1} = \mathrm{f}(u_n)$ for $n = 0, 1, 2, \ldots$, where

$$\mathrm{f}(x) = q + (x - p)x,$$

and p and q are given real numbers.

Show that there is no value of a for which the sequence is constant if and only if $\mathrm{f}(x) > x$ for all x.

Deduce that, if there is no value of a for which the sequence is constant, then there is no value of a for which the sequence has period 2.

Is it true that, if there is no value of a for which the sequence has period 2, then there is no value of a for which the sequence is constant?

Examiner's report

It was difficult to get full marks on this question, with most candidates struggling to correctly prove 'if and only if' statements in both directions.

Mostly, the two constant sequences were successfully found and then correctly rejected for sequences of period 2, but few thought to check that the other two solutions to the quartic did not also coincide with the constant sequences. Most candidates were able to use the discriminant to produce bounds on p, but many could not justify the strictness of the inequality, which was best done by considering the boundary cases separately.

The first request of the second part was answered well, with most using only the fact that it was a positive quadratic and a minority delving into the details of $\mathrm{f}(x)$. Most candidates who reached this part of the questions correctly used the result $\mathrm{f}(x) > x$ to show that $\mathrm{f}(\mathrm{f}(x))$ has no solutions, but many overlooked the connection between the final part and part (i).

Solution

In both constant sequences and sequences of period 2, we have $u_{n+2} = u_n$ however if a sequence is to have period 2 then we need to have different terms alternating, so we need $u_{n+1} \neq u_n$. A constant sequence looks like $a, a, a, \cdots$ and a sequence with period 2 looks like $a, b, a, b, \cdots$. When finding sequences which have period 2 it is important to discard the constant sequences.

Part (i) - 12 marks

Unusually for a STEP question, the first part of this question has more marks allocated to it than the second part. There are a couple of fiddly bits, including justifying why the limits on p are strict, which is why there are quite so many marks for this part.

If the sequence is to be constant then we need $u_0 = u_1 = u_2 = \cdots = a$. If $u_1 = u_0 = a$ then we have:

$$\begin{aligned} &a = p + (a-p)a \\ &a^2 - (p+1)a + p = 0 \end{aligned}$$

You can now use the quadratic formula to find a in terms of p, or you might notice that $a = p$ is a root of the equation (this is perhaps more easily seen in the first equation above). Similarly you might notice that $a = 1$ is a root.

We have:

$$(a-p)(a-1) = 0 \implies a = p \text{ or } a = 1$$

For the next part we have to be a little bit careful as the question states "if and only if". A sequence has period 2 if and only if $u_n = u_{n+2} \neq u_{n+1}$.

$$\begin{aligned} & u_n = u_{n+2} \\ \iff & a = \mathrm{f}\big(\mathrm{f}(a)\big) \\ \iff & a = \mathrm{f}(p + (a-p)a) \\ \iff & a = p + \Big[\cancel{p} + (a-p)a - \cancel{p}\Big]\big(p + (a-p)a\big) \\ \iff & 0 = (p-a) + (a-p)a\big(p + (a-p)a\big) \end{aligned}$$

There is a temptation now to expand the brackets, but instead notice that there is a common factor of $(a-p)$ on the RHS of the final equation above. Factorising this out gives:

$$\begin{aligned} & a = \mathrm{f}\big(\mathrm{f}(a)\big) \\ \iff & 0 = (a-p)\big[-1 + a(p + (a-p)a)\big] \\ \iff & 0 = (a-p)\big[-1 + pa + a^3 - pa^2\big] \end{aligned}$$

Note that $a = p$ is one of the 2 conditions for the sequence to be constant — if makes sense for this to be a solution as if a sequence is constant then we will have $a = \mathrm{f}\big(\mathrm{f}(a)\big)$. The other value of a which gives a constant sequence is $a = 1$, and if you substitute $a = 1$ into the above equation you will see that this gives another root.

$$\begin{aligned}
& a = \mathrm{f}\big(\mathrm{f}(a)\big) \\
\iff & 0 = (a-p)\big[-1 + pa + a^3 - pa^2\big] \\
\iff & 0 = (a-p)(a-1)(a^2 + (1-p)a + 1) \\
\iff & a = p \text{ or } a = 1 \text{ or } a^2 + (1-p)a + 1 = 0
\end{aligned}$$

The first two conditions on a ($a = p$ or $a = 1$) are the conditions for the sequence to be constant. Hence the sequence has period 2 if and only if $a^2 + (1-p)a + 1 = 0$. This quadratic has solutions if and only if $(1-p)^2 - 4 \geqslant 0$, i.e. $(1-p)^2 \geqslant 4$.

Note that the required answer has strict inequalities for p, whereas the discriminant condition has given us non-strict inequalities. This suggests that we may need to be careful!

$$\begin{aligned}
& (1-p)^2 \geqslant 4 \\
\iff & p^2 - 2p - 3 \geqslant 0 \\
\iff & (p-3)(p+1) \geqslant 0 \\
\iff & p \geqslant 3 \text{ or } p \leqslant -1
\end{aligned}$$

This isn't quite the required answer, so we need to look at the boundary cases in more detail. If $p = 3$ then the quadratic becomes $a^2 - 2a + 1 = 0$ which has solution $a = 1$, which is one of the two constant sequences again. Hence $p \neq 3$. If $p = -1$, then the quadratic becomes $a^2 + 2a + 1 = 0$ which has solution $a = -1$, and so $a = p$ and the sequence is constant again.

Therefore there is some value of a for which the sequence has period 2 (and is not constant) if and only if $p > 3$ or $p < -1$.

Part (ii) - 8 marks

There are no values of a for which the sequence is constant if and only if the equation $\mathrm{f}(a) = a$ has no solutions.

The temptation now is to jump into considering the discriminant of the quadratic — but that is not very helpful. Instead look at the condition required by the question[5].

$$\begin{aligned}
& \mathrm{f}(a) = a \quad \text{has no solutions} \\
\iff & \text{either } \mathrm{f}(x) > x \quad \text{for all } x \\
& \text{or } \mathrm{f}(x) < x \quad \text{for all } x
\end{aligned}$$

Since $\mathrm{f}(x) = x^2 - px + q$ is a quadratic with a positive coefficient of x it cannot be less than x for all x.

Hence $\mathrm{f}(a) = a$ has no solutions if and only if $\mathrm{f}(x) > x$ for all x.

If there are no values of a for which the sequence is constant then we have:

$$\begin{aligned}
& \mathrm{f}(x) > x \quad \text{for all } x \\
\implies & \mathrm{f}\big(\mathrm{f}(x)\big) > \mathrm{f}(x) > x \quad \text{for all } x
\end{aligned}$$

[5] There are often "clues" in a STEP question, for example the limits on an integral might give you an idea of what substitution might have been used. In this case the condition $\mathrm{f}(x) > x$ means that $\mathrm{f}(a) = a$ has no solutions, and so you cannot get a constant sequence — however the question was "if and only if" so a little more has to be done.

Hence we have $\mathrm{f}\big(\mathrm{f}(x)\big) > x$ for all x, and so $\mathrm{f}\big(\mathrm{f}(a)\big) = a$ has no solutions and there is no value of a for which the sequence has period 2.

It is not true that if there is no value of a for which the sequence has period 2, then there is no value of a for which the the sequence is constant. If we take the sequence $\mathrm{f}(x) = p + (x - p)x$ from part (i), and take $p = 3$ then this sequence has a constant sequence (when $a = p = 3$), but does not have a sequence of period 2.

Question 6

6 **Note**: You may assume that if the functions $y_1(x)$ and $y_2(x)$ both satisfy one of the differential equations in this question, then the curves $y = y_1(x)$ and $y = y_2(x)$ do not intersect.

(i) Find the solution of the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y + x + 1$$

that has the form $y = mx + c$, where m and c are constants.

Let $y_3(x)$ be the solution of this differential equation with $y_3(0) = k$. Show that any stationary point on the curve $y = y_3(x)$ lies on the line $y = -x - 1$. Deduce that solution curves with $k < -2$ cannot have any stationary points.

Show further that any stationary point on the solution curve is a local minimum.

Use the substitution $Y = y + x$ to solve the differential equation, and sketch, on the same axes, the solutions with $k = 0$, $k = -2$ and $k = -3$.

(ii) Find the two solutions of the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = x^2 + y^2 - 2xy - 4x + 4y + 3$$

that have the form $y = mx + c$.

Let $y_4(x)$ be the solution of this differential equation with $y_4(0) = -2$. (Do not attempt to find this solution.)

Show that any stationary point on the curve $y = y_4(x)$ lies on one of two lines that you should identify. What can be said about the gradient of the curve at points between these lines?

Sketch the curve $y = y_4(x)$. You should include on your sketch the two straight line solutions and the two lines of stationary points.

Examiner's report

Of the Pure questions, this was the question that had the lowest average mark, mainly due to the large number of attempts that did not manage to score any marks. Many candidates seemed uncomfortable with this question which asked them to look at what information can be gleaned about differential equations without directly solving them. Many candidates decided that the only way to proceed was to solve the differential equation, and almost invariably this led to long and convoluted methods. Candidates seemed to have very little idea that the differential equation can be interpreted as the gradient of a curve at different points – it was simply an object on which certain methods had to be applied. A surprisingly small number of candidates realised that setting $\frac{\mathrm{d}y}{\mathrm{d}x} = 0$ could (and should) be done directly in the differential equation to find the locus of stationary points.

This was also a question which required candidates to bring a lot of disparate information together in the final sketches. A large number of candidates said things like the gradient was negative between two lines, but their sketch showed something different. Some who said that there should be stationary points on the line $y = x - 1$ and $y = x - 3$ drew their curve tangentially to these two lines instead.

Overall this was a question which really benefited candidates who took a moment to stop and think about what was being suggested, rather than blindly applying methods.

Solution

The note at the top of the question was meant to help you when considering $y = -x - 2$ and $y = y_3$ in part **(i)**. When I was doing the question, I forgot about the note and tried to justify why the two curves could not intersect. I have left my original wording in, and have added a comment to show you what I could have said instead.

Part (i) - 12 marks
Another question where the first part has more marks than the second part!

Substituting $y = mx + c$ into the differential equation gives:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y + x + 1$$
$$m = mx + c + x + 1$$

Equating coefficients gives $m = -1$ and $c = -2$ so the solution is $y = -x - 2$.

$y_3(x)$ satisfies the differential equation, so at a stationary point we have $\frac{\mathrm{d}y}{\mathrm{d}x} = 0$ and so here we have $y + x + 1 = 0 \implies y = -x - 1$.

There is a *solution curve* $y = -x - 2$. Solutions of the differential equations cannot intersect (and so cannot cross) — this is hard to prove, but if two curves meet at the point (a, b) then they will both have gradient $\frac{\mathrm{d}y}{\mathrm{d}x} = b + a + 1$ at this point. This means they will both head off in the same direction to the point $(a + \delta a, b + \delta b)$ where they will still have the same gradient etc. Hence they can never separate into two curves.

Here I could have said that the solutions $y = -x - 2$ and $y_3(x)$ cannot intersect (from the note at the top of the question), and so if $y_3(x)$ lies below $y = -x - 2$ at some point it must stay below $y = -x - 2$.

Since there is a solution curve at $y = -x - 2$ the curve $y = y_3(x)$ cannot cross this line. The stationary points of $y_3(x)$ lie on $y = -x - 1$ which is parallel to the solution curve $y = -x - 2$. This means that if $y_3(x)$ lies below $y = -x - 2$ at any point it cannot reach $y = -x - 1$ and so has no stationary points.

The line $y = -x - 2$ has y-intercept $(0, -2)$ and we are told that $y_3(0) = k$. This means that if $k < -2$ then $y_3(x)$ lies below $y = -x - 2$ and so cannot reach $y = -x - 1$, therefore has no stationary points.

To show that stationary points have a minimum consider the sign of the second derivative.

$$\begin{aligned} \frac{\mathrm{d}y}{\mathrm{d}x} &= y + x + 1 \\ \implies \frac{\mathrm{d}^2y}{\mathrm{d}x^2} &= \frac{\mathrm{d}y}{\mathrm{d}x} + 1 \\ &= 1 \qquad \text{stationary point} \implies \frac{\mathrm{d}y}{\mathrm{d}x} = 0 \\ &> 0 \end{aligned}$$

Since the second derivative is positive then the stationary points are local minima.

Substituting $Y = y + x$, $\dfrac{\mathrm{d}Y}{\mathrm{d}x} = \dfrac{\mathrm{d}y}{\mathrm{d}x} + 1$ into the differential equation gives:

$$\begin{aligned} \frac{\mathrm{d}Y}{\mathrm{d}x} - 1 &= Y + 1 \\ \frac{\mathrm{d}Y}{\mathrm{d}x} &= Y + 2 \\ \int \frac{1}{Y+2}\, dY &= x + c \\ \ln(Y+2) &= x + c \\ Y + 2 &= A\mathrm{e}^x \\ y + x &= A\mathrm{e}^x - 2 \\ y &= A\mathrm{e}^x - 2 - x \end{aligned}$$

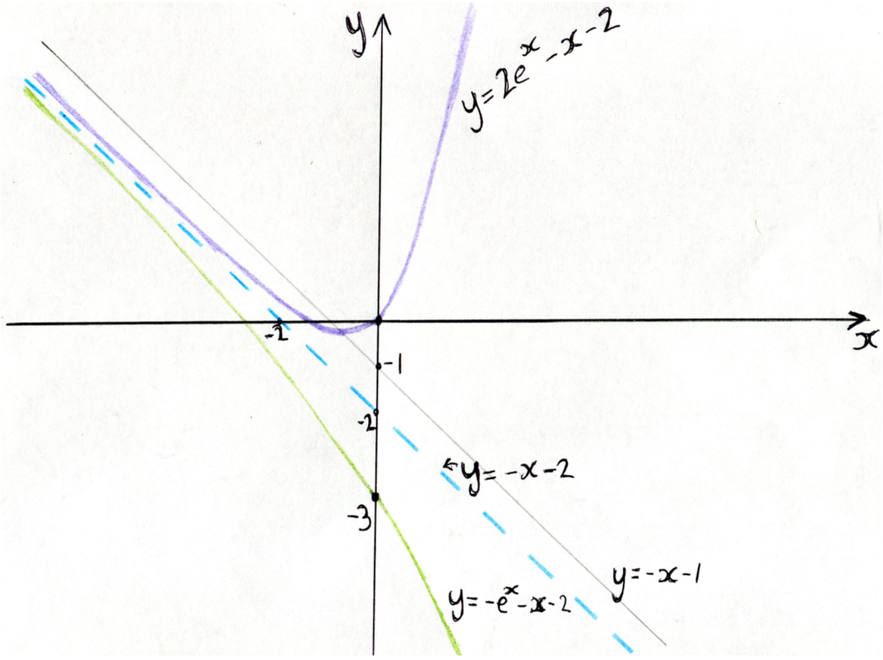
Using the condition $y_3(0) = k$ with $k = 0, -2$ and -3 gives the three solutions:

- $y = 2\mathrm{e}^x - x - 2$
- $y = -x - 2$
- $y = -\mathrm{e}^x - x - 2$

The first and third of these can be thought of as being a deviation from $y = -x-2$, and as $x \to -\infty$ they all converge. The third curve lies below $y = -x - 2$ so cannot have any stationary points where as the first curve will have one (as the gradient is negative for very large and negative x but positive for very large and positive x).

Finding the y-intercepts is easy (just substitute $x = 0$), so it is a good idea to do this and put them on your sketch. Finding the x-intercepts is much harder.

Your sketch might look something like this:



I have hand-drawn this sketch rather than using something like Desmos to give you a better idea of what a sketch in an exam might look like. In my efforts to ensure that it was clear after scanning I have probably managed to make it a little too neat to be a realistic example. However, if you produce something like this in an exam your markers will be grateful!

Part (ii) - 8 marks

Substituting $y = mx + c$ gives:

$$\begin{aligned} m &= x^2 + (mx+c)^2 - 2x(mx+c) - 4x + 4(mx+c) + 3 \\ m &= x^2 + m^2x^2 + 2mcx + c^2 - 2mx^2 - 2xc - 4x + 4mx + 4c + 3 \end{aligned}$$

Equating coefficients gives:

$$\begin{aligned} 0 &= 1 + m^2 - 2m \implies (m-1)^2 = 0 \\ 0 &= 2mc - 2c - 4 + 4m \\ m &= c^2 + 4c + 3 \end{aligned}$$

The first of these gives $m = 1$. Substituting $m = 1$ into the second one gives $0 = 2c - 2c - 4 + 4$ which is not very helpful. Substituting $m = 1$ into the third equation gives $c^2 + 4c + 2 = 0 \implies c = -2 \pm \sqrt{2}$. The two straight line solutions of this equation are $y = x - 2 \pm \sqrt{2}$.

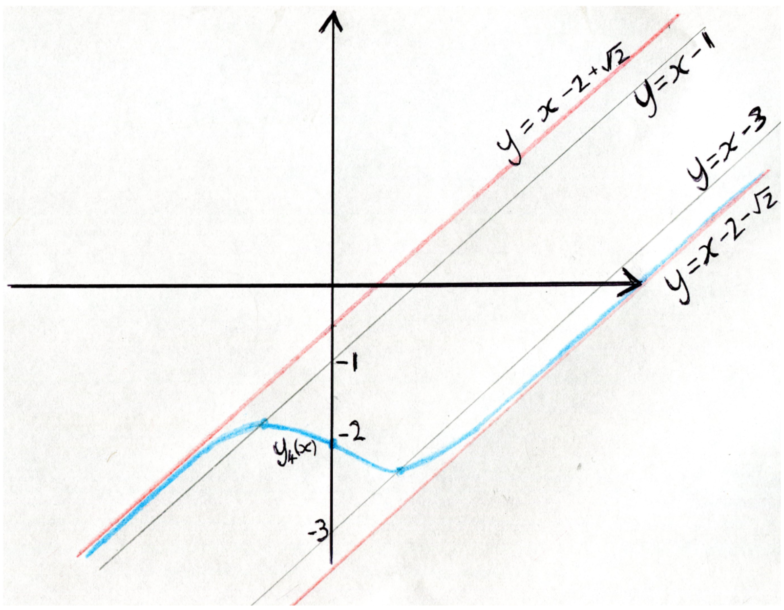
At the stationary points of this differential equation we have:

$$\begin{aligned} x^2 + y^2 - 2xy - 4x + 4y + 3 &= 0 \\ \implies (y - x + 1)(y - x + 3) &= 0 \end{aligned}$$

and so any stationary points lie on either $y = x - 1$ or $y = x - 3$.

We are told that the stationary points lie on one of two lines, so we might expect that we can factorise the equation to give two brackets each of which represents a straight line. I factorised by inspection, starting with a y at the start of each bracket and then realising that I needed a $-x$ in each so that I get x^2 and $-2xy$ when I expand.

Between the lines $y = x - 1$ and $y = x - 3$ the gradient must be negative as we have $\dfrac{\mathrm{d}y}{\mathrm{d}x} = (y - x + 1)(y - x + 3)$ and between the two lines one of these brackets will be negative and one positive. Above the two lines, or below, we will have positive gradient.



In an early draft of the question, there was another request in part **(ii)** which said "Show further that if the curve has a point of inflection, then this point lies on the line $y = x - 2$.". It was decided that this made the question too long and didn't provide much extra information for the graph.

Question 7

7 **(i)** The points A, B and C have position vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$, respectively. Each of these vectors is a unit vector (so $\mathbf{a} \cdot \mathbf{a} = 1$, for example) and

$$\mathbf{a} + \mathbf{b} + \mathbf{c} = \mathbf{0}.$$

Show that $\mathbf{a} \cdot \mathbf{b} = -\frac{1}{2}$. What can be said about the triangle ABC? You should justify your answer.

(ii) The four distinct points A_i ($i = 1, 2, 3, 4$) have unit position vectors $\mathbf{a}_i$ and

$$\sum_{i=1}^{4} \mathbf{a}_i = \mathbf{0}.$$

Show that $\mathbf{a}_1 . \mathbf{a}_2 = \mathbf{a}_3 . \mathbf{a}_4$.

(a) Given that the four points lie in a plane, determine the shape of the quadrilateral with vertices A_1, A_2, A_3 and A_4.

(b) Given instead that the four points are the vertices of a regular tetrahedron, find the length of the sides of this tetrahedron.

Examiner's report

This was the least popular of the Pure questions. Good solutions to this question often included clear diagrams to enable the angles being discussed to be identified easily. Many of the candidates were able to calculate the value of $\mathbf{a} . \mathbf{b}$ correctly, but often did not fully justify that the triangle ABC was equilateral. For the second part, many candidates were again able to establish the relationship between scalar products, but less success was seen in identifying the type of quadrilateral. In the final part there were a large number of different approaches taken and many of these were completed successfully by some of the candidates.

Solution

Part (i) - 8 marks

Vector questions are often a little off-putting. Often a good idea is to start by drawing a sketch to help you understand the situation. For this one we are given a fact connecting the three vectors, and want to find the dot product between two of them. About the only sensible thing that can be done to start with is to dot the given relationship with one of the vectors, probably either $\mathbf{a}$ or $\mathbf{b}$ so that we end up with an $\mathbf{a}\cdot\mathbf{b}$ term.

$$\begin{aligned} \mathbf{a}+\mathbf{b}+\mathbf{c} &= \mathbf{0} \\ \implies \mathbf{a}\cdot(\mathbf{a}+\mathbf{b}+\mathbf{c}) &= \mathbf{a}\cdot\mathbf{0} \\ \implies 1+\mathbf{a}\cdot\mathbf{b}+\mathbf{a}\cdot\mathbf{c} &= 0 \end{aligned}$$

Similarly we can dot with $\mathbf{b}$ and $\mathbf{c}$ to get the equations $\mathbf{a}\cdot\mathbf{b}+1+\mathbf{b}\cdot\mathbf{c}=0$ and $\mathbf{a}\cdot\mathbf{c}+\mathbf{b}\cdot\mathbf{c}+1=0$. Adding the first two equations gives:

$$2+2\mathbf{a}\cdot\mathbf{b}+\mathbf{a}\cdot\mathbf{c}+\mathbf{b}\cdot\mathbf{c}=0$$

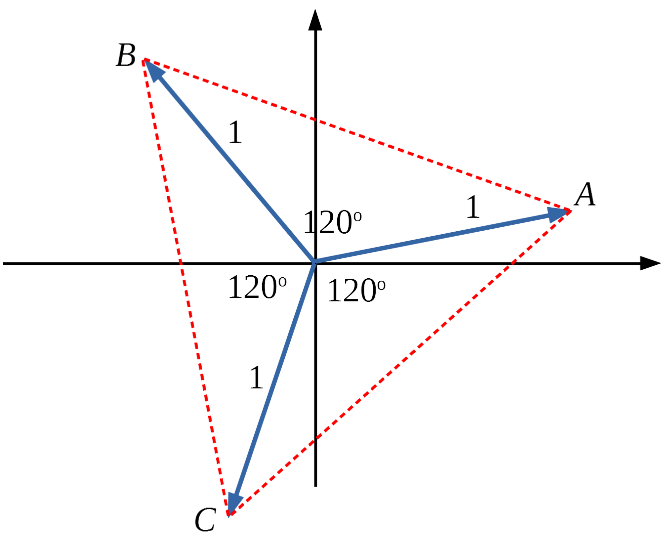
and then we can substitute the last equation to get:

$$2+2\mathbf{a}\cdot\mathbf{b}-1=0 \implies \mathbf{a}\cdot\mathbf{b}=-\frac{1}{2}$$

Since the original equation is symmetrical in $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$, we also have $\mathbf{b}\cdot\mathbf{c}=-\frac{1}{2}$ and $\mathbf{c}\cdot\mathbf{a}=-\frac{1}{2}$.

Since $\mathbf{a}\cdot\mathbf{b}=-\frac{1}{2}$, and $\mathbf{a}$ and $\mathbf{b}$ are both unit vectors, then the angle between them satisfies $\cos\theta=-\frac{1}{2}$ and so the angle between then is 120°. The same is true for $\mathbf{b}$ and $\mathbf{c}$, and for $\mathbf{c}$ and $\mathbf{a}$ so triangle ABC is an equilateral triangle.

It can be quite helpful to sketch a diagram showing what is happening, such as the one below. You can also refer to your diagram as part of your explanation (and this might well make your explanation clearer). You could simplify matters by putting one of the points on the x or y axes — e.g. "WLOG[6] let point A be at $(1,0)$".



[6] This means "Without Loss Of Generality", i.e. it doesn't change anything if we rotate our diagram so that A is on the x-axis.

Part (ii) stem - 3 marks Dotting the equation $\mathbf{a_1}+\mathbf{a_2}+\mathbf{a_3}+\mathbf{a_4}=0$ with each of the $\mathbf{a_i}$ in turn, and rearranging gives the equations:

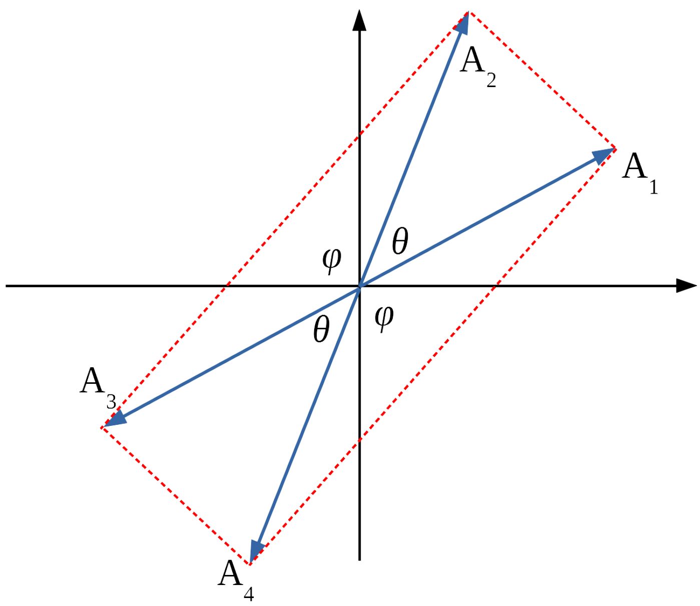
$$\mathbf{a_1}\cdot\mathbf{a_2}+\mathbf{a_1}\cdot\mathbf{a_3}+\mathbf{a_1}\cdot\mathbf{a_4}=-1 \tag{1}$$
$$\mathbf{a_2}\cdot\mathbf{a_1}+\mathbf{a_2}\cdot\mathbf{a_3}+\mathbf{a_2}\cdot\mathbf{a_4}=-1 \tag{2}$$
$$\mathbf{a_3}\cdot\mathbf{a_1}+\mathbf{a_3}\cdot\mathbf{a_2}+\mathbf{a_3}\cdot\mathbf{a_4}=-1 \tag{3}$$
$$\mathbf{a_4}\cdot\mathbf{a_1}+\mathbf{a_4}\cdot\mathbf{a_2}+\mathbf{a_4}\cdot\mathbf{a_3}=-1 \tag{4}$$

Using (1) + (2) − (3) − (4) gives $2\mathbf{a_1}\cdot\mathbf{a_2}-2\mathbf{a_3}\cdot\mathbf{a_4}=0$ and so we have $\mathbf{a_1}\cdot\mathbf{a_2}=\mathbf{a_3}\cdot\mathbf{a_4}$.

Part (ii) (a) - 3 marks

Since $\mathbf{a_1}\cdot\mathbf{a_2}=\mathbf{a_3}\cdot\mathbf{a_4}$ we know that $\angle A_1OA_2=\angle A_3OA_4$. By symmetry (as the given condition is symmetric in the $\mathbf{a_i}$) we also have $\mathbf{a_2}\cdot\mathbf{a_3}=\mathbf{a_4}\cdot\mathbf{a_1}\implies\angle A_2OA_3=\angle A_4OA_1$.

All four points are distinct (so none of the angles between them are equal to 0), and the vectors are all unit length. This means we have a picture that looks something like this:



The four points form the vertices of a rectangle.

Part (ii) (b) - 6 marks

This question is asking us to find the length of one of the sides of the tetrahedron. All the sides are the same, so lets consider $|A_1A_2|$, i.e. the length between the points A_1 and A_2.

Things are a little easier if we square to get:

$$\begin{aligned}|A_1A_2|^2&=(\mathbf{a_2}-\mathbf{a_1})^2\\&=(\mathbf{a_2}-\mathbf{a_1})\cdot(\mathbf{a_2}-\mathbf{a_1})\\&=\mathbf{a_1}^2+\mathbf{a_2}^2-2\mathbf{a_1}\cdot\mathbf{a_2}\\&=2-2\mathbf{a_1}\cdot\mathbf{a_2}\end{aligned}$$

It would be useful to know what $\mathbf{a_1} \cdot \mathbf{a_2}$ is. We have $\mathbf{a_1} \cdot \mathbf{a_2} + \mathbf{a_1} \cdot \mathbf{a_3} + \mathbf{a_1} \cdot \mathbf{a_4} = -1$. If we think about the positions of $\mathbf{a_2}$, $\mathbf{a_3}$ and $\mathbf{a_4}$ with respect to $\mathbf{a_1}$[7] then the angle that $\mathbf{a_1}$ makes with the others is the same for each of them (the question is careful to say that it is a *regular* tetrahedron). We also know that all of the $\mathbf{a_i}$ are unit vectors, so we have:

$$\mathbf{a_1} \cdot \mathbf{a_2} = \mathbf{a_1} \cdot \mathbf{a_3} = \mathbf{a_1} \cdot \mathbf{a_4}$$

Therefore $\mathbf{a_1} \cdot \mathbf{a_2} + \mathbf{a_1} \cdot \mathbf{a_3} + \mathbf{a_1} \cdot \mathbf{a_4} = -1 \implies \mathbf{a_1} \cdot \mathbf{a_2} = -\frac{1}{3}$.

So the lengths of the sides are given by:

$$\begin{aligned} |A_1A_2|^2 &= 2 - 2\mathbf{a_1} \cdot \mathbf{a_2} = \frac{8}{3} \\ \implies |A_1A_2| &= \frac{2\sqrt{2}}{\sqrt{3}} \end{aligned}$$

See MEI for a slightly different solution to this.

[7] I find this easier to picture if I have $\mathbf{a_1}$ on the z-axis (i.e. at the point $(0, 0, 1)$).

Question 8

8 The domain of the function f is the set of all 2×2 matrices and its range is the set of real numbers. Thus, if $\mathbf{M}$ is a 2×2 matrix, then $\mathrm{f}(\mathbf{M}) \in \mathbb{R}$.

The function f has the property that $\mathrm{f}(\mathbf{MN}) = \mathrm{f}(\mathbf{M})\mathrm{f}(\mathbf{N})$ for any 2×2 matrices $\mathbf{M}$ and $\mathbf{N}$.

(i) You are given that there is a matrix $\mathbf{M}$ such that $\mathrm{f}(\mathbf{M}) \neq 0$. Let $\mathbf{I}$ be the 2×2 identity matrix. By considering $\mathrm{f}(\mathbf{MI})$, show that $\mathrm{f}(\mathbf{I}) = 1$.

(ii) Let $\mathbf{J} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. You are given that $\mathrm{f}(\mathbf{J}) \neq 1$. By considering $\mathbf{J}^2$, evaluate $\mathrm{f}(\mathbf{J})$.

Using $\mathbf{J}$, show that, for any real numbers a, b, c and d,

$$\mathrm{f}\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = -\mathrm{f}\left(\begin{pmatrix} c & d \\ a & b \end{pmatrix}\right) = \mathrm{f}\left(\begin{pmatrix} d & c \\ b & a \end{pmatrix}\right).$$

(iii) Let $\mathbf{K} = \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}$ where $k \in \mathbb{R}$. Use $\mathbf{K}$ to show that, if the second row of the matrix $\mathbf{A}$ is a multiple of the first row, then $\mathrm{f}(\mathbf{A}) = 0$.

(iv) Let $\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. By considering the matrices $\mathbf{P}^2$, $\mathbf{P}^{-1}$, and $\mathbf{K}^{-1}\mathbf{PK}$ for suitable values of k, evaluate $\mathrm{f}(\mathbf{P})$.

Examiner's report

Many good solutions were seen to this question, but solutions often lacked clear enough justification to be awarded full marks. However, there were also a surprising number of candidates who did not manage to invert the 2×2 matrices successfully. Candidates who claimed that the function f was the determinant of the matrix were not able to score high marks as the solutions did not then demonstrate that the results were true of any function satisfying the property given. The first two parts of this question were largely done well. The third part was found more difficult, with few candidates realising that $\begin{pmatrix} a & b \\ ka & kb \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}\begin{pmatrix} a & b \\ a & b \end{pmatrix}$. Those who did were then often able to provide a full solution, although often these were not fully justified. Several candidates instead used $\begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}\begin{pmatrix} a & b \\ k^{-1}a & k^{-1}b \end{pmatrix} = \begin{pmatrix} a & b \\ a & b \end{pmatrix}$ to produce a solution which covered all cases apart from the one where $k = 0$. In some cases, candidates did not appear to consider $\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$ to be an example of a matrix in which the second row was a multiple of the first. In part (iv) many candidates made use of the fact that $\mathrm{f}(P) \neq 0$ without showing that this must be the case.

Solution
One thing to note is that although in general $\mathbf{MN} \neq \mathbf{NM}$, we do have $\mathrm{f}(\mathbf{M})\mathrm{f}(\mathbf{N}) = \mathrm{f}(\mathbf{N})\mathrm{f}(\mathbf{M})$ as $\mathrm{f}(\mathbf{M})$ and $\mathrm{f}(\mathbf{N})$ are real numbers. This means that we have $\mathrm{f}(\mathbf{MN}) = \mathrm{f}(\mathbf{NM})$.

Part (i) - 2 marks
From the question we know that $\mathrm{f}(\mathbf{MN}) = \mathrm{f}(\mathbf{M})\mathrm{f}(\mathbf{N})$, and so we have:

$$\begin{aligned} \mathrm{f}(\mathbf{MI}) &= \mathrm{f}(\mathbf{M})\mathrm{f}(\mathbf{I}) \\ \implies \mathrm{f}(\mathbf{M}) &= \mathrm{f}(\mathbf{M})\mathrm{f}(\mathbf{I}) \implies \mathrm{f}(\mathbf{M})\Big(1 - \mathrm{f}(\mathbf{I})\Big) = 0 \\ \implies \mathrm{f}(\mathbf{I}) &= 1 \qquad \text{as } \mathrm{f}(\mathbf{M}) \neq 0 \end{aligned}$$

Part (ii) - 7 marks
We have $\mathbf{J}^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, and so $\mathrm{f}(\mathbf{J}^2) = \mathrm{f}(\mathbf{I}) \implies \mathrm{f}(\mathbf{J})\mathrm{f}(\mathbf{J}) = 1$. Since $\mathrm{f}(\mathbf{J}) \neq 1$ we must have $\mathrm{f}(\mathbf{J}) = -1$.

We have $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \mathbf{J}\begin{pmatrix} c & d \\ a & b \end{pmatrix}$, and so since $\mathrm{f}(\mathbf{J}) = -1$ we have:

$$\mathrm{f}\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = -1 \times \mathrm{f}\left(\begin{pmatrix} c & d \\ a & b \end{pmatrix}\right).$$

Similarly we have $\begin{pmatrix} d & c \\ b & a \end{pmatrix} = \begin{pmatrix} c & d \\ a & b \end{pmatrix}\mathbf{J}$ and so we have:

$$\mathrm{f}\left(\begin{pmatrix} d & c \\ b & a \end{pmatrix}\right) = \mathrm{f}\left(\begin{pmatrix} c & d \\ a & b \end{pmatrix}\right) \times -1 = -\mathrm{f}\left(\begin{pmatrix} c & d \\ a & b \end{pmatrix}\right).$$

Hence $\mathrm{f}\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = -\mathrm{f}\left(\begin{pmatrix} c & d \\ a & b \end{pmatrix}\right) = \mathrm{f}\left(\begin{pmatrix} d & c \\ a & b \end{pmatrix}\right)$ as required.

Part (iii) - 4 marks
Let $\mathbf{A} = \begin{pmatrix} a & b \\ ka & kb \end{pmatrix}$ (which is a general matrix where the second row is a multiple of the first row). We then have:

$$\begin{aligned} \mathrm{f}(\mathbf{A}) &= \mathrm{f}\left(\begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}\right)\mathrm{f}\left(\begin{pmatrix} a & b \\ a & b \end{pmatrix}\right) \\ &= \mathrm{f}(K)\mathrm{f}\left(\begin{pmatrix} a & b \\ a & b \end{pmatrix}\right) \end{aligned}$$

Using $\mathrm{f}\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = -1 \times \mathrm{f}\left(\begin{pmatrix} c & d \\ a & b \end{pmatrix}\right)$ we have:

$$\mathrm{f}\left(\begin{pmatrix} a & b \\ a & b \end{pmatrix}\right) = -\mathrm{f}\left(\begin{pmatrix} a & b \\ a & b \end{pmatrix}\right)$$

and so we have $\mathrm{f}\left(\begin{pmatrix} a & b \\ a & b \end{pmatrix}\right) = 0$, and hence $\mathrm{f}(\mathbf{A}) = 0$.

Part (iv) - 7 marks

When doing this question I was initially a bit confused as to why I had been asked to consider $\mathbf{P}^{-1}$ as one of the matrices. As it turns out, we needed to find $\mathbf{P}^{-1}$ at the end of the question in order to eliminate one of the options for $\mathrm{f}(\mathbf{P})$.

We have $\mathbf{P}^2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ and $\mathbf{P}^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$. We also have $\mathbf{K}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{k} \end{pmatrix}$ so we have:

$$\begin{aligned} \mathbf{K}^{-1}\mathbf{PK} &= \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{k} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 0 & \frac{1}{k} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix} \\ &= \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \end{aligned}$$

We also have $\mathrm{f}(\mathbf{K})\mathrm{f}(\mathbf{K}^{-1}) = \mathrm{f}(\mathbf{KK}^{-1}) = \mathrm{f}(\mathbf{I}) = 1$.

Now:

$$\begin{aligned} \mathrm{f}(\mathbf{K}^{-1}\mathbf{PK}) &= \mathrm{f}(\mathbf{K}^{-1})\mathrm{f}(\mathbf{PK}) \\ &= \mathrm{f}(\mathbf{K}^{-1})\mathrm{f}(\mathbf{P})\mathrm{f}(\mathbf{K}) \\ &= \mathrm{f}(\mathbf{P})\mathrm{f}(\mathbf{K})\mathrm{f}(\mathbf{K}^{-1}) \\ &= \mathrm{f}(\mathbf{P}) \end{aligned}$$

Note that you need to use $\mathrm{f}(\mathbf{MN}) = \mathrm{f}(\mathbf{M})\mathrm{f}(\mathbf{N})$ twice — you cannot jump straight to the three term version without justification.

Since $\mathbf{K}^{-1}\mathbf{PK} = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ and $\mathbf{P}^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ take $k = 2$. We now have:

$$\begin{aligned} \mathbf{K}^{-1}\mathbf{PK} &= \mathbf{P}^2 \\ \mathrm{f}(\mathbf{K}^{-1}\mathbf{PK}) &= \mathrm{f}(\mathbf{P}^2) \\ \mathrm{f}(\mathbf{P}) &= \mathrm{f}(\mathbf{P})\mathrm{f}(\mathbf{P}) \\ 0 &= \mathrm{f}(\mathbf{P})\Big(\mathrm{f}(\mathbf{P}) - 1\Big) \end{aligned}$$

Therefore we have either $\mathrm{f}(\mathbf{P}) = 0$ or $\mathrm{f}(\mathbf{P}) = 1$. The question asks you to evaluate $\mathrm{f}(\mathbf{P})$, which suggests that there is probably just one answer. You can (and should) check that you haven't missed something in the question (such as a condition, or other piece of information). In this case we have not yet used $\mathbf{P}^{-1}$, so that might be helpful.

We know that $\mathbf{P}^{-1}$ exists, and we have $\mathrm{f}(\mathbf{P})\mathrm{f}(\mathbf{P}^{-1}) = \mathrm{f}(\mathbf{I}) = 1$ which means that $\mathrm{f}(\mathbf{P})$ cannot be 0 and must be equal to 1.

Question 9

9 A particle P is projected from a point O on horizontal ground with speed u and angle of projection α, where $0 < \alpha < \frac{1}{2}\pi$.

(i) Show that if $\sin\alpha < \dfrac{2\sqrt{2}}{3}$, then the distance OP is increasing throughout the flight.

Show also that if $\sin\alpha > \dfrac{2\sqrt{2}}{3}$, then OP will be decreasing at some time before the particle lands.

(ii) At the same time as P is projected, a particle Q is projected horizontally from O with speed v along the ground in the opposite direction from the trajectory of P. The ground is smooth. Show that if

$$2\sqrt{2}v > (\sin\alpha - 2\sqrt{2}\cos\alpha)u,$$

then QP is increasing throughout the flight of P.

Examiner's report

This was the most popular of the Mechanics and Statistics questions, but also one of the questions that attracted a large number of solutions that received no marks.

Students seemed relatively good at setting up the kinematics equations in this question and most had the useful idea of differentiating. Somewhat fewer thought about using either completing the square or the quadratic discriminant to decide where the derivative was positive. The logic of the question was very poorly understood, with many students seeing the given inequality as the end point rather than the starting point of the question.

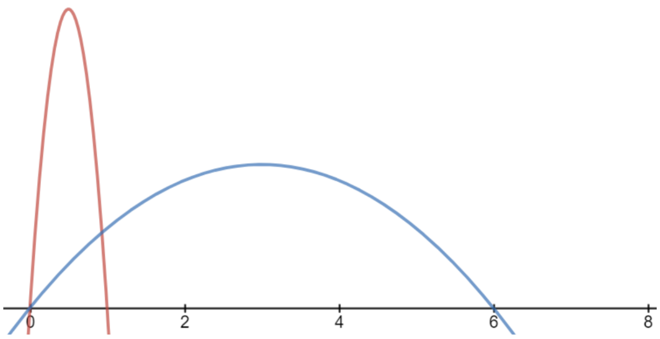
In the second part of part (i) it was important that students demonstrated not just that a time existed where the distance is decreasing, but that this time was in the acceptable domain of the question.

Part (ii) was conceptually very similar to part (i) but most students found the increased algebraic demand too much.

Solution

Part (i) - 11 marks

At first glance the request to show that the distance OP is increasing throughout the flight seems a little odd — there is no horizontal acceleration so the horizontal distance is always increasing. However if you think of a parabola which is very "high and narrow", then the distance OP at the top of the parabola will be greater than the distance OP when the particle lands. A couple of examples are shown below, one (blue) where the distance OP is always increasing and one (red) where the distance OP increases and then decreases.



The first thing to do is find the vertical and horizontal distances covered at a time t. Using $s = ut + \frac{1}{2}at^2$ and initial velocity $\mathbf{u} = \begin{pmatrix} u\cos\alpha \\ u\sin\alpha \end{pmatrix}$ we have displacement given by:

$$\begin{pmatrix} ut\cos\alpha \\ ut\sin\alpha - \frac{1}{2}gt^2 \end{pmatrix}$$

This gives the distance $(OP)^2$ as:

$$\begin{aligned}(ut\cos\alpha)^2 + \left(ut\sin\alpha - \tfrac{1}{2}gt^2\right) &= u^2t^2\cos^2\alpha + u^2t^2\sin^2\alpha - gut^3\sin\alpha + \tfrac{1}{4}g^2t^4 \\ &= u^2t^2 - gut^3\sin\alpha + \tfrac{1}{4}g^2t^4\end{aligned}$$

This will be increasing as long as the derivative of this is positive. Differentiating with respect to t gives us:

$$\begin{aligned}2u^2t - 3gut^2\sin\alpha + g^2t^3 &= t\left[2u^2 - 3gut\sin\alpha + g^2t^2\right] \\ &= t\left[2u^2 + \left(gt - \tfrac{3}{2}u\sin\alpha\right)^2 - \tfrac{9}{4}u^2\sin^2\alpha\right] \\ &= t\left[\left(gt - \tfrac{3}{2}u\sin\alpha\right)^2 + u^2\left(2 - \tfrac{9}{4}\sin^2\alpha\right)\right]\end{aligned}$$

Hence $\dfrac{\mathrm{d}}{\mathrm{d}t}(OP)^2$ is positive for all values of t if $2 - \frac{9}{4}\sin^2\alpha > 0$ i.e. if $\sin^2\alpha < \frac{8}{9}$. Since $0 < \alpha < \frac{1}{2}\pi$ this means that the derivative will be positive for all values of t if $\sin\alpha < \frac{2\sqrt{2}}{3}$.

Completing the square is often a useful technique when you want to show if something is positive (or negative).

If $\sin\alpha > \frac{2\sqrt{2}}{3}$ then the expression for $\dfrac{\mathrm{d}}{\mathrm{d}t}(OP)^2$ will become negative for some value of t, but it not immediately clear that this will happen before the particle lands. The particle lands when $ut\sin\alpha - \frac{1}{2}gt^2 = 0 \implies t = \frac{2u\sin\alpha}{g}$.

The expression for $\dfrac{\mathrm{d}}{\mathrm{d}t}(OP)^2$ will be certainly be negative when the completed square part is equal to 0. This happens when $gt - \frac{3}{2}u\sin\alpha = 0 \implies t = \frac{3u\sin\alpha}{2g}$. This is less than $t = \frac{2u\sin\alpha}{g}$, and so if $\sin\alpha > \frac{2\sqrt{2}}{3}$ then the distance OP will decrease at some point before it lands.

In this question we don't need to find the range of t for which OP is decreasing, just show that there exists a time before the particle lands where OP is decreasing.

Part (ii) - 9 marks

The position vector of P is given by:

$$\mathbf{p} = \begin{pmatrix} ut\cos\alpha \\ ut\sin\alpha - \frac{1}{2}gt^2 \end{pmatrix}$$

and the position vector of Q is given by:

$$\mathbf{q} = \begin{pmatrix} -vt \\ 0 \end{pmatrix}$$

The distance PQ satisfies:

$$\begin{aligned}(PQ)^2 &= (ut\cos\alpha + vt)^2 + \left(ut\sin\alpha - \tfrac{1}{2}gt^2\right)^2 \\ &= (u\cos\alpha + v)^2t^2 + u^2t^2\sin^2\alpha - ug\sin\alpha t^3 + \tfrac{1}{4}g^2t^4\end{aligned}$$

Differentiating gives:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(PQ)^2 &= 2t(u\cos\alpha + v)^2 + 2tu^2\sin^2\alpha - 3ug\sin\alpha t^2 + g^2t^3 \\ &= t\left(g^2t^2 - 3ugt\sin\alpha + 2(u\cos\alpha+v)^2 + 2u^2\sin^2\alpha\right) \\ &= t\Big[\left(gt - \tfrac{3}{2}u\sin\alpha\right)^2 - \tfrac{9}{4}u^2\sin^2\alpha + 2(u\cos\alpha+v)^2 + 2u^2\sin^2\alpha\Big] \\ &= t\Big[\left(gt - \tfrac{3}{2}u\sin\alpha\right)^2 + 2(u\cos\alpha+v)^2 - \tfrac{1}{4}u^2\sin^2\alpha\Big]\end{aligned}$$

Therefore the distance PQ will be increasing throughout the flight if:

$$\begin{aligned}2(u\cos\alpha+v)^2 - \tfrac{1}{4}u^2\sin^2\alpha &> 0 \\ 8(u\cos\alpha+v)^2 &> u^2\sin^2\alpha \\ 2\sqrt{2}(u\cos\alpha+v) &> u\sin\alpha \\ 2\sqrt{2}v &> u\sin\alpha - 2\sqrt{2}u\cos\alpha \\ 2\sqrt{2}v &> (\sin\alpha - 2\sqrt{2}\cos\alpha)u\end{aligned}$$

The first time I attempted this question I expanded $(u\cos\alpha+v)^2$ and simplified the $u^2\cos^2\alpha$ and $u^2\sin^2\alpha$ terms. This was not very helpful.

Question 10

10 A small light ring is attached to the end A of a uniform rod AB of weight W and length $2a$. The ring can slide on a rough horizontal rail.

One end of a light inextensible string of length $2a$ is attached to the rod at B and the other end is attached to a point C on the rail so that the rod makes an angle of θ with the rail, where $0 < \theta < 90°$. The rod hangs in the same vertical plane as the rail.

A force of kW acts vertically downwards on the rod at B and the rod is in equilibrium.

(i) You are given that the string will break if the tension T is greater than W. Show that (assuming that the ring does not slip) the string will break if

$$2k + 1 > 4 \sin \theta.$$

(ii) Show that (assuming that the string does not break) the ring will slip if

$$2k + 1 > (2k + 3)\mu \tan \theta,$$

where μ is the coefficient of friction between the rail and the ring.

(iii) You are now given that $\mu \tan \theta < 1$.

Show that, when k is increased gradually from zero, the ring will slip before the string breaks if

$$\mu < \frac{2 \cos \theta}{1 + 2 \sin \theta}.$$

Examiner's report

As with so many questions, the big stumbling block for students was drawing a good diagram from the information, including all the relevant forces.

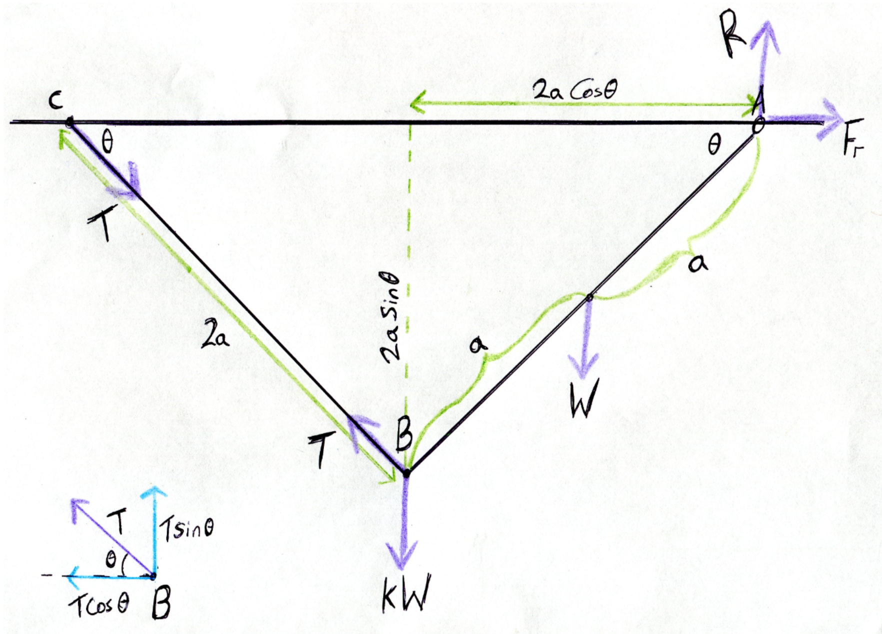
With "show that" questions it is beholden on candidates to explain their working. Equations which just appear and lead to the correct answer are not sufficient. In mechanics, it would be very helpful for students to say, for example, "Taking moments about point A for the rod" or "Resolving for the string-rod system vertically" to give some sense of where an equation arises.

The flow of logic is a fundamental idea in mathematics, but it was clear in this question that it was not familiar to the vast majority of students. The questions effectively asked "if *given condition* show that *mechanical outcome*". Most students reversed this to show that "if *mechanical outcome* then *given condition*". In this question, most arguments were reversible, but it still demonstrated a fundamental misunderstanding of what was being asked.

The other issue which flummoxed students was dealing with inequalities. There are different rules of algebra associated with inequalities and this is something which is frequently tested in STEP. Candidates would benefit from thinking carefully about things like when can one inequality be substituted into another, or when can an inequality be squared. The intuition from equalities was too often applied without thinking

Solution

First thing to do is to draw a diagram showing the forces. Note that the rod and string are the same length so ABC is an isosceles triangle.



I have also resolved the tension at B into horizontal and vertical components which may be useful later.

The layout below is not how I first attempted the question — I started by equating vertical and horizontal forces on the rod, before realising that I needed to consider moments as well. You do not need to rewrite solutions to make them "fit" the parts of the question, instead label the equations when you have written them down and then in part **(ii)** you can refer to equations (1) and (2) etc.

Part (i) - 6 marks

Taking moments about A we have:

$$\begin{aligned} W \times a\cos\theta + kW \times 2a\cos\theta &= T\sin\theta \times 2a\cos\theta + T\cos\theta \times 2a\sin\theta \\ W(1+2k)\cos\theta &= 4T\sin\theta\cos\theta \\ (1+2k)W &= 4T\sin\theta \qquad \text{as } \cos\theta \neq 0 \\ T &= \frac{1+2k}{4\sin\theta}W \end{aligned}$$

The string will break if $T > W$, i.e. the string will break if $\dfrac{1+2k}{4\sin\theta} > 1 \implies 1+2k > 4\sin\theta$.

Part (ii) - 6 marks

Resolving forces on the rod horizontally gives:

$$F_r = T\cos\theta$$

and vertically:

$$kW + W = R + T\sin\theta$$

The ring will slip if $T\cos\theta > F_{\text{max}} = \mu R$, i.e. it will slip if:

$$T\cos\theta > \mu((k+1)W - T\sin\theta)$$

We can use the relationship between T and W found by considering moments in part **(i)** to eliminate W. The ring will slip if:

$$\begin{aligned} T\cos\theta &> \mu\left((k+1)\frac{4\sin\theta}{1+2k}T - T\sin\theta\right) \\ 1 &> \mu\tan\theta\left(\frac{4(k+1)}{1+2k} - 1\right) && \text{note } \cos\theta > 0 \\ 1+2k &> \mu\tan\theta(4(k+1) - (1+2k)) && \text{note } 1+2k > 0 \\ 1+2k &> \mu\tan\theta(2k+3) \end{aligned}$$

Part (iii) - 8 marks

This part asks you to consider what happens as k is increased, so it might be an idea to make k the subject of each inequality.

The string breaks when $k > \dfrac{4\sin\theta - 1}{2}$ and the ring slips when:

$$\begin{aligned} 1+2k &> \mu\tan\theta(2k+3) \\ 2k(1-\mu\tan\theta) &> 3\mu\tan\theta - 1 \\ k &> \frac{3\mu\tan\theta - 1}{2(1-\mu\tan\theta)} \end{aligned}$$

Note this is ok as we are told that $\mu\tan\theta < 1$

So the ring will slip before the string breaks if:

$$\begin{aligned} \frac{3\mu\tan\theta - 1}{2(1-\mu\tan\theta)} &< \frac{4\sin\theta - 1}{2} \\ 3\mu\tan\theta - 1 &< (4\sin\theta - 1)(1-\mu\tan\theta) \end{aligned}$$

Note this is ok as we are told that $\mu\tan\theta < 1$

$$\begin{aligned} 3\mu\tan\theta - 1 &< 4\sin\theta - 4\mu\sin\theta\tan\theta - 1 + \mu\tan\theta \\ 2\mu\tan\theta(1+2\sin\theta) &< 4\sin\theta \\ \mu(1+2\sin\theta) &< 2\cos\theta \\ \mu &< \frac{2\cos\theta}{1+2\sin\theta} && \text{ok as } 1+2\sin\theta > 0 \end{aligned}$$

as required.

Question 11

11 **(i)** The three integers n_1, n_2 and n_3 satisfy $0 < n_1 < n_2 < n_3$ and $n_1 + n_2 > n_3$. Find the number of ways of choosing the pair of numbers n_1 and n_2 in the cases $n_3 = 9$ and $n_3 = 10$.

Given that $n_3 = 2n + 1$, where n is a positive integer, write down an expression (which you need not prove is correct) for the number of ways of choosing the pair of numbers n_1 and n_2. Simplify your expression.

Write down and simplify the corresponding expression when $n_3 = 2n$, where n is a positive integer.

(ii) You have N rods, of lengths 1, 2, 3, ... , N (one rod of each length). You take the rod of length N, and choose two more rods at random from the remainder, each choice of two being equally likely. Show that, in the case $N = 2n+1$ where n is a positive integer, the probability that these three rods can form a triangle (of non-zero area) is

$$\frac{n-1}{2n-1}.$$

Find the corresponding probability in the case $N = 2n$, where n is a positive integer.

(iii) You have $2M + 1$ rods, of lengths 1, 2, 3, ... , $2M + 1$ (one rod of each length), where M is a positive integer. You choose three at random, each choice of three being equally likely. Show that the probability that the rods can form a triangle (of non-zero area) is

$$\frac{(4M+1)(M-1)}{2(2M+1)(2M-1)}.$$

Note: $\sum_{k=1}^{K} k^2 = \frac{1}{6}K(K+1)(2K+1)$.

Examiner's report

Candidates got the correct number of pairs in the special cases $n_3 = 9$ and $n_3 = 10$ but sometimes the working was very unclear. A large majority found the expressions for general n, the most common error being a shift $n \to n + 1$ in the answer.

Those who could obtain the result given for odd n in part (ii) were generally able to find the corresponding result for even n too. A common error was to double count the number of pairs of rods and not to double the number of pairs which made a triangle. Many candidates failed to explain why the conditions of part (i) were relevant for forming triangles.

The most successful candidates in part (iii) counted the number of triples which make a triangle using a sum, and divided by $\binom{2M+1}{3}$, while those who conditioned on the largest rod and used conditional probability did less well. A common conceptual error was to assume that each integer was equally likely to appear as the largest rod, and candidates making this assumption lost many marks. Otherwise, algebraic errors were the most common. Candidates should remember that when an answer is given in the question, they need to take care to fully justify their answers.

Solution

For the first part of this question, working systematically is important and will help you find the general expressions.

Part (i) - 7 marks

If $n_3 = 9$ then we need to find two numbers less than $n_3 = 9$ which add to more than 9. Starting with the lowest possible value of n_1 (remembering that $n_1 < n_2$) then we have:

- $n_1 = 1$ means $n_2 \geqslant 9$ which is not possible
- $n_1 = 2$ then $n_2 = 8$
- $n_1 = 3$ then $n_2 = 7$ or $n_2 = 8$
- $n_1 = 4$ then $n_2 = 6, 7, 8$
- $n_1 = 5$ then $n_2 = 6, 7, 8$
- $n_1 = 6$ then $n_2 = 7, 8$
- $n_1 = 7$ then $n_2 = 8$
- $n_1 = 8$ then $n_2 \geqslant 9$ which is not possible

So there are 12 ways of doing this.

If $n_3 = 10$ we have:

- $n_1 = 1$ then $n_2 \geqslant 10$, not possible
- $n_1 = 2$ then $n_2 = 9$
- $n_1 = 3$ then $n_2 = 8, 9$
- $n_1 = 4$ then $n_2 = 7, 8, 9$
- $n_1 = 5$ then $n_2 = 6, 7, 8, 9$
- $n_1 = 6$ then $n_2 = 7, 8, 9$
- $n_1 = 7$ then $n_2 = 8, 9$
- $n_1 = 8$ then $n_2 = 9$
- $n_1 = 9$ then $n_2 \geqslant 10$, not possible

So there are 16 ways.

The formulae for general n_3 do not need any justification, so you can use the examples as a guide. It is also helpful if you know the formula for the sum of the integers $1+2+\cdots+n = \frac{1}{2}n(n+1)$.

When $n_3 = 2n+1$ the number of ways to pick n_1 and n_2 is:

$$(1+2+3+\cdots+(n-1)) \times 2 = n(n-1)$$

When $n_3 = 2n$ the number of ways is:

$$(1+2+3+\cdots(n-2)) \times 2 + (n-1) = (n-1)(n-2) + (n-1) = (n-1)^2$$

A quick check shows us that these give the correct answers when $n_3 = 9$ $(2n+1 = 9 \implies n = 4)$ and $n_3 = 10$.

Part (ii) - 4 marks

In order to make a triangle, you need the sum of the two smaller rods to be greater than N (i.e. we need to pick n_1 and n_2 so that $n_1 + n_2 > N$. If $N = 2n+1$ then the number of ways of doing this is $n(n-1)$. The total number of ways of picking 2 rods from the $N-1$ left is $\dfrac{(N-1)(N-2)}{2} = \dfrac{2n(2n-1)}{2} = n(2n-1)$ (dividing by 2 as the order of picking doesn't matter, i.e. $(1,2)$ is the same as $(2,1)$). The probability of making a triangle is therefore:

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

When $N = 2n$ the number of ways of picking two smaller rods is $\dfrac{(2n-1)(2n-2)}{2} = (n-1)(2n-1)$. The probability of being able to make a triangle is:

$$\frac{(n-1)^2}{(n-1)(2n-1)} = \frac{n-1}{2n-1}\,.$$

Part (iii) - 9 marks

Now we are picking all three rods rather than just two. The total number of ways of picking three rods is $\dfrac{(2M+1)(2M)(2M-1)}{6}$ (again, order doesn't matter).

Let the largest of the three rods be k, and so we have $3 \leqslant k \leqslant 2M+1$. For each value of k we can find the number of ways of picking two rod lengths so that the three rods make a triangle:

- If $k = 3 = 2\times 1+1$ there is no way to pick the rods (since $1+2 = 3$)
- If $k = 4 = 2\times 2$ there is one way to pick the rods $(2,3)$
- If $k = 5 = 2\times 2+1$ there are two ways to pick the rods $(2,4$ and $3,4)$

Using the number of ways found in part **(i)**, we know that if the largest rod is $2n+1$ there are $n(n-1)$ ways to pick two smaller ones so that you can make up a triangle, and if the largest rod is $2n$ then there are $(n-1)^2$ ways.

We need to sum up all the ways for each possible length of the largest rod. Separating out the even and odd numbers we have:

$$\text{Number of ways} = \sum_{n=1}^{M} n(n-1) + \sum_{n=2}^{M} (n-1)^2$$

The first sum here is for when the largest rod is an odd length (so $k = 3, 5, 7, \cdots, 2M+1$) and the second sum is for when the largest rod is an even length ($k = 4, 6, \cdots, 2M$). Note that there are M terms in the first sum and $M-1$ in the second given a total of $2M-1$ which is what we would expect as the largest rod can be $3, 4, \cdots, 2M+1$. Note that $(1-1)^2 = 0$, so we can actually start the second sum from $n = 1$ to make things easier.[8]

$$\begin{aligned}
\text{Number of ways} &= \sum_{n=1}^{M} n(n-1) + \sum_{n=1}^{M} (n-1)^2 \\
&= \sum_{n=1}^{M} \left[n(n-1) + (n-1)^2\right] \\
&= \sum_{n=1}^{M} (n-1)(2n-1) \\
&= \sum_{n=1}^{M} (2n^2 - 3n + 1) \\
&= 2\sum_{n=1}^{M} n^2 - 3\sum_{n=1}^{M} n + \sum_{n=1}^{M} 1 \\
&= \tfrac{2}{6}M(M+1)(2M+1) - \tfrac{3}{2}M(M+1) + M \\
&= \tfrac{1}{6}M\Big[(4M^2 + 6M + 2) - 9(M+1) + 6\Big] \\
&= \tfrac{1}{6}M(4M^2 - 3M - 1) \\
&= \tfrac{1}{6}M(4M+1)(M-1)
\end{aligned}$$

The probability that we make a triangle is therefore:

$$\frac{\cancel{\tfrac{1}{6}}\cancel{M}(4M+1)(M-1)}{\cancel{\tfrac{1}{6}}(2M+1)(2\cancel{M})(2M-1)} = \frac{(4M+1)(M-1)}{2(2M+1)(2M-1)}$$

as required.

[8] This is an example of "adding zero creatively" which can be quite useful!

Question 12

12 The random variable X has the probability density function on the interval $[0,1]$:

$$\mathrm{f}(x) = \begin{cases} nx^{n-1} & 0 \leqslant x \leqslant 1, \\ 0 & \text{elsewhere,} \end{cases}$$

where n is an integer greater than 1.

(i) Let $\mu = \mathrm{E}(X)$. Find an expression for μ in terms of n, and show that the variance, σ^2, of X is given by

$$\sigma^2 = \frac{n}{(n+1)^2(n+2)}.$$

(ii) In the case $n = 2$, show without using decimal approximations that the interquartile range is less than 2σ.

(iii) Write down the first three terms and the $(k+1)$th term (where $0 \leqslant k \leqslant n$) of the binomial expansion of $(1+x)^n$ in ascending powers of x.

By setting $x = \dfrac{1}{n}$, show that μ is less than the median and greater than the lower quartile.

Note: You may assume that

$$1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots < 4.$$

a

Examiner's report

Almost all candidates who attempted this question were able to achieve full marks on the first part. In the second part, the values of the interquartile range and 2σ were generally found correctly, but then many candidates did not realise that squaring would eliminate the square roots from the values to be compared.

In the final part of the question some candidates failed to recognise that the $(k+1)^{th}$ term of the expansion was the term in x^k and gave the term in x^{k+1} instead. A good number were successful in finding the lower quartile and the median, but only a minority realised that $\mu^{-n} = \left(1 + \frac{1}{n}\right)^n$. Those that did were more successful in proving that $\mu > \left(\frac{1}{4}\right)^n$ than $\mu < \left(\frac{1}{2}\right)^n$.

Solution

Note that we have $\int_0^1 nx^{n-1}\,\mathrm{d}x = [x^n]_0^1 = 1$, which is encouraging. This question asks about the interquartile range. In the same way that the median has 50% of the distribution lying above and below it, the quartiles split the distribution into 4. The lower quartile (which has 25% of the distribution below it) is sometimes given the symbol Q_1 and the upper quartile is Q_3 (Q_2 being the median). The interquartile range measures how spread out the middle 50% of the distribution is and is given by $Q_3 - Q_1$.

Part (i) - 5 marks

We have:

$$\begin{aligned}\mu &= \int_0^1 x \times nx^{n-1} \\ &= \int_0^1 nx^n\,\mathrm{d}x \\ &= \left[\frac{n}{n+1}x\right]_0^1 \\ &= \frac{n}{n+1}\end{aligned}$$

We also have:

$$\begin{aligned}\mathrm{E}(X^2) &= \int_0^1 nx^{n+1}\,\mathrm{d}x \\ &= \left[\frac{n}{n+2}x^{n+2}\right]_0^1 \\ &= \frac{n}{n+2}\end{aligned}$$

So using $\sigma^2 = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$ we have:

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

Part (ii) - 7 marks

When $n = 2$ the p.d.f. is given by $\mathrm{f}(x) = 2x$ for $0 \leqslant x \leqslant 1$. The lower quartile Q_1 is the point such that:

$$\int_0^{Q_1} \mathrm{f}(x)\,\mathrm{d}x = \frac{1}{4}$$

So we have:

$$\begin{aligned}\int_0^{Q_1} 2x\,\mathrm{d}x &= \frac{1}{4}\\ \left[x^2\right]_0^{Q_1} &= \frac{1}{4}\\ {Q_1}^2 &= \frac{1}{4}\\ Q_1 &= \frac{1}{2}\end{aligned}$$

Similarly, to find the upper quartile Q_3 use $\displaystyle\int_0^{Q_3} 2x\,\mathrm{d}x = \frac{3}{4}$ to get ${Q_3}^2 = \dfrac{3}{4} \implies Q_3 = \dfrac{\sqrt{3}}{2}$. The interquartile range is given by $Q_3 - Q_1 = \dfrac{\sqrt{3}-1}{2}$.

When $n = 2$ we have $\sigma^2 = \dfrac{2}{3^2 \times 4}$ so $\sigma = \dfrac{\sqrt{2}}{3 \times 2}$ and so $2\sigma = \dfrac{\sqrt{2}}{3}$. We want to show that the interquartile range is less than 2σ so we want to show that $\dfrac{\sqrt{2}}{3} > \dfrac{\sqrt{3}-1}{2}$ or equivalently that $\dfrac{\sqrt{2}}{3} - \dfrac{\sqrt{3}-1}{2} > 0$[9]. Consider:

$$\frac{\sqrt{2}}{3} - \frac{\sqrt{3}-1}{2} = \frac{2\sqrt{2} - 3\sqrt{3} + 3}{6}$$

So we want to show that $2\sqrt{2} + 3 > 2\sqrt{3}$. Squaring both sides gives $8 + 9 + 12\sqrt{2} > 12$, which is true so we have $2\sqrt{2} + 3 > 2\sqrt{3}$ and so the IQR is less than 2σ.

Part (iii) - 8 marks

We have:

$$1 + nx + \frac{n(n-1)}{2}x^2 + \cdots + \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}x^k + \cdots$$

We have $\mu = \dfrac{n}{n+1}$, and Q_1 is given by $\displaystyle\int_0^{Q_1} nx^{n-1}\,\mathrm{d}x = \frac{1}{4} \implies (Q_1)^2 = \frac{1}{4}$ and so $Q_1 = \left(\dfrac{1}{4}\right)^{\frac{1}{n}}$.

Similarly the median is given by $Q_2 = \left(\dfrac{2}{4}\right)^{\frac{1}{n}} = \left(\dfrac{1}{2}\right)^{\frac{1}{n}}$.

[9]When trying to show that an inequality is true it is often best to rearrange so that you are trying to show that something is positive or negative.

The question says "by setting $x = \frac{1}{n}$", so we should substitute $x = \frac{1}{n}$ in the binomial expansion to get:

$$\left(1+\frac{1}{n}\right)^n = 1 + n \times \frac{1}{n} + \frac{n(n-1)}{2} \times \left(\frac{1}{n}\right)^2 + \cdots$$

The next stage is to try to work out how this could be helpful. Note that $\dfrac{1}{\mu} = \dfrac{n+1}{n} = 1 + \dfrac{1}{n}$, and so the expansion above is the same as $\left(\dfrac{1}{\mu}\right)^n$.

We want to show that $\mu > Q_1$, i.e, we want to show that $\mu > \left(\dfrac{1}{4}\right)^{\frac{1}{n}}$, and we also want to show that $\mu < Q_2 = \left(\dfrac{1}{2}\right)^{\frac{1}{n}}$.

Starting with the median this condition can be re-written as:

$$\begin{aligned} \mu &< \left(\frac{1}{2}\right)^{\frac{1}{n}} \\ \mu^n &< \frac{1}{2} \\ \left(\frac{1}{\mu}\right)^n &> 2 \\ \left(1+\frac{1}{n}\right)^n &> 2 \end{aligned}$$

This is true because:

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

so we have $\mu < \left(\dfrac{1}{2}\right)^{\frac{1}{n}}$.

For the lower quartile condition we have:

$$\begin{aligned} \mu &> \left(\frac{1}{4}\right)^{\frac{1}{n}} \\ \mu^n &> \frac{1}{4} \\ \left(\frac{1}{\mu}\right)^{\frac{1}{n}} &< 4 \\ \left(1+\frac{1}{n}\right)^{\frac{1}{n}} &< 4 \end{aligned}$$

We have:

$$\begin{aligned}\left(1+\frac{1}{n}\right)^{\frac{1}{n}} &= 1+1+\frac{n(n-1)}{2!n^2}+\frac{n(n-1)(n-2)}{3!n^3}+\cdots+\frac{n(n-1)\cdots(n-k+1)}{k!n^k}+\cdots \\ &< 1+1+\frac{1}{2!}+\frac{1}{3!}+\cdots+\frac{1}{k!}+\cdots \\ &< 4 \qquad \text{using the given result}\end{aligned}$$

Therefore we have:

$$\left(\frac{1}{4}\right)^{\frac{1}{n}} < \mu < \left(\frac{1}{2}\right)^{\frac{1}{n}}$$

as required.