STEP Support Programme

2020 STEP 2 Worked Paper

General comments - In 2020 the STEP papers were delivered remotely, and were only available to students with offers involving STEP from Cambridge, Warwick or Imperial.

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 Cambridge Assessment Admissions Testing website. These are the general comments for the STEP 2020 exam from the Examiner's report:

"*There were just over 800 entries for this paper, and good solutions were seen to all of the questions. Candidates should be aware of the need to provide clear explanations of their reasoning throughout the paper (and particularly in questions where the result to be shown is given in the question). Short explanatory comments at key points in solutions can be very helpful in this regard, as can clearly drawn diagrams of the situation described in the question. The paper included a few questions where a statement of the form "A if and only if B" needed to be proven – candidates should be aware of the meaning of such statements and make sure that both directions of the implication are covered clearly.*
In general, candidates who performed better on the questions in this paper recognised the relationships between the different parts of each question and were able to adapt methods used in earlier parts when working on the later sections of the question."

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

Question 1 2
Question 2 6
Question 3 13
Question 4 17
Question 5 21
Question 6 24
Question 7 28
Question 8 31
Question 9 35
Question 10 39
Question 11 43
Question 12 47

Question 1

1 **(i)** Use the substitution $x = \frac{1}{1-u}$, where $0 < u < 1$, to find in terms of x the integral

$$\int \frac{1}{x^{\frac{3}{2}}(x-1)^{\frac{1}{2}}}\,\mathrm{d}x \quad \text{(where } x > 1\text{)}.$$

(ii) Find in terms of x the integral

$$\int \frac{1}{(x-2)^{\frac{3}{2}}(x+1)^{\frac{1}{2}}}\,\mathrm{d}x \quad \text{(where } x > 2\text{)}.$$

(iii) Show that

$$\int_2^{\infty} \frac{1}{(x-1)(x-2)^{\frac{1}{2}}(3x-2)^{\frac{1}{2}}}\,\mathrm{d}x = \tfrac{1}{3}\pi.$$

Examiner's report

This was the most popular question on the paper, and also the one on which candidates performed the best.

In general, candidates were confident in applying the substitution given in part **(i)** and completed the integration correctly, although there were a number of solutions in which careless errors were seen.

Successful completion of the remaining two parts required candidates to understand the reason why the substitution suggested in part **(i)** was helpful, and so candidates who continued to apply the same substitution from part **(i)** to later parts were often unable to make useful progress, particularly on part **(iii)**.

In part **(ii)** many candidates took the approach of making a sequence of two substitutions to reach the answer. Attempts involving just one substitution were more likely to include errors, although a number of these were also completed successfully. In some cases candidates recognised that the integral was going to be reduced to a form similar to that in part **(i)**, but then did not obtain exactly the correct function to be integrated.

In part **(iii)** candidates who were continuing with the same substitution as in part **(i)** often spent a lot of time rearranging the function to be integrated without success. Some attempts to apply partial fractions were seen in this section despite the fact that the factors were square roots of linear expressions. Many of those who were able to adapt the substitution from part **(i)** did recognise the form of the new function to be integrated and often selected an appropriate substitution.

Solution

(i) Using $x = \dfrac{1}{1-u} = (1-u)^{-1}$ we have $\dfrac{\mathrm{d}x}{\mathrm{d}u} = (1-u)^{-2}$. We also have $x - 1 = \dfrac{1}{1-u} - 1 = \dfrac{1-(1-u)}{1-u} = \dfrac{u}{1-u}$.

Substituting for x, $x-1$ and $\mathrm{d}x$ we have:

$$\begin{aligned}\int \frac{1}{x^{\frac{3}{2}}(x-1)^{\frac{1}{2}}}\,\mathrm{d}x &= \int \frac{1}{(1-u)^{-\frac{3}{2}}u^{\frac{1}{2}}(1-u)^{-\frac{1}{2}}} \times (1-u)^{-2}\mathrm{d}u \\ &= \int \frac{1}{\cancel{(1-u)^{-2}}u^{\frac{1}{2}}} \times \cancel{(1-u)^{-2}}\mathrm{d}u \\ &= \int u^{-\frac{1}{2}}\,\mathrm{d}u \\ &= 2u^{\frac{1}{2}} + c\end{aligned}$$

We need to find u in terms of x to complete the integration. Rearranging gives:

$$\begin{aligned}(1-u)x &= 1 \\ x - 1 &= ux \\ \frac{x-1}{x} &= u\end{aligned}$$

and so the final answer is $2\left(\dfrac{x-1}{x}\right)^{\frac{1}{2}} + c$.

(ii) Here we need to find a suitable substitution. Comparing to before it looks very similar to the integral in part **(i)**. Starting with a substitution of $t = x - 2$ (chosen so that we have $t^{\frac{3}{2}}$ in the denominator) we have:

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

This now looks more promising. We need to find another substitution which will have a similar affect to the one in part **(i)**. My first thought was to try $t = \dfrac{1}{u+3}$, but this did not give something nice when considering the $(t+3)^{\frac{1}{2}}$ term.

The next thing I tried was the same substitution as in part **(i)**, i.e. $t = \dfrac{1}{1-u}$. This gives $t + 3 = \dfrac{1}{1-u} + 3 = \dfrac{1+3-3u}{1-u}$, and the constant terms on the numerator still do not cancel out.

However, this led me to consider $t = \dfrac{a}{1-u}$ which gives $t + 3 = \dfrac{a+3-3u}{1-u}$, and so if I take $a = -3$ something useful might happen! If $a = -3$ we have $t = \dfrac{-3}{1-u} = \dfrac{3}{u-1}$.

You might have spotted a suitable substitution quicker than this!

Using $t = \dfrac{3}{u-1}$ we have $t + 3 = \dfrac{3u}{u-1}$ and $\dfrac{\mathrm{d}t}{\mathrm{d}u} = -3(u-1)^{-2}$. Substituting these we have:

$$\begin{aligned}
\int \frac{1}{t^{\frac{3}{2}}(t+3)^{\frac{1}{2}}}\,\mathrm{d}t &= \int \frac{1}{3^{\frac{3}{2}}(u-1)^{\frac{-3}{2}} \times (3u)^{\frac{1}{2}}(u-1)^{\frac{-1}{2}}} \times -3(u-1)^{-2}\mathrm{d}u \\
&= \int \frac{1}{3^2 u^{\frac{1}{2}}\cancel{(u-1)^{-2}}} \times -3\cancel{(u-1)^{-2}}\mathrm{d}u \\
&= \frac{1}{3}\int u^{-\frac{1}{2}}\,\mathrm{d}u \\
&= -\frac{1}{3} \times 2u^{\frac{1}{2}} + c \\
&= -\frac{2}{3}\left(\frac{t+3}{t}\right)^{\frac{1}{2}} + c \\
&= -\frac{2}{3}\left(\frac{x+1}{x-2}\right)^{\frac{1}{2}} + c
\end{aligned}$$

The penultimate line here uses $t = \dfrac{3}{u-1} \implies tu - t = 3 \implies u = \dfrac{t+3}{t}$, and the last line uses $t = x - 2$.

(iii) The substitutions so far have been of the form of rational expressions where the top and bottom have been linear (or constant) polynomials of u. Lets consider $x = \dfrac{a+u}{b+u}$ and see if this can help us find a suitable substitution.

If $x = \dfrac{a+u}{b+u}$ then:

$$\begin{aligned}
x - 1 &= \frac{a-b}{b+u} \\
x - 2 &= \frac{a-2b-u}{b+u} \\
3x - 2 &= \frac{3a-2b+u}{b+u} \qquad \text{and} \\
\frac{\mathrm{d}x}{\mathrm{d}u} &= \frac{b-a}{(b+u)^2} \qquad \text{(using the quotient rule).}
\end{aligned}$$

This gives:

$$\begin{aligned}
\int \frac{1}{(x-1)(x-2)^{\frac{1}{2}}(3x-2)^{\frac{1}{2}}}\,\mathrm{d}x &= \int \frac{(b+u)^2}{(a-b)(a-2b-u)^{\frac{1}{2}}(3a-2b+u)^{\frac{1}{2}}} \times -\frac{(a-b)}{(b+u)^2}\,\mathrm{d}u \\
&= \int \frac{\cancel{(b+u)^2}}{\cancel{(a-b)}(a-2b-u)^{\frac{1}{2}}(3a-2b+u)^{\frac{1}{2}}} \times -\frac{\cancel{(a-b)}}{\cancel{(b+u)^2}}\,\mathrm{d}u \\
&= -\int \frac{1}{(a-2b-u)^{\frac{1}{2}}(3a-2b+u)^{\frac{1}{2}}}\,\mathrm{d}u
\end{aligned}$$

Note that we cannot have $a = b$. This looks as if it has been quite useful, ans the next stage is to try putting in a couple of values for a and b. Take $a = 1$ and $b = 0$ (which is ok as $a \neq b$, and these are the simplest values I can think of). This means we are using the substitution $x = \dfrac{1+u}{u}$, and the limits are:

$$\begin{aligned} x = 2 &\implies \frac{1+u}{u} = 2 \\ &\implies 1 + u = 2u \\ &\implies u = 1 \\ x \to \infty &\implies \frac{1+u}{u} \to \infty \\ &\implies u \to 0 \end{aligned}$$

Using these limits, and the values $a = 1$, $b = 0$, gives:

$$\begin{aligned} -\int_{x=2}^{x\to\infty} \frac{1}{(a-2b-u)^{\frac{1}{2}}(3a-2b+u)^{\frac{1}{2}}}\,\mathrm{d}u &= -\int_1^0 \frac{1}{(1-u)^{\frac{1}{2}}(3+u)^{\frac{1}{2}}}\,\mathrm{d}u \\ &= \int_0^1 \frac{1}{(3-2u-u^2)^{\frac{1}{2}}}\,\mathrm{d}u \end{aligned}$$

Note that the negative sign has been used to flip the limits. Completing the square inside the square root gives:

$$\int_0^1 \frac{1}{(3-2u-u^2)^{\frac{1}{2}}}\,\mathrm{d}u = \int_0^1 \frac{1}{(4-(u+1)^2)^{\frac{1}{2}}}\,\mathrm{d}u$$

This is now in a form where we can use a sin or cos substitution. Take $u + 1 = 2\sin t$, which gives $\dfrac{\mathrm{d}u}{\mathrm{d}t} = 2\cos t$. The limits becomes $u = 0 \implies \sin t = \frac{1}{2} \implies t = \frac{\pi}{6}$ and $u = 1 \implies \sin t = 1 \implies t = \frac{\pi}{2}$. This gives:

$$\begin{aligned} \int_0^1 \frac{1}{(4-(u+1)^2)^{\frac{1}{2}}}\,\mathrm{d}u &= \int_{\frac{\pi}{6}}^{\frac{\pi}{2}} \frac{1}{(4-4\sin^2 t)^{\frac{1}{2}}} \times 2\cos t\,\mathrm{d}t \\ &= \int_{\frac{\pi}{6}}^{\frac{\pi}{2}} 1\,\mathrm{d}t \\ &= \frac{\pi}{2} - \frac{\pi}{6} = \frac{\pi}{3} \end{aligned}$$

Question 1 is supposed to be “accessible”, but that does not mean easy! In this question it is relatively easy to see what is required, i.e. you need to find and use substitutions in order to evaluate the integral, but it is still quite a long question, and you might have found that you tried a few substitutions which did not help massively.

Question 2

2 The curves C_1 and C_2 both satisfy the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{kxy - y}{x - kxy},$$

where $k = \ln 2$.

All points on C_1 have positive x and y co-ordinates and C_1 passes through $(1, 1)$. All points on C_2 have negative x and y co-ordinates and C_2 passes through $(-1, -1)$.

(i) Show that the equation of C_1 can be written as $(x - y)^2 = (x + y)^2 - 2^{x+y}$.

Determine a similar result for curve C_2.

Hence show that $y = x$ is a line of symmetry of each curve.

(ii) Sketch on the same axes the curves $y = x^2$ and $y = 2^x$, for $x \geqslant 0$. Hence show that C_1 lies between the lines $x + y = 2$ and $x + y = 4$.

Sketch curve C_1.

(iii) Sketch curve C_2.

Examiner's report

This was a popular question and candidates in general achieved good marks. Most candidates approached part **(i)** with reasonable confidence and followed the question's intended path of separating variables and integrating directly. Of these, the majority failed to consider the integral of $\frac{1}{x}$ as $\ln|x|$. This was entirely understandable for work on C_1 as the question specifies that it lies entirely in the first quadrant i.e. where both x and y are positive, so $|x| = x$ etc. While $\ln(xy)$ was okay to deal with in the case where both x and y were negative, it would be good for candidates to be clear about the way in which the modulus function is being dealt with here.

Around 1 in 5 candidates took the alternate route of differentiating the given answer to show that it fitted the differential equation in the question. Unfortunately, not only is this much more demanding work, but almost all such attempts failed to realise the need to check the initial condition $x = y = 1$ as part of the solution. Candidates should be clear about the distinction between "show" and "verify" in such questions.

Additional note here — if a question wants you to substitute in something and check that it works, it will use the word "verify". Usually questions are designed so that "verifying" is not easier than actually solving, and in this case solving for C_1 will help you find C_2. There are some reasons why you might be given a "show that", one is so that you write the result in a useful form for a next part, and another is so that you can still gain marks for knowing how to do the next parts (it is very rare for the last part to be a "show that").

In part **(ii)**, the sketching of two fairly straightforward functions caused unexpected difficulties when it came to putting them together suitably on the same diagram; it was important to show that the two curves intersect twice and many sketches failed to have them crossing more than once, if at all. This made the subsequent reasoning and sketch of C_1 very awkward. Even for those who got this far entirely successfully, it was a common problem to find the sketch of C_1 drawn without the helpful guidance supplied within the question and any results arising from correct working to date. In particular, it was important for candidates to demonstrate the symmetry in the line $y = x$ and the restrictions provided by the lines $x + y = 2$ and $x + y = 4$, all of which really should have appeared on the diagram.

Many attempts petered out by the time of part **(iii)**, and very few candidates attempted to consider a graphical method (directly analogous to the method promoted in part **(ii)**) to show that the curve of C_2 was constrained by the line $x + y = -2$. By the time they came to draw this second solution to the original differential equation, many candidates had forgotten either or both of the given bits of information; namely, that there was symmetry in $y = x$ and that C_2 existed only in the 3rd-quadrant. Many candidates just assumed that C_2 was the reflection of C_1.

Candidates should be advised not to attempt to sketch curves by plotting points, as was seen in a number of cases. Instead, the information established in earlier parts of the question should be used to ensure that the key points are marked and that the shape is correct.

Solution

(i) Factorising the numerator and denominator and then separating variables gives:

$$\begin{aligned}
\frac{\mathrm{d}y}{\mathrm{d}x} &= \frac{y(kx-1)}{x(1-ky)} \\
\int \frac{1-ky}{y}\,\mathrm{d}y &= \int \frac{kx-1}{x}\,\mathrm{d}x \\
\int \frac{1}{y} - k\,\mathrm{d}y &= \int k - \frac{1}{x}\,\mathrm{d}x \\
\ln|y| - ky &= kx - \ln|x| + c \\
\ln|x| + \ln|y| &= k(x+y) + c \\
\ln(|x||y|) &= k(x+y) + c
\end{aligned}$$

We are told that all the points on curve C_1 have positive x and y coordinates, and all the points of C_2 have negative x and y coordinates. This means that in each case we have $|x||y| = xy$ (as if x and y are negative then $|x| = -x$ and $|y| = -y$ which gives $|x||y| = -x \times -y = xy$). Hence we can replace $\ln(|x||y|)$ with $\ln xy$[1].

[1]This is a lot more detail that was needed in an exam — a statement along the lines of "x and y are either both negative or both positive so $\ln(|x||y|) = \ln xy$" would have been absolutely fine.

We have:

$$\begin{aligned}\ln(xy) &= k(x+y)+c\\ xy &= \mathrm{e}^{k(x+y)+c}\\ &= \mathrm{e}^{c}\times e^{k(x+y)}\\ &= A\mathrm{e}^{k(x+y)}\\ &= A\mathrm{e}^{(x+y)\ln 2}\\ &= A\mathrm{e}^{\ln\left(2^{x+y}\right)}\\ &= A\times 2^{x+y}\end{aligned}$$

For C_1 the curve passes through $(1,1)$, which means that:

$$1 = A\times 2^2 \implies A = \frac{1}{4}\,.$$

For C_1 the equation is:

$$\begin{aligned}xy &= \frac{1}{4}\times 2^{x+y}\\ 4xy &= 2^{x+y}\\ x^2+2xy+y^2-x^2+2xy-y^2 &= 2^{x+y}\\ (x+y)^2-(x-y)^2 &= 2^{x+y}\\ \implies (x-y)^2 &= (x+y)^2-2^{x+y} \quad \text{as required}\end{aligned}$$

For C_2 the curve passes through $(-1,-1)$, so we have:

$$1 = A\times 2^{-2} \implies A = 4\,.$$

In a similar way to C_1, the equation of C_2 is:

$$\begin{aligned}xy &= 4\times 2^{x+y}\\ 4xy &= 16\times 2^{x+y}\\ (x+y)^2-(x-y)^2 &= 2^4\times 2^{x+y}\\ \implies (x-y)^2 &= (x+y)^2-2^{x+y+4}\end{aligned}$$

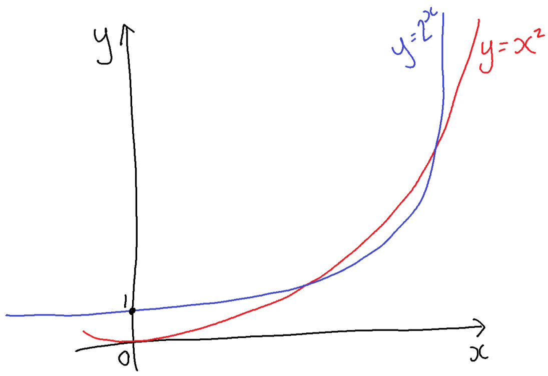
The last request in this part is to show that the curves have a line of symmetry in $y=x$. All you need to do here is to state that "as the equations are symmetrical in x and y[2] therefore the curves have a line of symmetry in $y=x$", or even just "Symmetrical in x, $y \implies y=x$ is a line of symmetry".

[2]This means that is we swap all the x's for y's and vice-versa the equation is unchanged.

(ii) The first request here seems a little odd, but it is actually connected to the RHS of the equation for C_1, which is $(x+y)^2 - 2^{x+y}$, so it has the form $z^2 - 2^z$. As $x \to \infty$, I know that $2^x > x^2$ (exponents beat polynomials is a general rule of thumb), and I know that when $x = 0$, $2^x = 1$ and $x^2 = 0$. It wasn't immediately clear to me that the graphs intersect, but if we take $x = 2$ we know that $2^x = x^2$ here, so they must intersect when $x = 2$.

Working out a few values of each will give you an idea of how the curves behave, and it will even show you where they intersect which is nice. However, try to avoid the urge to plot points and join them up!

After a couple of attempts at using a graphics tablet to sketch the graph, I ended up with this one, which is good enough! The $y = 2^x$ graph is rather exaggerated, and it does look as though it have a vertical asymptote, but the trend of the lines is clear.

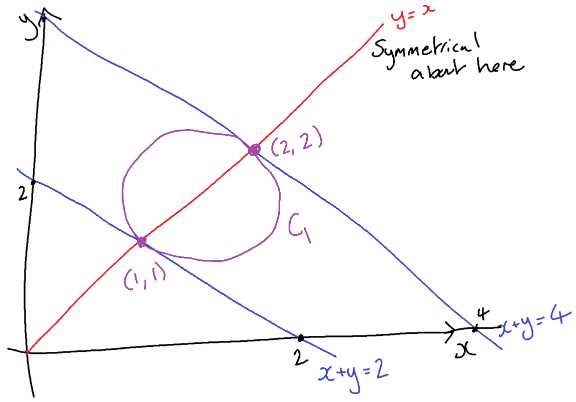


From this sketch, you can see that the two graphs intersect only twice, and these will be at the points $(2, 4)$ and $(4, 16)$ (as $4^2 = 16 = 2^4$). These are the only two points of intersection. We also know that $x^2 > 2^x$ for $2 < x < 4$ (there will be some negative values of x as well, but we are considering C_1 so only need to think about positive x).

We were asked to find the equation of C_1 in the form $(x-y)^2 = (x+y)^2 - 2^{x+y}$, and since we know that $(x-y)^2 \geqslant 0$ we must have $(x+y)^2 \geqslant 2^{x+y}$. From the paragraph above we know that this means that $2 \leqslant x+y \leqslant 4$, and so C_1 lies between the lines $x+y = 2$ and $x+y = 4$.

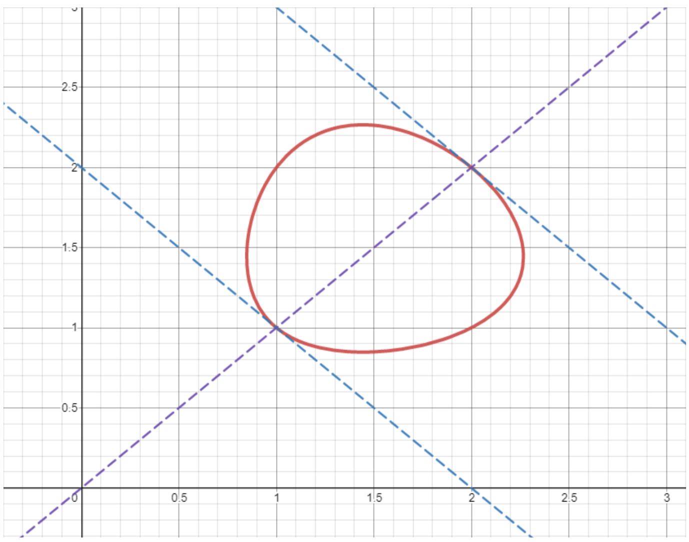
We also know that the curve passes through $(1, 1)$ and it is symmetrical about $y = x$. We also know that when it lies on $y = x$ we have $(x+y)^2 = 2^{x+y}$, which is only true when $x+y = 2$ and $x+y = 4$, so we also know that it intersects the line $y = x$ at the point $(2, 2)$.

This is enough to sketch a picture:



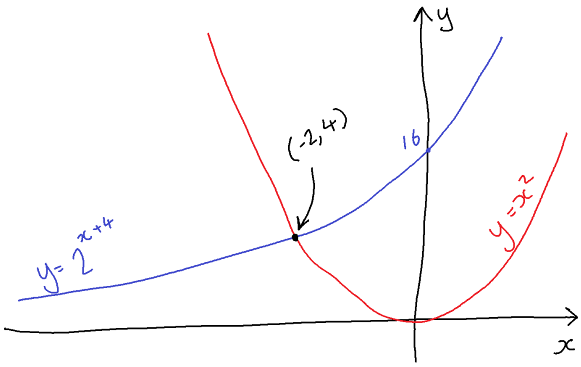
Note that I have included on the diagram a comment to say that it is symmetrical in $x = y$, and I have also marked on the two points where it meets $y = x$. I should have used a ruler for the axes and the straight lines — please do this in an examination! The question did not ask you to specifically find the points where it crosses the $y = x$ line, but it is a good idea to label or otherwise indicate what important points on the graph are (this also helps if your graph is a bit wobbly - it indicates to the examiner that you did intend the correct diagram!).

If you use Desmos to plot the graph you will find that it looks like the one below, slightly less circular than mine but otherwise fairly similar.



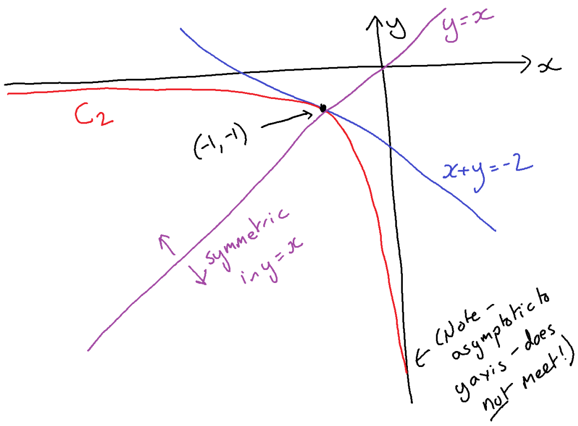
(iii) Here we are left to our own devices a little more, but it is a good idea to look back at the previous part and see if we can adapt the method.

The equation for C_2 is $(x-y)^2 = (x+y)^2 - 2^{x+y+4}$, and we know that C_2 is contained in the quadrant where both x and y are always negative. We know that we need $(x+y)^2 \geqslant 2^{x+y+4}$, so consider the graphs of $y = x^2$ and $y = 2^{x+4}$ where $x < 0$.[3]



These only cross once in this range, when $x = -2$ (as then we have $2^{-2+4} = 4 = (-2)^2$ — you can just evaluate the first few values of each graph to find this). This means that for $(x+y)^2 \geqslant 2^{x+y+4}$ we need to have $x + y \leqslant -2$, so the curve C_2 is bounded above by $x + y = -2$. We also know that the graph is symmetrical in $y = x$ and that it passes through $(-1, -1)$.

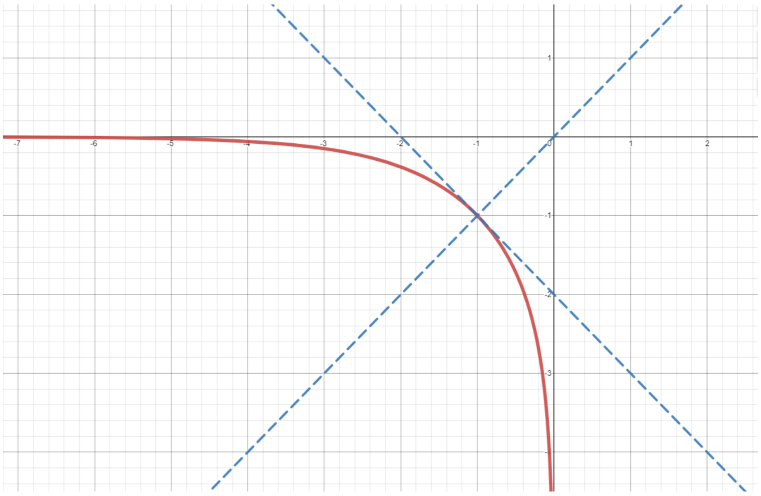
As $x \to -\infty$, we have $2^{x+y+4} \to 0$ and so $(x-y)^2 \approx (x+y)^2$, which means that $y \to 0$, and similarly as $y \to -\infty$, $x \to 0$. This now gives us enough to finish the graph.



Note that I have put some words on my graph as well, especially to point out that I know it doesn't intersect the y axis. If your graph goes a little wobbly then you can do this (if there was a very obvious crossing of the y axis, or your graph was very different to intended then a few words might not be enough).

[3] If it makes you feel more comfortable you could consider a substitution $t = x + y$ and then consider the graphs $\mathrm{f}(t) = t^2$ and $\mathrm{g}(t) = 2^{t+4}$.

Below is the picture that Desmos produces:



Question 3

3 A sequence $u_1, u_2, \ldots, u_n$ of positive real numbers is said to be unimodal if there is a value k such that

$$u_1 \leqslant u_2 \leqslant \ldots \leqslant u_k$$

and

$$u_k \geqslant u_{k+1} \geqslant \ldots \geqslant u_n.$$

So the sequences $1, 2, 3, 2, 1$; $1, 2, 3, 4, 5$; $1, 1, 3, 3, 2$ and $2, 2, 2, 2, 2$ are all unimodal, but $1, 2, 1, 3, 1$ is not.

A sequence $u_1, u_2, \ldots, u_n$ of positive real numbers is said to have property L if $u_{r-1}u_{r+1} \leqslant u_r^2$ for all r with $2 \leqslant r \leqslant n-1$.

(i) Show that, in any sequence of positive real numbers with property L,

$$u_{r-1} \geqslant u_r \implies u_r \geqslant u_{r+1}.$$

Prove that any sequence of positive real numbers with property L is unimodal.

(ii) A sequence $u_1, u_2, \ldots, u_n$ of real numbers satisfies $u_r = 2\alpha u_{r-1} - \alpha^2 u_{r-2}$ for $3 \leqslant r \leqslant n$, where α is a positive real constant. Prove that, for $2 \leqslant r \leqslant n$,

$$u_r - \alpha u_{r-1} = \alpha^{r-2}(u_2 - \alpha u_1)$$

and, for $2 \leqslant r \leqslant n-1$,

$$u_r^2 - u_{r-1}u_{r+1} = (u_r - \alpha u_{r-1})^2.$$

Hence show that the sequence consists of positive terms and is unimodal, provided $u_2 > \alpha u_1 > 0$.

In the case $u_1 = 1$ and $u_2 = 2$, prove by induction that
$u_r = (2-r)\alpha^{r-1} + 2(r-1)\alpha^{r-2}$.

Let $\alpha = 1 - \dfrac{1}{N}$, where N is an integer with $2 \leqslant N \leqslant n$.

In the case $u_1 = 1$ and $u_2 = 2$, prove that u_r is largest when $r = N$.

Examiner's report

This was the second least attempted of the pure questions. Relatively few candidates made a complete attempt at all of the parts and only 4 achieved full marks for the question. The question consisted of a succession of given results which were to be established. Thus, candidates needed to be more aware of the importance of providing careful and thorough explanations and justifications for each step that they took along the way. Many marks were lost as a result of carelessness in providing all of the necessary details.

A significant number of candidates thought that the implication in **(i)** showed that the sequence was either increasing or decreasing and so got little or no credit. Establishing the given relations in **(i)** was generally done quite well, with candidates demonstrating a considerable range of algebraic skills in their working. But then a lot of candidates failed to show that the sequence was positive, which undermined their attempts to deduce that the sequence was unimodal.

Many candidates used an induction proof for the first proof in part **(ii)** despite the fact that a more direct approach was possible and considerably simpler. The "asked-for" induction proof was usually handled well, though establishing the baseline case was often flawed; many either overlooked the need to establish both of the cases $n = 1$ and $n = 2$ or, when giving a one-step induction proof with the help of the previously-established result, chose an incorrect baseline case.

The final part of the question was often avoided, though full attempts often gained full credit. Again, the usual oversight was to fail to establish positivity. Many of those who produced only a faltering solution here overlooked the need to compare successive terms, usually merely working with an expression for u_r only, often with the use of differentiation - attempts along such lines invariably lost all of the final 7 marks allocated, primarily because the required result is based on discrete values of r while calculus works with continuous values.

Solution

In the stem of the question, the word "and" is very important! Basically the terms are increasing in value up to a point (u_k) and then decrease again after that (where "increasing" and "decreasing" means not strictly, so includes staying the same!). *Unimodal* means "one-humped", so there is at most one peak. Sometimes when a definition which might be a bit tricky is given in a question there are some examples to help clarify what we mean, as in this case where we have 4 examples of what is allowed (starting with a classic "one-hump", followed by some other cases which also count as unimodal), followed by a "two-humped" non-example.

The final bit of the stem gives a condition for another possible property of a sequence, which for convenience when writing the question is called "L" (and you can use this label in your solution). You are not asked to do anything in the stem (it is always worth re-reading it to make sure that there are no requests there!)

(i) Here we know that the sequence has property L, so we have $u_{r-1}u_{r+1} \leqslant u_r^2$. If we assume that $u_{r-1} \geqslant u_r$ then we have:

$$\begin{aligned} u_{r-1}u_{r+1} &\leqslant u_r^2 \\ &\leqslant u_r u_{r-1} \end{aligned}$$

and **as the sequence is positive** we can divide the inequality to get $u_{r+1} \leqslant u_r$. This forms the inductive step of a proof by induction, as we have shown that if the r^{th} term of a sequence of positive numbers is less than the $(r-1)^{\text{th}}$, then the $(r+1)^{\text{th}}$ will be smaller than the r^{th}. This means that once the sequence starts decreasing then it will keep on decreasing indefinitely, and so the sequence is unimodal.

(ii) We are given $u_r = 2\alpha u_{r-1} - \alpha^2 u_{r-2}$ (for $3 \leqslant r \leqslant n$), and so:

$$\begin{aligned} u_r - \alpha u_{r-1} &= 2\alpha u_{r-1} - \alpha^2 u_{r-2} - u_{r-1} \\ &= \alpha u_{r-1} - \alpha^2 u_{r-2} \\ &= \alpha(u_{r-1} - \alpha u_{r-2}) \end{aligned}$$

If we keep applying this we get:

$$\begin{aligned} u_r - \alpha u_{r-1} &= \alpha(u_{r-1} - \alpha u_{r-2}) \\ &= \alpha^2(u_{r-2} - \alpha u_{r-3}) \\ &= \alpha^3(u_{r-3} - \alpha u_{r-4}) \end{aligned}$$

and after $r-2$ applications we have $u_r - \alpha u_{r-1} = \alpha^{r-2}(u_2 - \alpha u_1)$ (note that substituting $r = 2$ also makes this statement true, so this is true for $2 \leqslant r \leqslant n$).

From the property $u_r = 2\alpha u_{r-1} - \alpha^2 u_{r-2}$, we have $u_{r+1} = 2\alpha u_r - \alpha^2 u_{r-1}$ (for $2 \leqslant r \leqslant n-1$) and so:

$$\begin{aligned} u_r^2 - u_{r-1}u_{r+1} &= u_r^2 - u_{r-1}(2\alpha u_r - \alpha^2 u_{r-1}) \\ &= u_r^2 - 2\alpha u_r u_{r-1} + (\alpha u_{r-1})^2 \\ &= (u_r - \alpha u_{r-1})^2 \end{aligned}$$

Using $u_r - \alpha u_{r-1} = \alpha^{r-2}(u_2 - \alpha u_1)$, then using the fact that α is positive and that $u_2 > \alpha u_1 > 0$ (i.e. $u_2 - \alpha u_1 > 0$), we know that $u_1 > 0$, $u_2 > 0$ and that $u_r - \alpha u_{r-1} > 0 \implies u_r > \alpha u_{r-1}$ so $u_r > 0$ for all r.

Using $u_r^2 - u_{r-1}u_{r+1} = (u_r - \alpha u_{r-1})^2 \geqslant 0$, we have $u_r^2 \geqslant u_{r-1}u_{r+1}$, which is property L and so the sequence is unimodal.

Since the relationship for the sequence depends on the two previous terms, we need to consider the first two terms of the sequence as our "base case".

$$\begin{aligned} \text{substituting } r = 1 \text{ gives } u_1 &= (2-1)\alpha^{1-1} + \cancel{2(1-1)\alpha^{1-2}} \\ &= 1 \text{ as required} \\ \text{substituting } r = 2 \text{ gives } u_2 &= \cancel{(2-2)\alpha^{2-1}} + 2(2-1)\alpha^{2-2} \\ &= 2 \text{ as required} \end{aligned}$$

Assuming the given form for u_r is true when $r = k$ and when $r = k-1$,[4] we have
$u_k = (2-k)\alpha^{k-1} + 2(k-1)\alpha^{k-2}$ and $u_{k-1} = (2-(k-1))\alpha^{(k-1)-1} + 2((k-1)-1)\alpha^{(k-1)-2}$.

Considering $r = k+1$ we have:

$$\begin{aligned} u_{k+1} &= 2\alpha u_k - \alpha^2 u_{k-1} \\ &= 2\alpha\big[(2-k)\alpha^{k-1} + 2(k-1)\alpha^{k-2}\big] - \alpha^2\big[(3-k)\alpha^{k-2} + 2(k-2)\alpha^{k-3}\big] \\ &= 2(2-k)\alpha^k + 4(k-1)\alpha^{k-1} - (3-k)\alpha^k - 2(k-2)\alpha^{k-1} \\ &= \big[4 - 2k - 3 + k\big]\alpha^k + \big[4k - 1 - 2k + 4\big]\alpha^{k-1} \\ &= (1-k)\alpha^k + (2k)\alpha^{k-1} \\ &= [2-(k+1)]\alpha^k + 2[(k+1)-1]\alpha^{k-1} \end{aligned}$$

[4] If you need the two previous results to complete the proof by induction then you will need to show that the first two cases are true.

which is the required for for u_{k+1}, and since it is true for u_1 and u_2, and being true for u_{k-1} and u_k it is true for u_{k+1} then it is true for all u_r.

For the very last part, we are trying to find a maximum of a set of discrete values (rather than of a continuous function where differentiation is appropriate). Instead we can look at the differences between successive values, i.e $u_r - u_{r-1}$. We have already shown that the sequence is unimodal, so we are looking for a value k, such that $u_k \geqslant u_{k-1} \implies (u_k - u_{k-1}) \geqslant 0$ and $u_k \geqslant u_{k+1} \implies (u_{k+1} - u_k) \leqslant 0$. Consider $u_r - u_{r-1}$:

$$\begin{aligned}
u_r - u_{r-1} &= \left[(2-r)\alpha^{r-1} + 2(r-1)\alpha^{r-2}\right] - \left[(3-r)\alpha^{r-2} + 2(r-2)\alpha^{r-3}\right] \\
&= \alpha^{r-3}\left[(2-r)\alpha^2 + (3r-5)\alpha + (4-2r)\right] \\
&= \alpha^{r-3}\left[(2-r)\frac{(N-1)^2}{N^2} + (3r-5)\frac{N-1}{N} + (4-2r)\right] \\
&= \frac{\alpha^{r-3}}{N^2}\left[(N-1)^2(2-r) + N(N-1)(3r-5) + N^2(4-2r)\right] \\
&= \frac{\alpha^{r-3}}{N^2}\left[(N^2-2N+1)(2-r) + (N^2-N)(3r-5) + N^2(4-2r)\right] \\
&= \frac{\alpha^{r-3}}{N^2}\left[N^2\big((2-r) + (3r-5) + (4-2r)\big) - N\big(2(2-r) + (3r-5)\big) + (2-r)\right] \\
&= \frac{\alpha^{r-3}}{N^2}\left[N^2 + (1-r)N + (2-r)\right]
\end{aligned}$$

So when $r = N$ we have:

$$\begin{aligned}
u_N - u_{N-1} &= \frac{\alpha^{r-3}}{N^2}\left[N^2 + (1-N)N + (2-N)\right] \\
&= \frac{\alpha^{r-3}}{N^2}\left[N^2 + N - N^2 + 2 - N\right] \\
&= \frac{\alpha^{r-3}}{N^2} \times 2 > 0
\end{aligned}$$

and so we know that $u_N > u_{N-1}$.

When $r = N+1$ we have:

$$\begin{aligned}
u_{N+1} - u_N &= \frac{\alpha^{r-3}}{N^2}\left[N^2 + (1-(N+1))N + (2-(N+1))\right] \\
&= \frac{\alpha^{r-3}}{N^2}\left[N^2 - N^2 + 1 - N\right] \\
&= \frac{\alpha^{r-3}}{N^2} \times (1-N)
\end{aligned}$$

and as $N \geqslant 2$ we know that $u_{N+1} - u_N < 0$. Hence we have $u_{N-1} < u_N > u_{N+1}$, and as the sequence is unimodal this means that u_N is the greatest value in the sequence.

Question 4

4 **(i)** Given that a, b and c are the lengths of the sides of a triangle, explain why $c < a + b$, $a < b + c$ and $b < a + c$.

(ii) Use a diagram to show that the converse of the result in part **(i)** also holds: if a, b and c are positive numbers such that $c < a+b$, $a < b+c$ and $b < c+a$ then it is possible to construct a triangle with sides of length a, b and c.

(iii) When a, b and c are the lengths of the sides of a triangle, determine in each case whether the following sets of three lengths can

- always
- sometimes but not always
- never

form the sides of a triangle. Prove your claims.

(A) $a + 1$, $b + 1$, $c + 1$.

(B) $\frac{a}{b}$, $\frac{b}{c}$, $\frac{c}{a}$.

(C) $|a - b|$, $|b - c|$, $|c - a|$.

(D) $a^2 + bc$, $b^2 + ca$, $c^2 + ab$.

(iv) Let f be a function defined on the positive real numbers and such that, whenever $x > y > 0$,

$$\mathrm{f}(x) > \mathrm{f}(y) > 0 \text{ but } \frac{\mathrm{f}(x)}{x} < \frac{\mathrm{f}(y)}{y}.$$

Show that, whenever a, b and c are the lengths of the sides of a triangle, then $\mathrm{f}(a)$, $\mathrm{f}(b)$ and $\mathrm{f}(c)$ can also be the lengths of the sides of a triangle.

Examiner's report

Most candidates could justify the triangle inequality in part **(i)** (as well as arguing that the shortest distance is the straight line, there were successful uses of the cosine rule or $c = a\cos B + b\cos A$), but were less confident in proving the converse in part **(ii)**. Successful approaches were to consider two circular loci for the SSS construction, or to fix two sides and vary the angle between them; in both cases care was needed to ensure all three inequalities were actually used, for example checking that neither circle can contain the other, which was often omitted. There was one elegant solution using three pairwise tangent circles.

A reasonable number of candidates obtained correct answers of "always", "sometimes" (by examples) and "never" for part **(iii)** A, B and C respectively, although it was surprisingly common to forget that a, b, c must be the sides of a triangle. However, B caused some confusion as many candidates spent some time trying to prove that the new lengths did satisfy the triangle inequality.

Parts **(iii)** D and **(iv)** were found much harder and many candidates did not attempt them. A reasonable number of candidates were able to make some progress, but there were few complete solutions to either of these parts. Common errors for **(iii)** D were showing that the sum of the three inequalities is a true statement, or attempting to prove a positive result by examples. Most substantial attempts at **(iv)** used a different approach of fixing the order of a, b, c and reducing the problem to proving one of the three inequalities

Solution

(i) Let the triangle have vertices A, B, C (and the length AB is c etc.). The direct route between A and B must be shorter than travelling from A to C to B. Hence we have:

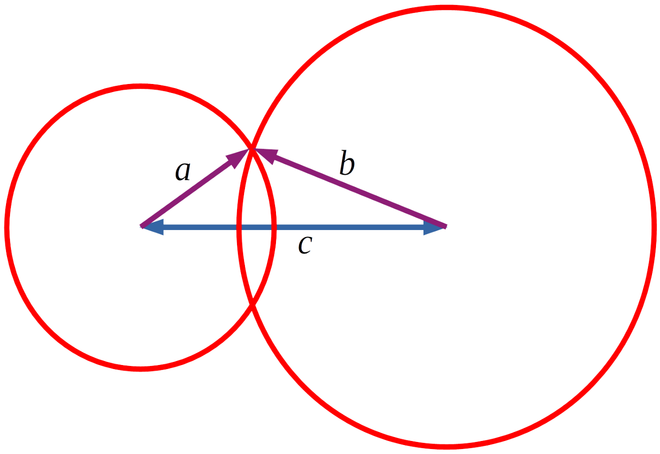
$$AB < AC + CB \implies c < a + b$$

and by symmetry we also have $a < b + c$ and $b < c + a$.

Here there was nothing special about the order we pick the points, and so we can argue the other two cases "by symmetry". I would probably have also drawn a sketch of a triangle with the points labelled.

(ii) WLOG[5] let $a \leqslant b \leqslant c$, so that c is the greatest length. If c is the greatest then we will have $a < b + c$ and $b < a + c$ for all $a \leqslant b \leqslant c$.

Consider two circles with radius a and radius b, positioned so that their centres are a distance c apart.



As long as we have $c < a + b$ then the two circles intersect and we can draw a triangle with lengths a, b and c.

By symmetry, if all of $c < a + b$, $a < b + c$ and $b < c + a$ then we can construct a triangle with sides of length a, b and c.

(iii) For this part a reasonable initial strategy is to try a couple of sets of values and see if they work or not. This might help you work out which classifications might be applicable in each case.

(A) If $a + b > c$ then we have:

$$\begin{aligned}(a+1)+(b+1) &= (a+b)+2 \\ &> c+2 \\ &> c+1\end{aligned}$$

so we have $(a + 1) + (b + 1) > (c + 1)$. The other two results follow in exactly the same way, so if a, b and c are the sides of a triangle then $(a + 1), (b + 1)$ and $(c + 1)$ are the sides of a triangle (i.e. the answer is ALWAYS).

(B) If we take $a = b = c = 1$ (which can form a triangle) then we get $\frac{a}{b} = \frac{b}{c} = \frac{c}{a} = 1$ which is also a triangle.

If we take $a = 1$ and $b = c = 2$ then we get $\frac{a}{b} = \frac{1}{2}, \frac{b}{c} = 1$ and $\frac{c}{a} = 2$. These cannot form a triangle as $\frac{1}{2} + 1 < 2$.

Since we have found one example that works and one that does not the answer is SOMETIMES.

Be careful to make sure that your initial values of a, b and c can form a triangle (e.g. $a = b = 1$, $c = 2$ would **not** be a suitable set of values).

[5] This means "without loss of generality". With three lengths we can arrange them in order so that there is a largest one, and we can label this as "c".

(C) WLOG take $a \leqslant b \leqslant c$. Then we have:

$$|a-b| = b-a$$
$$|b-c| = c-b$$
$$|c-a| = c-a$$

This means that we have $|c-a| = |b-c| + |a-b|$, and as the sum of two sides is equal to the third then these cannot be used to make a triangle. The answer is NEVER.

(D) Consider $(a^2+bc)+(b^2+ca)$, which we want to show is greater than c^2+ab. We know that $a+b>c$

$$\begin{aligned}(a^2+bc)+(b^2+ca) &= a^2+b^2+c(a+b)\\ &= (a-b)^2+2ab+c(a+b)\\ &> 0+2ab+c^2 \qquad \text{as } (a-b)^2 \geqslant 0 \text{ and } a+b>c\\ &> c^2+ab \qquad \text{as } 2ab>ab\end{aligned}$$

By symmetry, we also have $(b^2+ca)+(c^2+ab) > (a^2+bc)$ and $(c^2+ab)+(a^2+bc) > (b^2+ac)$. The answer is ALWAYS.

(iv) We have $a+b>c$, so $\text{f}(a+b) > \text{f}(c)$. We would like to be able to show that $\text{f}(a)+\text{f}(b) > \text{f}(c)$.

We know that $a+b>a$, so we have $\dfrac{\text{f}(a+b)}{a+b} < \dfrac{\text{f}(a)}{a}$ which can be written as $a\text{f}(a+b) < (a+b)\text{f}(a)$. Similarly we have $a+b>b$, and so

$$\frac{\text{f}(a+b)}{a+b} < \frac{\text{f}(b)}{b} \implies b\text{f}(a+b) < (a+b)\text{f}(b)\,.$$

Adding these together gives:

$$(a+b)\text{f}(a+b) < (a+b)\text{f}(a) + (a+b)\text{f}(b) \implies \text{f}(a+b) < \text{f}(a)+\text{f}(b)\,.$$

Since $\text{f}(a+b) > \text{f}(c)$ this gives $\text{f}(a)+\text{f}(b) > \text{f}(c)$.

So if $a+b>c$ then we have $\text{f}(a)+\text{f}(b) > \text{f}(c)$, and the other results follow by symmetry. Hence if a, b and c are the sides of a triangle then $\text{f}(a), \text{f}(b)$ and $\text{f}(c)$ are also the sides of a triangle.

Question 5

5 If x is a positive integer, the value of the function $\mathrm{d}(x)$ is the sum of the digits of x in base 10. For example, $\mathrm{d}(249) = 2 + 4 + 9 = 15$.

An n-digit positive integer x is written in the form $\displaystyle\sum_{r=0}^{n-1} a_r \times 10^r$, where $0 \leqslant a_r \leqslant 9$ for all $0 \leqslant r \leqslant n-1$ and $a_{n-1} > 0$.

(i) Prove that $x - \mathrm{d}(x)$ is non-negative and divisible by 9.

(ii) Prove that $x - 44\mathrm{d}(x)$ is a multiple of 9 if and only if x is a multiple of 9.

Suppose that $x = 44\mathrm{d}(x)$. Show that if x has n digits, then $x \leqslant 396n$ and $x \geqslant 10^{n-1}$, and hence that $n \leqslant 4$.

Find a value of x for which $x = 44\mathrm{d}(x)$. Show that there are no further values of x satisfying this equation.

(iii) Find a value of x for which $x = 107\mathrm{d}\left(\mathrm{d}(x)\right)$. Show that there are no further values of x satisfying this equation.

Examiner's report

This was a popular question and many of the solutions made good progress on the early parts of the question.

The majority of candidates gained full marks for part **(i)**, but some candidates did not mention that $x - d(x) \geqslant 0$.

There was a wide range of marks achieved on part **(ii)**. The proof that $x - 44d(x)$ is a multiple of 9 if and only if x is a multiple of 9 was completed well by those who managed to prove the result, but the majority of other attempts seen did not score any marks. In a small number of cases only the "if" direction was proved. Those who were unable to prove the first result in this part were often able to continue and find the required bounds on x however. Candidates who had completed both of these parts generally managed to find the correct answer $x = 792$, but did not necessarily fully justify that it was the only one.

Most candidates scored low marks on part **(iii)**. It was very common to see an insufficient proof that $9|x$ (this means "9 divides x"). Without guidance from the question as to how to find bounds on x, students produced a wide range of approaches; better bounds were needed if the student only used $107|x$, but the simple bound $d(d(x)) \leqslant d(x)$ together with divisibility by 963 was sufficient.

Solution
With summation questions I often find it easier to write out the sum explicitly rather than working with the sigma signs, however with part **(i)** here I decided not to.
Note that $249 = 2 \times 10^2 + 4 \times 10^1 + 9 \times 10^0$ which has the form $a_2 \times 10^2 + a_1 \times 10^1 + a_0 \times 10^0$.

(i) If n is non-negative then we have $n \geqslant 0$.

We have:

$$\begin{aligned} x - \mathrm{d}(x) &= \sum_{r=0}^{n-1} a_r \times 10^r - \sum_{r=0}^{n-1} a_r \\ &= \sum_{r=0}^{n-1} a^r (10^r - 1) \end{aligned}$$

We know that $10^r - 1 \geqslant 0$ for all r, and we are told that $a_r \geqslant 0$ so $x - \mathrm{d}(x)$ is the sum of non-negative products, so is non-negative.

We also have $10^r - 1 = (10-1)(10^{r-1} + 10^{r-2} + \cdots + 10r^0)$, so $10 - 1 = 9$ divides $10^r - 1$.

(ii) We have $x - 44\mathrm{d}(x) = 44x - 44\mathrm{d}(x) - 43x = 44(x - \mathrm{d}(x)) - 43x$. We know that $x - \mathrm{d}(x)$ is a multiple of 9, so $x - 44\mathrm{d}(x)$ is a multiple of 9 if and only if $43x$ is a multiple of 9. 43 is co-prime to 9, so $43x$ is a multiple of 9 if and only if x is a multiple of 9.

Alternatively, from part **(i)** we know that $x - \mathrm{d}(x) = 9k$ for some integer k. Hence we have:

$$\begin{aligned} x - 44\mathrm{d}(x) &= x - 44(x - 9k) \\ &= 44 \times 9k - 43x \end{aligned}$$

44×9 is divisible by 9, so $x - 44\mathrm{d}(x)$ is divisible by 9 if and only if $43x$ is divisible by 9. The the last step is as above.

If x has n digits then $\mathrm{d}(x) \leqslant 9n$ (as the maximum value of $\mathrm{d}(x)$ is $9 + 9 + \cdots + 9$). Since $x = 44\mathrm{d}(x)$ then we must have $x \leqslant 44 \times 9n \implies x \leqslant 396n$.

If x is an n digit number then we have $x = a_0 \times 10^0 + a_1 \times 10^1 + \cdots + a_{n-1} \times 10^{n-1}$, where $a_{n-1} > 0$ and so $x \geqslant 10^{n-1}$.

We need $10^{n-1} \leqslant x \leqslant 396n$ which is possible for $n = 1, 2, 3, 4$ (the last case gives $1000 \leqslant x \leqslant 1584$) but is not possible for $n \geqslant 5$ as $10^4 = 10,000 > 5 \times 396 = 1980$.

We have $x - 44\mathrm{d}(x) = 0$, which is a multiple of 9, so we must have x being a multiple of 9. We also have $x = 44\mathrm{d}(x)$, so x must be a multiple of 44. 44 and 9 are co-prime, so x must be a multiple of $44 \times 9 = 396$, i.e. $x = 396k$ (note that k might be different to n).

We are told that x is positive, so $x > 0$. We have also shown that x is a one, two, three or four digit number. Since $x = 396k$, and x has to be a four digit number or less, the possible values of x are $396, 792, 1188, 1584$. After this we would have $x > 1584$ which is not allowed for a four digit number, as earlier we showed that $x \leqslant 396n$ which in the case of a 4 digit number means we have $x \leqslant 1584$. Note that all four possible values of x have a digit sum of 18.

Checking these four values we have:

$$\begin{aligned} x = 396\,, &\quad 44\mathrm{d}(x) = 44 \times 18 = 792 \quad \text{X} \\ x = 792\,, &\quad 44\mathrm{d}(x) = 44 \times 18 = 792 \quad \checkmark \\ x = 1188\,, &\quad 44\mathrm{d}(x) = 44 \times 18 = 792 \quad \text{X} \\ x = 1584\,, &\quad 44\mathrm{d}(x) = 44 \times 18 = 792 \quad \text{X} \end{aligned}$$

and so the only solution is $x = 792$.

(iii) If x is an n digit number then we know that $\mathrm{d}(x) \leqslant 9n$ (as before). We also know that $\mathrm{d}(\mathrm{d}(x)) \leqslant \mathrm{d}(x)$, so we have $x \leqslant 963n$. We also know that $x \geqslant 10^{n-1}$, so we have $10^{n-1} \leqslant x \leqslant 963n$.

For $n = 1, 2, 3, 4$ this is fine but when $n = 5$ this becomes $10,000 \leqslant x \leqslant 4815$, so we have $n \leqslant 4$. We also know that since $x = 107\mathrm{d}(\mathrm{d}(x))$ that x must be a multiple of 107 (which is prime), but it would be nice if we could also show that it has to be a multiple of 9 like in the last part.

Consider $x - 107\mathrm{d}(\mathrm{d}(x))$. This can be rewritten as:

$$107(x - \mathrm{d}(x)) + 107(\mathrm{d}(x) - \mathrm{d}(\mathrm{d}(x))) - 106x$$

We know that $x - \mathrm{d}(x)$ is a multiple of 9, and so is $\mathrm{d}(x) - \mathrm{d}(\mathrm{d}(x))$ (in this second case x has been replaced by $\mathrm{d}(x)$, which is another integer).
This means that $107(x - \mathrm{d}(x)) + 107(\mathrm{d}(x) - \mathrm{d}(\mathrm{d}(x))) - 106x$ is a multiple of 9 if and only if $106x$ is a multiple of 9.

Therefore if $x = 107\mathrm{d}(\mathrm{d}(x))$ we have $x - 107\mathrm{d}(\mathrm{d}(x)) = 0$, which is a multiple of 9, and therefore $106x$ is a multiple of 9. $106 = 2 \times 53$ is co-prime to 9, and so x must be a multiple of 9.

Hence x is a multiple of 107 and a multiple of 9, and 107 and 9 are co-prime, hence x must be a multiple of 963 and so we have $x = 963k$ for some k, and since $x \leqslant 963n$ where $n \leqslant 4$, the only possible values of x are $963, 1926, 2889$ and 3852.

Trying these out gives:

$$\begin{array}{rllll} x = 963: & \mathrm{d}(x) = 18 & \mathrm{d}(\mathrm{d}(x) = 9 & 107\mathrm{d}(\mathrm{d}(x) = 963 & \checkmark \\ x = 1926: & \mathrm{d}(x) = 18 & \mathrm{d}(\mathrm{d}(x) = 9 & 107\mathrm{d}(\mathrm{d}(x) = 963 & \times \\ x = 2889: & \mathrm{d}(x) = 27 & \mathrm{d}(\mathrm{d}(x) = 9 & 107\mathrm{d}(\mathrm{d}(x) = 963 & \times \\ x = 3852: & \mathrm{d}(x) = 18 & \mathrm{d}(\mathrm{d}(x) = 9 & 107\mathrm{d}(\mathrm{d}(x) = 963 & \times \end{array}$$

So the only solution is $x = 963$.

Question 6

6 A 2×2 matrix $\mathbf{M}$ is real if it can be written as $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where a, b, c and d are real.

In this case, the *trace* of matrix $\mathbf{M}$ is defined to be $\text{tr}(\mathbf{M}) = a + d$ and $\det(\mathbf{M})$ is the determinant of matrix $\mathbf{M}$. In this question, $\mathbf{M}$ is a real 2×2 matrix.

(i) Prove that

$$\text{tr}(\mathbf{M}^2) = \text{tr}(\mathbf{M})^2 - 2\det(\mathbf{M}).$$

(ii) Prove that

$$\mathbf{M}^2 = \mathbf{I} \text{ but } \mathbf{M} \neq \pm\mathbf{I} \iff \text{tr}(\mathbf{M}) = 0 \text{ and } \det(\mathbf{M}) = -1,$$

and that

$$\mathbf{M}^2 = -\mathbf{I} \iff \text{tr}(\mathbf{M}) = 0 \text{ and } \det(\mathbf{M}) = 1.$$

(iii) Use part **(ii)** to prove that

$$\mathbf{M}^4 = \mathbf{I} \iff \mathbf{M}^2 = \pm\mathbf{I}.$$

Find a necessary and sufficient condition on $\det(\mathbf{M})$ and $\text{tr}(\mathbf{M})$ so that $\mathbf{M}^4 = -\mathbf{I}$.

(iv) Give an example of a matrix $\mathbf{M}$ for which $\mathbf{M}^8 = \mathbf{I}$, but which does not represent a rotation or reflection. [Note that the matrices $\pm\mathbf{I}$ are both rotations.]

Examiner's report

This was the second most popular question on the paper. In general, candidates need to be careful when proving statements of the form "A if and only if B" and should be aware that in some cases it may not be possible to prove both directions in one go. Candidates should also be aware that, in some cases, the algebra is not sufficient on its own to demonstrate the reasoning and explanations of the steps are often helpful.

Part **(i)** of the question was generally completed well. In part **(ii)** many largely successful attempts were seen, but few candidates picked up all of the marks for this section. The main errors arose from not adequately considering cases and so dividing by 0, and from not noticing that $a^2 = d^2 = 1$ could result in $a = -d = \pm 1$. The most successful attempts in this part were the ones that separated the two directions of the implications. Many candidates misused the condition $\mathbf{M} \neq \pm I$ in trying to prove the implication in one direction or did not check this condition when proving the implication in the opposite direction.

Few attempts at part **(iii)** were seen and a common mistake was to do the component-wise algebra to find $\mathbf{M}^4$ instead of using the results from previous parts. In general, those who had understood the previous parts and attempted part **(iv)** were able to solve the final part of the question.

Solution

(i) The *trace* of a Matrix is the sum of the elements on the *leading diagonal*, which in the case of a 2×2 matrix $\mathbf{M}$ as above gives $\text{tr}(\mathbf{M}) = a + d$.

If $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then $\mathbf{M^2} = \begin{pmatrix} a^2 + bc & b(a+d) \\ c(a+d) & d^2 + bc \end{pmatrix}$.

This means that we have:

$$\begin{aligned} \text{tr}(\mathbf{M^2}) &= (a^2 + bc) + (d^2 + bc) \\ &= (a+d)^2 - 2ad + 2bc \\ &= (a+d)^2 - 2(ad - bc) \\ &= \left[\text{tr}(\mathbf{M})\right]^2 - 2\text{det}(\mathbf{M}) \end{aligned}$$

(ii) First we will work on the $\implies$ direction. Assume that $\mathbf{M^2} = \mathbf{I}$, but $\mathbf{M} \neq \pm\mathbf{I}$.

If $\mathbf{M^2} = \mathbf{I}$ then we have:

$$\begin{aligned} a^2 + bc &= 1 \\ d^2 + bc &= 1 \\ b(a+d) &= 0 \\ c(a+d) &= 0 \end{aligned}$$

The last two equations tell us that either $a + d = 0$, or $b = c = 0$.

If we have $b = c = 0$, then the first two equations tell us that $a^2 = d^2 = 1$, and so $a = \pm d = \pm 1$. The only cases which do not contradict $\mathbf{M} \neq \pm\mathbf{I}$ are $a = -1, d = 1$ and $a = 1, d = -1$. These both satisfy $a + d = 0$.

Therefore we must have $a + d = 0$, and so $\text{tr}(\mathbf{M}) = 0$.

Using the result from part **(i)**, and using $\text{tr}(\mathbf{I}) = 1 + 1 = 2$, we have:

$$\begin{aligned} \text{tr}(\mathbf{M^2}) &= \left[\text{tr}(\mathbf{M})\right]^2 - 2\text{det}(\mathbf{M}) \\ 2 &= 0^2 - 2\text{det}(\mathbf{M}) \\ \implies \text{det}(\mathbf{M}) &= -1 \end{aligned}$$

Working in the opposite direction we have:

$$\begin{aligned} \text{tr}(\mathbf{M}) = 0 &\quad \text{which gives } a + d = 0 \quad \text{i.e. } d = -a \\ \text{det}(\mathbf{M}) = -1 &\quad \text{which gives } ad - bc = -1 \implies -a^2 - bc = -1 \text{ and } -d^2 - bc = -1 \end{aligned}$$

Substituting these into the form for $\mathbf{M}^2$ gives:

$$\mathbf{M^2} = \begin{pmatrix} a^2 + bc & b(a+d) \\ c(a+d) & d^2 + bc \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and since $a = -d$ we have $\mathbf{M} \neq \pm\mathbf{I}$.

For the second statement, start with $\mathbf{M}^2 = -\mathbf{I}$. This gives:

$$\begin{aligned} a^2 + bc &= -1 \\ d^2 + bc &= -1 \\ b(a+d) &= 0 \\ c(a+d) &= 0 \end{aligned}$$

As before, either $a + d = 0$ or $b = c = 0$. In the second case, this gives $a^2 = d^2 = -1$, which contradicts $\mathbf{M}$ being a real matrix. Therefore we have $a+d = 0$, and so $\text{tr}(\mathbf{M}) = 0$, and using the result from part **(i)**, and using $\text{tr}(-\mathbf{I}) = -1 - 1 = -2$, we have:

$$\begin{aligned} \text{tr}(\mathbf{M^2}) &= \big[\text{tr}(\mathbf{M})\big]^2 - 2\det(\mathbf{M}) \\ -2 &= 0^2 - 2\det(\mathbf{M}) \\ \implies \det(\mathbf{M}) &= 1 \end{aligned}$$

Working in the opposite direction we have:

$$\begin{aligned} \text{tr}(\mathbf{M}) = 0 &\quad \text{which gives } a + d = 0 \quad \text{i.e. } d = -a \\ \det(\mathbf{M}) = 1 &\quad \text{which gives } ad - bc = 1 \implies -a^2 - bc = 1 \text{ and } -d^2 - bc = 1 \end{aligned}$$

Substituting these into the form for $\mathbf{M}^2$ gives:

$$\mathbf{M^2} = \begin{pmatrix} a^2 + bc & b(a+d) \\ c(a+d) & d^2 + bc \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

and so $\mathbf{M}^2 = -\mathbf{I}$.

(iii) If we have $\mathbf{M}^2 = \pm\mathbf{I}$ then we have $\mathbf{M}^4 = \mathbf{I}$, i.e. we have

$$\mathbf{M}^2 = \pm\mathbf{I} \implies \mathbf{M}^4 = \mathbf{I}$$

If we have $\mathbf{M}^4 = \mathbf{I}$, but $\mathbf{M}^2 \neq \pm\mathbf{I}$ then we must have $\det(\mathbf{M}^2) = -1$ by part **(ii)**. However since $\det(\mathbf{M}^2) = \big[\det \mathbf{M}\big]^2$, and $\mathbf{M}$ is real, this is impossible, and so we have:

$$\mathbf{M}^4 = \mathbf{I} \implies \mathbf{M}^2 = \pm\mathbf{I}$$

Hence we have $\mathbf{M}^4 = \mathbf{I} \iff \mathbf{M}^2 = \pm\mathbf{I}$.

From part **(ii)** we have

$$\mathbf{M}^2 = -\mathbf{I} \iff \text{tr}(\mathbf{M}) = 0 \text{ and } \det(\mathbf{M}) = 1$$

Replacing $\mathbf{M}$ with $\mathbf{M}^2$ gives:

$$\mathbf{M}^4 = -\mathbf{I} \iff \text{tr}(\mathbf{M}^2) = 0 \text{ and } \det(\mathbf{M}^2) = 1$$

If $\det(\mathbf{M}^2) = 1$ then $\det(\mathbf{M}) = \pm 1$. Using part **(i)** we have:

$$\begin{aligned} \text{tr}(\mathbf{M}^2) &= \text{tr}(\mathbf{M})^2 - 2\det(\mathbf{M}) \\ 0 &= \big[\text{tr}(\mathbf{M})\big]^2 - 2\det(\mathbf{M}) \\ \big[\text{tr}(\mathbf{M})\big]^2 &= 2\det(\mathbf{M}) \end{aligned}$$

Since the matrix is real, we must have $\big[\text{tr}(\mathbf{M})\big]^2 \geqslant 0$, and so we have:

$$\begin{aligned} \det(\mathbf{M}) &= 1 \\ \text{tr}(\mathbf{M}) &= \pm\sqrt{2} \end{aligned}$$

(iv) For this last part we could find a matrix so that $(\mathbf{M})^2 = \mathbf{I}$ (or $(\mathbf{M})^2 = -\mathbf{I}$ or $(\mathbf{M})^4 = -\mathbf{I}$).

If we try to find a matrix so that $(\mathbf{M})^2 = \mathbf{I}$ then we need $\text{tr}(\mathbf{M}) = 0$ and $\det(\mathbf{M}) = -1$. Let $a = 5$ and $d = -5$ (as that we have $\text{tr}(\mathbf{M}) = 0$). We then need to satisfy $ad - bc = -1 \implies bc = -24$, so take $b = -2$, $c = 12$. This gives:

$$\mathbf{M} = \begin{pmatrix} 5 & -2 \\ 12 & -5 \end{pmatrix}$$

This satisfies $\mathbf{M}^2 = \mathbf{I}$, but does not have the form of a rotation or a reflection matrix (see STEP 2 matrices topic notes).

Question 7

7 In this question, $w = \dfrac{2}{z-2}$.

(i) Let z be the complex number $3+t\text{i}$, where $t \in \mathbb{R}$. Show that $|w-1|$ is independent of t. Hence show that, if z is a complex number on the line $\text{Re}(z) = 3$ in the Argand diagram, then w lies on a circle in the Argand diagram with centre 1.

Let V be the line $\text{Re}(z) = p$, where p is a real constant not equal to 2. Show that, if z lies on V, then w lies on a circle whose centre and radius you should give in terms of p. For which z on V is $\text{Im}(w) > 0$?

(ii) Let H be the line $\text{Im}(z) = q$, where q is a non-zero real constant. Show that, if z lies on H, then w lies on a circle whose centre and radius you should give in terms of q. For which z on H is $\text{Re}(w) > 0$?

Examiner's report

This was the least popular of the pure questions and also the one on which marks were lowest on average.

Many candidates were able to show the first result, that $|w-1|$ is independent of t. However, candidates often did not explain well enough the connection between the form of z and the line $\mathbf{Re}(z) = 3$.

The next part of part **(i)** required noting that the centre lies on the real axis and working out $|w-c|$. Some candidates guessed the value of c. Common mistakes here included guessing $c = 1, p-2$, or failing to note conditions in which $|w-c|$ is independent of t. In many solutions the absolute value sign on the radius was forgotten.

Part **(ii)** was similar to the previous part but required noting that the centre lies on the imaginary axis and working out $|w-c\text{i}|$. In both parts a common attempt was the guess the centre to be at a point $z = a + b\text{i}$, few candidates were successful using this method. Again, absolute value signs on the radius were regularly forgotten.

Another successful method employed by candidates in all parts of the question was to use the substitution $t = \tan\frac{\theta}{2}$, $t = (p-2)\tan\frac{\theta}{2}$, $t = q\tan\frac{\theta}{2} + 2$ and using various trig identities to achieve the centre and radius. A few students also expressed t in terms of $\mathbf{Re}(w)$ and $\mathbf{Im}(w)$ and used that to obtain the equation of a circle in $\mathbb{R}^2$.

Solution

(i) If $w = \dfrac{2}{z-2}$ then $|w-1| = \left|\dfrac{2}{z-2} - 1\right| = \left|\dfrac{2-(z-2)}{z-2}\right| = \left|\dfrac{4-z}{z-2}\right|$. Substituting $z = 3 + t\text{i}$ gives:

$$\begin{aligned}
|w-1| &= \left|\frac{4-(3+t\text{i})}{3+t\text{i}-2}\right| \\
&= \left|\frac{1-t\text{i}}{1+t\text{i}}\right| \\
\implies |w-1|^2 &= \left|\frac{1-t\text{i}}{1+t\text{i}}\right|^2 \\
&= \left(\frac{1-t\text{i}}{1+t\text{i}}\right) \times \left(\frac{1+t\text{i}}{1-t\text{i}}\right) \qquad \text{using } |z|^2 = zz^* \\
&= 1
\end{aligned}$$

which is independent of t, and hence $|w-1|$ is independent of t.

If z lies on the line $\text{Re}(z) = 3$, then it has the form $z = 2 + t\text{i}$, where t is a real number. We know that if $z = 3 + t\text{i}$ then $|w-1| = 1$, which is a circle centre at $1 + 0\text{i}$ and with radius 1.

If z lies on the line $\text{Re}(z) = p$, then we have $z = p + t\text{i}$, where t is real. Consider $|w-c|^2$, where c is real:

$$\begin{aligned}
|w-c|^2 &= \left|\frac{2}{z-2} - c\right|^2 \\
&= \left|\frac{2-c(z-2)}{z-2}\right|^2 \\
&= \left|\frac{2+2c-cz}{z-2}\right|^2 \\
&= \left|\frac{(2+2c-pc)-ct\text{i}}{(p-2)+t\text{i}}\right|^2 \\
&= \frac{(2+2c-pc)^2 + c^2t^2}{(p-2)^2+t^2}
\end{aligned}$$

For this to be independent of t we need the numerator to be a multiple of the denominator. We can see that the t terms have been multiplied by c^2, we need:

$$\begin{aligned}
(2+2c-pc)^2 &= c^2(p-2)^2 \\
(2-c(p-2))^2 &= c^2(p-2)^2 \\
4 - 4c(p-2) + \cancel{c^2(p-2)^2} &= \cancel{c^2(p-2)^2} \\
\implies c(p-2) &= 1 \\
c &= \frac{1}{p-2}
\end{aligned}$$

So the centre of the circle is at $c = \dfrac{1}{p-2}$, and the radius is given by

$$\begin{aligned} r^2 &= \frac{(2+2c-pc)^2 + c^2t^2}{(p-2)^2+t^2} \\ &= c^2 \\ \implies r &= \frac{1}{|p-2|} \end{aligned}$$

We have $w = \dfrac{2}{(p-2)+t\mathrm{i}} = \dfrac{2\big[(p-2)-t\mathrm{i}\big]}{(p-2)^2+t^2}$, and so we have $\mathrm{Im}(w) > 0$ when $t < 0$, i.e. when $\mathrm{Im}(z) < 0$.

(ii) In this case H is a horizontal line (a line where the imaginary competent is constant). If z lies on H then z has the form $z = t + q\mathrm{i}$, where t is real (and t varies). Comparing this case to the ones in part **(i)** a reasonable assumption would be that the centre of the circle lies on the imaginary axes.

Consider $|w - d\mathrm{i}|^2$, where d is real:

$$\begin{aligned} |w - d\mathrm{i}|^2 &= \left|\frac{2}{z-2} - d\mathrm{i}\right|^2 \\ &= \left|\frac{2 - d\mathrm{i}(z-2)}{z-2}\right|^2 \\ &= \left|\frac{2 + 2d\mathrm{i} - d\mathrm{i}z}{z-2}\right|^2 \\ &= \left|\frac{2 + 2d\mathrm{i} - d\mathrm{i}(t+q\mathrm{i})}{t+q\mathrm{i}-2}\right|^2 \\ &= \left|\frac{(2+dq) + d\mathrm{i}(2-t)}{(t-2)+q\mathrm{i}}\right|^2 \\ &= \frac{(2+dq)^2 + d^2(2-t)^2}{(t-2)^2+q^2} \\ &= \frac{d^2(t-2)^2 + (2+dq)^2}{(t-2)^2+q^2} \end{aligned}$$

This is independent of t if $(2+dq)^2 = d^2q^2$, which gives $4 + 4dq = 0 \implies c = -\dfrac{1}{q}$. The centre of the circle is at $0 - \dfrac{1}{q}\mathrm{i}$ and the radius satisfies $r^2 = d^2 \implies r = \dfrac{1}{|q|}$.

In this case $w = \dfrac{2}{(t-2)+q\mathrm{i}} = \dfrac{2\big[(t-2)-q\mathrm{i}\big]}{((t-2)^2+q^2}$, and so we have $\mathrm{Re}(w) > 0$ when $t > 2$, i.e. when $\mathrm{Re}(z) > 2$.

Question 8

8 In this question, $\mathrm{f}(x)$ is a quartic polynomial where the coefficient of x^4 is equal to 1, and which has four real roots, 0, a, b and c, where $0 < a < b < c$.

$\mathrm{F}(x)$ is defined by $\mathrm{F}(x) = \displaystyle\int_0^x \mathrm{f}(t)\,\mathrm{d}t$.

The area enclosed by the curve $y = \mathrm{f}(x)$ and the x-axis between 0 and a is equal to that between b and c, and half that between a and b.

(i) Sketch the curve $y = \mathrm{F}(x)$, showing the x co-ordinates of its turning points.

Explain why $\mathrm{F}(x)$ must have the form $\mathrm{F}(x) = \frac{1}{5}x^2(x-c)^2(x-h)$, where $0 < h < c$.

Find, in factorised form, an expression for $\mathrm{F}(x) + \mathrm{F}(c-x)$ in terms of c, h and x.

(ii) If $0 \leqslant x \leqslant c$, explain why $\mathrm{F}(b) + \mathrm{F}(x) \geqslant 0$ and why $\mathrm{F}(b) + \mathrm{F}(x) > 0$ if $x \neq a$. Hence show that $c - b = a$ or $c > 2h$.

By considering also $\mathrm{F}(a) + \mathrm{F}(x)$, show that $c = a + b$ and that $c = 2h$.

(iii) Find an expression for $\mathrm{f}(x)$ in terms of c and x only.

Show that the points of inflection on $y = \mathrm{f}(x)$ lie on the x-axis.

Examiner's report

For the first part candidates were asked to sketch a curve $y = \mathrm{F}(x)$ based on some information about the function $\mathrm{f}(x)$. A not insignificant number of candidates instead sketched $y = \mathrm{f}(x)$ but those who sketched the correct curve often earned most of the marks. When justifying the given form of $\mathrm{F}(x)$ some good explanations were provided, but in many cases the repeated roots at $x = 0$ and $x = c$ were not explained. The final section of part **(i)** was generally completed well by those who reached it.

The next part was found to be more difficult with many candidates mistakenly using the local maximum of $\mathrm{F}(x)$ at $x = b$ to justify the first inequality instead of the local minimum at $x = a$. It was common to see justification such as $|\mathrm{F}(x)| < \mathrm{F}(b)$ without showing first that $\mathrm{F}(b) = -\mathrm{F}(a)$. Candidates who spotted the connection with part **(i)** and substituted $x = b$ into their expression for $\mathrm{F}(x) + \mathrm{F}(c-x)$ were usually able to show $c - b = a$ or $c > 2h$. For the last section of part **(ii)**, those who realised the connection with the first paragraph had no issues.

Candidates who reached the final part of the question were often able to obtain expression for $\mathrm{f}(x)$ and most realised that they needed to calculate $\mathrm{f}''(x)$ in order to find the inflection points. However, the final mark for spotting that the roots of $\mathrm{f}''(x) = 0$ are necessarily roots of $\mathrm{f}(x) = 0$ without explicitly calculating them (and thereby wasting time) eluded the majority of candidates who reached this part.

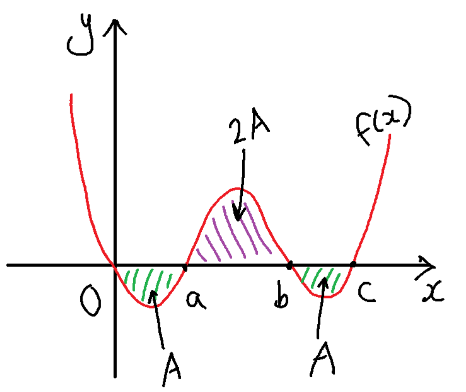
Solution

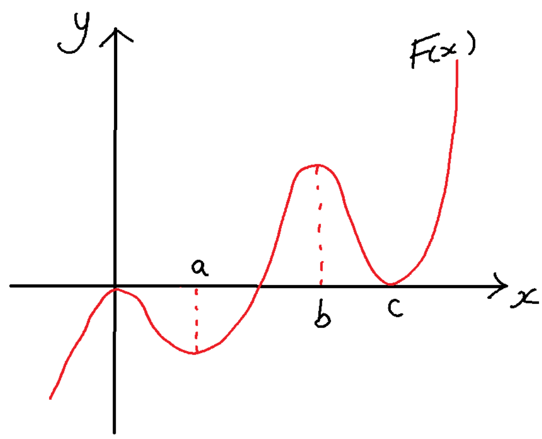
(i) It's not a bad idea to start by sketching $y = \mathrm{f}(x)$. Some points to note:

1. Since $\mathrm{f}(x)$ is a quartic we know that $\mathrm{F}(x)$ is a quintic
2. $\displaystyle\int_0^0 \mathrm{f}(x)\,\mathrm{d}x = 0$, so we know that $\mathrm{F}(0) = 0$.
3. $\displaystyle\int_0^c \mathrm{f}(x)\,\mathrm{d}x = \int_0^a \mathrm{f}(x)\,\mathrm{d}x + \int_a^b \mathrm{f}(x)\,\mathrm{d}x + \int_b^c \mathrm{f}(x)\,\mathrm{d}x = 0$, so $\mathrm{F}(c) = 0$
4. If $x = -a$ (where $a > 0$) then:
$$\int_0^{-a} \mathrm{f}(x)\,\mathrm{d}x = -\int_{-a}^0 \mathrm{f}(x)\,\mathrm{d}(x)$$
so when x is negative then $\mathrm{F}(x)$ is negative.
5. As x increases in the range $0 < x < a$, more and more negative area is included, so $\mathrm{F}(x)$ has a minimum when $x = a$. Similarly there is a maximum when $x = b$.

This gives us enough information to sketch the graph as shown below:

It would not be necessary to write all of these out in the exam!





For the form of $\mathrm{F}(x)$, we know that $\mathrm{f}(x) = x^4 + O(x^3)$, so when we integrate this we get $\mathrm{F}(x) = \frac{1}{5}x^5 + O(x^4)$, so the coefficient of x^5 is $\frac{1}{5}$.

There is a repeated root when $x = 0$, as there is a turning point on the x-axis here, so x^2 must be a factor of $\mathrm{F}(x)$. Similarly there is a repeated root when $x = c$, as there is a turning point where the graph meets the x-axis here as well so $(x - c)^2$ must be a factor of $\mathrm{F}(x)$. There is one more root of $\mathrm{F}(x)$, which occurs in the interval (a, b)[6] so there is a single root at $x = h$, where $h \in (a, b)$ (and so we must have $h \in (0, c)$), and therefore $(x - h)$ is a factor of $\mathrm{F}(x)$. This means we have:

$$\mathrm{F}(x) = \tfrac{1}{5}x^2(x - c)^2(x - h)$$

[6] Interval notation can be useful, here (a, b) means a value in the range $a < x < b$. $[a, b)$ would means a value in the range $a \leqslant x < b$. $x \in [0, \infty)$ means that x is non-negative.

Considering $\mathrm{F}(x)+\mathrm{F}(c-x)$ we have:

$$\begin{aligned}\mathrm{F}(x)+\mathrm{F}(c-x) &= \tfrac{1}{5}x^2(x-c)^2(x-h)+\tfrac{1}{5}(c-x)^2[(c-x)-c]^2[(c-x)-h]\\ &= \tfrac{1}{5}x^2(x-c)^2(x-h)+\tfrac{1}{5}(c-x)^2[-x]^2[c-x-h]\\ &= \tfrac{1}{5}x^2(x-c)^2[(x-h)+(c-x-h)]\\ &= \tfrac{1}{5}x^2(x-c)^2(c-2h)\end{aligned}$$

(ii) Using the graph in part **(i)** we have $\mathrm{F}(a)=-A$, $\mathrm{F}(b)=2A-A=A$ and $\mathrm{F}(c)=0$. The minimum of $\mathrm{F}(x)$ in the range $0\leqslant x\leqslant c$ occurs when $x=a$, and $\mathrm{F}(a)=-A$ so in this range we have:

$$\begin{aligned}\mathrm{F}(b)+\mathrm{F}(x) &\geqslant F(b)+F(a)\\ \mathrm{F}(b)+\mathrm{F}(x) &\geqslant A-A\\ \mathrm{F}(b)+\mathrm{F}(x) &\geqslant 0\end{aligned}$$

We also know that $\mathrm{F}(b)+\mathrm{F}(x)=0$ if and only if $x=a$, so if $0\leqslant x\leqslant c$ and $x\neq a$ we have $\mathrm{F}(b)+\mathrm{F}(x)>0$.

Substituting $x=b$ into our expression for $\mathrm{F}(x)+\mathrm{F}(c-x)$ we have:

$$\mathrm{F}(b)+\mathrm{F}(c-b)=\tfrac{1}{5}b^2(b-c)^2(c-2h)$$

We know that if $c-b\neq a$ then $\mathrm{F}(b)-\mathrm{F}(c-b)>0$ which means that we must have $c-2h>0$ (as the other brackets are squared and non-zero, and hence are positive).

Therefore either $c>2h$, or we have $c=2h$ and $c-b=a$.

Considering $\mathrm{F}(a)+\mathrm{F}(x)$ for $0\leqslant x\leqslant c$, we know that the maximum value of $\mathrm{F}(x)$ in this range is $\mathrm{F}(b)=A$. Hence we have:

$$\begin{aligned}\mathrm{F}(a)+\mathrm{F}(x) &\leqslant F(a)+F(b)\\ \mathrm{F}(a)+\mathrm{F}(x) &\leqslant -A+A\\ \mathrm{F}(a)+\mathrm{F}(x) &\leqslant 0\end{aligned}$$

and $\mathrm{F}(a)+\mathrm{F}(x)=0$ iff $x=b$.

Substituting $x=a$ into our expression for $\mathrm{F}(x)+\mathrm{F}(c-x)$ we have:

$$\mathrm{F}(a)+\mathrm{F}(c-a)=\tfrac{1}{5}a^2(a-c)^2(c-2h)$$

We know that if $c-a\neq b$ then $\mathrm{F}(a)-\mathrm{F}(c-a)<0$ which means that we must have $c-2h<0$ (as the other brackets are squared and non-zero, and hence are positive). Otherwise we have $c-a=b$ and $c=2h$.

It is not possible to have both $c-2h<0$ and $c-2h>0$ true simultaneously, so we must have $c=2h$ and $c=a+b$.

(iii) Using $c = 2h$ in our expression for $\mathrm{F}(x)$ we have:

$$\begin{aligned}\mathrm{F}(x) &= \tfrac{1}{5}x^2(x-c)^2\left(x-\tfrac{1}{2}c\right)\\ &= \tfrac{1}{10}x^2(x-c)^2(2x-c)\end{aligned}$$

Differentiating[7] this gives:

$$\begin{aligned}\mathrm{f}(x) &= \tfrac{1}{10}\Big[2x(x-c)^2(2x-c)+2x^2(x-c)(2x-c)+2x^2(x-c)^2\Big]\\ &= \tfrac{1}{5}x(x-c)\Big[(x-c)(2x-c)+x(2x-c)+x(x-c)\Big]\\ &= \tfrac{1}{5}x(x-c)\Big[5x^2-5cx+c^2\Big]\end{aligned}$$

We know that the roots of $\mathrm{f}(x)$ are $0, a, b, c$, so the roots of the quadratic $5x^2 - 5cx + c^2$ must be a and b.

At a point of inflection of $y = \mathrm{f}(x)$ we must have $\mathrm{f}''(x) = 0$. Differentiating gives:

$$\begin{aligned}\mathrm{f}'(x) &= \tfrac{1}{5}\Big[(x-c)(5x^2-5cx+c^2)+x(5x^2-5cx+c^2)+x(x-c)(10x-5c)\Big]\\ &= \tfrac{1}{5}\Big[20x^3-30cx^2+12c^2x-c^3\Big]\\ \mathrm{f}''(x) &= \tfrac{1}{5}\Big[60x^2-60cx+12c^2\Big]\\ &= \tfrac{12}{5}\Big[5x^2-5cx+c^2\Big]\end{aligned}$$

and so the roots of $\mathrm{f}''(x) = 0$ are the roots of $5x^2 - 5cx + c^2$, and so they must be $x = a, x = b$. Looking at the sketch of $\mathrm{f}(x)$ from the start of the question we can see that the points where $x = a$ and $x = b$ are not maxima or minima. Hence the points $(a, 0)$ and $(b, 0)$ are points of inflection.

In retrospect, I should have expanded the expression for $\mathrm{f}(x)$ before differentiating, it would have been less work! Note that there is a small mistake in $\mathrm{f}'(x)$ in the Cambridge Assessment published solutions.

[7] A useful result is the "three term product rule". If differentiating $\mathrm{f}(x)\mathrm{g}(x)\mathrm{h}(x)$ we get $\mathrm{f}'(x)\big[\mathrm{g}(x)\mathrm{h}(x)\big] + \mathrm{f}(x)\big[\mathrm{g}(x)\mathrm{h}(x)\big]' = \mathrm{f}'(x)\mathrm{g}(x)\mathrm{h}(x) + \mathrm{f}(x)\mathrm{g}'(x)\mathrm{h}(x) + \mathrm{f}(x)\mathrm{g}(x)\mathrm{h}'(x)$

Question 9

9 Point A is a distance h above ground level and point N is directly below A at ground level. Point B is also at ground level, a distance d horizontally from N. The angle of elevation of A from B is β. A particle is projected horizontally from A, with initial speed V. A second particle is projected from B with speed U at an acute angle θ above the horizontal. The horizontal components of the velocities of the two particles are in opposite directions. The two particles are projected simultaneously, in the vertical plane through A, N and B.

Given that the two particles collide, show that

$$d\sin\theta - h\cos\theta = \frac{Vh}{U}$$

and also that

(i) $\theta > \beta$;

(ii) $U\sin\theta \geqslant \sqrt{\dfrac{gh}{2}}$;

(iii) $\dfrac{U}{V} > \sin\beta$.

Show that the particles collide at a height greater than $\frac{1}{2}h$ if and only if the particle projected from B is moving upwards at the time of collision.

Examiner's report

This was a question that was found to be difficult. In general, this question was not attempted well, with very few candidates progressing past the first section. Most candidates managed to pick up all the marks in the initial section of the question. However, a significant minority of students could not set up the problem correctly or knew a lot of linear acceleration (suvat) equations but could not apply them correctly (for example mistaking displacement for position) and received zero marks. Some candidates eliminated t in favour of x and could not progress to the last calculation.

Around half the candidates picked up full marks for part **(i)**. However, many candidates tried to reason with words – almost always unsuccessfully, often believing that the particle projected from point A could not pass through the line AB.

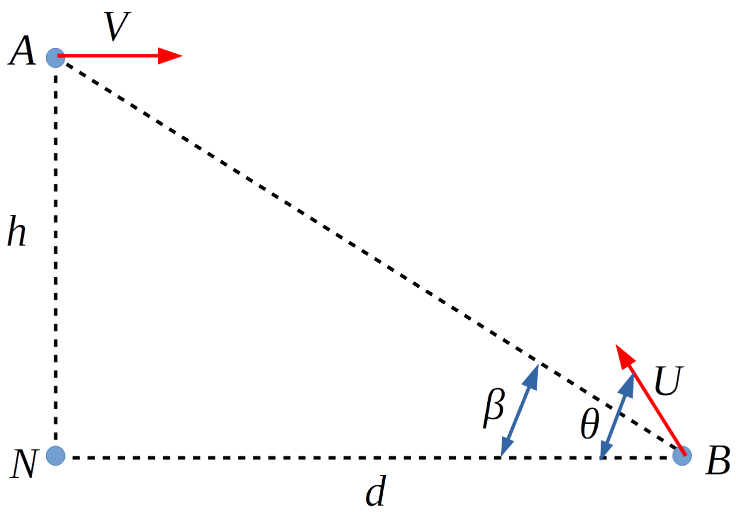
Most of the candidates received zero marks for part **(ii)**, failing to realise that the result follows from the height of the particle at the time of collision being non-negative. Some tried to use conservation of momentum or energy, or the equation $v^2 = u^2 + 2as$ due to the answer being suggestive of velocity squared. Candidates who were able to progress well on this part generally achieved all of the marks.

Very few candidates progressed to part **(iii)** and the attempts were often poor. Candidates who did know how to proceed to the result often did not justify bounds they used to obtain inequalities.

A significant number of candidates attempted the final part of the question having omitted earlier parts. In many cases candidates did not fully appreciate the requirements when asked to show a statement of the form "A if and only if B".

Solution

Start by putting the information into a diagram:



Note that in my diagram I have $\theta > \beta$, but I have not shown this to be true yet.

The position of the particle projected from A at time t is given by:

$$s_A = \begin{pmatrix} Vt \\ h - \frac{1}{2}gt^2 \end{pmatrix}$$

and the position of the particle projected from B is:

$$s_B = \begin{pmatrix} d - Ut\cos\theta \\ Ut\sin\theta - \frac{1}{2}gt^2 \end{pmatrix}$$

When the particles meet the x and y displacements are the same, so we have:

$$Vt = d - Ut\cos\theta$$
$$h - \tfrac{1}{2}gt^2 = Ut\sin\theta - \tfrac{1}{2}gt^2$$

From the second equation we have $t = \dfrac{h}{U\sin\theta}$, and substituting this into the first equation we have:

$$V \times \frac{h}{U\sin\theta} = d - U\cos\theta \times \frac{h}{U\sin\theta}$$
$$\frac{Vh}{U} = d\sin\theta - h\cos\theta$$

(i) Since we have $V, h, U > 0$ then we have:

$$\begin{aligned} d\sin\theta - h\cos\theta &> 0 \\ d\sin\theta &> h\cos\theta \\ \frac{\sin\theta}{\cos\theta} &> \frac{h}{d} \\ \tan\theta &> \tan\beta \qquad \left(\text{from the diagram } \frac{h}{d} = \tan\beta\right) \end{aligned}$$

Since both θ and β are in the range $(0, \frac{1}{2}\pi)$ and we know that $\tan\theta$ is an increasing function in this range then we have $\tan\theta > \tan\beta \implies \theta > \beta$.

(ii) The point at which they meet must have a non-negative vertical displacement, so considering the y component of s_B we have

$$\begin{aligned} Ut\sin\theta - \tfrac{1}{2}gt^2 &\geqslant 0 \\ U\sin\theta &\geqslant \tfrac{1}{2}gt \\ U\sin\theta &\geqslant \tfrac{1}{2}g \times \frac{h}{U\sin\theta} \\ (U\sin\theta)^2 &\geqslant \tfrac{1}{2}gh \\ U\sin\theta &\geqslant \sqrt{\frac{gh}{2}} \qquad \text{as both sides are positive} \end{aligned}$$

(iii) Considering $d\sin\theta - h\cos\theta$, let $d = R\cos\alpha$ and $h = R\sin\alpha$. Then we have $\tan\alpha = \dfrac{h}{d} = \tan\beta$, and so $\alpha = \beta$. We also have $R^2 = d^2 + h^2$. Using this in the stem result gives:

$$\begin{aligned} \frac{Vh}{U} &= d\sin\theta - h\cos\theta \\ \frac{Vh}{U} &= \sqrt{d^2 + h^2}(\cos\beta\sin\theta - \cos\theta\sin\beta) \\ \frac{Vh}{U} &= \sqrt{d^2 + h^2}\sin(\theta - \beta) \end{aligned}$$

Since $0 < \theta - \beta < \frac{1}{2}\pi$ we have $0 < \sin(\theta - \beta) < 1$ and so:

$$\begin{aligned} \frac{Vh}{U} &< \sqrt{d^2 + h^2} \\ \frac{U}{V} &> \frac{h}{\sqrt{d^2 + h^2}} \\ \frac{U}{V} &> \sin\beta \qquad \text{(using the right angled triangle in the diagram)} \end{aligned}$$

The last request is an if and only if, so it needs careful handling. Note that the particle projected from B is moving upwards if and only if the vertical component of the velocity is positive, i.e. we have $U\sin\theta - gt > 0$.

Substituting in the value of t when they collide, we have that B is moving upwards if and only if

$$\begin{aligned} U\sin\theta - \frac{gh}{U\sin\theta} &> 0 \\ \iff (U\sin\theta)^2 &> gh \end{aligned}$$

The height at which they collide is given by $h - \frac{1}{2}gt^2 = h - \frac{1}{2}g\left(\frac{h}{U\sin\theta}\right)^2$. This is greater than $\frac{1}{2}h$ if and only if:

$$\begin{aligned} & h - \tfrac{1}{2}g\left(\frac{h}{U\sin\theta}\right)^2 > \tfrac{1}{2}h \\ \iff & \tfrac{1}{2}h > \tfrac{1}{2}g\left(\frac{h}{U\sin\theta}\right)^2 \\ \iff & h > \frac{gh^2}{(U\sin\theta)^2} \\ \iff & (U\sin\theta)^2 > gh \end{aligned}$$

This is the same condition as before, i.e. we have:

$$\begin{aligned} B \text{ moving upwards when collide} &\iff (U\sin\theta)^2 > gh \\ \text{Collide at a height greater than } \tfrac{1}{2}h &\iff (U\sin\theta)^2 > gh \end{aligned}$$

Hence we have:

$$B \text{ moving upwards when collide} \iff \text{Collide at a height greater than } \tfrac{1}{2}h$$

Question 10

10 A particle P of mass m moves freely and without friction on a wire circle of radius a, whose axis is horizontal. The highest point of the circle is H, the lowest point of the circle is L and angle $PHL = \theta$. A light spring of modulus of elasticity λ is attached to P and to H. The natural length of the spring is l, which is less than the diameter of the circle.

(i) Show that, if there is an equilibrium position of the particle at $\theta = \alpha$, where $\alpha > 0$, then $\cos\alpha = \dfrac{\lambda l}{2(a\lambda - mgl)}$.

Show also that there will only be such an equilibrium position if $\lambda > \dfrac{2mgl}{2a - l}$.

When the particle is at the lowest point L of the circular wire, it has speed u.

(ii) Show that, if the particle comes to rest before reaching H, it does so when $\theta = \beta$, where $\cos\beta$ satisfies

$$(\cos\alpha - \cos\beta)^2 = (1 - \cos\alpha)^2 + \frac{mu^2}{2a\lambda}\cos\alpha,$$

where $\cos\alpha = \dfrac{\lambda l}{2(a\lambda - mgl)}$.

Show also that this will only occur if $u^2 < \dfrac{2a\lambda}{m}(2 - \sec\alpha)$.

Examiner's report

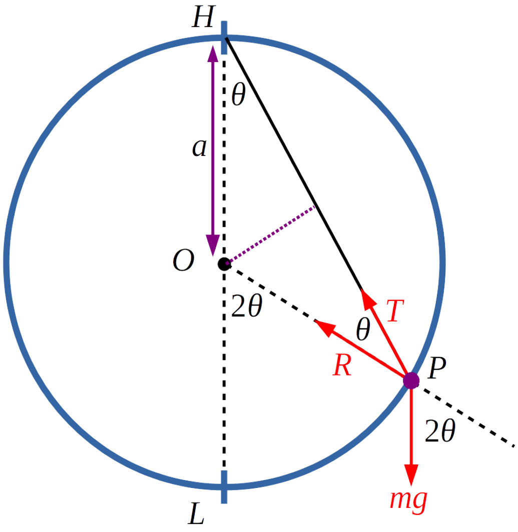
This was the least popular question on the paper and was found to be quite difficult by those who did attempt it. It is very useful in questions of this type to produce a good sketch of the situation described and, unfortunately, many attempts did not do this. The candidates who made good progress did tend to have good diagrams.

When resolving forces, it is useful to consider the different directions that could be chosen. In the case of this question, many candidates chose to resolve horizontally and vertically and as a result produced more complicated equations to deal with. While in most cases they were able to simplify the equations that they obtained, this method did result in more work than was necessary. When dividing equations and inequalities by expressions, candidates should be careful to consider whether the expressions are known to be non-zero (and in the case of inequalities that it is known whether it is positive or negative).

In the second part of the question, candidates struggled with the more complicated algebraic expressions that had to be manipulated and many candidates gave up before reaching the end of the question.

Solution

As this is a mechanics question, a good starting point is to put the information from the stem into a diagram.



We are told that $\angle PHL = \theta$, and if O is the centre of the circle then we have $\angle LOP = 2\theta$ (since $\triangle OPH$ is an isosceles triangle we have $\angle OHL = \angle OPH = \theta$).

There are three forces acting on P, weight, the tension in the spring and the reaction force that the wire circle exerts on P, which acts towards the centre of the circle. There is no frictional force (we are told this in the stem!).

The distance HP can be found by using right angled trigonometry to get $HP = 2a\cos\theta$. (You can also use the Sine rule or Cosine rule, but this takes a little more work).

(i) If the particle is in equilibrium when $\theta = \alpha$ then resolving tangentially gives:

$$T\sin\alpha = mg\sin 2\alpha \qquad (*)$$

Using Hooke's law, $T = \dfrac{\lambda x}{l}$, where x is the extension of the spring we have:

$$T = \frac{\lambda(2a\cos\alpha - l)}{l}$$

Substituting this into (*) gives:

$$\begin{aligned}\frac{\lambda(2a\cos\alpha - l)}{l}\sin\alpha &= mg\sin 2\alpha\\ \frac{\lambda(2a\cos\alpha - l)}{l}\sin\alpha &= 2mg\sin\alpha\cos\alpha\\ \sin\alpha\left(\frac{\lambda(2a\cos\alpha - l)}{l} - 2mg\cos\alpha\right) &= 0\end{aligned}$$

We are told that $\alpha > 0$ (and we must have $\alpha < \frac{\pi}{2}$), and so we know that $\sin\alpha \neq 0$.

Hence we have:

$$\begin{aligned}\frac{\lambda(2a\cos\alpha - l)}{l} - 2mg\cos\alpha &= 0\\ \lambda(2a\cos\alpha - l) &= 2mgl\cos\alpha\\ \cos\alpha(2a\lambda - 2mgl) &= \lambda l\\ \cos\alpha &= \frac{\lambda l}{2(a\lambda - mgl)}\end{aligned}$$

For this to be an achievable equilibrium point, we need $0 < \cos\alpha < 1$.
Using $\cos\alpha > 0$ means we must have $a\lambda - mgl > 0$.
Using $\cos\alpha < 1$ we have:

$$\begin{aligned}\frac{\lambda l}{2(a\lambda - mgl)} &< 1\\ \lambda l &< 2(a\lambda - mgl) \quad \text{(since } a\lambda - mgl > 0 \text{ we can do this!)}\\ 2mgl &< (2a - l)\lambda\\ \frac{2mgl}{2a - l} &< \lambda\end{aligned}$$

The last step here is fine as we are told that l is less than the diameter of the circle ($2a$), and so we can divide by $2a - l$ whilst keeping the direction of the inequality.

Hence there is an equilibrium point only if $\lambda > \dfrac{2mgl}{2a - l}$.

(ii) Looking at the required result, we have a "u^2" term, which suggests that conservation of energy might be a useful thing to consider. When P is at a point so that $\angle PHL = \theta$ assume it has speed v. I will take the horizontal line through L as the reference height for zero potential energy.

Using conservation of energy (KE, PE and EPE) we have:

$$\tfrac{1}{2}mu^2 + \tfrac{\lambda}{2l}(2a - l)^2 = \tfrac{1}{2}mv^2 + \tfrac{\lambda}{2l}(2a\cos\theta - l)^2 + mg(a - a\cos 2\theta)$$

For the potential energy you need to find the vertical height of point P. The easiest thing to do is to consider a right angled triangle with vertices at O and P.

If the particle comes to rest at $\theta = \beta$ then at this point we will have $v = 0$ and so:

$$\begin{aligned}\tfrac{1}{2}mu^2 + \tfrac{\lambda}{2l}(2a - l)^2 &= \tfrac{\lambda}{2l}(2a\cos\beta - l)^2 + mga(1 - \cos 2\beta)\\ lmu^2 + \lambda(2a - l)^2 &= \lambda(2a\cos\beta - l)^2 + 2mgal(2 - 2\cos^2\beta)\\ lmu^2 + \lambda(2a - l)^2 &= (4a^2\lambda - 4mgal)\cos^2\beta - 4al\lambda\cos\beta + \lambda l^2 + 4mgal\\ lmu^2 + 4a^2\lambda - 4al\lambda + \cancel{\lambda l^2} &= 2a(2a\lambda - 2mgl)\cos^2\beta - 4al\lambda\cos\beta + \cancel{\lambda l^2} + 4mgal\\ \frac{lmu^2}{2(a\lambda - mgl)} + \frac{4a\lambda(a - l)}{2(a\lambda - mgl)} &= 2a\cos^2\beta - \frac{4al\lambda}{2(a\lambda - mgl)}\cos\beta + \frac{4mgal}{2(a\lambda - mgl)}\end{aligned}$$

We have $\cos\alpha = \dfrac{\lambda l}{2(a\lambda - mgl)}$. This means we can write:

$$\begin{aligned}
\frac{\cos\alpha}{\lambda} \times mu^2 + \frac{\cos\alpha}{l} \times 4a(a-l) &= 2a\cos^2\beta - 4a\cos\alpha\cos\beta + \frac{\cos\alpha}{\lambda} \times 4mga \\
\frac{mu^2}{2a\lambda}\cos\alpha + \frac{\cos\alpha}{l} \times 2(a-l) &= \cos^2\beta - 2\cos\alpha\cos\beta + \frac{\cos\alpha}{2\lambda} \times 4mg \\
\frac{mu^2}{2a\lambda}\cos\alpha + \frac{2a\cos\alpha}{l} - 2\cos\alpha &= \cos^2\beta - 2\cos\alpha\cos\beta + \frac{\cos\alpha}{2\lambda} \times 4mg \\
\frac{mu^2}{2a\lambda}\cos\alpha + \frac{2a\cos\alpha}{l} - \frac{4mg\cos\alpha}{2\lambda} - 2\cos\alpha &= \cos^2\beta - 2\cos\alpha\cos\beta \\
\frac{mu^2}{2a\lambda}\cos\alpha + \cos\alpha \times \frac{4a\lambda - 4mgl}{2l\lambda} - 2\cos\alpha &= (\cos\alpha - \cos\beta)^2 - \cos^2\alpha \\
\frac{mu^2}{2a\lambda}\cos\alpha + \cos\alpha \times \frac{2(a\lambda - mgl)}{l\lambda} - 2\cos\alpha &= (\cos\alpha - \cos\beta)^2 - \cos^2\alpha \\
\frac{mu^2}{2a\lambda}\cos\alpha + \cos\alpha \times \frac{1}{\cos\alpha} - 2\cos\alpha &= (\cos\alpha - \cos\beta)^2 - \cos^2\alpha \\
\frac{mu^2}{2a\lambda}\cos\alpha + 1 - 2\cos\alpha + \cos^2\alpha &= (\cos\alpha - \cos\beta)^2 \\
\frac{mu^2}{2a\lambda}\cos\alpha + (1 - \cos\alpha)^2 &= (\cos\alpha - \cos\beta)^2
\end{aligned}$$

There was a lot of algebraic manipulation here, and a certain amount of "brute force and ignorance" required, i.e. just keep going and try to force some $\cos\alpha$ terms where possible. It's a nice moment when things start to simplify and you can see you will actually get to the required answer!

If this is going to happen then we need $0 < \cos\beta < 1$, and we already have $0 < \cos\alpha < 1$. This means that $(\cos\alpha - \cos\beta)^2 < (\cos\alpha)^2$, and so we have:

$$\begin{aligned}
\cos^2\alpha &> \frac{mu^2}{2a\lambda}\cos\alpha + (1 - \cos\alpha)^2 \\
0 &> \frac{mu^2}{2a\lambda}\cos\alpha + 1 - 2\cos\alpha \\
u^2 &< (2\cos\alpha - 1) \times \frac{2a\lambda}{m\cos\alpha} \\
u^2 &< \frac{2a\lambda}{m}(2 - \sec\alpha)
\end{aligned}$$

Question 11

11 A coin is tossed repeatedly. The probability that a head appears is p and the probability that a tail appears is $q = 1 - p$.

(i) A and B play a game. The game ends if two successive heads appear, in which case A wins, or if two successive tails appear, in which case B wins.

Show that the probability that the game never ends is 0.

Given that the first toss is a head, show that the probability that A wins is $\dfrac{p}{1 - pq}$.

Find and simplify an expression for the probability that A wins.

(ii) A and B play another game. The game ends if three successive heads appear, in which case A wins, or if three successive tails appear, in which case B wins.

Show that

$$\text{P(A wins} \,|\, \text{the first toss is a head)} = p^2 + (q + pq)\,\text{P(A wins} \,|\, \text{the first toss is a tail)}$$

and give a similar result for P(A wins | the first toss is a tail).

Show that

$$\text{P(A wins)} = \frac{p^2(1 - q^3)}{1 - (1 - p^2)(1 - q^2)}.$$

(iii) A and B play a third game. The game ends if a successive heads appear, in which case A wins, or if b successive tails appear, in which case B wins, where a and b are integers greater than 1.

Find the probability that A wins this game.

Verify that your result agrees with part **(i)** when $a = b = 2$.

Examiner's report

This was the more popular of the probability questions and many good attempts were seen, although the majority were incomplete and only attempted the first one or two parts. Some candidates made errors when dealing with the conditional probabilities, often thinking for example that P(A wins) could be obtained by adding P(A wins | H first) and P(A wins | T first). In general, those that were able to work confidently with the conditional probabilities were able to perform very well on this question.

In the first part, a number of candidates failed to consider that games could begin with either heads or tails when showing that the probability that the game never ends is 0. Additionally, some candidates assumed that $p = q = \frac{1}{2}$ for the first part of the question, although they often did then use correct expressions in terms of p and q in the later parts of the question.

In part **(ii)** some candidates tried to find a way to enumerate all possible sequences for any total number of flips, but this approach almost always resulted in some of the possible cases being omitted.

Many candidates failed to spot that the solution to the third part could be found by an analogous method to that used in part **(ii)** and so in many cases no attempt was made at this final part.

Solution

(i) If the game never ends then we must have either $HTHTHT\cdots$ or $THTHTH\cdots$. The probability that after $2n$ tosses we have an alternating sequence is therefore $2(pq)^n$. As n tends to infinity we have $\lim_{n\to\infty} 2(pq)^n \to 0$, and so the probability that the game never ends is 0.

Given the first toss is a head, then A will win if the following tosses are H, or THH, or $THTHH$, or $THTHTHH$ etc. Therefore we have

$$\begin{aligned} \text{P(A wins} \mid \text{H first)} &= p + (qp)p + (qp)^2p + (qp)^3p + \cdots \\ &= p\Big(1 + qp + (qp)^2 + (qp)^3 + \cdots\Big) \\ &\frac{p}{1-pq} \end{aligned}$$

If the first toss is a tail then A will win if the next tosses are HH, or $HTHH$, or $HTHTHH$ etc. Note that the second toss has to be a head or B will win.

Therefore we have

$$\begin{aligned} \text{P(A wins} \mid \text{T first)} &= p^2 + (pq)p^2 + (pq)^2p^2 + (pq)^3p^2 + \cdots \\ &= p^2\Big(1 + pq + (pq)^2 + (pq)^3 + \cdots\Big) \\ &= \frac{p^2}{1-pq} \end{aligned}$$

The probability that A wins is given by:

$$\text{P(H first)} \times \text{P(A wins} \mid \text{H first)} + \text{P(T first)} \times \text{P(A wins} \mid \text{T first)}$$

$$\begin{aligned} &= p \times \frac{p}{1-pq} + q \times \frac{p^2}{1-pq} \\ &= \frac{p^2(1+q)}{1-pq} \end{aligned}$$

A tree diagram can make the first line above more obvious. The first pair of branches would be "H first" and "T first". Note that if we substitute $p = q = \frac{1}{2}$ the above expression gives $\text{P(A)} = \frac{1}{2}$, which is what we would expect if the coin was a fair coin.

It is often useful to use specific values as a test, and in this case if you had mistakenly used "$\text{P(A)} = \text{P(A wins} \mid \text{H first)} + \text{P(A wins} \mid \text{T first)}$" then using $p = q = \frac{1}{2}$ would give you $\text{P(A)} = 1$, which is a little suspicious.

(ii) Let P(A wins | H first) $= P_H(\text{A})$, and let P(A wins | T first) $= P_T(\text{A})$.
If the first toss is a Head, then if the next two are also Heads then A wins (probability p^2).
If the second toss is a Tail, then after this A will win with probability $P_T(\text{A})$.
If the second and third tosses are HT, then after this A will win with probability $P_T(\text{A})$.

This means that:

$$P_H(\text{A}) = p^2 + qP_T(\text{A}) + pqP_T(\text{A}) = p^2 + (q + pq)P_T(\text{A}) \qquad (*)$$

If the first toss is a Tail then, then if the second toss is Heads, A will win with probability $P_H(\text{A})$
If the second and third toss are TH, then after this A will win with probability $P_H(\text{A})$
If the second and third tosses are Tails the B wins.

This means that:

$$\text{P}_T(\text{A}) = pP_H(\text{A}) + qpP_H(\text{A}) = (p + pq)P_H(\text{A}) \qquad (\dagger)$$

Substituting for $\text{P}_T(\text{A})$ in (*) gives:

$$\begin{aligned} P_H(\text{A}) &= p^2 + (q + pq)(p + pq)P_H(\text{A}) \\ P_H(\text{A})\Big(1 - (q + pq)(p + pq)\Big) &= p^2 \\ P_H(\text{A}) &= \frac{p^2}{1 - pq(1 + p)(1 + q)} \end{aligned}$$

Substituting for $P_H(\text{A})$ into (†) gives:

$$\begin{aligned} \text{P}_T(\text{A}) &= (p + pq)P_H(\text{A}) \\ &= p(1 + q) \times \frac{p^2}{1 - pq(1 + p)(1 + q)} \\ &= \frac{p^3(1 + q)}{1 - pq(1 + p)(1 + q)} \end{aligned}$$

and so:

$$\begin{aligned} \text{P}(\text{A}) &= p \times P_H(\text{A}) + q \times P_T(\text{A}) \\ &= \frac{p^3}{1 - pq(1 + p)(1 + q)} + \frac{qp^3(1 + q)}{1 - pq(1 + p)(1 + q)} \\ &= \frac{p^3 + qp^3(1 + q)}{1 - pq(1 + p)(1 + q)} \end{aligned}$$

This doesn't look like the required form yet, so a bit of work is needed, including use of $q = 1 - p$ and $p = 1 - q$.

$$\begin{aligned} \text{P}(\text{A}) &= \frac{p^3\Big(1 + q(1 + q)\Big)}{1 - pq(1 + p)(1 + q)} \\ &= \frac{p^3(1 + q + q^2)}{1 - (1 - q)(1 - p)(1 + p)(1 + q)} \\ &= \frac{p^2(1 - q)(1 + q + q^2)}{1 - (1 - p)(1 + p)(1 - q)(1 + q)} \\ &= \frac{p^2(1 - q^3)}{1 - (1 - p^2)(1 - q^2)} \end{aligned}$$

(iii) In this game, A will win if there are a successive heads before there are b successive tails. Let $P_H(\mathrm{A})$ and $P_T(\mathrm{A})$ be as before.

If the first toss is a head then the next tosses could be T, HT, HHT, up to $a-2$ Heads followed by a tail and $a-1$ Heads (in which case A has won as there have been a successive heads. Hence we have:

$$\begin{aligned} P_H(\mathrm{A}) &= p^{a-1} + \left(q + pq + p^2q + \cdots + p^{a-2}q\right) P_T(\mathrm{A}) \\ &= p^{a-1} + q\left(1 + p + p^2 + \cdots + p^{a-2}\right) P_T(\mathrm{A}) \\ &= p^{a-1} + \cancel{q}\left(\frac{1-p^{a-1}}{\cancel{1-p}}\right) P_T(\mathrm{A}) \end{aligned}$$

If there is a tail first, then the next tosses could be H, TH, TTH up to and including $b-2$ tails followed by a head. We have:

$$\begin{aligned} P_T(\mathrm{A}) &= (p + qp + q^2p + \cdots + q^{b-2}p)P_H(\mathrm{A}) \\ &= p(1 + q + q^2 + \cdots + q^{b-2})P_H(\mathrm{A}) \\ &= \cancel{p}\left(\frac{1-q^{b-1}}{\cancel{1-q}}\right) \mathrm{P}_H(\mathrm{A}) \end{aligned}$$

Substituting for $P_T(\mathrm{A})$ gives:

$$\begin{aligned} P_H(\mathrm{A}) &= p^{a-1} + (1-p^{a-1})(1-q^{b-1})\mathrm{P}_H(\mathrm{A}) \\ P_H(\mathrm{A}) &= \frac{p^{a-1}}{1-(1-p^{a-1})(1-q^{b-1})} \end{aligned}$$

and also:

$$P_T(\mathrm{A}) = \frac{p^{a-1}(1-q^{b-1})}{1-(1-p^{a-1})(1-q^{b-1})}$$

Then using $\mathrm{P}(\mathrm{A}) = p \times P_H(\mathrm{A}) + q \times P_T(\mathrm{A})$ we have:

$$\begin{aligned} \mathrm{P}(\mathrm{A}) &= \frac{p \times p^{a-1}}{1-(1-p^{a-1})(1-q^{b-1})} + \frac{q \times p^{a-1}(1-q^{b-1})}{1-(1-p^{a-1})(1-q^{b-1})} \\ &= \frac{p^a + qp^{a-1} - p^{a-1}q^b}{1-(1-p^{a-1})(1-q^{b-1})} \\ &= \frac{p^{a-1}(p+q-q^b)}{1-(1-p^{a-1})(1-q^{b-1})} \\ &= \frac{p^{a-1}(1-q^b)}{1-(1-p^{a-1})(1-q^{b-1})} \end{aligned}$$

Substituting in $a = b = 2$ gives:

$$\begin{aligned} \mathrm{P}(\mathrm{A}) &= \frac{p^{2-1}(1-q^2)}{1-(1-p^{2-1})(1-q^{2-1})} \\ &= \frac{p(1-q^2)}{1-(1-p)(1-q)} \\ &= \frac{p(1-q)(1+q)}{1-qp} \\ &= \frac{p^2(1+q)}{1-pq} \end{aligned}$$

which agrees with the answer in part **(i)**.

Question 12

12 The score shown on a biased n-sided die is represented by the random variable X which has distribution $\mathrm{P}(X = i) = \dfrac{1}{n} + \varepsilon_i$ for $i = 1, 2, \ldots, n$, where not all the ε_i are equal to 0.

(i) Find the probability that, when the die is rolled twice, the same score is shown on both rolls. Hence determine whether it is more likely for a fair die or a biased die to show the same score on two successive rolls.

(ii) Use part **(i)** to prove that, for any set of n positive numbers x_i $(i = 1, 2, \ldots, n)$,

$$\sum_{i=2}^{n} \sum_{j=1}^{i-1} x_i x_j \leqslant \frac{n-1}{2n} \left(\sum_{i=1}^{n} x_i \right)^2 .$$

(iii) Determine, with justification, whether it is more likely for a fair die or a biased die to show the same score on three successive rolls.

Examiner's report

This question had the second smallest number of attempts on the paper. Many of these successfully completed the first part of the question, but then made little progress in the later sections.

Part **(i)** was generally well done, although some candidates did not appreciate that $\sum \varepsilon_i = 0$ when determining whether a fair or biased die is more likely to show the same score on two successive rolls. Where candidates were able to see the connection between the first and second parts of the question, the required result was generally proven clearly.

Part **(iii)** could be approached in a similar way to part **(i)**, but many of the candidates who reached this point failed to deal with the more complicated terms that arise from the expansion. Correct solutions to this part were generally very well set out.

Solution

In more modern British English, “dice” is probably the more common singular noun for the spotty cube used when playing board games, but STEP often uses the more archaic version “die”. In American English “die” is more commonly used.

Note that if all of the $\varepsilon_i = 0$ then the dice would not be biased!

(i) Let R_1 be the score on roll 1, and R_2 be the score on roll 2. We have:

$$\begin{aligned}\mathrm{P}(R_1 = R_2) &= \mathrm{P}(1,1) + \mathrm{P}(2,2) + \cdots + \mathrm{P}(n,n) \\ &= \left(\frac{1}{n} + \varepsilon_1\right)^2 + \left(\frac{1}{n} + \epsilon_2\right)^2 + \cdots + \left(\frac{1}{n} + \varepsilon_n\right)^2 \\ &= \sum_{i=1}^{n} \left(\frac{1}{n} + \varepsilon_i\right)^2\end{aligned}$$

We can simplify this a little. We know that the sum of the probabilities of X has to be equal to 1, so we have:

$$\begin{aligned}\sum_{i=1}^{n}\left(\frac{1}{n} + \varepsilon_i\right) &= 1 \\ n \times \frac{1}{n} + \sum_{i=1}^{n} \varepsilon_i &= 1 \\ \implies \sum_{i=1}^{n} \varepsilon_i &= 0\end{aligned}$$

Hence the probability that the two scores are the same can be written as:

$$\begin{aligned}\sum_{i=1}^{n}\left(\frac{1}{n} + \varepsilon_i\right)^2 &= \sum_{i=1}^{n}\left[\left(\frac{1}{n}\right)^2 + \frac{2}{n}\varepsilon_i + \varepsilon_i^2\right] \\ &= n \times \left(\frac{1}{n}\right)^2 + \frac{2}{n}\sum_{i=1}^{n}\varepsilon_i + \sum_{i=1}^{n}\varepsilon_i^2 \\ &= \frac{1}{n} + \sum_{i=1}^{n}\varepsilon_i^2\end{aligned}$$

Since we have $\varepsilon_i^2 \geqslant 0$, then this probability is greater than $\dfrac{1}{n}$ (which is the probability of getting the same score on both tosses of a fair dice) and so this probability is greater for a biased dice than a fair dice.

(ii) First thing to do is have a think about the result we are being asked to prove. We have a double sum here, which can look intimidating. Note though that we can take the "x_i" term out of the inner sum (a bit like removing a constant outside of a sum or integral):

$$\begin{aligned}\sum_{i=2}^{n}\sum_{j=1}^{i-1} x_i x_j &= \sum_{i=2}^{n} x_i \left(\sum_{j=1}^{i-1} x_j\right) \\ &= x_2 \times \sum_{j=1}^{1} x_j + x_3 \times \sum_{j=1}^{2} x_j + \cdots + x_n \times \sum_{j=1}^{n-1} x_j \\ &= x_2 x_1 + x_3(x_1 + x_2) + \cdots + x_n(x_1 + x_2 + \cdots x_{n-1})\end{aligned}$$

This is the sum of half of the cross terms on $(x_1 + x_2 + x_3 + \cdots + x_n)^2$, as shown below:

	x_1	x_2	x_3	$\cdots$	x_{n-1}	x_n
x_1	x_1^2	x_1x_2	x_1x_3	$\cdots$	x_1x_{n-1}	x_1x_n
x_2	x_2x_1	x_2^2	x_2x_3	$\cdots$	x_2x_{n-1}	x_2x_n
x_3	x_3x_1	x_3x_2	x_3^2	$\cdots$	x_3x_{n-1}	x_3x_n
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_{n-1}	$x_{n-1}x_1$	$x_{n-1}x_2$	$x_{n-1}x_3$	$\cdots$	x_{n-1}^2	$x_{n-1}x_n$
x_n	x_nx_1	x_nx_2	x_nx_3	$\cdots$	x_nx_{n-1}	x_n^2

It would not be necessary to replicate the above in an exam, it is there as a motivation for what I will do next!

We also know by symmetry that the sum of the terms below the diagonal is equal to the sum of the terms above the diagonal, so we have:

$$\sum_{i=2}^{n}\sum_{j=1}^{i-1} x_ix_j = \frac{1}{2}\left[(x_1+x_2+\cdots+x_n)^2 - \sum_{i=1}^{n} x_i^2\right]$$

Let F_1 be the score on roll 1 with a fair dice, and F_2 be the score on roll 2 of a fair dice. From part **(i)** we know that $\mathrm{P}(R_1 = R_2) \geqslant \mathrm{P}(F_1 = F_2)$, and we also know that $\mathrm{P}(R_1 > R_2) = \mathrm{P}(R_1 < R_2)$ by symmetry (and similar for the fair dice). We also have $\mathrm{P}(R_1 > R_2) + \mathrm{P}(R_1 = R_2) + \mathrm{P}(R_1 < R_2) = 1$ and so:

$$\mathrm{P}(R_1 > R_2) = \frac{1}{2}\Big(1 - \mathrm{P}(R_1 = R_2)\Big)$$

Since we have $\mathrm{P}(R_1 = R_2) \geqslant \mathrm{P}(F_1 = F_2)$ this means that $\mathrm{P}(R_1 > R_2) \leqslant \mathrm{P}(F_1 > F_2)$.

Consider a fair n sided dice. The number of combinations which have $\mathrm{P}(F_1 > F_2)$ is equal to the number of entries below the diagonal in the above table, and so there are $\frac{n(n-1)}{2}$ different ways of doing this. Each individual entry has probability $\frac{1}{n^2}$ and so we have:

$$\mathrm{P}(F_1 > F_2) = \frac{n(n-1)}{2n^2} = \frac{n-1}{2n}$$

Alternatively, and probably more simply:

$$\begin{aligned}\mathrm{P}(F_1 > F_2) &= \frac{1}{2}\Big(1 - \mathrm{P}(F_1 = F_2)\Big)\\ &= \frac{1}{2}\left(1 - \frac{1}{n}\right)\\ &= \frac{n-1}{2n}\end{aligned}$$

Let $\displaystyle\sum_{i=1}^{n} x_i = N$, and then let $p_i = \dfrac{x_i}{N}$. We then have $\sum p_i = 1$ and so consider a (possibly biased) dice with $\mathrm{P}(X = i) = p_i$.

We then have:

$$\begin{aligned}P(R_1 > R_2) &= \sum_{i=2}^{n}\sum_{j=1}^{i-1}\frac{x_i x_j}{N^2}\\ &= \frac{1}{N^2}\sum_{i=2}^{n}\sum_{j=1}^{i-1}x_i x_j\end{aligned}$$

Using $P(R_1 > R_2) \leqslant P(F_1 > F_2)$ we have:

$$\begin{aligned}\frac{1}{N^2}\sum_{i=2}^{n}\sum_{j=1}^{i-1}x_i x_j &\leqslant \frac{n-1}{2n}\\ \sum_{i=2}^{n}\sum_{j=1}^{i-1}x_i x_j &\leqslant \frac{n-1}{2n}\times N^2\\ \sum_{i=2}^{n}\sum_{j=1}^{i-1}x_i x_j &\leqslant \frac{n-1}{2n}\times\left(\sum_{i=1}^{n}x_i\right)^2\end{aligned}$$

as we set $\displaystyle\sum_{i=1}^{n}x_i = N$ earlier.

(iii) For the fair dice we have:

$$P(F_1 = F_2 = F_3) = \sum_{i=1}^{n}\frac{1}{n^3} = \frac{1}{n^2}$$

For the unbiased dice we have:

$$\begin{aligned}P(R_1 = R_2 = R_3) &= \sum_{i=1}^{n}\left(\frac{1}{n}+\varepsilon_i\right)^3\\ &= \sum_{i=1}^{n}\frac{1}{n^3} + \sum_{i=1}^{n}\frac{3\varepsilon_i}{n^2} + \sum_{i=1}^{n}\frac{3\varepsilon_i^2}{n} + \sum_{i=1}^{n}\varepsilon_i^3\\ &= \frac{1}{n^2} + \frac{3}{n^2}\sum_{i=1}^{n}\varepsilon_i + \frac{3}{n}\sum_{i=1}^{n}\varepsilon_i^2 + \sum_{i=1}^{n}\varepsilon_i^3\end{aligned}$$

Considering $P(R_1 = R_2 = R_3) - P(F_1 = F_2 = F_3)$ and noting that $\sum \epsilon_i = 0$ (as before) we have:

$$\begin{aligned}P(R_1 = R_2 = R_3) - P(F_1 = F_2 = F_3) &= \frac{3}{n}\sum_{i=1}^{n}\varepsilon_i^2 + \sum_{i=1}^{n}\varepsilon_i^3\\ &= \sum_{i=1}^{n}\left[\varepsilon_i^2\left(\tfrac{3}{n}+\varepsilon_i\right)\right]\end{aligned}$$

We know that $\varepsilon_i^2 \geqslant 0$, and since $p_i = \frac{1}{n}+\varepsilon_i \geqslant 0$ we must have $\frac{3}{n}+\varepsilon_i > 0$ (as this is $\frac{2}{n}$ greater than p_i). Hence we have a sum of non-negative terms, or which at least some are positive (as not all of the ε_i are equal to 0).

Hence we have $P(R_1 = R_2 = R_3) - P(F_1 = F_2 = F_3) > 0$ and the probability of scoring three scores the same is greater for a biased dice.