Advanced Problems

in

Core Mathematics

Stephen Siklos

Fourth edition, October 2008

ABOUT THIS BOOKLET

This booklet is intended to help you to prepare for STEP examinations. It should also be useful as preparation for any undergraduate mathematics course, even if you do not plan to take STEP.

The questions are all based on recent STEP questions. I chose the questions either because they are 'nice' – in the sense that you should get a lot of pleasure from tackling them – or because I felt I had something interesting to say about them. In this booklet, I have restricted myself (reluctantly) to the syllabus for Papers I and II, which is the A-level core (i.e. C1 to C4) with a few additions. This material should be familiar to you if you are taking the International Baccalaureate, Scottish Advanced Highers or other similar courses.

The first two questions (the sample worked questions) are in a 'stream of consciousness' format. They are intended to give you an idea how a trained mathematician would think when tackling them. This approach is much too long-winded to sustain, but it should help you to see what sort of questions you should be asking yourself as you work through the later questions.

I have given each of the subsequent questions a difficulty rating ranging from $(*)$ to $(*\,*\,*)$. A question labelled $(*)$ might be found on STEP I; a question labelled $(**)$ might be found on STEP II; a question labelled $(***)$ might be found on STEP III. But difficulty in mathematics is in the eye of the beholder: you might find a question difficult simply because you overlooked some key step, which on another day you would not have hesitated over. You should not therefore be discouraged if you are stuck on a one-star question; though you should probably be encouraged if you get through three-star questions without mishap.

Each question is followed by a comment and a full solution. The comment will contain hints and/or direct your attention to key points; or it might put the question in its true mathematical context. The solution should give sufficient working for you to read it through and pick up the gist of the method. It is of course just one way of producing the required result: there may be many other equally good or better ways.

Calculators are not required for any of the problems in this booklet; and calculators are not permitted in STEP examinations. In the early days of STEP, calculators were permitted but they were not required for any question. It was found that candidates who tried to use calculators sometimes ended up missing the point of the question or getting a silly answer. My advice is to remove the battery so that you are not tempted.

I hope that you will use the comments and solutions as springboards rather than feather beds. You will only really benefit from this booklet if you have a good go at each question before looking at the comment and another good go before looking at the solution. You should gain by reading both even if you have completed the question on your own.

If you find errors or have suggestions for ways in which this booklet can be improved, please e-mail me: stcs@cam.ac.uk. I update it regularly.

I hope you enjoy using this booklet as much as I have enjoyed putting it together.

STCS

STEP MATHEMATICS

What is STEP?

STEP (Sixth Term Examination Paper) is an examination used by Cambridge colleges as the basis for conditional offers. There were STEPs in many subjects but after 2001 the mathematics papers only remain. They are used not just by Cambridge, but also by Warwick, and many other university mathematics departments recommend that their appicants practise on the past papers even if they do not take the examination. In 2008, 2100 scripts were marked, only about 750 of which were written by Cambridge applicants.

There are three STEPs, called papers I, II and III. Papers I and II are based the core A-level syllabus, with some minor additions, and Paper III is based on a 'typical' further mathematics mathematics syllabus (for which there is no recognised core). The questions on Papers II and III are supposed to be of about the same difficulty, and are harder than those on Paper I. Each paper has thirteen questions, including three on mechanics and two on probability/statistics. Candidates are assessed on six questions only.

What is the purpose of STEP?

From the point of view of admissions to a university mathematics course, STEP has three purposes. First, it acts as a hurdle: success in STEP is thought to be a good indicator of potential to do well on a difficult course. Second, it acts as preparation for the course, because the style of mathematics found in STEP questions is similar to that of undergraduate mathematics. Thirdly, it tests motivation. It is important to prepare for STEP (by working through old papers, for example), which can require considerable dedication. Those who are not willing to make the effort are unlikely to thrive on a difficult mathematics course.

STEP *vs* A-level

A-level tests mathematical knowledge and technique by asking you to tackle fairly stereotyped problems. STEP asks you to apply the same knowledge and technique to problems that are, ideally, unfamiliar. Here is a quotation from Roger Porkess[1] which illustrates why A-level examinations do not test satisfactorily the skills required to tackle university mathematics courses at the highest levels:

> Students following modular syllabuses work much harder and the system rewards them for doing so. This means that there is a new sort of A–C grade student: not necessarily a particularly intuitive mathematician but someone who has shown the ability to learn the subject successfully by dint of hard work and, quite possibly, good teaching.

No one believes that modular A-levels are uniformly bad; anything that encourages more students to study mathematics must be basically good. However, they are not good preparation for the students who plan to take a degree in mathematics from one of the top universities.

[1] From *Every cloud has a silver lining*: Roger Porkess (1995).

Here is an A-level question:

> By using the substitution $u = 2x - 1$, or otherwise, find
> $$\int \frac{2x}{(2x-1)^2}\,\mathrm{d}x.$$

Here is a STEP question, which requires both competence in basic mathematical techniques and mathematical intuition. Note that help is given for the first integral, so that everyone starts at the same level. Then, for the second integral, candidates have to show that they understand why the substitution used in the first part worked, and how it can be adapted.

> Use the substitution $x = 2 - \cos\theta$ to evaluate the integral
> $$\int_{3/2}^{2} \left(\frac{x-1}{3-x}\right)^{\frac{1}{2}} \mathrm{d}x.$$
> Show that, for $a < b$,
> $$\int_{p}^{q} \left(\frac{x-a}{b-x}\right)^{\frac{1}{2}} \mathrm{d}x = \frac{(b-a)(\pi + 3\sqrt{3} - 6)}{12},$$
> where $p = (3a+b)/4$ and $q = (a+b)/2$.

The differences between STEP and A-level are:

1. There is considerable choice on STEP papers (6 out of 13). This is partly to allow for different A-level syllabuses.
2. STEP questions are much longer. Candidates completing four questions in three hours will almost certainly get a grade 1.
3. STEP questions are much less routine.
4. STEP questions may require considerable dexterity in performing mathematical manipulations.
5. Individual STEP questions may require knowledge of different areas of mathematics (especially the mechanics and statistics questions, which will often require advanced pure mathematical techniques).
6. The marks available for each part of the question are not disclosed on the paper.
7. Calculators are not permitted.

The Advanced Extension Award

Advanced Extension Awards were taken for the first time in 2002. The last papers, in all subjects except mathematics, will be sat in 2010. Mathematics has a reprieve for at least a few years. The AEA is not a replacement for STEP. It is aimed at the top 10% or so of the A-level cohort; i.e. at the top 7000 or soi whereas STEP is aimed at the top 2000 or so. This difference matters, because in mathematics (though not perhaps in other subjects) it is important to match the difficulty of the question and the ability of the candidates exactly.[2] The AEA contains questions on only the core A-level syllabus: no statistics and (more important for us) no mechanics. Many of my colleagues believe that, apart from the resulting reduction of choice, this is undesirable because it sends out the wrong signals about the importance of these other areas of mathematics.

Setting STEP

STEP is produced under the auspices of the Cambridge Assessment examining board. The setting procedure starts 18 months before the date of the examination, when the three examiners (one for each paper, from schools or universities) are asked to produce a draft paper. The first drafts are then vetted by the STEP coordinator, who tries to enforce uniformity of difficulty, checks suitability of material and style and tries to reduce overlap between the papers. Examiners then produce a second draft, based on the coordinator's suggestions and taking into account his/her comments. The second drafts are circulated to the three moderators (normally school teachers), and to the other examiners, who produce written comments and discuss the drafts in a two-day meeting. The examiners then produce third drafts, taking into account the consensus at the meetings. These drafts are sent to a vetter, who works through the papers, pointing out mistakes and infelicities. The final draft is checked by a second vetter. At each stage, the drafts are produced camera-ready, using a special mathematical word-processing package called LaTeX so no typesetting errors can arise after the second vetting.

STEP Questions

STEP questions do not fall into any one category. Typically, there will be a range of types on each papers. An excellent type involves two parts: in the first part, candidates are asked to perform some task and are told how to do it (integration by substitution, or expressing a quadratic as the algebraic sum of two squares, for example); for the second part, candidates are expected to demonstrate that they have understood and learned from the first part by applying the method to a new and perhaps more complicated task. Another good type of question requires candidates to do some preliminary special-case work and then guess and prove a general result. In another type, candidates have to show that they can understand and use new notation or a new theorem to perform perhaps standard tasks in disguise. Multipart questions of the A-level type are generally avoided; but if they occur, there tends to be a 'sting in the tail' involving putting all the parts together in some way.

As a general rule, questions on standard material tend to be harder than questions on material that candidates are not expected to have seen before. Sometimes, the first question on the paper is of a basic nature, and on material that will be familiar to all candidates.

There is a syllabus for STEP, which is given at the end of this book.

Often there is at least one question that does not rely on any part of the syllabus: it might be a 'common sense' question, involving counting or seeing patterns, or it might involve some aspect of GCSE mathematics with an unusual slant.

Some questions are devised to check that you do not simply apply routine methods blindly. Remember, for example, that a function can have a maximum value at which the derivative is non-zero

[2] You could reasonably ask the question 'Was Henry VIII a good king' on examination papers from GCSE to PhD level. The answers would (or should) differ according to the level. Mathematics does not work like this.

(it would have to be at the end of a range), so finding the maximum in such a case is not simply a question of routine differentiation.

Some questions take you through a new idea and then ask you to apply it to something else. Some questions try to test your capacity for clear thought without using much mathematical knowledge (like the calendar question mentioned in the next section or questions concerning islands populated by toads 'who always tell the truth' and frogs 'who always fib').

There are always questions specifically on integration or differentiation, and many others (including mechanics and probability) that use calculus as a means to an end, Graph-sketching is regarded by most mathematicians as a fundamental skill and there are nearly always one or more sketching questions or parts of questions. Some questions simply involve difficult algebraic or trigonometric manipulation. Basic ideas from analysis (such as $0 \leqslant \mathrm{f}(x) \leqslant k \Rightarrow 0 \leqslant \int_a^b \mathrm{f}(x)\mathrm{d}x \leqslant (b-a)k$ or $\mathrm{f}'(x) > 0 \Rightarrow \mathrm{f}(b) > \mathrm{f}(a)$ for $b > a$ or a relationship between an integral and a sum) often come up, though knowledge of such results is never assumed — a sketch 'proof' of an 'obvious' result may be asked for.

Some areas are difficult to examine and questions tend to come up rarely. Complex numbers fall into this category, and the same applies to (pure mathematical applications of) vectors.

The mechanics and probability questions often involve significant ideas from pure mathematics.

Questions on pure statistical tests are rare, because questions that require real understanding (rather than 'cookbook' methods) tend to be too difficult. More often, the questions in this section are about probability.

The mechanics questions normally require a firm understanding of the basic principles (when to apply conservation of momentum and energy, for example) and may well involve a differential equation. Projectile questions are often set, but are never routine.

Advice to candidates

First appearances

Your first impression on looking at a STEP paper is likely to be that it looks very hard. Don't be discouraged! Its difficult appearance is largely due to it being very different in style from A-levels. If you are considering studying mathematics at Cambridge (or Oxford or Warwick), it is likely that you will manage to work through most or all of an A-level module in the time available: maybe 10 questions in 75 minutes (i.e. 7.5 minutes per question). In STEP, you are only supposed to do six questions in three hours, and you are doing very well if you manage four; that means that each question is designed to take 45 minutes. Thus although the questions are long and arduous, you have much more time to work on them.

You may be put off by the number of subjects covered on the paper. You should not be. STEP is supposed to provide sufficient questions for all candidates, no matter which A-level mathematics syllabus they have covered. It would be a very exceptional candidate who had the knowledge required to do all the questions.

Once you get used to the idea that STEP is very different from A-level, it becomes much less daunting.

Preparation

The best preparation for STEP is to work slowly through old papers.[3] Hints and answers for the some years are available, but you should use these with discretion: doing a question with hints and answers in front of you is nothing like doing it yourself, and may well miss the whole point of the question (which is to make you think about mathematics). In general, thinking about the problem is much more important than getting the answer.

Should you try to learn up areas of mathematics that are not in your syllabus in preparation for STEP? The important thing to know is that it is much better to be very good at your syllabus than to have have a sketchy knowledge of lots of additional topics: depth rather than breadth is what matters. It may conceivably be worth your while to round out your knowledge of a topic you have already studied to fit in with the STEP syllabus; it is probably not worth your while to learn a new topic for the purposes of the exam. It is worth emphasising that there is no 'hidden agenda': a candidate who did two complete probability questions and one complete mechanics question will obtain the same grade as one who did four complete pure questions.

Just as the examiners have no hidden agenda concerning syllabus, so they have no hidden agenda concerning your method of answering the question. If you can get to the end of a question correctly you will get full marks whatever method you use. Some years ago one of the questions asked candidates to find the day of the week of a given date (say, the 5th of June 1905). A candidate who simply counted backwards day by day from the date of the exam would have received full marks for that question (but would not have had time to do any other questions). You may be worried that the examiners expect some mysterious thing called rigour. Do not worry: STEP is an exam for schools, not university students, and the examiners understand the difference. Nevertheless, it is extremely important that you present ideas clearly and show working at all stages.

Presentation

You should set out your answer legibly and logically (don't scribble down the first thought that comes into your head) – this not only helps you to avoid silly mistakes but also signals to the examiner that you know what you are doing (which can be effective even if you haven't the foggiest idea what you are doing).

Examiners are not as concerned with neatness as you might fear. However if you receive complaints from your teachers that your answers are difficult to follow then you should listen. Remember that more space often means greater legibility. Try writing on alternate lines (this leaves a blank line for corrections).

Try to read your answers with a hostile eye. Have you made it clear when you have come to the end of a particular argument? Try underlining your conclusions. Have you explained what you are trying to do? (If a question asks 'Is A true?' try beginning your answer by writing 'A is true' — if it is true — so that the examiner knows which way your argument leads.) If you used some idea (for example, integration by substitution) have you told the examiner that this is what you are doing?

[3] These and the other publications mentioned below are obtainable from the STEP web site www.stepmathematics.org.uk.

What to do if you cannot get started on a problem

Try the following, in order.

- Reread the question to check that you understand what is wanted;
- Reread the question to look for clues – the way it is phrased, or the way a formula is written, or other relevant parts of the question. (You may think that the setters are trying to set difficult questions or to catch you out. Usually, nothing could be further from the truth: they are probably doing all in their power to make it easy for you by trying to tell you what to do).
- Try to work out exactly what it is that you don't understand.
- Simplify the notation – e.g. by writing out sums explicitly.
- Look at special cases (choose special values which simplify the problem) in order to try to understand why the result is true.
- Write down your thoughts – in particular, try to express the exact reason why you are stuck.
- Go on to another question and go back later.
- If you are preparing for the examination (but not in the actual examination!) take a short break.[4]
- Discuss it with a friend or teacher (again, better not do this in the actual examination) or consult the hints and answers, but make sure you still think it through yourself. REMEMBER: following someone else's solution is not remotely the same thing as doing the problem yourself. Once you have seen someone else's solution to an example, then you are deprived, for ever, of most of the benefit that could have come from working it out yourself. Even if, ultimately, you get stuck on a particular problem, you derive vastly more benefit from seeing a solution to something with which you have already struggled, than by simply following a solution to something to which you've given very little thought.

What if a problem isn't coming out?

If you have got started but the answer doesn't seem to be coming out then:

- Check your algebra. In particular, make sure that what you have written works in special cases. For example: if you have written the series for $\log(1+x)$ as

  $$1 - x + x^2/2 - x^3/3 + \cdots$$

  then a quick check will reveal that it doesn't work for $x = 0$; clearly, the 1 should not be there.[5]

[4] Littlewood (distinguished Cambridge mathematician and author of the highly entertaining *Mathematicians Miscellany*) used to work seven days a week until an experiment revealed that when he took Sundays off the good ideas had a way of coming on Mondays.

[5] Another check will reveal that for very small positive x, log is positive (since its argument is bigger than 1) whereas the series is negative, so there is clearly something else wrong.

- Make sure that what you have written makes sense. For example, in a problem which is dimensionally consistent, you cannot add x (with dimensions length, say) to x^2 or $\exp x$ (which itself does not make sense — the argument of exp has to be dimensionless). Even if there are no dimensions in the problem, it is often possible to mentally assign dimensions and hence enable a quick check. Be wary of applying familiar processes to unfamiliar objects (very easy to do when you are feeling at sea): for example, it is all too easy, if you are not sure where your solution is going, to solve the vector equation $\mathbf{a}.\mathbf{x} = 1$ by dividing both sides by $\mathbf{a}$.
- Analyse exactly what you are being asked to do; try to understand the hints (explicit and implicit); remember to distinguish between terms such as explain/prove/define/etc. (There is essentially no difference between 'prove' and 'show': the former tends to be used in more formal situations, but if you are asked to 'show' something, a proper proof is required.)
- Remember that different parts of a question are often linked (it is sometimes obvious from the notation and choice of names of variables).
- If you get irretrievably stuck in the exam, state in words what you are trying to do and move on (at A-level, you don't get credit for merely stating intentions, but STEP examiners are generally grateful for any sign of intelligent life).

What to do after completing a question

It is a natural instinct to consider that you have finished with a question once you have done got to an answer. However this instinct should be resisted both mathematically and from the much narrower view of preparing for an examination. Instead, when you have completed a question you should stop for a few minutes and think about it. Here is a check list for you to run through.

- Look back over what you have done, checking that the arguments are correct and making sure that they work for any special cases you can think of. It is surprising how often a chain of completely spurious arguments and gross algebraic blunders leads to the given answer.
- Check that your answer is reasonable. For example if the answer is a probability p then you should check that $0 \le p \le 1$. If your answer depends on an integer n does it behave as it should when $n \to \infty$? Is it dimensionally correct?

  If, in the exam, you find that your answer is not reasonable, but you don't have time to do anything about it, then write a brief phrase showing that you understand that you answer is unreasonable (e.g. 'This is wrong because mass must be positive').
- Check that you have used all the information given. In many ways the most artificial aspect of examination questions is that you are given exactly the amount of information required to answer the question. If your answer does not use everything you are given then either it is wrong or you are faced with a very unusual examination question.
- Check that you have understood the point of the question. It is, of course, true that not all exam questions have a point but many do. What idea did the examiners want you to have? Which techniques did they want you to demonstrate? Is the result of the question interesting in some way? Does it generalise? If you can see the point of the question would your working show the point to someone who did not know it in advance?

- Make sure that you are not unthinkingly applying mathematical tools which you do not fully understand.[6]
- In preparation for the examination, try to see how the problem fits into the wider context and see if there is a special point which it is intended to illustrate. You may need help with this.
- In preparation for the examination, make sure that you actually understand not only what you have done, but also why you have done it that way rather than some other way. This is particularly important if you have had to use a hint or solution.
- In the examination, check that you have given the detail required. There often comes a point in a question where, if we could show that A implies B, the result follows. If after a lot of thought you suddenly see that A does indeed imply B the natural thing to do is to write triumphantly 'But A implies B so the result follows'[7]. Unfortunately unscrupulous individuals who have no idea why A should imply B (apart from the fact that it would complete the question) could and do write the exactly the same thing. Go back through the major points of the question making sure that you have not made any major unexplained leaps.

[6] Mathematicians should feel as insulted as Engineers by the following joke.

A mathematician, a physicist and an engineer enter a mathematics contest, the first task of which is to prove that all odd number are prime. The mathematician has an elegant argument: '1's a prime, 3's a prime, 5's a prime, 7's a prime. *Therefore*, by mathematical induction, all odd numbers are prime. It's the physicist's turn: '1's a prime, 3's a prime, 5's a prime, 7's a prime, 11's a prime, 13's a prime, so, to within experimental error, all odd numbers are prime.' The most straightforward proof is provided by the engineer: '1's a prime, 3's a prime, 5's a prime, 7's a prime, 9's a prime, 11's a prime . . . '.

[7] There is a standard anecdote about the distinguished Professor X. In the middle of a lecture she writes 'It is obvious that A', suddenly falls silent and after a few minutes walks out of the room. The awed students hear her pacing up and down outside. Then after twenty minutes she returns says 'It is obvious' and continues the lecture.

Sample worked question

Let

$$f(x) = ax - \frac{x^3}{1+x^2} ,$$

where a is a constant.

Show that, if $a \geqslant 9/8$, then

$$f'(x) \geqslant 0$$

for all x.

First thoughts

When I play tennis and I see a ball that I think I can hit, I rush up to it and smack it into the net. This is a tendency I try to overcome when I am doing mathematics. In this question, for example, even though I'm pretty sure that I can find $f'(x)$, I'm going to pause for a moment before I do so. I am going to use the pause to think about two things. First, I'm going to think about what to do when I have found $f'(x)$. Second, I'm going to try to decide what the question is really about. Of course, I may not be able to decide what to do with $f'(x)$ until I actually see what it looks like; and I may not be able to see what the question is really about until I have finished it; maybe not even then. (In fact, some questions are not really about anything in particular.) I am also going to use the pause to think a bit about the best way of performing the differentiation. Should I simplify first? Should I make some sort of substitution? Clearly, if knowing how to tackle the rest of the question might guide me in deciding the best way to do the differentiation. Or it may turn out that the differentiation is fairly straight forward, so that it doesn't matter how I do it.

Two more points occur to me as I reread the question. I notice that the inequalities are not strict (they are $\geqslant$ rather than $>$). Am I going to have to worry about the difference? I also notice that there is an 'If ... then', and I wonder if this is going to cause me trouble. I will need to be careful to get the implication the right way round. I mustn't try to prove that if $f'(x) \geqslant 0$ then $a \geqslant 9/8$. It will be interesting to see why the implication is only one way — why it is not an 'if and only if' question.

Don't turn over until you have spent a little time thinking along these lines.

Doing the question

Looking ahead, it is clear that the real hurdle is going to be showing that $\mathrm{f}'(x) \geqslant 0$. How am I going to do that? Two ways suggest themselves. First, if $\mathrm{f}'(x)$ turns out to be a quadratic function, or an obviously positive multiple of a quadratic function, I should be able to use some standard method: looking at the discriminant ('$b^2 - 4ac$'), etc; or, better, completing the square. But if $\mathrm{f}'(x)$ is not of this form, I will have to think of something else: maybe I will have to sketch a graph. I'm hoping that I won't have to do this, because otherwise I could be here all night.

I've just noticed that f is an odd function, i.e. $\mathrm{f}(x) = -\mathrm{f}(-x)$ (did you notice that?). That is is helpful, because it means that $\mathrm{f}'(x)$ is an even function.[8] If it had been odd, there could have been a difficulty, because odd functions (at least, those with no vertical asymptotes) always cross the horizontal axis $y = 0$ at least once (they start in one corner of the graph paper for x large and negative and end in the diagonally opposite corner for x large and positive) and cannot therefore be positive for all x.

Now, how should I do the differentiation? I could divide out the fraction, giving

$$\frac{x^3}{1+x^2} = x - \frac{x}{1+x^2} \qquad \text{and} \qquad \mathrm{f}(x) = (a-1)x + \frac{x}{1+x^2} = bx + \frac{x}{1+x^2}.$$

(Is this right? I'll just check that it works when $x = 2$ — yes, it does: $\mathrm{f}(2) = 2a - 8/5 = 2(a-1) + 2/5$.) This might save a bit of writing, but I don't at the moment see this helping me towards a positive function. On balance, I think I'll stick with the original form.

Another thought: should I differentiate the fraction using the quotient formula or should I write it as $x^3(1+x^2)^{-1}$ and use the product rule? I doubt if there is much in it. I never normally use the quotient rule — it's just extra baggage to carry round — but on this occasion, since the final form I am looking for is a single fraction, I will.

One more thought: since I am trying to prove an inequality, I must be careful throughout not to cancel any quantity which might be negative; or at least if I do cancel a negative quantity, I must remember to reverse the inequality.

Here goes (at last):

$$\mathrm{f}'(x) = a - \frac{3x^2(1+x^2) - 2x(x^3)}{(1+x^2)^2} = \frac{a + (2a-3)x^2 + (a-1)x^4}{(1+x^2)^2}\,.$$

This is working out as I had hoped: the denominator is certainly positive and the numerator is a quadratic function of x^2. I can finish this off most elegantly by completing the square. There are two ways of doing this, and it is marginally easier to do it the unnatural way:

$$\begin{aligned} a + (2a-3)x^2 + (a-1)x^4 &= a(1 + (2 - 3a^{-1})x^2) + (a-1)x^4 \\ &= a\left[1 + \tfrac{1}{2}(2 - 3a^{-1})x^2\right]^2 + \left[(a-1) - \tfrac{1}{4}a(2 - 3a^{-1})^2\right]x^4. \end{aligned}$$

We certainly need $a \geqslant 0$ to make the first term non-negative. Simplifying the second square bracket gives $[2 - 9a^{-1}/4]$, so it is true that if $a \geqslant 9/8$ then $\mathrm{f}'(x) \geqslant 0$ for all x.

[8] You can see this easily from a by sketching a 'typical' odd function.

Post-mortem

Now I can look back and analyse what I have done.

On the technical side, it seems I was right not to use the simplified form of $x^3/(1+x^2)$. This would have lead to the quadratic $(b+1)+(2b-1)x^2+bx^4$, which isn't any easier to handle than the quadratic involving a and just gives an extra opportunity to make an algebraic error.

Actually, I see on re-reading my solution that I have made a bit of a meal out of the ending. I needn't have completed the square at all; I could used the inequality $a \geqslant 9/8$ immediately after finding $f'(x)$, since a appears in $f'(x)$ with a plus sign always:

$$\mathrm{f}'(x) = \frac{a+(2a-3)x^2+(a-1)x^4}{(1+x^2)^2} \geqslant \frac{9/8+(9/4-3)x^2+(9/8-1)x^4}{(1+x^2)^2} = (x^2-3)^2/8 \geqslant 0$$

as required.

I wonder why the examiner wanted me to investigate the sign of $\mathrm{f}'(x)$. The obvious reason is to see what the graph looks like. We can now see what this question is about. It is clear that the examiners really wanted to set the question: 'Sketch the graphs of the function $ax - x^3/(1+x^2)$ in the different cases that arise according to the value of a' but it was thought too long or difficult. It is worth looking back over my working to see what can be said about the shape of the graph of $\mathrm{f}\,(x)$ when $a < 9/8$. My unnecessarily elaborate proof, involving completing the square, makes the role of a a bit clearer than it is in the shorter alternative. (I leave that to you to think about!)

Sample worked question 2

The n positive numbers $x_1, x_2, \ldots, x_n$, where $n \geqslant 3$, satisfy

$$x_1 = 1 + \frac{1}{x_2}, \quad x_2 = 1 + \frac{1}{x_3}, \quad \ldots, \quad x_{n-1} = 1 + \frac{1}{x_n},$$

and also

$$x_n = 1 + \frac{1}{x_1}.$$

Show that

(i) $\quad x_1, x_2, \ldots, x_n > 1$,

(ii) $\quad x_1 - x_2 = -\dfrac{x_2 - x_3}{x_2 x_3}$,

(iii) $\quad x_1 = x_2 = \cdots = x_n$.

Hence find the value of x_1.

First thoughts

My first thought is that this question has an unknown number of variables: $x_1, \ldots, x_n$. That makes it seem rather complicated. I might, if necessary, try to understand the result by choosing an easy value for n (maybe $n = 3$). If I manage to prove some of the results in this special case, I will certainly go back to the general case: doing the special case might help me tackle the general case, but I don't expect to get many marks in an exam if I just prove the result in one special case.

Next, I see that the question has three sub-parts, then a final one. The final one begins 'Hence ...'. This must mean that the final part will depend on at least one of the previous parts. It is not clear from the structure of the question whether the three subparts are independent; the proof of (ii) and (iii) may require the previous result(s). However, I expect to use the results of parts (i), (ii) and (iii) later in the question.

Actually, I think I can see how to do the very last part. If I assume that (iii) holds, so that $x_1 = x_2 = \cdots = x_n = x$ (say), then each of the equations given in the question is identical and each gives a simple equation for x.

I am a bit puzzled about part (i). How can an *inequality* help to derive the *equalities* in the later parts? I can think of a couple of ways in which the result $x_i > 1$ could be used. One is that I may need to cancel, say, x_1 from both sides of an equation in which case I would need to know that $x_1 \neq 0$. But looking back at the question, I see I already know that $x_1 > 0$ (it is given), so this cannot be the right answer. Maybe I need to cancel some other factor, such as $(x_1 - 1)$. Another possibility is that I get two or more solutions by putting $x_1 = x_2 = \cdots = x_n = x$ and I need the one with $x > 1$. This may be the answer: looking back at the question again, I see that it asks for *the* value of x_1 — so I am looking for a single value. I'm still puzzled, but I will remember in later parts to keep a sharp look out for ways of using part (i).

One more thing strikes me about the question. The equations satisfied by the x_i are given on two lines ('and also'). This could be for typographic reasons (the equations would not all fit on one line) but more likely it is to make sure that I have noticed that the last equation is a bit different: all the other equations relate x_i to x_{i+1}, whereas the last equation relates x_n to x_1. It goes back to the beginning, completing the cycle. I'm pleased that I thought of this, because this circularity must be important.

Doing the question

I think I will do the very last part first, and see what happens.

Suppose, assuming the result of part (iii), that $x_1 = x_2 = \cdots = x_n = x$. Then substituting into any of the equations given in the question gives

$$x = 1 + \frac{1}{x}$$

i.e. $x^2 - x - 1 = 0$. Using the quadratic formula gives

$$x = \frac{1 \pm \sqrt{5}}{2} \,,$$

which does indeed give two answers (despite the fact that the question asks for just one). However, I see that one is negative and can therefore be eliminated by the condition $x_i > 0$ which was given in the question (not, I note, the condition $x_1 > 1$ from part (i); I still have to find a use for this).

I needed only part (iii) to find x_1, so I expect that either I need both (i) and (ii) directly to prove (iii), or I need (i) to prove (ii), and (ii) to prove (iii).

Now that I have remembered that $x_i > 0$ for each i, I see that part (i) is obvious. Since $x_2 > 0$ then $1/x_2 > 0$ and the first equation given in the question, $x_1 = 1 + 1/x_2$ show immediately that $x_1 > 1$ and the same applies to x_2, x_3, etc.

Now what about part (ii)? The given equation involves x_1, x_2 and x_3, so clearly I must use the first two equations given in the question:

$$x_1 = 1 + \frac{1}{x_2} \,, \qquad x_2 = 1 + \frac{1}{x_3} \,.$$

Since I want $x_1 - x_2$, I will see what happens if I subtract the two equations:

$$x_1 - x_2 = (1 + \frac{1}{x_2}) - (1 + \frac{1}{x_3}) = \frac{1}{x_2} - \frac{1}{x_3} = \frac{x_3 - x_2}{x_2 x_3} \,. \qquad (*)$$

That seems to work!

One idea that I haven't used so far is what I earlier called the circularity of the equations: the way that x_n links back to x_1. I'll see what happens if I extend the above result. Since there is nothing special about x_1 and x_2, the same result must hold if I add 1 to each of the suffices:

$$x_2 - x_3 = \frac{x_4 - x_3}{x_3 x_4} \,.$$

I see that I can combine this with the previous result:

$$x_1 - x_2 = \frac{x_3 - x_2}{x_2 x_3} = \frac{x_3 - x_4}{x_2 x_3^2 x_4} \,.$$

I now see where this is going. The above step can be repeated to give

$$x_1 - x_2 = \frac{x_3 - x_4}{x_2 x_3^2 x_4} = \frac{x_5 - x_4}{x_2 x_3^2 x_4^2 x_5} = \cdots$$

and eventually I will get back to $x_1 - x_2$:

$$x_1 - x_2 = \frac{x_5 - x_4}{x_2 x_3^2 x_4^2 x_5} = \cdots = \pm \frac{x_1 - x_2}{x_2 x_3^2 x_4^2 x_5^2 \ldots x_n^2 x_1^2 x_2}$$

i.e.

$$(x_1 - x_2)\left(1 \mp \frac{1}{x_1^2 x_2^2 x_3^2 x_4^2 x_5^2 \dots x_n^2}\right) = 0\,.$$

I have put in a $\pm$ because each step introduces a minus sign and I'm not sure yet whether the final sign should be $(-1)^n$ or $(-1)^{n-1}$. I can check this later (for example, by working out one simple case such as $n = 3$); but I may not need to.

I deduce from this last equation that either $x_1^2 x_2^2 x_3^2 x_4^2 x_5^2 \dots x_n^2 = \pm 1$ or $x_1 = x_2$ (which is what I want). At last I see where to use part (i): I know that $x_1^2 x_2^2 x_3^2 x_4^2 x_5^2 \dots x_n^2 \neq \pm 1$ because $x_1 > 1$, $x_2 > 1$, etc. Thus $x_1 = x_2$, and since there was nothing special about x_1 and x_2, I deduce further that $x_2 = x_3$, and so on, as required.

Postmortem

There were a number of useful points in this question.

The first point concerns using the information given in the question. The process of teasing information from what is given is fundamental to the whole of mathematics. It is very important to study what is given (especially seemingly unimportant conditions, such as $x_i > 0$) to see why they have been given. If you find you reach the end of a question without apparently using some condition, then you should look back over your work: it is very unlikely that a condition has been given that is not used in some way. (It may not be a necessary condition — and we will see that the condition $x_i > 0$ is not, in a sense, necessary in this question — but it should be sufficient.) The other piece of information in the question which you might easily have overlooked is the use of the singular (' ... find the value of ...'), implying that there is just one value, despite the fact that the final equation is quadratic..

The second point concerns using the structure of the question. Here, the position of the word 'Hence' suggested strongly that none of the separate parts were stand-alone results; each had to be used for a later proof. Understanding this point made the question much easier, because I was always on the look out for an opportunity to use the earlier parts. Of course, in some problems (without that 'hence') some parts may be stand-alone; though this is rare in STEP questions. You may think that this is like playing a game according to hidden (STEP) rules, but that is not the case. Precision writing and precision reading is vital in mathematics and in many professions (law, for example). Mathematicians have to be good at it, which is the reason that everyone wants to employ people with mathematical training.

The third point was the rather inconclusive speculation about the way inequalities might help to derive an equality. It turned out that what was actually required was $x_1 x_2 \dots x_n \neq \pm 1$. I was a bit puzzled by this possibility in my first thoughts, because it seemed that the result ought to hold under conditions different from those given; for example, $x_i < 0$ for all i (does this condition work??). Come to think of it, why are conditions given on all x_i when they are all related by the given equations? This makes me think that there ought to be a better way of proving the result which would reveal exactly the conditions under which it holds.

Fourth was the idea (which I didn't actually use) that I might try to prove the result for, say, $n = 3$. This would not have counted as a proof of the result (or anything like it), but it might have given me ideas for tackling the question.

Fifth was the observation that the equations given in the question are 'circular'. It was clear that the circularity was essential to the question and it turned out to be the key the most difficult part.

Having identified it early on, I was ready to use it when the opportunity arose.

Final thoughts

It occurs to me only now, after my post-mortem, that there is another way of obtaining the final result. Suppose I start with the idea of circularity (as indeed I might have, had I not been otherwise directed by the question) and use the given equations to find x_1 in terms of first x_2, then x_3, then x_4 and eventually in terms of x_1 itself. That should give me an equation I can solve, and I should be able to find out what conditions are needed on the x_i. Try it. You may need to guess a formula for x_1 in terms of x_i from a few special cases, then prove it by induction. You will find it useful to define a sequence of numbers F_i such that $F_0 = F_1 = 1$ and $F_{n+1} = F_n + F_{n-1}$. (These numbers are called *Fibonacci numbers*[9].)

You should find that if $x_n = x_1$ for any n (greater than 1), then $x_n = \frac{1 \pm \sqrt{5}}{2}$. No conditions are required for this result to hold, except that all the x_i exist, so it was not necessary to set the condition $x_1 > 1$. (The condition on x_1 for all x_i to exist is that $x_1 \neq F_k/F_{k-1}$ for any $k \leqslant n$, as you will see if you obtain the general formula for x_1 in terms of x_n.)

[9]Fibonacci (short for filius Bonacci — son of Bonacci) was called the greatest European mathematician of the middle ages. He was born in Pisa (Italy) in about 1175 AD. He introduced the series of numbers named after him in his book of 1202 called Liber Abbaci (Book of the Abacus). It was the solution to the problem of the number of pairs of rabbits produced by an initial pair: *A pair of rabbits are put in a field and, if rabbits take a month to become mature and then produce a new pair every month after that, how many pairs will there be in twelve months time?*

Question 1(∗)

Find all the solutions of the equation

$$|x+1| - |x| + 3|x-1| - 2|x-2| = x+2\ .$$

Comments

This looks more difficult than it is. There is no easy way to deal with the modulus function: you have to look at the different cases individually (for example, for $|x|$ you have to look at $x \leqslant 0$ and $x \geqslant 0$). The most straightforward approach would be to solve the equation in each of the different regions determined by the modulus signs: $x \leqslant -1$, $-1 \leqslant x \leqslant 0$, $0 \leqslant x \leqslant 1$, etc. You might find a graphical approach helps you to picture what is going on (I didn't).

Solution to question 1

Let

$$\mathrm{f}(x) = |x+1| - |x| + 3|x-1| - 2|x-2| - (x+2).$$

We have to solve $\mathrm{f}(x) = 0$ in the five regions of the x-axis determined by the modulus functions, namely

$$x \leqslant -1\,; \quad -1 \leqslant x \leqslant 0\,; \quad 0 \leqslant x \leqslant 1\,; \quad 1 \leqslant x \leqslant 2\,; \quad 2 \leqslant x\,.$$

In the separate regions, we have

$$\mathrm{f}(x) = \begin{cases} (x+1) - x + 3(x-1) - 2(x-2) - (x+2) & = & 0 & \text{for} & 2 \leqslant x < \infty \\ (x+1) - x + 3(x-1) + 2(x-2) - (x+2) & = & 4x-8 & \text{for} & 1 \leqslant x \leqslant 2 \\ (x+1) - x - 3(x-1) + 2(x-2) - (x+2) & = & -2x-2 & \text{for} & 0 \leqslant x \leqslant 1 \\ (x+1) + x - 3(x-1) + 2(x-2) - (x+2) & = & -2 & \text{for} & -1 \leqslant x \leqslant 0 \\ -(x+1) + x - 3(x-1) + 2(x-2) - (x+2) & = & -2x-4 & \text{for} & -\infty < x \leqslant -1 \end{cases}$$

Solving in each region gives:

(i) $x \geqslant 2$

Here, $\mathrm{f}(x) = 0$ so the equation is satisfied for all values of x. $\mathrm{f}(-2) = 0$ and $\mathrm{f}(x) = 0$ for any $x \geqslant 2$.

(ii) $1 \leqslant x \leqslant 2$

Here, $\mathrm{f}(x) = 0$ only if $x = 2$.

(iii) $0 \leqslant x \leqslant 1$

Here, $\mathrm{f}(x) = 0$ only if $x = -1$. This is not a solution since the point $x = -1$ does not lie in the region $0 \leqslant x \leqslant 1$.

(iv) $-1 \leqslant x \leqslant 0$

Here, $\mathrm{f}(x) = -2$ so there is no solution.

(v) $x \leqslant -1$

Here, $\mathrm{f}(x) = 0$ only if $x = -2$. This is a solution since the point $x = -2$ does lie in the region $x \leqslant -1$.

The equation $\mathrm{f}(x) = 0$ is therefore satisfied by $x = -2$ and by any x greater than, or equal to, 2.

Question 2$(*)$

(i) Find the coefficient of x^6 in

$$(1 - 2x + 3x^2 - 4x^3 + 5x^4)^3.$$

You should set out your working clearly.

(ii) By considering the binomial expansions of $(1 + x)^{-2}$ and $(1 + x)^{-6}$, or otherwise, find the coefficient of x^6 in

$$(1 - 2x + 3x^2 - 4x^3 + 5x^4 - 6x^5 + 7x^6)^3.$$

Comments

Your first instinct is probably just to multiply out the brackets. A moment's thought will show that it is going to be much quicker to concentrate only on those terms with powers of x that multiply together to give x^6. The difficulty with this would be making sure you have got them all, so a systematic approach is called for. Alternatively, you could work out $(1 - 2x + 3x^2 - 4x^3 + 5x^4)^2$ first. Squaring an expression which is the sum of 5 terms leads to an expression with 15 terms after simplification (why?) but in this case, simplification will lead to at most 9 terms (why?). 9 terms shouldn't prove too much of a challenge, but I wouldn't fancy it myself.

Obviously, the second part can be done 'otherwise' by multiplying out the brackets as in the first part. However, it is much quicker to follow the hint. If you can't see how considering the expansion of $(1 + x)^{-2}$ helps, you should write it out.

Solution to question 2

(i) Writing out the cube explicitly makes it easier to collect up the terms:

$$(1-2x+3x^2-4x^3+5x^4)(1-2x+3x^2-4x^3+5x^4)(1-2x+3x^2-4x^3+5x^4)$$

To get the coefficient of x^6, we need to find all ways of choosing one term from each bracket in such a way that, when multiplied together, the result is a multiple of x^6.

Start by taking the first term (i.e. 1) from the first bracket. This gives

$$1 \times [(3x^2) \times (5x^4) + (-4x^3) \times (-4x^3) + (5x^4) \times (3x^2)]$$

which sums to $46x^6$.

The second term from the first bracket gives a contribution of

$$(-2x) \times [(-2x) \times (5x^4) + (3x^2) \times (-4x^3) + (-4x^3) \times (3x^2) + (5x^4) \times (-2x)] = 88x^6 \ .$$

The third term from the first bracket gives a contribution of

$$(3x^2) \times [(1) \times (5x^4) + (-2x) \times (-4x^3) + (3x^2) \times (3x^2) + (-4x^3) \times (-2x) + (5x^4) \times (1)] = 105x^6 \ .$$

The fourth term from the first bracket gives a contribution of

$$(-4x^3) \times [(1) \times (-4x^3) + (-2x) \times (3x^2) + (3x^2) \times (-2x) + (-4x^3) \times (1)] = 80x^6 \ .$$

The fifth term from the first bracket gives a contribution of

$$(5x^4) \times [(1) \times (3x^2) + (-2x) \times (-2x) + (3x^2) \times (1)] = 50x^6 \ .$$

The grand total is therefore $369x^6$ and 369 is the required coefficient.

(ii) The given expression is the beginning of the expansion of $\big((1+x)^{-2}\big)^3$, i.e. of $(1+x)^{-6}$. But

$$(1+x)^{-6} = 1 - 6x + \frac{(-6)\times(-7)}{2!}x^2 + \cdots + \frac{11!}{6! \times 5!}x^6 + \cdots + (-1)^n \frac{(n+6-1)!}{n! \times 5!}x^n + \ldots$$

so the coefficient of x^6 is $11!/6!5! = 462$.

Postmortem

There was not much to this question. There are two points worth emphasising. The first is the need for a systematic approach (and also for neat presentation and clear explanation) in the first part. Otherwise, there is a real danger of confusing yourself and a certainty of confusing anyone else reading your answer. The second point is the use of the hint in the second part of the question. Although you already knew one sure way of reaching the answer — which would have got full marks because of the 'or otherwise' — it paid to persevere with the hint, because it led to considerable saving of effort.

You might ask yourself why the method of the second part doesn't work for the first part.

There were two 'why?'s in the comments section. Why does squaring an expression with five terms give a total of 15 terms when simplified? Of course, multiplying two different 5-term expressions gives 5×5 terms. If the two expressions are the same, then 5×4 terms occur twice, giving a total of $5 + 10$. Why does the expression we are squaring in this question result in 9 terms? The expression is a polynomial of degree 4, so the result of squaring is a polynomial of degree 8, which can have at most 9 different powers of x and hence (after simplification) at most 9 terms.

Question 3($*$)

Show that you can make up 10 pence in eleven ways using 10p, 5p, 2p and 1p coins.

In how many ways can you make up 20 pence using 20p, 10p, 5p, 2p and 1p coins?

Comments

I don't really approve of this sort of question, but I thought I'd better include one in this collection. The one given above seems to me to be a particularly bad example, because there are a number of neat and elegant mathematical ways of approaching it, none of which turn out to be any use. The quickest instrument is the bluntest: just write out all the possibilities. Two things are important: first you must be systematic or you will get hopelessly confused; second, you must lay out your solution, with careful explanations, in a way which allows other people (examiners, for example) to understand exactly what you are doing.

Here are a couple of the red herrings. For the first part, what is required is the coefficient of x^{10} in the expansion of

$$\frac{1}{(1-x^{10})(1-x^5)(1-x^2)(1-x)} \,.$$

You can see why this is the case by using the binomial expansion (compare question 1). Although this is neat, it doesn't help, because there is no easy way of obtaining the required coefficient. And the second part would be even worse.

The second red herring concerns the relationship between the two parts. Since different parts of STEP questions are nearly always related, you might be led to believe that the result of the second part follows from the first: you divide the required twenty pence into two tens and then use the result of the first part to give the number of ways of making up each 10. This would give an answer of 66 (why?) plus one for a single 20p piece. This would also be neat, but the true answer is less than 67 because some arrangements are counted twice by this method — and it is not easy to work out which ones.

Solution to question 3

Probably the best approach is to start counting with the arrangements which use as many high denomination coins as possible, then work down.

We can make up 10p as follows:

10;

5+5 (one way using two 5p coins);

5+2+2+1, 5+2+1+1+1, 5+1+1+1+1+1, (three ways using one 5p coin);

2+2+2+2+2, 2+2+2+2+1+1, etc, (six ways using no 5p coins);

making a total of 11 ways.

We can make up 20p as follows:

20;

10 + any of the 11 arrangements in the first part of the question;

5+5+5+5;

5+5+5+2+2+1, etc (3 ways using three 5p coins);

5+5+2+2+2+2+2, 5+5+2+2+2+2+1+1 etc (6 ways using two 5p coins);

5+2+2+2+2+2+2+2+1, etc (8 ways using one 5 and making 15 out of 2p and 1p coins);

2+2+2+2+2+2+2+2+2+2, etc (11 ways of making 20p with 2p and 1p coins).

Grand total = 41.

Postmortem

As in the previous question, the most important lesson to be learnt here is the value of a systematic approach and clear explanations. You should not be happy just to obtain the answer: there is no virtue in that. You should only be satisfied if you displayed your working at least as systematically as I have, above.

Question 4($*$)

Consider the system of equations

$$\begin{aligned} 2yz + zx - 5xy &= 2 \\ yz - zx + 2xy &= 1 \\ yz - 2zx + 6xy &= 3. \end{aligned}$$

Show that

$$xyz = \pm 6$$

and find the possible values of x, y and z.

Comments

At first sight, This looks forbidding. A closer look reveals that the variables x, y and z occur only in pairs yz, zx and xy. The problem therefore boils down to solving three simultaneous equations in these variables, then using the solution to find x, y and z individually.

What do you make of the $\pm$ in the equation $xyz = \pm 6$? Does this give you a clue?

There are two ways of tackling simultaneous equations. You could use the first equation to find an expression for one variable (yz say) in terms of the other two variables, then substitute this into the other equations to eliminate the zx from the system. Then use the second equation (in its new form) to find an expression for one of the two remaining variables (zx say), then substitute this into the third equation (in its new form) to obtain a linear equation for the third variable (xy).[10] Having solved this equation, you can then substitute back to find the other variables. This method is called Gauss elimination.

Alternatively, you could eliminate one variable (yz say) from the first two of equations by multiplying the first equation by something suitable and the second equation by something suitable and subtracting. You then eliminate yz from the second and third equations similarly. That leaves you with two simultaneous equations in two variables which you can solve by your favorite method.

There is another way of solving the simultaneous equations, which is very good in theory by not at all good in practice. You write the equations in matrix form $\mathbf{Mx} = \mathbf{c}$, where in this case

$$\mathbf{M} = \begin{pmatrix} 2 & 1 & -5 \\ 1 & -1 & 2 \\ 1 & -2 & 6 \end{pmatrix}, \qquad \mathbf{x} = \begin{pmatrix} yz \\ zx \\ xy \end{pmatrix}, \qquad \mathbf{c} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}.$$

The solution is then $\mathbf{x} = \mathbf{M}^{-1}\mathbf{c}$.

[10] A *linear* equation in the variable x is one that does not involve and powers of x except the first power, or any other functions of x. Suppose the equation is $\mathrm{f}(x) = 0$ and the solution is $x = a$. Then if you replace x by $2x$, so that $\mathrm{f}(2x) = 0$, then the solution is $x = a/2$, and this can be thought of as the defining property of a linearity in this context.

Solution to question 4

Start by labelling the equations:

$$\begin{aligned} 2yz + zx - 5xy &= 2\,, & (1)\\ yz - zx + 2xy &= 1\,, & (2)\\ yz - 2zx + 6xy &= 3\,. & (3) \end{aligned}$$

We use Gaussian elimination. Rearranging equation (1) gives

$$yz = -\tfrac{1}{2}zx + \tfrac{5}{2}xy + 1\,, \tag{4}$$

which we substitute back into equations (2) and (3) :

$$-\tfrac{3}{2}zx + \tfrac{9}{2}xy = 0\,, \tag{5}$$

$$-\tfrac{5}{2}zx + \tfrac{17}{2}xy = 2\,. \tag{6}$$

Thus $zx = 3xy$ (using equation (5)). Substituting into equation (6) gives $xy = 2$ and $zx = 6$. Finally, substituting back into equation (1) shows that $yz = 3$.

The question is now plain sailing. Multiplying the three values together gives $(xyz)^2 = 36$ and taking the square root gives $xyz = \pm 6$ as required.

Now it remains to solve for x, y and z individually. We know that $yz = 3$, so if $xyz = +6$ then $x = +2$, and if $xyz = -6$ then $x = -2$. The solutions are therefore either $x = +2$, $y = 1$, and $z = 3$ or $x = -2$, $y = -1$, and $z = -3$.

Postmortem

There were two key observations which allowed us to do this question quite easily. Both came from looking carefully at the question. The first was that the given equations, although non-linear in x, y and z (they are quadratic, since they involve products of these variables) could be thought of as three linear equations in yz, zx and xy. That allowed us to make a start on the question. The second observation was that the equation $xyz = \pm 6$ is almost certain to come from $(xyz)^2 = 36$ and that gave us the next step after solving the simultaneous equations. (Recall the next step was to multiply all the variables together.)

There was a point of technique in the solution: it is often very helpful in this sort of problem (and many others) to number your equations. This allows you to refer back clearly and quickly, for your benefit as well as for the benefit of your readers.

The three simultaneous quadratic equations (1)–(3) have a geometric interpretation. First consider the case of *linear* equations. Each such equation can be interpreted as the equation of a plane, of the form $\mathbf{k}.\mathbf{x} = a$ (which has normal $\mathbf{k}$ and is at distance $a/|\mathbf{k}|$ from the origin). The solution of the three equations represents a point which is the intersection of all three planes. Clearly, there may not be such a point: for example, the planes may intersect in three pairs of lines forming a Toblerone. In this case, there is no point common to the three planes and therefore no solution to the equations, unless the lines happen to coincide (a Toblerone of zero volume) in which case every point on the line is a solution.

The simultaneous equations in the question are quadratic and each is the equation of a hyperboloid (i.e. a surface for which two sets of sections are hyperbolae and the other set of sections are ellipses) — like a radar dish. The number of intersections of three infinite radar dishes is not easy to calculate in general. In this case, we know the answer must be at least 2: replacing x, y and z in the equations by $-x$, $-y$ and $-z$, respectively, makes no difference to the equations. If $x = a$, $y = b$ and $z = c$ satisfies the equations, then so also must $x = -a$, $y = -b$ and $z = -c$.

Question 5(∗)

Suppose that

$$3 = \frac{2}{x_1} = x_1 + \frac{2}{x_2} = x_2 + \frac{2}{x_3} = x_3 + \frac{2}{x_4} = \cdots .$$

Guess an expression, in terms of n, for x_n. Then, by induction or otherwise, prove the correctness of your guess.

Comments

Wording this sort of question is a real headache for the examiners. Suppose you guess wrong; how can you then prove your guess by induction (unless you get that wrong too)? How else can the question be phrased? In the end, we decided to assume that you are all so clever that your guesses will all be correct.

To guess the formula, you need to work out x_1, x_2, x_3, etc and look for a pattern. You should not need to go beyond x_4.

Proof by induction is not in the core A-level syllabus. I decided to include it in the syllabus for STEP I and II because the idea behind it is not difficult and it is very important both as a method of proof and also as an introduction to more sophisticated mathematical thought.

Solution to question 5

First we put the equations into a more manageable form. Each equality can be written in the form

$$3 = x_n + \frac{2}{x_{n+1}}, \qquad \text{i.e.} \qquad x_{n+1} = \frac{2}{3 - x_n}\ .$$

We find $x_1 = 2/3$, $x_2 = 6/7$, $x_3 = 14/15$ and $x_4 = 30/31$. The denominators give the game away. We guess

$$x_n = \frac{2^{n+1} - 2}{2^{n+1} - 1}\ .$$

For the induction, we need a starting point: our guess certainly holds for $n = 1$ (and 2, 3, and 4!).

For the inductive step, we suppose our guess also holds for $n = k$, where k is any integer. If we can show that it then also holds for $n = k + 1$, we are done.

We have, from the equation given in the question,

$$x_{k+1} = \frac{2}{3 - x_k} = \frac{2}{3 - \frac{2^{k+1}-2}{2^{k+1}-1}} = \frac{2(2^{k+1} - 1)}{3(2^{k+1} - 1) - (2^{k+1} - 2)} = \frac{2^{k+2} - 2}{2^{k+2} - 1}\ ,$$

as required.

Postmortem

There's not much to say about this. By STEP standards, it is fairly easy and short. Nevertheless, you are left to your own devices from the beginning, so you should be pleased if you got it out.

Question 6$(*)$

Show that, if $\tan^2\theta = 2\tan\theta + 1$, then $\tan 2\theta = -1$.

Find all solutions of the equation

$$\tan\theta = 2 + \tan 3\theta$$

which satisfy $0 < \theta < 2\pi$, expressing your answers as rational multiples of π.

Find all solutions of the equation

$$\cot\phi = 2 + \cot 3\phi$$

which satisfy $-\dfrac{3\pi}{2} < \phi < \dfrac{\pi}{2}$, expressing your answers as rational multiples of π.

Comments

There are three distinct parts. It is pretty certain that they are related, but it is not obvious what the relationship is. The first part must surely help with the second part in some way that will only be apparent once the second part is under way.

In the absence of any other good ideas, it looks right to express the double and triple angle tans and cots in terms of single angle tans and cots. You should remember the formula

$$\tan(A+B) = \frac{\tan A + \tan B}{1 - \tan A \tan B}\,.$$

You can use this for $\tan 3\theta$ and hence $\cot 3\theta$. If you have forgotten the $\tan(A+B)$ formula, you can quickly work it out by from the corresponding sin and cos formulae:

$$\frac{\sin(A+B)}{\cos(A+B)} = \frac{\sin A\cos B + \cos A\sin B}{\cos A\cos B - \sin A\sin B}\,.$$

You are are not expected to remember the more complicated triple angle formulae. (I certainly don't.)

You may well find yourself solving cubic equations at some stage in this question.

Solution to question 6

We will write t for $\tan\theta$ (or $\tan\phi$) throughout.

First part : $\tan 2\theta = \dfrac{2t}{1-t^2} = -1$ (since $t^2 = 2t+1$ is given).

For the second part, we first work out $\tan 3\theta$. We have

$$\tan 3\theta = \tan(\theta + 2\theta) = \frac{\tan\theta + \tan 2\theta}{1 - \tan\theta\tan 2\theta} = \frac{t + 2t/(1-t^2)}{1 - t(2t/(1-t^2))} = \frac{3t - t^3}{1 - 3t^2}\ ,$$

so the equation becomes

$$\frac{3t - t^3}{1 - 3t^2} = t - 2, \quad \text{i.e.} \quad t^3 - 3t^2 + t + 1 = 0.$$

One solution (by inspection) is $t = 1$. Thus one set of roots is given by $\theta = n\pi + \pi/4$.

There are no other obvious integer roots, but we can reduce the cubic equation to a quadratic equation by dividing out the known factor $(t-1)$. I would write $t^3 - 3t^2 + t + 1 = (t-1)(t^2 + at - 1)$ since the coefficients of t^2 and of t^0 in the quadratic bracket are obvious. Then I would multiply out the brackets to find that $a = -2$. Now we see the connection with the first part: $t^2 + at - 1 = 0 \Rightarrow \tan 2\theta = -1$, and hence $2\theta = n\pi - \pi/4$.

The roots are therefore $\theta = n\pi + \pi/4$ and $\theta = n\pi/2 - \pi/8$. The multiples of $\pi/8$ in the given range are $\{2, 3, 7, 10, 11, 15\}$.

For the last part, we could set $\cot\phi = 1/\tan\phi$ and $\cot 3\phi = 1/\tan 3\phi$ thereby obtaining

$$\frac{1}{t} = 2 + \frac{1 - 3t^2}{3t - t^3}\ .$$

This simplifies to the cubic equation $t^3 + t^2 - 3t + 1 = 0$. There is an integer root $t = 1$, and the remaining quadratic is $t^2 + 2t - 1 = 0$. Learning from the first part, we write this as $\dfrac{2t}{1-t^2} = 1$, which means that $\tan 2\phi = n\pi + \pi/4$. Proceeding as before gives (noting the different range) the following multiples of $\pi/8$: $\{2, 1, -3, -6, -7, -11\}$.

Postmortem

There was a small but worthwhile notational point in this question: it is often possible to use the abbreviation t for tan (or s for sin, etc), which can save a great deal of writing.

There are two other points worth recalling. First is the way that the first part fed into the second part, but had to be mildly adapted for the third part. This is a typical device used in STEP questions aimed to see how well you learn new ideas. Second is what to do when faced with a cubic equation. There is a formula for the roots of a cubic, but no one knows it nowadays. Instead, you have to find at least one root by inspection. Having found one root, you have a quick look to see if there are any other obvious roots, then divide out the know factor to obtain a quadratic equation.

The real detectives among you might have wondered whether there was as deep reason for the peculiar choice of range ($-3\pi/2 \leqslant \phi \leqslant \pi/2$) for the third part. You will not be surprised to know that there was. The reciprocal relation between tan and cot used for the first part is not the only relation. We could have instead used $\cot A = \tan(\pi/2 - A)$.

The second equation transforms into the first equation if we set $\phi = \pi/2 - \theta$. Furthermore, the given range of ϕ corresponds exactly to the range of θ given in the second part. We can therefore write down the solution directly from the solution to the second part.

Question 7(∗)

Show, by means of a change of variable or otherwise, that

$$\int_0^\infty \mathrm{f}\big((x^2+1)^{1/2}+x\big)\,\mathrm{d}x = \frac{1}{2}\int_1^\infty (1+t^{-2})\mathrm{f}(t)\,\mathrm{d}t\ ,$$

for any given function f .

Hence, or otherwise, show that

$$\int_0^\infty \big((x^2+1)^{1/2}+x\big)^{-3}\,\mathrm{d}x = \frac{3}{8}\ .$$

Comments

There are two things to worry about when you are trying to find a change of variable to convert one integral to another: you need to make the integrand match up and you need to make the limits match up. Sometimes, the limits give the clue to the change of variable. (For example, if the limits on the original integral were 0 and 1 and the limits on the transformed integral were 0 and $\pi/4$, then an obvious possibility would be to make the change $t = \tan x$). Here, the change of variable is determined by the integrand, since it must work for all choices of f.

Perhaps you are worried about the infinite upper limit of the integrals. If you are trying to prove some rigorous result about infinite integrals, you might use the definition

$$\int_0^\infty \mathrm{f}\,(x)\mathrm{d}x = \lim_{a\to\infty}\int_0^a \mathrm{f}\,(x)\mathrm{d}x\ ,$$

but for present purposes you just do the integral and put in the limits. The infinite limit will not normally present problems. For example,

$$\int_1^\infty (x^{-2}+\mathrm{e}^{-x})\mathrm{d}x = \big(-x^{-1}-\mathrm{e}^{-x}\big)\Big|_1^\infty = -\frac{1}{\infty} - \mathrm{e}^{-\infty} + \frac{1}{1} + \mathrm{e}^1 = 1+\mathrm{e}\ .$$

Don't be afraid of writing things like $1/\infty = 0$. It is perfectly OK to use this as shorthand for $\lim_{x\to\infty} 1/x = 0$; but ∞/∞ and $0/0$ are definitely not OK, because of their ambiguity.

Solution to question 7

Clearly, to get the argument of f right, we must set

$$t = (x^2 + 1)^{1/2} + x \ .$$

We must check that the new limits are correct (if not, we are completely stuck). When $x = 0$, $t = 1$ as required. Also, $(x^2 + 1)^{1/2} \to \infty$ as $x \to \infty$, so $t \to \infty$. Thus the upper limit is still ∞, again as required.

The transformed integral is

$$\int_1^\infty \mathrm{f}\,(t) \frac{\mathrm{d}x}{\mathrm{d}t}\, \mathrm{d}t$$

so the next task is to find $\dfrac{\mathrm{d}x}{\mathrm{d}t}$. This we can do in two ways: we find x in terms of t and differentiate it; or we could find $\dfrac{\mathrm{d}t}{\mathrm{d}x}$ and turn it upside down. The snag with the second method is that the answer will be in terms of x, so we will have to express x in terms of t anyway — in which case, we may as well use the first method.

We start by finding x in terms of t:

$$t = (x^2 + 1)^{1/2} + x \Rightarrow (t - x)^2 = (x^2 + 1) \Rightarrow x = \frac{t^2 - 1}{2t} = \frac{t}{2} - \frac{1}{2t} \ .$$

Then we differentiate:

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

which agrees exactly the factor that appears in the second integrand.

For the last part, we take $\mathrm{f}\,(t) = t^{-3}$. Thus

$$\begin{aligned} \int_0^\infty \left((x^2+1)^{1/2} + x\right)^{-3} \mathrm{d}x &= \frac{1}{2} \int_1^\infty (t^{-3} + t^{-5})\, \mathrm{d}t \\ &= \frac{1}{2} \left(\frac{t^{-2}}{-2} + \frac{t^{-4}}{-4} \right) \Bigg|_1^\infty \\ &= \frac{1}{2} \left(\frac{1}{2} + \frac{1}{4} \right) = \frac{3}{8} \ . \end{aligned}$$

as required.

Postmortem

This is not a difficult question conceptually once you realise the significance of the fact that the change of variable must work whatever function f is in the integrand.

There was a useful point connected with calculating $\dfrac{\mathrm{d}x}{\mathrm{d}t}$. It was a good idea not to plunge into the algebra without first thinking about alternative methods; in particular, use of the result

$$\frac{\mathrm{d}x}{\mathrm{d}t} = 1 \Big/ \frac{\mathrm{d}t}{\mathrm{d}x} \ .$$

Finally, there was the infinite limit of the integrand, which I hope you saw was not something to worry about (even though infinite limits are excluded from the A-level core). If you were setting up a formal definition of what an integral is, you would have to use finite limits, but if you are merely calculating the value of an integral, you just go ahead and do it, with whatever limits you are given.

Question 8($*$)

Which of the following statements are true and which are false? Justify your answers.

(i) $a^{\ln b} = b^{\ln a}$ for all positive numbers a and b.

(ii) $\cos(\sin\theta) = \sin(\cos\theta)$ for all real θ.

(iii) There exists a polynomial P such that $|\mathrm{P}(x) - \cos x| \leqslant 10^{-6}$ for all (real) x.

(iv) $x^4 + 3 + x^{-4} \geqslant 5$ for all $x > 0$.

Comments

You have to decide first whether the statement is true of false. If true, the justification has to be a proof. If false, you could prove that it is false, though it is often better to find a simple counterexample.

Part (iii) might look a bit odd. I suppose it relates to the standard approximation $\cos x \approx 1 - x^2/2$ which holds when x is small. This result can be improved by using a polynomial of higher degree: the next term is $x^4/4!$. It can be proved that, given any number k, $\cos x$ can be approximated, for $-k < x < k$, by a polynomial of the form

$$\sum_{n=0}^{N} (-1)^n \frac{x^{2n}}{(2n)!}\,.$$

You have to use more terms of the approximation (i.e. a larger value of N) if you want either greater accuracy or a larger value of k. In Part (iii) of this question, you are being asked if there is a polynomial such that the approximation is good for all values of x.

Solution to question 8

(i) True. The easiest way to see this is to log both sides. For the left hand side, we have

$$\ln(a^{\ln b}) = (\ln b)(\ln a)$$

and for the right hand side we have

$$\ln(a^{\ln b}) = (\ln a)(\ln b) \ ,$$

which agree.

Note that we have to be a bit careful with this sort of argument. The argument used is that $A = B$ because $\ln A = \ln B$. This requires the property of the ln function that $\ln A = \ln B \Rightarrow A = B$. You can easily see that this property holds because ln is a strictly increasing function; if $A > B$, then $\ln A > \ln B$. The same would not hold for (say) sin (i.e. $\sin A = \sin B \not\Rightarrow A = B$).

(ii) False. $\theta = \pi/2$ is an easy counterexample.

(iii) False. Roughly speaking, any polynomial can be made as large as you like by taking x to be very large (provided it is of degree greater than zero), whereas $|\cos x| \leqslant 1$. There is obviously no polynomial of degree zero (i.e. no constant number) for which the statement holds.

(iv) True: $x^4 + 3 + x^{-4} = (x^2 - x^{-2})^2 + 5 \geqslant 5$.

Postmortem

The important point here is that if you want to show a statement is true, you have to give a formal proof, whereas if you want to show that it is false, you only need give one counterexample. It does not have to be an elaborate counterexample — in fact, the simpler the better.

There may still be a question mark in your mind over part (iii). Given a polynomial $\mathrm{P}(x)$, how do you actually prove that it is possible to find x such that $|\mathrm{P}(x)| > 1 + 10^{-6}$? This is one of those many mathematical results that seems very obvious but is irritatingly difficult to prove. (Often, the most irritatingly-difficult-to-prove obvious results are false, though this one is true.)

You should try to cobble together a proof for yourself. I would start by choosing a general polynomial of degree N. Let

$$\mathrm{P}(x) = \sum_{n=0}^{N} a_n x^n \ ,$$

where $a_N \neq 0$. Then I would assume that $a_N = 1$ since if I could prove the result for this case, I'm sure I could easily modify the proof to deal with the general case. Now what?

A simple way forward would be to make use of the fundamental theorem of algebra, which says that $\mathrm{P}(x)$ can be written in the form

$$\mathrm{P}(x) \equiv (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n)$$

where α_1, α_2, ..., α_n are the roots of the equation $\mathrm{P}(x) = 0$. (The theorem says that this equation has n roots.) Then

$$|\mathrm{P}(x)| \equiv |x - \alpha_1||x - \alpha_2| \cdots |x - \alpha_n|$$

which can be made larger than any given number, M say, by choosing x so that $|x - \alpha_i| > M^{1/n}$ for each i.

Actually, one one or two things have been brushed under the carpet in this proof (what if some of the roots are complex?), but I'm confident that there is nothing wrong that couldn't be easily fixed,

Question 9$(*)$

Prove that the rectangle of greatest perimeter which can be inscribed in a given circle is a square.

The result changes if, instead of maximising the sum of lengths of sides of the rectangle, we seek to maximise the sum of nth powers of the lengths of those sides for $n \geqslant 2$. What happens if $n = 2$? What happens if $n = 3$? Justify your answers.

Comments

Obviously, the perimeter of a *general* rectangle has no maximum (it can be as long as you like), so the key to this question is to use the constraint that the rectangle lies in the circle.

The word 'greatest' in the question immediately suggests differentiating something. The perimeter of a rectangle is expressed in terms of two variables, length x and breadth y, so you must find a way of using the constraint to eliminate one variable. This could be done in two ways: express y in terms of x, or express both x and y in terms of another variable (an angle, say).

Finally, there is the matter of deciding whether your solution is the greatest or least value (or neither). This can be done either by considering second derivatives, or by trying to understand the different situations. Often the latter method, or a combination, is preferable.

Finding stationary values of functions subject to constraints is very important: it has widespread applications (for example, in theoretical physics, financial mathematics and in fact in almost any area to which mathematics is applied). It forms a whole branch of mathematics called optimisation. Normally, the methods described above (eliminating one variable) are not used: a clever idea, the methods of Lagrange multipliers, is used instead,

Solution to question 9

Let the circle have diameter d and let the length of one side of the rectangle be x and the length of the adjacent side be y.

Then, by Pythagoras's theorem,

$$y = \sqrt{d^2 - x^2} \qquad (\dagger)$$

and the perimeter P is given by

$$P = 2x + 2\sqrt{d^2 - x^2}\,.$$

We can find the largest possible value of P as x varies by calculus. We have

$$\frac{\mathrm{d}P}{\mathrm{d}x} = 2 - 2\frac{x}{\sqrt{d^2 - x^2}}\,,$$

so for a stationary point, we require (cancelling the factor of 2 and squaring)

$$1 = \frac{x^2}{d^2 - x^2} \qquad \text{i.e.} \qquad 2x^2 = d^2\,.$$

Thus $x = d/\sqrt{2}$ (ignoring the negative root for obvious reasons). Substituting this into ($\dagger$) gives $y = d/\sqrt{2}$, so the rectangle is indeed a square, with perimeter $2\sqrt{2}\,d$.

But is it the maximum perimeter? The easiest way to investigate is to calculate the second derivative, which is easily seen to be negative for all values of x, and in particular when $x = d/\sqrt{2}$. The stationary point is certainly a maximum.

For the second part, we consider

$$\mathrm{f}\,(x) = x^n + (d^2 - x^2)^{n/2}\,.$$

The first thing to notice is that f is constant if $n = 2$, so in this case, the largest (and smallest) value is d^2. For $n = 3$, we have

$$\mathrm{f}'(x) = 3x^2 - 3x(d^2 - x^2)^{1/2},$$

so $\mathrm{f}\,(x)$ is stationary when $x^4 = x^2(d^2 - x^2)$, i.e. when $2x^2 = d^2$ as before or when $x = 0$. The corresponding stationary values are $\sqrt{2}d^3$ and $2d^3$, so this time the largest value occurs when $x = 0$.

Postmortem

You will almost certainly want to investigate the situation for other values of n yourself, just to see what happens.

Were you happy with the proof given above that the square has the *largest* perimeter? And isn't it a bit odd that we did not discover a stationary point corresponding to the smallest value (which can easily be seen to occur at $x = 0$ or $x = d$)? To convince yourself that the maximum point gives the largest value, it is a good idea to sketch a graph: you can check that $\mathrm{d}P/\mathrm{d}x$ is positive for $0 < x < d/\sqrt{2}$ and negative for $d/\sqrt{2} < x < d$. The smallest values of the perimeter occur at the endpoints of the interval $0 \leqslant x \leqslant d$ and therefore do not have to be turning points.

A much better way of tacking the problem is to set $x = d\cos\theta$ and $y = d\sin\theta$ and find the perimeter as a function of θ. Because θ can take any value (there are no end points such as $x = 0$ to consider), the largest and smallest perimeters correspond to stationary points. Try it.

Question 10$(*)$

Use the first four terms of the binomial expansion of $(1 - 1/50)^{1/2}$, to derive the approximation $\sqrt{2} \approx 1.414214$.

Calculate similarly an approximation to the cube root of 2 to six decimal places by considering $(1 + N/125)^{1/3}$, where N is a suitable number.

[You need not justify the accuracy of your approximations.]

Comments

Although you do not have to justify your approximations, you do need to think carefully about the number of terms required in the expansions. You first have to decide how many you will need to obtain the given number of decimal places; then you have to think about whether the next term is likely to affect the value of the last decimal.

It is not at all obvious where the $\sqrt{2}$ in the first part comes from until you write $(1 - 1/50)^{1/2} = \sqrt{(49/50)}$.

In the second part, you have to choose N in such a way that $125 + N$ has something to do with a power of 2.

It is a bit easier to do the binomial expansions by making the denominator a power of 10 (so that the expansion in the first part becomes $(1 - 2/100)^{1/2}$).

Solution to question 10

First we expand binomially:

$$\begin{aligned}
&\left(1-\frac{2}{100}\right)^{\frac{1}{2}}\\
\approx\ \ & 1+\left(\frac{1}{2}\right)\left(-\frac{2}{100}\right)+\left(\frac{1}{2!}\right)\left(\frac{1}{2}\right)\left(-\frac{1}{2}\right)\left(-\frac{2}{100}\right)^2-\left(\frac{1}{3!}\right)\left(\frac{1}{2}\right)\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\left(-\frac{2}{100}\right)^3+\cdots\\
=\ \ & 1-\frac{1}{100}-\frac{0.5}{10^4}-\frac{0.5}{10^6}=0.9899495\,.
\end{aligned}$$

It is clear that the next term in the expansion would introduce the seventh and eighth places of decimals, which it seems we do not need. Of course, after further manipulations we might find that the above calculation does not supply the 6 decimal places we need, in which case we will work out the next term in the expansion.

But $\left(\dfrac{98}{100}\right)^{\frac{1}{2}}=\dfrac{7\sqrt{2}}{10}$, so $\sqrt{2}\approx 9.899495/7\approx 1.414214\,.$

Second part:

$$\begin{aligned}
&\left(1+\frac{3}{125}\right)^{\frac{1}{3}}\\
\approx\ \ & 1+\left(\frac{1}{3}\right)\left(\frac{3}{125}\right)+\left(\frac{1}{2!}\right)\left(\frac{1}{3}\right)\left(-\frac{2}{3}\right)\left(\frac{3}{125}\right)^2+\left(\frac{1}{3!}\right)\left(\frac{1}{3}\right)\left(-\frac{2}{3}\right)\left(-\frac{5}{3}\right)\left(-\frac{3}{125}\right)^3+\cdots\\
=\ \ & 1-\frac{8}{1000}-\frac{64}{10^6}+\frac{5}{3}\frac{8^3}{10^9}\\
=\ \ & 1.007936+\frac{256}{3}\frac{1}{10^8}=1.007937\ .
\end{aligned}$$

Successive terms in the expansion decrease by a factor of about 1000, so this should give the right number of decimal places.

But $\left(\dfrac{128}{125}\right)^{\frac{1}{3}}=\dfrac{4\sqrt[3]{2}}{5}=\dfrac{8\sqrt[3]{2}}{10}$ so $\sqrt[3]{2}\approx 10.0793/8\approx 1.259921\,.$

Postmortem

This question required a bit of intuition, and some accurate arithmetic. You don't have to be brilliant at arithmetic to be a good mathematician, but most mathematicians aren't bad at it. There have in the past been children who were able to perform extroardinary feats of arithmetic. For example, Zerah Colburn, a 19th century American, toured Europe at the age of 8. He was able to multiply instantly any two four digit numbers given to him by the audiences. George Parker Bidder (the Calculating Boy) could perform similar feats, though unlike Colburn he became a distinguished mathematician and scientist. One of his brothers knew the bible by heart.

The error in the approximation is the weak point of this question. Although it is clear that the next term in the expansion is too small to affect the accuracy, it is not obvious that the sum of all the next hundred (say) terms of the expansion is negligible (though in fact it is). What is needed is an estimate of the truncation error in the binomial expansion. Such an estimate is not hard to obtain (first year university work) and is typically of the same order of magnitude as the first neglected term in the expansion.

Question 11($*$)

How many integers greater than or equal to zero and less than 1000 are not divisible by 2 or 5? What is the average value of these integers?

How many integers greater than or equal to zero and less than 9261 are not divisible by 3 or 7? What is the average value of these integers?

Comments

There are a number of different ways of tackling this problem, but it should be clear that whatever way you choose for the first part will also work for the second part (especially when you realise the significance of the number 9261). A key idea for the first part is to think in terms of blocks of 10 numbers, realising that all blocks of 10 are the same for the purposes of the problem.

Solution to question 11

Only integers ending in 1, 3, 7, or 9 are not divisible by 2 or 5. This is 4/10 of the possible integers, so the total number of such integers is 4/10 of 1,000, i.e. 400.

The integers can be added in pairs:

$$\text{sum} = (1+999)+(3+997)+\cdots+(499+501)\,.$$

There are 200 such pairs, so the sum is $1,000 \times 200$ and the average is 500. A simpler argument would be to say that this is obvious by symmetry: there is nothing in the problem that favours an answer greater (or smaller) than 500.

Alternatively, we can find the number of integers divisible by both 2 and 5 by adding the number divisible by 2 (i.e. 500) to the number divisible by 5 (i.e. 200), and subtracting the number divisible by both 2 and 5 (i.e. 100) since these have been counted twice. To find the sum, we can sum those divisible by 2 (using the formula for a geometric progression), add the sum of those divisible by 5 and subtract the sum of those divisible by 10.

For the second part, either of these methods will work. In the first method you consider integers in blocks of 21 (essentially arithmetic to base 21): there are 12 integers in each such block that are not divisible by 3 or 7 (namely 1, 2, 4, 5, 8, 10, 11, 13, 16, 17, 19, 20) so the total number is $9261\times 12/21 = 5292\,$. The average is $9261/2$ as can be seen using the pairing argument $(1+9260)+(2+9259)+\cdots$ or the symmetry argument.

Postmortem

I was in year 7 at school when I had my first encounter with lateral thinking in mathematics. The problem was to work out how many houses a postman delivers to in a street of houses numbered from 1 to 1000, given that he refuses to deliver to houses with the digit 9 in the number. It seemed impossible to do it systematically, until the idea of counting in base 9 occurs; and then it seemed brilliantly simple. I didn't know I was counting in base 9; in those days, school mathematics was very traditional and base 9 would have been thought very advanced. All that was required was the idea of working out how many numbers can be made out of nine digits (i.e. 9^3) instead of 10; and of course it doesn't matter which nine.

Did you spot the significance of the number 9261? It is 21^3, i.e. the number written as 1000 in base 21. Note how carefully the question is written and laid out to suggest a connection between the two paragraphs (and to highlight the difference); there is not even a part (i) and Part (ii) to break the symmetry. For the examination, the number 1,000,000 was used instead of 1,000 to discourage candidates from spending the first hour of the examination writing down the numbers from 1 to 1000.

Having realised the that the first part involves $(2\times 5)^3$ and the second part involves $(3\times 7)^3$ you are probably now wondering what the general result is, i.e:

How many integers greater than or equal to zero and less than $(pq)^3$ are not divisible by p or q? What is the average value of these integers?

I leave this to you, with the hint that you need to think about only pq, rather than $(pq)^3$, to start with.

Question 12(*)

Sketch the following subsets of the x-y plane:

(i) $|x| + |y| \leqslant 1$;

(ii) $|x - 1| + |y - 1| \leqslant 1$;

(iii) $|x - 1| - |y + 1| \leqslant 1$;

(iv) $|x|\,|y - 2| \leqslant 1$.

Comments

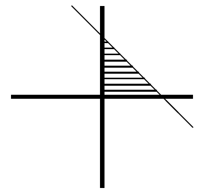
Often with modulus signs, it is easiest to consider the cases separately, so for example is part (i), you would first work out the case $x > 0$ and $y > 0$, then $x > 0$ and $y < 0$, and so on. Here there is a simple geometric understanding of the different cases: once you have worked out the first case, the other three can be deduced by symmetry.

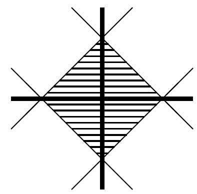
Another geometric idea should be in your mind when tackling this question, namely the idea of translations in the plane.

Solution to question 12

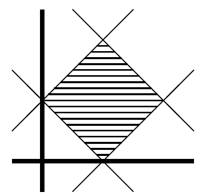
The way to deal with the modulus signs in this question is to consider first the case when the things inside the modulus signs are positive, and then get the full picture by symmetry, shifting the origin as appropriate.

For part (i), consider the the first quadrant $x \geqslant 0$ and $y \geqslant 0$. In this quadrant, the inequality is $x + y \leqslant 1$. Draw the line $x + y = 1$ and then decide which side of the line is described by the inequality. It is obviously (since x has to be smaller than something) the region to the left of the line; or (since y also has to be smaller than something) the region below the line, which is the same region.



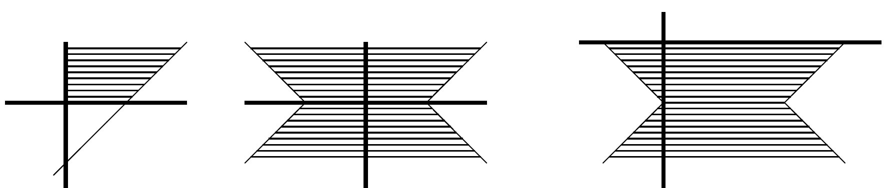


Similar arguments could be used in the other quadrants, but it is easier to note that the inequality $|x| + |y| \leqslant 1$ unchanged when x is replaced by $-x$, or y is replaced by $-y$, so the sketch should have reflection symmetry in both axes.

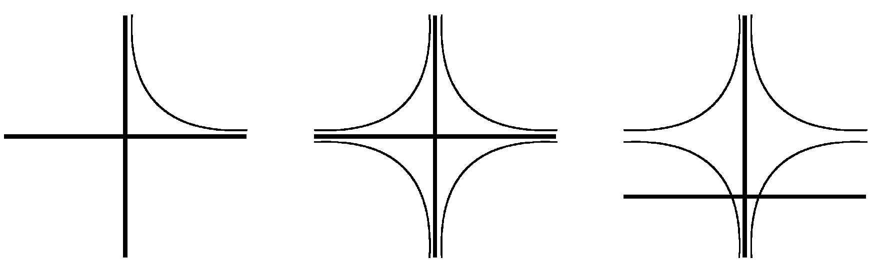


Part (ii) is the same as part (i), except that the origin is translated to $(1, 1)$.

For part (iii), consider first $x - y \leqslant 1$, where $x > 0$ and $y > 0$, which give an area that is infinite in extent in the positive y direction. Then reflect this in both axes and translate one unit down the y axis and 1 unit along the positive x axis.



Part (iv) is the rectangular hyperbola $xy = 1$ reflected in both axes and the region required is enclosed by the four hyperbolas. The four hyperbolas are then translated 2 units up the y axis.



Question 13$(**)$

Show that

$$x^2 - y^2 + x + 3y - 2 = (x - y + 2)(x + y - 1)$$

and hence, or otherwise, indicate by means of a sketch the region of the x-y plane for which

$$x^2 - y^2 + x + 3y > 2.$$

Sketch also the region of the x-y plane for which

$$x^2 - 4y^2 + 3x - 2y < -2.$$

Give the coordinates of a point for which both inequalities are satisfied or explain why no such point exists.

Comments

This question gets a two star difficulty rating because inequalities always need to be handled with care.

For the very first part, you could either factorise the left hand side to obtain the right hand side, or multiply out the right hand side to get the left hand side. Obviously, it is much easier to do the multiplication than the factorisation. But is that 'cheating' or taking a short cut that might lose marks? The answer to this is no: it doesn't matter if you start from the given answer and work backwards — it is still a mathematical proof and any proof will get the marks. (But note that if there is a 'hence' in the question, you will lose marks if you do not do it using the thing you have just proved.)

In the second part, you have to do the factorisation yourself, so you should look carefully at where the terms in the (very similar) first part come from when you do the multiplication. It will help to spot the similarities bewteen $x^2 - y^2 + x + 3y - 2$ and $x^2 - 4y^2 + 3x - 2y + 2$

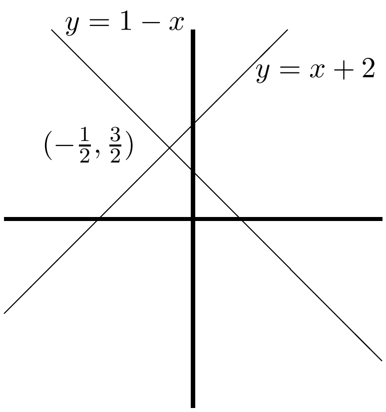
Since no indication is given as to what detail should appear on the sketch, you are have to use your judgement: it is clearly important to know where the regions lie relative to the coordinate axes.

Solution to question 13

To do the multiplication, it pays to be systematic and to set out the algebra nicely:

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

as required.



For the first inequality, we need $x-y+2$ and $x+y-1$ both to be either positive or negative. To sort out the inequalities (or inequations as they are sometimes horribly called), the first thing to do is to draw the lines corresponding to the corresponding equalities. These lines divide the plane into four regions and we then have to decide which regions are relevant.

The diagonal lines $x-y=-2$ and $x+y=1$ intersect at $(-\frac{1}{2}, \frac{3}{2})$; this is the important point to mark in on the sketch. The required regions are the left and right quadrants formed by these diagonal lines (since the inequalities mean that the regions are to the right of both lines or to the left of both lines).

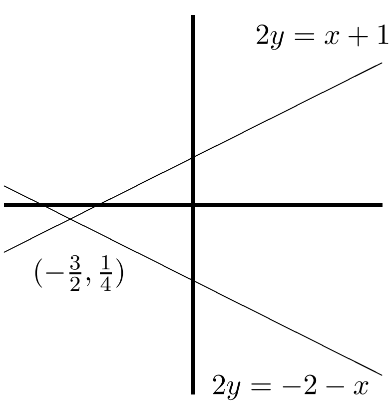
For the second part, the first thing to do is to factorise $x^2-4y^2+3x-2y+2$. The key similarity with the first part is the absence of 'cross' terms of the form xy . This allows a difference-of-two-squares factorisation of the first two terms: $x^2-4y^2=(x-2y)(x+2y)$. Following the pattern of the first part, we can then try a factorisation of the form

$$x^2-4y^2+3x-2y+2=(x-2y+a)(x+2y+b)$$

where $ab=2$. Considering the terms linear in x and y gives $a+b=3$ and $2a-2b=-2$ which quickly leads to $a=1$ and $b=2$. Note that altogether there were three equations for a and b, so we had no right to expect a consistent solution (except for the fact that this is a STEP question for which we had every right to believe that the first part would guide us through the second part).

Alternatively, we could have completed the square in x and in y and then used difference of two squares:

$$\begin{aligned}x^2-4y^2+3x-2y+2 &= (x+\tfrac{3}{2})^2-\tfrac{9}{4}-(2y+\tfrac{1}{2})^2+\tfrac{1}{4}+2\\ &= (x+\tfrac{3}{2})^2-(2y+\tfrac{1}{2})^2=(x+\tfrac{3}{2}-2y-\tfrac{1}{2})(x+\tfrac{3}{2}+2y+\tfrac{1}{2})\ .\end{aligned}$$



As in the previous case, the required area is formed by two intersecting lines; this time, they intersect at $(-\frac{3}{2}, -\frac{1}{4})$ and the upper and lower regions are required, since the inequality is the other way round.

It is easy to see from the sketches that there are points that satisfy both inequalities: for example $(1,2)$.

Question 14($**$)

Solve the inequalities

(i) $\quad 1 + 2x - x^2 > 2/x \qquad (x \neq 0)$,

(ii) $\quad \sqrt{(3x + 10)} > 2 + \sqrt{(x + 4)} \qquad (x \geqslant -10/3)$.

Comments

The two parts are unrelated (unusually for STEP questions), except that they deal with inequalities. Both parts need care, having traps for the unwary. In the first part you have to watch out when you multiply an inequality: if the thing you muliply is negative then the inequality reverses. In the second part you have to consider the possibility that your algebraic manipulations have created extra spurious solutions. In the first part, you might find sketching a graph helpful.

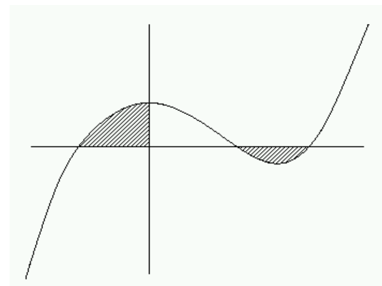
For the second part, it is worth (after you have finished the question) sketching the graphs of $\sqrt{(3x + 10)}$, $2 + \sqrt{(x + 4)}$ and $2 - \sqrt{(x + 4)}$. You can get them by translations of $\sqrt{x}$ (which you can get by reflection $y = x^2$ in the line $y = x$. I would have drawn them for you in the postmortem comment overleaf, but there wasn't room on the page.

Solution to question 14

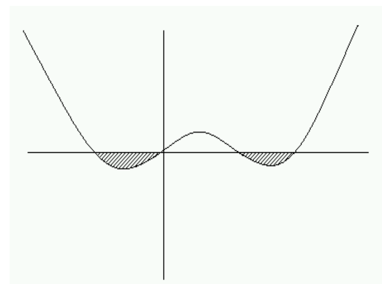
(i) We would like to multiply both sides of the inequality by x in order to obtain a nice cubic expression. However, we have to allow for the possibility that x is negative (which would reverse the inequality). One way is to consider the cases $x > 0$ and $x < 0$ separately.

For $x > 0$, we multiply the whole equation by x without changing the inequality:

$$\begin{aligned} 1 + 2x - x^2 > 2/x &\Rightarrow x + 2x^2 - x^3 > 2 \\ &\Rightarrow x^3 - 2x^2 - x + 2 < 0 \\ &\Rightarrow (x-1)(x+1)(x-2) < 0 \\ &\Rightarrow 1 < x < 2, \end{aligned} \qquad (*)$$



discarding the possibility $x < -1$, since we have assumed that $x > 0$. The easiest way of obtaining the result $(*)$ from the previous line is to sketch the graph of $(x-1)(x+1)(x-2) < 0$.
For $x < 0$, we must reverse the inequality when we multiply by x so in this case, $(x-1)(x+1)(x-2) > 0$, which gives $-1 < x < 0$.



The smart way to do $x > 0$ and $x < 0$ in one step is to multiply by x^2 (which is never negative and hence never changes the direction of the inequality) and analyse $x(x-1)(x+1)(x-2) < 0$. Again, a sketch is useful.

(ii) First square both sides the inequality $\sqrt{(3x+10)} > 2 + \sqrt{(x+4)}$:

$$3x + 10 > 4 + 4(x+4)^{\frac{1}{2}} + (x+4) \qquad \text{i.e.} \qquad x + 1 > 2(x+4)^{\frac{1}{2}}\,.$$

Note that the both sides of the original inequality are positive or zero (i.e. non-negative), so the direction of the inequality is not changed by squaring. Now consider the new inequality $x + 1 > 2(x+4)^{\frac{1}{2}}$. If both sides are non-negative, that is if $x > -1$, we can square both sides again without changing the direction of the inequality. But, if $x < -1$, the inequality cannot be satisfied since the right hand side is always non-negative. Squaring gives

$$x^2 + 2x + 1 > 4(x+4) \qquad \text{i.e.} \qquad (x-5)(x+3) > 0\,.$$

Thus, $x > 5$ or $x < -3$. However, we must reject $x = -3$ because of the condition $x > -1$. Therefore the inequality holds for $x > 5$.

Postmortem

The spurious result $x < -3$ at the end of the second part arises through loss of information in the process of squaring. (If you square an expression, you lose its sign.) After squaring twice, the resulting inequality is the same as would have resulted from $\sqrt{(3x+10)} > 2 - \sqrt{(x+4)}$ and it is to this inequality that $x < -3$ is the solution.

Question 15($**$)

Sketch, without calculating the stationary points, the graph of the function $\mathrm{f}(x)$ given by

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

where $p < q < r$. By considering the quadratic equation $\mathrm{f}'(x) = 0$, or otherwise, show that

$$(p + q + r)^2 > 3(qr + rp + pq) \,.$$

By considering $(x^2 + gx + h)(x - k)$, or otherwise, show that $g^2 > 4h$ is a sufficient condition but not a necessary condition for the inequality

$$(g - k)^2 > 3(h - gk)$$

to hold.

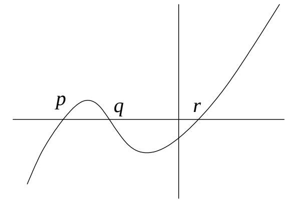
Comments

For the sketch it is not necessary to do more than think about the behaviour for x large and positive, and for x large and negative, and the points at which the graph crosses the x-axis.

The argument required for the second part is very similar to that of the first part, except that it also tests understanding of the meaning of the terms necessary and sufficient.

Solution to question 15

From the sketch, we see that $\mathrm{f}(x)$ has two turning points, so the equation $\mathrm{f}'(x) = 0$ has two real roots.



Now $\mathrm{f}(x) = (x-p)(x-q)(x-r) = x^3 - (p+q+r)x^2 + (qr+rp+pq)x - pqr$ so

$$\mathrm{f}'(x) = 3x^2 - 2(p+q+r)x + (qr+rp+pq) = 0\,.$$

Using the condition '$b^2 > 4ac$' for the quadratic to have two real root gives

$$4(p+q+r)^2 > 12(qr+rp+pq)\,,$$

which is the required result.

For the second part, we can use a similar argument.

In this case, the cubic expression is factorised into a linear factor $x-k$ and a quadratic factor x^2+gx+h. If the quadratic equation $x^2+gx+h=0$ has two distinct real roots (i.e. if $g^2 > 4h$), then the function $(x^2+gx+h)(x-k)$ has three zeros and therefore two turning points. Thus if $g^2 > 4h$ the equation $\mathrm{f}'(x) = 0$ has two real roots.

Expanding $\mathrm{f}(x)$ gives $\mathrm{f}(x) = (x^2+gx+h)(x-k) = x^3 + (g-k)x^2 + (h-gk)$ and differentiating gives

$$\mathrm{f}'(x) = 3x^2 + 2(g-k)x + (h-gk) = 0\,.$$

The condition for two real roots is $4(g-k)^2 > 12(h-gk)$. Thus $g^2 > 4h$ is a sufficient condition for this inequality to hold.

Why is it not necessary? That is to say, why is it too strong a condition? How could the condition for two turning points hold without $\mathrm{f}(x)$ having three zeros. The answer is that both turning points could be above the x–axis or both could be below. Saying this would get all the marks. Or you could give a counterexample: the easy counterexample is $g^2 = 4h$; if this holds, there are still two turning points so $(g-k)^2 > 3(h-gk)$. (Careful of this counterexample: it will not work if the double root of the quadratic (at $x = -g/2$ coincides with the root of the linear factor (at $x = k$), so we need the extra condition that $g/2 + k \neq 0$.)

Postmortem

I got considerable satisfaction from the way that non-obvious inequalities are derived from understanding simple graphs and the quadratic formula when I set this question. But are the inequalities really obscure? The answer is unfortunately no.

The inequality of the first part is equivalent to $(q-r)^2 + (r-p)^2 + (p-q)^2 > 0$ (the inequality coming from the fact the squares of real numbers are non-negative and in this case the given condition $p < q < r$ means that the expression is strictly positive (not equal to zero).

If we rewrite the second inequality as $(k+g/2)^2 + 3(g^2/4 - h) > 0$ we see that it is certainly holds if $(g^2/4 - h) > 0$. Clearly, this is not a necessary condition: there is the obvious counterexample $(g^2/4 - h) = 0$ and $(k+g/2) \neq 0$. But there are lots of other easy counterexamples — for example, $g = 0$ and $k^2 > 3h > 0$.

Question 16($**$)

In a cosmological model, the radius R of the universe is a function of the age t of the universe. The function R satisfies the three conditions:

$$\mathrm{R}(0) = 0, \qquad \mathrm{R}'(t) > 0 \text{ for } t > 0, \qquad \mathrm{R}''(t) < 0 \text{ for } t > 0, \tag{$*$}$$

where R$''$ denotes the second derivative of R. The function H is defined by

$$\mathrm{H}(t) = \frac{\mathrm{R}'(t)}{\mathrm{R}(t)} \,.$$

(i) Sketch a graph of $\mathrm{R}(t)$. By considering a tangent to the graph, show that $t < 1/\mathrm{H}(t)$.

(ii) Observations reveal that $\mathrm{H}(t) = a/t$, where a is constant. Derive an expression for $\mathrm{R}(t)$. What range of values of a is consistent with the three conditions ($*$)?

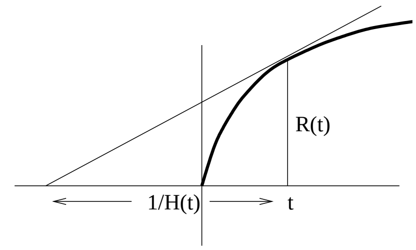
(iii) Suppose, instead, that observations reveal that $\mathrm{H}(t) = bt^{-2}$, where b is constant. Show that this is not consistent with conditions ($*$) for any value of b.

Comments

The sketch just means any graph starting at the origin and increasing, but with decreasing gradient. The second part of (i) needs a bit of thought (where does the tangent intersect the x-axis?) so don't despair if you don't see it immediately. Parts (ii) and (iii) are perhaps easier than (i).

Solution to question 16

(i) The height of the right-angled triangle in the figure is $\mathrm{R}(t)$ and the slope of the hypotenuse is $\mathrm{R}'(t)$. The length of the base is therefore $\mathrm{R}(t)/\mathrm{R}'(t)$, i.e.$1/\mathrm{H}(t)$. The figure shows that the intersection of the tangent to the graph for $t > 0$ intercepts the negative t–axis, so $1/\mathrm{H}(t) > t$.



(ii) One way to proceed is to integrate the differential equation:

$$\mathrm{H}(t) = \frac{a}{t} \Rightarrow \frac{\mathrm{R}'(t)}{R(t)} = \frac{a}{t} \Rightarrow \int \frac{\mathrm{R}'(t)}{\mathrm{R}(t)}\mathrm{d}t = \int \frac{a}{t}\mathrm{d}t \Rightarrow \ln \mathrm{R(t)} = a \ln t + \text{constant} \Rightarrow \mathrm{R}(t) = At^a \,.$$

The first two conditions $(*)$ are satisfied if $a > 0$ and $A > 0$. For the third condition, we have $\mathrm{R}''(t) = a(a-1)At^{a-1}$ which is negative provided $a < 1$. The range of a is therefore $0 < a < 1$,

(iii) The obvious way to do this just follows part (ii). This time, we have

$$\mathrm{H}(t) = \frac{b}{t^2} \Rightarrow \frac{\mathrm{R}'(t)}{\mathrm{R}(t)} = \frac{b}{t^2} \Rightarrow \int \frac{\mathrm{R}'(t)}{\mathrm{R}(t)}\mathrm{d}t = \int \frac{b}{t^2}\mathrm{d}t \Rightarrow \ln \mathrm{R}(t) = -\frac{b}{t} + \text{constant} \Rightarrow \mathrm{R}(t) = Ae^{-b/t} \,.$$

Thus $\mathrm{R}'(t) = \mathrm{H}(t)\mathrm{R}(t) = Abt^{-2}e^{-b/t}$. Clearly $A > 0$ since $\mathrm{R}(t) > 0$ for $t > 0$ so $\mathrm{R}'(t) > 0$. Furthermore $\mathrm{R}(0) = 0$ (think about this!), so only the condition $\mathrm{R}''(t) < 0$ remains to be checked. Differentiating $\mathrm{R}'(t)$ gives

$$\mathrm{R}''(t) = Ab(-2t^{-3})e^{-b/t} + Abt^{-2}e^{-b/t}(bt^{-2}) = Abt^{-4}e^{-b/t}(-2t + b)$$

This is positive when $t < b/2$, which contradicts the condition $\mathrm{R}'' < 0$.

Instead of solving the differential equation, we could proceed as follows. We have

$$\frac{\mathrm{R}'}{\mathrm{R}} = bt^{-2} \Rightarrow \mathrm{R}' = b\mathrm{R}t^{-2} \Rightarrow \mathrm{R}'' = b\mathrm{R}'t^{-2} - 2b\mathrm{R}t^{-3} = b^2t^{-4}\mathrm{R} - 2b\mathrm{R}t^{-3} = b(b-2t)t^{-4}\mathrm{R}\,.$$

This is positive when $t < b/2$, which contradicts the condition $\mathrm{R}'' < 0$.

Postmortem

The interpretation of the conditions is as follows. The first condition $\mathrm{R}(0) = 0$ says that the universe started from zero radius — in fact, from the 'big bang'. The second condition says that the universe is expanding. This was one of the key discoveries in cosmology in the last century. The quantity called here $\mathrm{H}(t)$ is called *Hubble's constant* (though it varies with time). It measures the rate of expansion of the universe. Its present day value is a matter of great debate (it is of the order of 75 megaparsecs per century). The last condition says that the expansion of the universe is slowing down. This is expected on physical grounds because of the gravitational attraction of the galaxies on each other.

Another observational uncertainly is the value of the deceleration. It is not clear yet whether the universe will expand for ever or recollapse.

Question 17$(*)$

Show that the equation of any circle passing through the points of intersection of the ellipse

$$(x+2)^2 + 2y^2 = 18$$

and the ellipse

$$9(x-1)^2 + 16y^2 = 25$$

can be written in the form

$$x^2 - 2ax + y^2 = 5 - 4a\ .$$

Comments

When this question was set (as question 1 on STEP I) it seemed too easy. But quite a high proportion of the candidates made no real progress. Of course, is is not easy to keep a cool head under examination conditions. But surely it is obvious that either x or y has to be eliminated from the two equations for the ellipses; and having decided that, it is obvious which one to eliminate.

It is worth thinking about how many points of intersection of the ellipses we are expecting: a sketch might help.

Solution to question 17

First we have to find the intersections of the two ellipses by solving the simultaneous equations

$$\begin{aligned}(x+2)^2+2y^2 &= 18\\ 9(x-1)^2+16y^2 &= 25\ .\end{aligned}$$

We can eliminate y by multiplying the first equation by 8 and subtracting the two equations:

$$8(x+2)^2-9(x-1)^2=144-25 \qquad \text{i.e.} \qquad x^2-50x+96=0 \qquad \text{i.e.} \qquad (x-2)(x-48)=0\ .$$

The two possible values for x at the intersections are therefore 2 and 48.

Next we find the values of y at the intersection. Taking $x=2$ and substituting into the first simultaneous equation, we have $16+2y^2=18$, so $y=\pm1$. Taking $x=48$ gives $50^2+2y^2=18$ which has no (real) roots. Thus there are two points of intersection, at $(2,1)$ and $(2,-1)$.

Suppose a circle through these points has centre (p,q) and radius R. Then the equation of the circle is $(x-p)^2+(y-q)^2=R^2$. Setting $(x,y)=(2,1)$ and $(x,y)=(2,-1)$ gives two equations:.

$$(2-p)^2+(1-q)^2=R^2\ , \qquad (2-p)^2+(-1-q)^2=R^2\ .$$

Subtracting gives $q=0$ (it is obvious anyway that the centre of the circle must lie on the y axis). Thus the equation of any circle passing through the intersections is

$$(x-p)^2+y^2=(2-p)^2+1\ ,$$

which simplifies to the given result with $p=a$.

Question 18$(***)$

Let $f(x) = x^m(x-1)^n$, where m and n are both integers greater than 1. Show that

$$f'(x) = \left(\frac{m}{x} - \frac{n}{1-x}\right) f(x)\,.$$

Show that the curve $y = f(x)$ has a stationary point with $0 < x < 1$. By considering $f''(x)$, show that this stationary point is a maximum if n is even and a minimum if n is odd.

Sketch the graphs of $f(x)$ in the four cases that arise according to the values of m and n.

Comments

There is quite a lot in this question, but it is one of the best STEP questions on basic material that I came across (I looked through all 15 years of STEP's history).

First, you have to find $f'(x)$. The very first part (giving a form of $f'(x)$) did not occur in the original STEP question. This particular expression is extremely helpful when if comes to finding the value of $f''(x)$ at the stationary point. In the actual exam, a handful of candidates successfully found a general expression for $f''(x)$ in terms of x then evaluated it at the stationary point: first class work.

You should find that the value of $f''(x)$ at the stationary point can be expressed in the form $(\cdots)f(x)$, where the factor in the brackets is quite simple and always negative, so that the nature of the stationary point depends only on the sign of $f(x)$. (Actually, this is intuitively obvious: $f(0) = f(1) = 0$ so the one stationary point must be, for example, a maximum if $f(x) > 0$ for $0 < x < 1$.)

For the sketches, you really have only to think about the behaviour of $f(x)$ when $|x|$ is large and the sign of $f(x)$ between $x = 0$ and $x = 1$. That allows you to piece together the graph, knowing that there is only one stationary point between $x = 0$ and $x = 1$.

Note that, very close to $x = 0$, $f(x) \approx x^m(-1)^n$ so it is easy to see how the graph there depends on n and m; you can use this to check that you have got the graphs right.

Postmortem

[This didn't fit on the next page: read it after you have done the question.]

You may have noticed a little carelessness at the top of the next page: what happens in the logarithmic differentiation if any of $f(x)$ or x or $x-1$ are negative? The answer is that it doesn't matter. One way to deal with the problem of logs with negative arguments is to put modulus signs everywhere using the correct result

$$\frac{d}{dx}\ln|f(x)| = \frac{f'(x)}{f(x)}\,.$$

If you are not sure of this, try it on $f(x) = x$ taking the two cases $x > 0$ and $x < 0$ separately.

A more sophisticated way of dealing with logs with negative arguments is to note that $\ln(-f(x)) = \ln f(x) + \ln(-1)$. We don't have to worry about $\ln(-1)$ because it is a constant (of some sort) and so won't affect the differentiation. Actually, a value of $\ln(-1)$ can be obtained by taking logs of Euler's famous formula $e^{i\pi} = -1$ giving $\ln(-1) = i\pi$.

Solution to question 18

The first result can be established by differentiating $\mathrm{f}(x)$ directly, but the neat way to do it is to start with $\ln \mathrm{f}(x)$:

$$\ln \mathrm{f}(x) = m \ln x + n \ln(x-1) \Longrightarrow \frac{\mathrm{f}'(x)}{\mathrm{f}(x)} = \frac{m}{x} + \frac{n}{x-1} \, ,$$

as required.

Thus $\mathrm{f}(x)$ has stationary points at $x = 0$ and $x = 1$ (since $m - 1 > 0$ and $n - 1 > 0$) and when

$$\frac{m}{x} - \frac{n}{1-x} = 0 \, ,$$

i.e. when $m(1-x) - nx = 0$. Solving this last equation for x gives $x = m/(m+n)$, which lies between 0 and 1 since m and n are positive.

Next we calculate $\mathrm{f}''(x)$. Starting with

$$\mathrm{f}'(x) = \left(\frac{m}{x} - \frac{n}{1-x}\right) \mathrm{f}(x) \, ,$$

we obtain

$$\mathrm{f}''(x) = \left(-\frac{m}{x^2} - \frac{n}{(1-x)^2}\right) \mathrm{f}(x) + \left(\frac{m}{x} - \frac{n}{1-x}\right) \mathrm{f}'(x) \, .$$

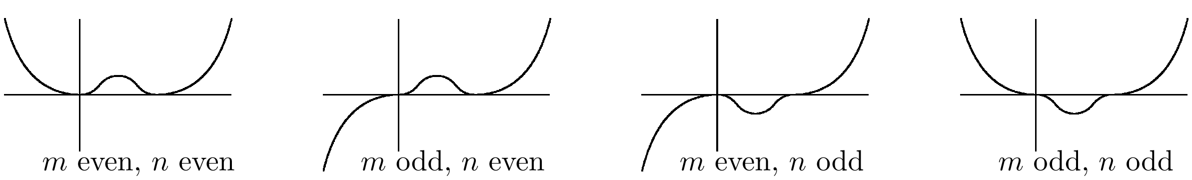
At the stationary point, the second of these two terms is zero because $\mathrm{f}'(x) = 0$. Thus

$$\mathrm{f}''(x) = \left(-\frac{m}{x^2} - \frac{n}{(x-1)^2}\right) \mathrm{f}(x) \, .$$

The bracketed expression is negative so at the stationary point $\mathrm{f}''(x) < 0$ if $\mathrm{f}(x) > 0$ and $\mathrm{f}''(x) > 0$ if $\mathrm{f}(x) < 0$.

The sign of $\mathrm{f}(x)$ for $0 < x < 1$ is the same as the sign of $(x-1)^n$, since $x^m > 0$ when $x > 0$. The stationary point is between 0 and 1, so $x - 1 < 0$ and $(x-1)^n > 0$ if n is even and $(x-1)^n < 0$ if n is odd. Thus $\mathrm{f}''(x) < 0$ if n is even and $\mathrm{f}''(x) > 0$ if n is odd, which is the required result.

The four cases to sketch are determined by whether m and n are even or odd. The easiest way to understand what is going on is to consider the graphs of x^m and $(x-1)^n$ separately, then try to join them up at the stationary point between 0 and 1. If m is odd, then $x = 0$ is a point of inflection, but if m is even it is a maximum or minimum according to the sign of $(x-1)^n$. You should also think about the behaviour for large $|x|$. All the various bits of information (including the nature of the turning point investigated above) should all piece neatly together.



m even, n even　　m odd, n even　　m even, n odd　　m odd, n odd

Question 19$(***)$

Give a sketch of the curve $y = \dfrac{1}{1+x^2}$, for $x \geqslant 0$.

Find the equation of the line that intersects the curve at $x = 0$ and is tangent to the curve at some point with $x > 0$. Prove that there are no further intersections between the line and the curve. Draw the line on your sketch.

By considering the area under the curve for $0 \leqslant x \leqslant 1$, show that $\pi > 3$.

Show also, by considering the volume formed by rotating the curve about the y axis, that $\ln 2 > 2/3$.

[**Note**: $\displaystyle\int_0^1 \frac{1}{1+x^2}\,\mathrm{d}x = \frac{\pi}{4}$.]

Comments

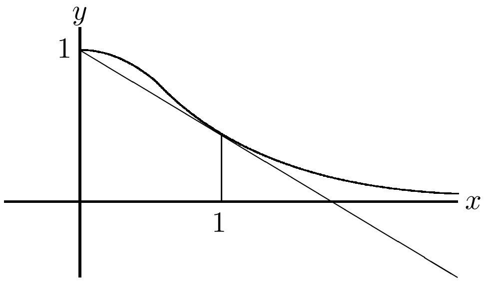
There is quite a lot to this question, so three stars even though none of it is particularly difficult. The most common mistake in finding the equation of the tangent is to muddle the y and x that occur in the equation of the line ($y = mx + c$) with the coordinates of the point at which the tangent meets the curve, getting 'constants' m and c that depend on x. I'm sure you wouldn't make this rather elementary mistake normally, but it is surprising what people do under examination conditions.

The reason for sketching the curve lies in the last parts: the shape of the curve relative to the straight line provides the inequality.

The note at the end of the question had to be given because integrals of that form (giving inverse trigonometric functions) are not in the core A-level syllabus.

Solution to question 19

The sketch should show a graph which has gradient 0 at $(0,1)$ and which asymptotes to the x-axis for large x. The line is the tangent to the graph from the point $(0,1)$ which is required later.



First we find the equation of the tangent to the curve $y = (1+x^2)^{-1}$ at the point (p,q). The gradient of the curve at the point (p,q) is $-2p(1+p^2)^{-2} = -2pq^2$, so the equation of the tangent is

$$y = -2pq^2x + c$$

where c is given by $q = -2pq^2p + c$. This line is supposed to pass through the point $(0,1)$, so $1 = c$. Thus

$$1 = \frac{1}{1+p^2} + \frac{2p^2}{(1+p^2)^2}$$

which simplifies to

$$(1+p^2)^2 = 1 + 3p^2 \qquad \text{i.e.} \qquad p^4 = p^2 \; .$$

The only positive solution is $p = 1$, so the equation of the line is $y = -x/2 + 1$.

We can check that this line does not meet the curve again by solving the equation

$$-\frac{x}{2} + 1 = \frac{1}{1+x^2} \; .$$

Multiplying by $(1+x)^2$ gives

$$(1 - x/2)(1+x^2) = 1 \qquad \text{i.e.} \qquad x^3 - 2x^2 + x = 0 \; .$$

Factorising shows that $x = 1$ and $x = 0$ satisfy this equation (these are the known roots, $x = 1$ being a double root corresponding to the tangency) and that there are no more roots.

The area under the curve for $0 \leqslant x \leqslant 1$ is $\pi/4$ as given. The sketch shows that this area is greater than the area under the tangent line for $0 \leqslant x \leqslant 1$, which is $\frac{1}{2} + \frac{1}{4}$ (the area of the rectangle plus the area of the triangle above it). Comparing with with $\pi/4$ gives the required result.

The volume formed by rotating the curve about the y axis is

$$\int_0^1 2\pi y x \mathrm{d}x = \pi \int_0^1 \frac{2x}{1+x^2} \mathrm{d}x = \pi \ln 2 \; .$$

This is greater than the volume formed by rotating the line about the y axis, which is

$$\int_0^1 2\pi y x \mathrm{d}x = 2\pi \int_0^1 (1 - x/2) x \mathrm{d}x = 2\pi/3 \; .$$

Comparison gives the required result.

Question 20($**$)

Let
$$I=\int_0^a \frac{\cos x}{\sin x+\cos x}\,\mathrm{d}x \qquad \text{and} \qquad J=\int_0^a \frac{\sin x}{\sin x+\cos x}\,\mathrm{d}x\,,$$
where $0\leqslant a<3\pi/4$. By considering $I+J$ and $I-J$, show that $2I=a+\ln(\sin a+\cos a)$.

Find also:

(i) $\displaystyle\int_0^{\pi/2} \frac{\cos x}{p\sin x+q\cos x}\,\mathrm{d}x$, where p and q are positive numbers;

(ii) $\displaystyle\int_0^{\pi/2} \frac{\cos x+4}{3\sin x+4\cos x+25}\,\mathrm{d}x$.

Comments

I love this question. When I first saw it (one of my colleagues proposed the first part for STEP Paper I a few years ago) I was completely taken by surprise. You might like to think how you would evaluate these integrals without using this method.

In order to use it again for STEP in 2002, I added parts (i) and (ii). It took me some time to think of a suitable extension. I was disappointed to find that the basic idea is more or less a one-off: there are very few denominators, besides the ones given, that lead to integrands amenable to this trick. But I was pleased with what I came up with. The question leads you through the opening paragraph and the extra parts depend very much on your having understood why the opening paragraph works.

In the examination, candidates who were successful in parts (i) and (ii) nearly always started off with the statement 'Now let $I=\cdots$ and $J=\cdots$.' The choice of significant numbers in part (ii) (3, 4 and 25) should be a clue.

Finally, you will have noticed that $0\leqslant a<3\pi/4$ was given in the question. Understanding the reason for this restriction does not help to do the question, but you should try to work out its purpose.

Solution to question 20

We have

$$\begin{aligned} I+J &= \int_0^a \frac{\cos x+\sin x}{\sin x+\cos x}\mathrm{d}x = \int_0^a \mathrm{d}x = a \\ I-J &= \int_0^a \frac{\cos x-\sin x}{\sin x+\cos x}\mathrm{d}x = \ln(\cos a+\sin a) \end{aligned}$$

(in the second integral, the numerator is the derivative of the denominator). Adding these equations gives the required expression for $2I$.

(i) Similarly, let $I = \displaystyle\int_0^{\pi/2} \frac{\cos x}{p\sin x+q\cos x}\mathrm{d}x$ and $J = \displaystyle\int_0^{\pi/2} \frac{\sin x}{p\sin x+q\cos x}\mathrm{d}x$. Then

$$\begin{aligned} qI+pJ &= \int_0^{\pi/2} \frac{q\cos x+p\sin x}{p\sin x+q\cos x}\mathrm{d}x = \frac{\pi}{2} \\ pI-qJ &= \int_0^{\pi/2} \frac{p\cos x-q\sin x}{p\sin x+q\cos x}\mathrm{d}x = \ln(p\sin\pi/2+q\cos\pi/2)-\ln(p\sin 0+q\cos 0) = \ln(p/q)\ . \end{aligned}$$

Now we multiply the first of these equations by q and the second by p and add the resulting equations:

$$(p^2+q^2)I = \frac{q\pi}{2} + p\ln(p/q)\ .$$

(ii) This time, let $I = \displaystyle\int_0^{\pi/2} \frac{\cos x+4}{3\sin x+4\cos x+25}\mathrm{d}x$ and $J = \displaystyle\int_0^{\pi/2} \frac{\sin x+3}{3\sin x+4\cos x+25}\mathrm{d}x$. Then

$$\begin{aligned} 4I+3J &= \int_0^{\pi/2} \frac{4\cos x+3\sin x+25}{3\sin x+4\cos x+25}\mathrm{d}x = \frac{\pi}{2} \\ 3I-4J &= \int_0^{\pi/2} \frac{3\cos x-4\sin x}{3\sin x+4\cos x+25}\mathrm{d}x = \ln(28/29)\ . \end{aligned}$$

Now we multiply the first of these equations by 4 and the second by 3 and add the resulting equations:

$$25I = 2\pi + 3\ln(28/29)\ .$$

Postmortem

The reason for restriction $0 \leqslant a < 3\pi/4$ is that the denominator should not be 0 for any value of x in the range of integration — otherwise, the integral is undefined. Writing $\sin x+\cos x = \sqrt{2}\sin(x+\pi/4)$ shows that the denominator is first zero when $x = 3\pi/4$.

Perhaps the easiest way of doing this sort of integral is to use the substitution $t = \tan(x/2)$ which converts the denominator to a quadratic in x.

Question 21($**$)

In this question, you may assume that if $k_1, \ldots, k_n$ are distinct positive real numbers, then

$$\frac{1}{n}\sum_{r=1}^{n} k_r > \left(\prod_{r=1}^{n} k_r\right)^{\frac{1}{n}},$$

i.e. their arithmetic mean is greater than their geometric mean.

Suppose that a, b, c and d are positive real numbers such that the polynomial

$$\mathrm{f}(x) = x^4 - 4ax^3 + 6b^2x^2 - 4c^3x + d^4$$

has four distinct positive roots.

(i) By considering the relationship between the coefficients of f and its roots, show that $c > d$.

(ii) By differentiating f, show that $b > c$.

(iii) Show that $a > b$.

Comments

This result is both surprising and pleasing. I'm sorry to say that I don't understand it. I've looked at it from various points of view, and asked a few selected colleagues who are usually good at this sort of thing, but so far no one has come up with a good explanation. Maybe the penny will drop in time for the next edition. (I haven't even worked out if the converse of this result holds, namely that the above inequalities are sufficient for the roots to lie in the given interval; I don't think this is hard, but I thought I would leave it until I undersood it better.)

There are general theorems on the number of roots in any given interval. This problem was studied by Descartes (1596–1650), who came up with his rule of signs. This relates the number of roots to the signs of terms in the polynomial, but it only gives an upper bound for the number of roots. The problem was finally cracked by Sturm in about 1835, but his solution is quite complicated, and does not seem to throw any light on the special case we are looking at here (where the interval is $(0, \infty)$).

It will be clear from the proof that it can be generalised to any equation of the form

$$x^N + \sum_{k=0}^{N-1} \binom{N}{k} (-a_k)^{N-k} x^k = 0\,,$$

where the numbers a_k are distinct and positive, which has positive distinct roots.

The question looks difficult, but you don't have to go very far before you come across something to substitute into the arithmetic/geometric inequality. Watch out for the condition in the arithmetic/geometric inequality that the numbers are distinct: you will have to show that any numbers (or algebraic expressions) you use for the inequality are distinct.

To obtain the relationship between the coefficients and the roots, you need to write the quartic equation in the form $(x-p)(x-q)(x-r)(x-s) = 0\,$.

Solution to question 21

(i) First write

$$\mathrm{f}(x) \equiv (x-p)(x-q)(x-r)(x-s)$$

where p, q, r and s are the four roots of the equation (known to be real, positive and distinct). Multiplying out the brackets and comparing with $x^4 - 4ax^3 + 6b^2x^2 - 4c^3x + d^4$ shows that $pqrs = d^4$ and $pqr + qrs + rsp + spq = 4c^3$.

The required result, $c > d$, follows immediately by applying the arithmetic/geometric mean inequality to the positive real numbers pqr, qrs, rsp and spq:

$$c^3 = \frac{pqr + qrs + rsp + spq}{4} > [(pqr)(qrs)(rsp)(spq)]^{1/4} = [p^3q^3r^3s^3]^{1/4} = [d^{12}]^{1/4} .$$

Taking the cube root (c and d are real) preserves the inequality.

The arithmetic/geometric mean inequality at the beginning of the question is stated only for the case when the numbers are distinct (though in fact it holds provided at least two of the numbers are distinct). To use the inequality, we have therefore to show that no two of pqr, qrs, rsp and spq are equal. This follows immediately from the fact that the roots are distinct and non-zero. For if, for example, $pqr = qrs$ then $qr(p-s) = 0$ which means that $q = 0$, $r = 0$ or $p = s$ any of which is a contradiction.

(ii) The polynomial $\mathrm{f}'(x)$ is cubic so it has three zeros (roots). These are at the turning points of $\mathrm{f}(x)$ (which lie between the zeros of $\mathrm{f}(x)$) and are therefore distinct and positive.

Now

$$\mathrm{f}'(x) = 4x^3 - 12ax^2 + 12b^2x - 4c^3$$

At the turning points,

$$x^3 - 3ax^2 + 3b^2x - c^3 = 0 .$$

Suppose that the roots of this cubic equation are u, v and w all real, distinct and positive. Then comparing $x^3 - 3ax^2 + 3b^2x - c^3 = 0$ and $(x-u)(x-v)(x-w)$ shows that $c^3 = uvw$ and $vw + wu + uv = 3b^2$.

The required result, $b > c$, follows immediately by applying the arithmetic/geometric mean inequality to the positive real numbers vw, wu and uv.

(iii) Apply similar arguments to $\mathrm{f}''(x)/12$.

Question 22(∗∗)

The nth Fermat number, F_n, is defined by

$$F_n = 2^{2^n} + 1\,, \qquad n = 0,\ 1,\ 2,\ \ldots\ ,$$

where 2^{2^n} means 2 raised to the power 2^n. Calculate F_0, F_1, F_2 and F_3. Show that, for $k = 1$, $k = 2$ and $k = 3$,

$$F_0F_1\ldots F_{k-1} = F_k - 2\,. \qquad (*)$$

Prove, by induction, or otherwise, that $(*)$ holds for all $k \geqslant 1$. Deduce that no two Fermat numbers have a common factor greater than 1.

Hence show that there are infinitely many prime numbers.

Comments

Fermat (1601—1665) conjectured that every number of the form $2^{2^n} + 1$ is a prime number. F_0 to F_4 are indeed prime, but Euler showed in 1732 that F_5 (4294967297) is divisible by 641. As can be seen, Fermat numbers get very big and not many more have been investigated; but those that have been investigated have been found not to be prime numbers. It is now conjectured that in fact only a finite number of Fermat numbers are prime numbers.

The Fermat numbers have a geometrical significance as well: Gauss proved that a regular polygon of n sides can be inscribed in a circle using a Euclidean construction (i.e. only a straight edge and a compass) if and only if n is a power of 2 times a product of distinct Fermat primes.

This rather nice proof that there are infinitely many prime numbers comes from *Proofs from the Book*[11], a set of proofs thought by Paul Erdös to be heaven-sent. (Most of them are much less elementary this one.) Erdös was an extraordinarily prolific mathematician. He had almost no personal belongings and no home[12]. He spent his life visiting other mathematicians and proving theorems with them. He collaborated with so many people that every mathematician is (jokingly) assigned an *Erdös number*; for example, if you wrote a paper with someone who wrote a paper with Erdös, you are given the Erdös number 2. It is conjectured that no mathematician has an Erdös number greater than 7. Mine is 8, but that is because I am more a theoretical physicist than a mathematician.[13]

[11] by M. Aigler and G.M. Ziegler, published by Springer, 1999.

[12] See *The man who loved only numbers* by Paul Hoffman published by Little Brown and Company in 1999 – a wonderfully readable and interesting book.

[13] Since I wrote this, I discovered 'collaborator distance' on http://www.ams.org/mathscinet, which gives the distance between any two researchers. I found that I am not after all counterexample to the conjecture: my Erdös is 5 (via Einstein). Sorry to bore you with all this personal stuff.

Solution to question 22

To begin with, we have by direct calculation that $F_0 = 3$, $F_1 = 5$, $F_2 = 17$ and $F_3 = 257$, so it is easy to verify the displayed result for $k = 1, 2, 3$.

For the induction, we start by assuming that the result holds for $k = m$ so that

$$F_0F_1 \dots F_{m-1} = F_m - 2 \ .$$

We need to show that this implies that that the result holds for $k = m + 1$, i.e. that

$$F_0F_1 \dots F_m = F_{m+1} - 2 \ .$$

Starting with the left hand side of this equation, we have

$$F_0F_1 \dots F_{m-1}F_m = (F_m - 2)F_m = (2^{2^m} - 1)(2^{2^m} + 1) = \left(2^{2^m}\right)^2 - 1 = 2^{2^{m+1}} - 1 = F_{m+1} - 2,$$

as required. Since we know that the result holds for $k = 1$, the induction is complete.

For the rest of the question: if p divides F_l and F_m, where $l < m$, then p divides $F_0F_1 \dots F_l \dots F_{m-1}$ and therefore p divides $F_m - 2$, by $(*)$. But all Fermat numbers are odd so no number other than 1 divides both $F_m - 2$ and F_m, which gives the contradiction. Hence no two Fermat numbers have a common factor greater than 1.

For the last part, note that every number can be written uniquely as a product of prime factors and that because the Fermat numbers are co-prime, each prime can appear in at most one Fermat number. Thus, since there are infinitely many Fermat numbers, there must be infinitely many primes.

Postmortem

This question is largely about proof. The above solution has a proof by induction and a proof by contradiction. The last part could have been written as a proof by contradiction too.

The difficulty is in presenting the proof. It is not enough to understand your own proof: you have to be able to set it out so that it is clear to the reader. This is a vital part of mathematics and its most useful transferable skill.

Question 23$(**)$

Show that the sum S_N of the first N terms of the series

$$\frac{1}{1.2.3}+\frac{3}{2.3.4}+\frac{5}{3.4.5}+\cdots+\frac{2n-1}{n(n+1)(n+2)}+\cdots$$

is

$$\frac{1}{2}\left(\frac{3}{2}+\frac{1}{N+1}-\frac{5}{N+2}\right).$$

What is the limit of S_N as $N \to \infty$?

The numbers a_n are such that

$$\frac{a_n}{a_{n-1}}=\frac{(n-1)(2n-1)}{(n+2)(2n-3)}.$$

Find an expression for $\dfrac{a_n}{a_1}$ and hence, or otherwise, evaluate $\sum\limits_{n=1}^{\infty} a_n$ when $a_1 = \dfrac{2}{9}$.

Comments

If you haven't the faintest idea how to do the sum, then look at the first line of the solution; but don't do this without first having a long think about it, and a long look at the form of the answer given in the question.

Limits form an important part of first year university mathematics. The definition of a limit is one of the basic ideas in *analysis* which is the rigorous study of calculus. Here no definition is needed: you just see what happens when N gets very large (some terms get very small and eventually go away).

The second part of the question looks as if it might be some new idea. Since this is STEP, you will probably realise that the series in the second part must be closely related to the series in the first part. The peculiar choice for a_1 $(= 2/9)$ should make you suspect that the sum will come out to some nice round number (not in fact round in this case, but straight and thin).

Having decided how to do the first part, please don't use the 'cover up' rule unless you understand why it works: mathematics at this level is not a matter of applying learned recipes.

Solution to question 23

The first thing to do is to put the general term of the series into partial fractions in the hope that something good might then happen. (The form of the answer suggests partial fractions have been used; and it is difficult to think of anything else to do.)

The general term of the series is

$$\frac{2n-1}{n(n+1)(n+2)} \equiv \frac{A}{n} + \frac{B}{n+1} + \frac{C}{n+2} \,.$$

(The 'equivalent' sign, $\equiv$, indicates an identity (something that holds for all values of n) rather than an equation to solve for n. The easiest way to deal with this is to multiply both sides of the identity by one of the terms in the denominator and then set this term equal to zero (since the identity must hold for all values of n). For example, multiplying by n gives

$$\frac{2n-1}{(n+1)(n+2)} \equiv A + \frac{Bn}{n+1} + \frac{Cn}{n+2} \,.$$

and setting $n = 0$ shows immediately that $A = -\frac{1}{2}$. Similarly, multiplying by $n+1$ and $n+2$ respectively and setting $n = -1$ and $n = -2$ respectively gives $B = 3$ and $C = -\frac{5}{2}$. Thus

$$\frac{2n-1}{n(n+1)(n+2)} = -\frac{1/2}{n} + \frac{3}{n+1} - \frac{5/2}{n+2}$$

The series can now be written

$$\left(-\frac{1/2}{1} + \frac{3}{2} - \frac{5/2}{3}\right) + \left(-\frac{1/2}{2} + \frac{3}{3} - \frac{5/2}{4}\right) + \left(-\frac{1/2}{3} + \frac{3}{4} - \frac{5/2}{5}\right) +$$
$$+ \cdots + \left(-\frac{1/2}{N-2} + \frac{3}{N-1} - \frac{5/2}{N}\right) + \left(-\frac{1/2}{N-1} + \frac{3}{N} - \frac{5/2}{N+1}\right) + \left(-\frac{1/2}{N} + \frac{3}{N+1} - \frac{5/2}{N+2}\right).$$

The first term has 3 as the last number in the denominator so the Nth term has $N+2$. All the terms in the series cancel, except those with denominators 1, 2, $N+1$ and $N+2$. The sum of these exceptions is the sum of the series.

The limit as $N \to \infty$ is $\frac{3}{4}$ since the other two terms obviously tend to zero.

For the second paragraph, we have $\dfrac{a_n}{a_{n-1}} = \dfrac{b_n}{b_{n-1}}$, where b_n is the general term of the series in the first part of the question. Thus

$$\frac{a_n}{a_1} = \frac{b_n}{b_1} \qquad \text{and} \qquad \sum_1^\infty a_n = \frac{a_1}{b_1}\sum_1^\infty b_n = \frac{2/9}{1/6} \times \frac{3}{4} = 1.$$

Alternatively,

$$\begin{aligned}
a_n &= \frac{(n-1)(2n-1)}{(n+2)(2n-3)} a_{n-1} = \frac{(n-1)(2n-1)}{(n+2)(2n-3)} \frac{(n-2)(2n-3)}{(n+1)(2n-5)} a_{n-2} \\
&= \frac{(n-1)(2n-1)}{(n+2)(2n-3)} \frac{(n-2)(2n-3)}{(n+1)(2n-5)} \frac{(n-3)(2n-5)}{(n)(2n-7)} \cdots \frac{5\times 11}{8\times 9} \frac{4\times 9}{7\times 7} \frac{3\times 7}{6\times 5} \frac{2\times 5}{5\times 3} \frac{1\times 3}{4\times 1} a_1 \\
&= \frac{2n-1}{n(n+1)(n+2)} \frac{3\times 2\times 1}{1} a_1 = \frac{12}{9} \frac{2n-1}{n(n+1)(n+2)},
\end{aligned}$$

all other terms cancelling.

Now using the result of the first part gives $\sum\limits_1^\infty a_n = 1$.

Question 24$(***)$

(i) Show that, if m is an integer such that

$$(m-3)^3 + m^3 = (m+3)^3, \qquad (*)$$

then m^2 is a factor of 54. Deduce that there is no integer m which satisfies the equation $(*)$.

(ii) Show that, if n is an integer such that

$$(n-6)^3 + n^3 = (n+6)^3, \qquad (**)$$

then n is even. By writing $n = 2m$ deduce that there is no integer n which satisfies the equation $(**)$.

(iii) Show that, if n is an integer such that

$$(n-a)^3 + n^3 = (n+a)^3\ , \qquad (***)$$

where a is a non-zero integer, then n is even and a is even. Deduce that there is no integer n which satisfies the equation $(***)$.

Comments

I slightly simplified the first two parts of the question, which comprised the whole of the original STEP question, and added the third part. This last part is conceptually tricky and very interesting: hence the three star rating.

It is an example of a method of proof by contradiction, called the *method of infinite descent*. You will understand why it is called infinite descent when you have finished the question (in fact, the name may serve as a hint). It was used by Fermat (1601 – 1665) to prove special cases of his Last Theorem[14], which of course is exactly what you are doing in this question. The method was probably invented by him and his faith in it sometimes led him astray. It is even possible that he thought he could use it to prove his Last Theorem in full. In fact, the proof of this theorem was only given in 1994 by Andrew Wiles; and it is 150 pages long.

[14]The theorem says that the equation $x^n + y^n = z^n$, where x, y, z and n are positive integers, can only hold if $n = 1$ or 2

Solution to question 24

(i) Simplifying gives $m^2(m-18)=54$.

Both m^2 and $(m-18)$ must divide 54, which is impossible since the only squares that divide 54 are 1 and 9, and neither $m=1$ nor $m=3$ satisfies $m^2(m-18)=54$.

You could also argue that $m-18$ must be positive so $m \geqslant 19$ and $m^2 \geqslant 361$ which is a contradiction.

(ii) Suppose n is odd. The two terms on the left hand side are both odd, which means that that the left hand side is even. But the right hand side is odd so the equation cannot balance.

Setting $n=2m$ gives

$$(2m-6)^3+(2m)^3=(2m+6)^3.$$

Taking a factor of 2^3 out of each term leaves $(m-3)^3+m^3=(m+3)^3$ which is the same as the equation that was shown in part (i) to have no solutions.

(iii) First we show that n is even, dealing with the cases a odd and a even separately. The first case is n odd and a odd. In that case, $(n-a)^3$ is even, n^3 is odd and $(n+a)^3$ is even, so the equation does not balance. In the second case, n is odd and a is even, the three terms are all odd and again the equation does not balance. Therefore n cannot be odd.

Next, we investigate the case n is even and a odd. This time the three terms are odd, even and odd respectively so there is no contradiction. But multiplying out the brackets and simplifying gives

$$n^2(n-6a)=2a^3\ .$$

Since n is even the left hand side is divisible by 4, because of the factor n^2. That means that a^3 is divisible by 2 and hence a is even.

Now that we know n and a are both even, we can follow the method used in part (ii) and set $n=2m$ and $a=2b$. This gives

$$(2m-2b)^3+(2m)^3=(2m+2b)^3$$

from which a factor of 2^3 can be cancelled from each term. Thus m and b satisfy the same equation as n and a. They are therefore both even and we can repeat the process.

Repeating the process again and again will eventually result in an integer that is odd which will therefore not satisfy the equation that it is supposed to satisfy: a contradiction. There is therefore no integer n that satisfies $(***)$.

Question 25$(***)$

The function f satisfies $0 \leqslant \mathrm{f}(t) \leqslant K$ when $0 \leqslant t \leqslant 1$. Explain by means of a sketch, or otherwise, why

$$0 \leqslant \int_0^1 \mathrm{f}(t)\,\mathrm{d}t \leqslant K\ .$$

By considering $\displaystyle\int_0^1 \frac{t}{n(n-t)}\,\mathrm{d}t$, or otherwise, show that, if $n > 1$,

$$0 \leqslant \ln\left(\frac{n}{n-1}\right) - \frac{1}{n} \leqslant \frac{1}{n-1} - \frac{1}{n}$$

and deduce that

$$0 \leqslant \ln N - \sum_{n=2}^{N} \frac{1}{n} \leqslant 1\ .$$

Deduce that $\displaystyle\sum_{n=1}^{N} \frac{1}{n} \to \infty$ as $N \to \infty$.

Noting that $2^{10} = 1024$, show also that if $N < 10^{30}$ then $\displaystyle\sum_{n=1}^{N} \frac{1}{n} < 101$.

Comments

Quite a lot of different ideas are required for this question; hence three stars. The first hurdle is to decide what the constant K in the first part has to be to make the second part work. You might try to maximise the integrand using calculus, but that would be the wrong thing to do (you should find if you do this that the integrand has no turning point in the given range: it increases throughout the range. The next hurdle is the third displayed equation, which follows from the preceding result. Then there is still more work to do.

It is not at all obvious at first sight that the series $\displaystyle\sum_1^N \frac{1}{n}$ (which is called the *harmonic series* tends to infinity as N increases, though there are easier ways of proving this. This way tells us two very interesting things; First it tells us that the sum increases very slowly indeed: the first 10^{30} terms only get to 100. Second it tells us that

$$0 \leqslant \sum_1^N \frac{1}{n} - \ln N \leqslant 1$$

for all N. (This is the third displayed equation in the question slightly rewritten.) Not only does the sum diverge, it does so logarithmically. In fact, in the limit $N \to \infty$,

$$\sum_1^N \frac{1}{n} - \ln N \to \gamma$$

where γ is *Euler's constant*. Its value is about 1/2.

Solution to question 25

Any sketch showing a squiggly curve all of which lies beneath the line $x = K$ and above the x-axis will do (area under curve is less than the area of the rectangle).

Evaluate the integral, noting that $t/(n-t) = n/(n-t) - 1$:

$$\int_0^1 \frac{t}{n(n-t)}\mathrm{d}t = \int_0^1 \left(\frac{1}{n-t} - \frac{1}{n}\right)\mathrm{d}t = \ln\left(\frac{n}{n-1}\right) - \frac{1}{n}\,.$$

Now note that the largest value of the integrand in the interval $[0,1]$ occurs at $t = 1$, since the numerator increases as t increases and the denominator decreases as t increases (remember that $n > 1$) so

$$0 \leqslant \ln\left(\frac{n}{n-1}\right) - \frac{1}{n} \leqslant \frac{1}{n(n-1)} = \frac{1}{n-1} - \frac{1}{n}\,.$$

Summing both sides from 2 to N and cancelling lots of terms in pairs gives

$$0 \leqslant \ln N - \sum_2^N \frac{1}{n} \leqslant 1 - \frac{1}{N}.$$

Note that $1-1/N < 1$. Since $\ln N \to \infty$ as $N \to \infty$, so also must $\displaystyle\sum_2^N \frac{1}{n}$ (it differs from the logarithm by less than 1).

Finally, rearranging the inequality gives $\displaystyle\sum_2^N \frac{1}{n} \leqslant \ln N$, i.e. $\displaystyle\sum_1^N \frac{1}{n} \leqslant \ln N + 1$. Setting $N = 10^{30}$, and using the identity given and also the inequality $\mathrm{e} > 2$ gives:

$$\sum_1^{10^{30}} \frac{1}{n} \leqslant \ln(10^{30}) + 1 < \ln(1024^{10}) + 1 = 100\ln 2 + 1 < 100 \ln \mathrm{e} + 1 = 100 + 1\,.$$

Postmortem

This question seems very daunting at first, because you are asked to prove a sequence of completely unfamiliar results. However, you should learn from this question that if you keep cool and follow the implicit hints, you can achieve some surprisingly sophisticated results. You might think that this situation is very artificial: in real life, you do not receive hints to guide you. But often in mathematical research, hints *are* buried deep in the problem if only you can recognise them.

Question 26$(***)$

Find $\dfrac{\mathrm{d}y}{\mathrm{d}x}$ if

$$y = \frac{ax+b}{cx+d}\,. \qquad (*)$$

By using changes of variable of the form $(*)$, or otherwise, show that

$$\int_0^1 \frac{1}{(x+3)^2}\,\ln\left(\frac{x+1}{x+3}\right)\mathrm{d}x = \frac{1}{6}\ln 3 - \frac{1}{4}\ln 2 - \frac{1}{12}\,,$$

and evaluate the integrals

$$\int_0^1 \frac{1}{(x+3)^2}\,\ln\left(\frac{x^2+3x+2}{(x+3)^2}\right)\mathrm{d}x \quad \text{and} \quad \int_0^1 \frac{1}{(x+3)^2}\,\ln\left(\frac{x+1}{x+2}\right)\mathrm{d}x\,.$$

By changing to the variable y defined by

$$y = \frac{2x-3}{x+1}\,,$$

evaluate the integral

$$\int_2^4 \frac{2x-3}{(x+1)^3}\,\ln\left(\frac{2x-3}{x+1}\right)\mathrm{d}x\,.$$

Evaluate the integral

$$\int_9^{25} \left(2z^{-3/2} - 5z^{-2}\right)\ln\left(2 - 5z^{-1/2}\right)\,\mathrm{d}z\,.$$

Comments

If you are thinking that this question looks extremely long, you are right. It is two questions — in fact, two versions of one question — separated by the horizontal line. Beneath the line is the first draft of a STEP question; above the line is the question as it was actually set. I thought you would be interested to see the evolution of a question. Note in particular the way that the ideas in the final draft are closely knit and better structured: the first change of variable is strongly signalled, but the remaining two parts, though based on the same idea, require increasing ingenuity. The first draft just required two unrelated changes of variable. Note also that the integral in the last part of the first draft has an unpleasant contrived appearance, whereas the the integrals of the final draft are rather pleasing: beauty matters to mathematicians.

A change of variable of the form $(*)$ is called a *linear fractional* or *Möbius* transformation. It is of great importance in the theory of the geometry of the complex plane.

Solution to question 26

Differentiating gives $\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{a(cx+d)-c(ax+b)}{(cx+d)^2} = \dfrac{ad-bc}{(cx+d)^2}$.

For the first integral, set $y = \dfrac{x+1}{x+3}$:

$$\int_0^1 \frac{1}{(x+3)^2}\ln\left(\frac{x+1}{x+3}\right)\mathrm{d}x = \frac{1}{2}\int_{1/3}^{1/2}\ln y\;dy = \frac{1}{2}\Big[y\ln y - y\Big]_{1/3}^{1/2},$$

which gives the required answer. (The integral of $\ln y$ is a standard integral; it can be done by parts, first substituting $z = \ln y$, if you like.)

The second integral can be expressed as the sum of two integrals of the same form as the first integral, since

$$\ln\left(\frac{x^2+3x+2}{(x+3)^2}\right) = \ln\left(\frac{(x+1)(x+2)}{(x+3)^2}\right) = \ln\left(\frac{x+1}{x+3}\right) + \ln\left(\frac{x+2}{x+3}\right) .$$

We have already done the first of these integrals. Using the substitution $y = \dfrac{x+2}{x+3}$ in the second of these integrals gives

$$\int_{2/3}^{3/4}\ln y\;\mathrm{d}y = \frac{17}{12}\ln 3 - \frac{13}{6}\ln 2 - \frac{1}{12} .$$

The required integral is therefore

$$\frac{19}{12}\ln 3 - \frac{29}{12}\ln 2 - \frac{1}{6} .$$

For the next integral, note that

$$\ln\left(\frac{x+1}{x+2}\right) = \ln\left(\frac{x+1}{x+3}\right) - \ln\left(\frac{x+2}{x+3}\right) ,$$

so the required integral is the difference of the two integrals that we summed in the previous part, i.e.

$$-\frac{5}{4}\ln 3 + \frac{23}{12}\ln 2 .$$

The change of variable gives

$$\frac{1}{5}\int_{1/3}^{1} y\ln y\;\mathrm{d}y$$

which can be integrated by parts. The result is $\dfrac{1}{90}\ln 3 - \dfrac{2}{45}$.

For the last part, you have to guess the substitution. There are plenty of clues, but the most obvious place to start is the log. The argument of the log can be written as

$$\frac{2(x+1)-5}{x+1} = 2 - \frac{5}{x+1}$$

so it looks as if we should take $z = (x+1)^2$. Making this transformation gives the previous integral almost exactly (note especially the the limits transform as they should). The only difference is a factor of 2 which comes from $\dfrac{\mathrm{d}z}{\mathrm{d}x}$ so the answer is twice the previous answer.

Question 27(∗)

The curve C has equation

$$y = \frac{x}{\sqrt{(x^2 - 2x + a)}}\,,$$

where the square root is positive. Show that, if $a > 1$, then C has exactly one stationary point.

Sketch C when (i) $a = 2$ and (ii) $a = 1$.

Comments

There is something you must know in order to do this question, and that is the definition of $\sqrt{(x^2 - 2x + a)}$. You might think that this is ambiguous, because the square root could be positive of negative, but by convention it means the positive square root. Thus $\sqrt{(x^2)} = |x|$ (not x).

For the sketches, you just need the position of any stationary points, any other points that the graph passes through, behaviour as $x \to \pm\infty$ and any vertical asymptotes. It should not be necessary in such a simple case to establish the nature of the stationary point(s).

You should at some stage think about the condition given on a. In fact, I doubt if many of you will want to leave it there. It is clear that the examiner would really have liked you to do is to sketch C in the different cases that arise according to the value of a but was told that this would be to long and/or difficult for the exam; I am sure that this is what you will do (or maybe what you have already done by the time you read this).

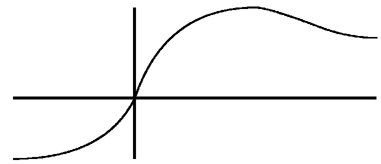
Please do not use your calculator for these sketches (except perhaps to check your answers): there is no point — the graph will either be wrong, in which case you will be confused, or it will be right and the value that you might have gained from thinking about the question yourself will be lost to you for ever.

Solution to question 27

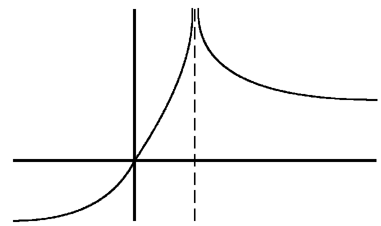
First differentiate to find the stationary points:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{(x^2 - 2x + a) - x(x-1)}{(x^2 - 2x + a)^{3/2}} = \frac{a - x}{(x^2 - 2x + a)^{3/2}}$$

which is only zero when $x = a$. [15]



(i) When $a = 2$, the stationary point is at $(2, \sqrt{2})$ (which can be seen to be a maximum by evaluating the second derivative at $x = a$, though this is not necessary). The curve passes through $(0, 0)$. As $x \to \infty$, $y \to 1$; as $x \to -\infty$, $y \to -1$.



(ii) When $a = 1$, we can rewrite y as $\dfrac{x}{|x-1|}$. The modulus signs arise because the square root is taken to be positive (or zero — but that can be discounted here since it is in the denominator). The graph is as before, except that the maximum point has been stretched into a vertical asymptote (like a volcanic plug) at $x = 1$.

Postmortem

The first thing to do after finishing the question is to try to understand the condition on a. Part (ii) above is the key. Writing $\mathrm{f}(x) = x^2 - 2x + a = (x-1)^2 + a - 1$ we see that $\mathrm{f}(x) > 0$ for all x if $a > 1$ but if $a < 1$ there are values of x for which $\mathrm{f}(x) < 0$. is negative. The borderline value is $a = 1$, for which $\mathrm{f}(x) \geqslant 0$ and $\mathrm{f}(1) = 0$.

If $a < 1$, we find $\mathrm{f}(x) = 0$ when $x = 1 \pm \sqrt{(1-a)}$ and $\mathrm{f}(x) < 0$ between these two values of x. The significance of this is that the square root in the denominator of the function we are trying to sketch is imaginary, so there is a gap in the graph between $x = 1 - \sqrt{(1-a)}$ and $x = 1 + \sqrt{(1-a)}$ and there are vertical asymptotes at these values. The point $x = a$ lies between those values for $a < 1$.

Is that the whole story? Well, no. We now have to decide whether the two significant points of our graph, namely $x = 0$ where the curve crosses the axis and $x = a$ where the curve has a stationary point, lie inside or outside the forbidden zone $1 - \sqrt{(1-a)} \leqslant x \leqslant 1 + \sqrt{(1-a)}$. Since $\mathrm{f}(0) < 0$ if $a < 0$ and $\mathrm{f}(a) < 0$ if $0 < a < 1$ we see that a has another critical value besides $a = 1$; if $a < 0$ the graph is going to look qualitatively different.

[15] My drawing package is absolutely useless, so I'm not going to attempt to improve these sketches, or label them fully (which you should certainly do). Maybe in the next draft ...

Question 28$(***)$

Prove that

$$\sum_{k=0}^{n} \sin k\theta = \frac{\cos\frac{1}{2}\theta - \cos(n+\frac{1}{2})\theta}{2\sin\frac{1}{2}\theta}\,. \qquad (*)$$

(i) Deduce that, for $n \geqslant 1$,

$$\sum_{k=0}^{n} \sin\left(\frac{k\pi}{n}\right) = \cot\left(\frac{\pi}{2n}\right)\,.$$

(ii) By differentiating $(*)$ with respect to θ, or otherwise, show that, for $n \geqslant 1$,

$$\sum_{k=0}^{n} k\sin^2\left(\frac{k\pi}{2n}\right) = \frac{(n+1)^2}{4} + \frac{1}{4}\cot^2\left(\frac{\pi}{2n}\right)\,.$$

Comments

The very first part can be done by multiplying both sides of $(*)$ by $\sin\frac{1}{2}\theta$ and using the identity

$$2\sin A\sin B = \cos(B-A) - \cos(B+A)\,.$$

It can also be done by induction (worth a try even if you do it by the above method) or by considering the imaginary part of $\sum \exp(ik\theta)$ (summing this as a geometric progression), if you know about de Moivre's theorem.

Before setting pen to paper for part (ii), it pays to think very hard about simplifications of $(*)$ that might make the differentiation more tractable – perhaps bearing in mind the calculations involved with part (i).

In the original question, you were asked to show that for large n

$$\sum_{k=0}^{n} \sin\left(\frac{k\pi}{n}\right) \approx \frac{2n}{\pi}$$

and

$$\sum_{k=0}^{n} k\sin^2\left(\frac{k\pi}{2n}\right) \approx \left(\frac{1}{4} + \frac{1}{\pi^2}\right)n^2\,,$$

using the approximations, valid for small θ, $\sin\theta \approx \theta$ and $\cos\theta \approx 1 - \frac{1}{2}\theta^2$. Of course, these follow quickly from the exact results; but if you only want approximate results you can save yourself a bit of work by approximating early to avoid doing some of the trigonometric calculations. I thought that the exact result was nicer than the approximate result (though approximations are an important part of mathematics).

Solution to question 28

For the first part, we will show that

$$\sum_{k=0}^{n} 2\sin k\theta \sin \tfrac{1}{2}\theta = \cos \tfrac{1}{2}\theta - \cos(n+\tfrac{1}{2})\theta \,.$$

Starting with the left hand side, we have

$$\sum_{k=0}^{n} 2\sin k\theta \sin \tfrac{1}{2}\theta = \sum_{k=0}^{n} \left[-\cos(k+\tfrac{1}{2})\theta + \cos(k-\tfrac{1}{2})\theta \right]$$

which gives the result immediately, since almost all the terms in the sum cancel. (Write our a few terms if you are not certain of this.)

(i) Let $\theta = \pi/n$. Then

$$\sum_{k=0}^{n} \sin\left(\frac{k\pi}{n}\right) = \frac{\cos\frac{1}{2}(\pi/n) - \cos(\pi + \frac{1}{2}(\pi/n))}{2\sin\frac{1}{2}(\frac{\pi}{n})} = \frac{\cos\frac{1}{2}(\pi/n) + \cos\frac{1}{2}(\pi/n))}{2\sin\frac{1}{2}(\pi/n)} = \cot\tfrac{1}{2}(\pi/n)$$

as required.

(ii) Differentiating the left hand side of $(*)$ and using a double angle trig. formula gives

$$\sum_{k=0}^{n} k\cos k\theta = \sum_{k=0}^{n} k\left(1 - 2\sin^2\tfrac{1}{2}k\theta\right) = \tfrac{1}{2}n(n+1) - 2\sum_{k=0}^{n} k\sin^2\tfrac{1}{2}k\theta \,. \qquad (\dagger)$$

Before attempting to differentiate the right hand side of $(*)$ it is a good idea to write it in a form that gets rid of some of the fractions. Omitting for the moment the factor 1/2, we have

$$\text{RHS of } (*) \;= (1-\cos n\theta)\cot\tfrac{1}{2}\theta + \sin n\theta$$

and differentiating gives $-\frac{1}{2}(1-\cos n\theta)\operatorname{cosec}^2\frac{1}{2}\theta + n\sin n\theta\cot\frac{1}{2}\theta + n\cos n\theta$. Now setting $\theta = \pi/n$ gives

$$-\operatorname{cosec}^2\tfrac{1}{2}(\pi/n) - n \,. \qquad (\ddagger)$$

Setting $\theta = \pi/n$ in equation $(\dagger)$ and comparing with $(\ddagger)$ (remembering that there is a factor of 1/2 missing) gives the required result.

Postmortem

Now that I have had another go at this question it does not seem terribly interesting. At first, I thought I might ditch it. Then I thought that it perhaps was worthwhile: keeping cool under the pressure of the differentiation for part (ii) – it should just be a few careful lines – and getting it out is a good confidence booster. Anyway, now that I have slogged[16] through it, I am going to leave it in.

(Next day) I recall now that the reason that I included this question in the first place was because in its original form (with approximate answers), the result for part (ii) can be obtained very quickly from part (i) by differentiating with respect to π. The snag with this is that it is not obviously going to work, because the derivatives of terms such as $\sin\pi$ that are omitted from (i) might feature in (ii). It is disappointing but not surprising that this trick cannot be used to obtain the exact formulae.

[16] I should come clean at this point: I just differentiated $(*)$ on autopilot, without considering whether I could simplify the task by rewriting the formula as suggested in my hint on the previous page. It was hard work. Serves me right for not following my own general advice.

Question 29$(***)$

Consider the cubic equation

$$x^3 - px^2 + qx - r = 0\ ,$$

where $p \neq 0$ and $r \neq 0$.

(i) If the three roots can be written in the form ak^{-1}, a and ak for some constants a and k, show that one root is q/p and that $q^3 - rp^3 = 0$.

(ii) If $r = q^3/p^3$, show that q/p is a root and that the product of the other two roots is $(q/p)^2$. Deduce that the roots are in geometric progression.

(iii) Find a necessary and sufficient condition involving p, q and r for the roots to be in arithmetic progression.

Comments

The Fundamental Theorem of Algebra says that a polynomial of degree n can be written as the product of n linear factors, so we can write

$$x^3 - px^2 + qx - r = (x-\alpha)(x-\beta)(x-\gamma)\,, \qquad (*)$$

where α, β and γ are the roots of the equation $x^3 - px^2 + qx - r = 0$. The basis of this question is the comparison between the left hand side of $(*)$ and the right hand side, multiplied out, of $(*)$. Some of the roots may not be real, but you don't have to worry about that here.

The 'necessary and sufficient' in part (iii) looks a bit forbidding, but if you just repeat the steps of parts (i) and (ii), it is straightforward.

Solution to question 29

(i) We have

$$(x - ak^{-1})(x-a)(x-ak) \equiv x^3 - px^2 + qx - r$$

i.e.

$$x^3 - a(k^{-1}+1+k)x^2 + a^2(k^{-1}+1+k)x - a^3 = x^3 - px^2 + qx - r\ .$$

Thus $p = a(k^{-1}+1+k)$, $q = a^2(k^{-1}+1+k)$, and $r = a^3$. Dividing gives $q/p = a$, which is a root, and $q^3/p^3 = a^3 = r$ as required.

(ii) Set $r = q^3/p^3$. Substituting $x = q/p$ into the cubic shows that it is a root:

$$(q/p)^3 - p(q/p)^2 + q(q/p) - (q/p)^3 = 0\ .$$

Since q/p is one root, and the product of the three roots is q^3/p^3 ($= r$ in the original equation), the product of the other two roots must be q^2/p^2. The two roots can therefore be written in the form $k^{-1}(q/p)$ and $k(q/p)$ for some number k , which shows that they are in geometric progression.

(iii) The roots are in arithmetic progression if and only if they are of the form $(a-d)$, a and $(a+d)$.

If the roots are in this form, then the equation is

$$\big(x - (a-d)\big)\big(x-a\big)\big(x-(a+d)\big) \equiv x^3 - px^2 + qx - r$$

i.e.

$$x^3 - 3ax^2 + (3a^2-d^2)x - a(a^2-d^2) \equiv x^3 - px^2 + qx - r\ .$$

Thus $p = 3a$, $q = 3a^2 - d^2$, $r = a(a^2-d^2)$. A necessary condition is therefore $r = (p/3)(q - 2p^2/9)$. Note that one root is $p/3$.

Conversely, if $r = (p/3)(q-2p^2/9)$, the equation is

$$x^3 - px^2 + qx - (p/3)(q-2p^2/9) = 0.$$

We can verify that $p/3$ is a root by substitution. Since the roots sum to p and one of them is $p/3$, the others must be of the form $p/3 - d$ and $p/3 + d$ for some d. They are therefore in arithmetic progression.

A necessary and sufficient condition is therefore $r = (p/3)(q - 2p^2/9)$.

Postmortem

As usual, it is a good idea to give a bit of thought to the conditions given, namely $p \neq 0$ and $r \neq 0$.

Clearly, if the roots are in geometric progression, we cannot have a zero root. We therefore need $r \neq 0$. However, we don't need the condition $p \neq 0$. If the roots are in geometric progression with $p = 0$, then $q = 0$, but there is no contradiction: the roots are $be^{-2\pi i/3}$, b and $be^{2\pi i/3}$ where $b^3 = r$, and these are certainly in geometric progression. So the necessary and sufficient conditions are $r \neq 0$ and $p^3 r = q^3$.

Neither of these conditions is required if the roots are in arithmetic progression.

Therefore, the condition $p \neq 0$ is only there as a convenience — one thing less for you to worry about. The condition $r \neq 0$ should really have been given only for the first part. It should be said that this sort of question is incredibly difficult to word, which is why the examiner was a bit heavy handed with the conditions.

Question 30$(* * *)$

(i) Let $\mathrm{f}(x) = (1+x^2)e^x - k$, where k is a given constant. Show that $\mathrm{f}'(x) \geqslant 0$ and sketch the graph of $\mathrm{f}(x)$ in the cases $k < 0$, $0 < k < 1$ and $k > 1$.

Hence, or otherwise, show that the equation

$$(1+x^2)\mathrm{e}^x - k = 0\,,$$

has exactly one real root if $k > 0$ and no real roots if $k \leqslant 0$.

(ii) Determine the number of real roots of the equation

$$(\mathrm{e}^x - 1) - k\tan^{-1} x = 0$$

in the different cases that arise according to the value of the constant k.

Note: If $y = \tan^{-1} x$, then $\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{1}{1+x^2}$.

Comments

In good STEP style, this question has two related parts. In this case, the first part not only gives you guidance for doing the second part, but also provides a result that is useful for the second part.

In the original question you were asked to determine the number of real roots in the cases (a) $0 < k \leqslant 2/\pi$ and (b) $2/\pi < k < 1$. I thought that you would like to work out all the different cases for yourself — but it makes the question considerably harder, especially if you think about the special values of k as well as the ranges of k.

For part (i) you need to know that $x\mathrm{e}^x \to 0$ as $x \to -\infty$. This is just a special case of the result that exponentials go to zero faster than any power of x.

For part (ii), it is not really necessary to do all the sketches, but I think you should (though you perhaps wouldn't under examination conditions) because it gives you a complete understanding of the way the function depends on the parameter k. For the graphs, you may want to consider the sign of $\mathrm{f}'(0)$; this will give you an important clue. Remember that, by definition, $-\pi/2 < \tan^{-1} x < \pi/2$.

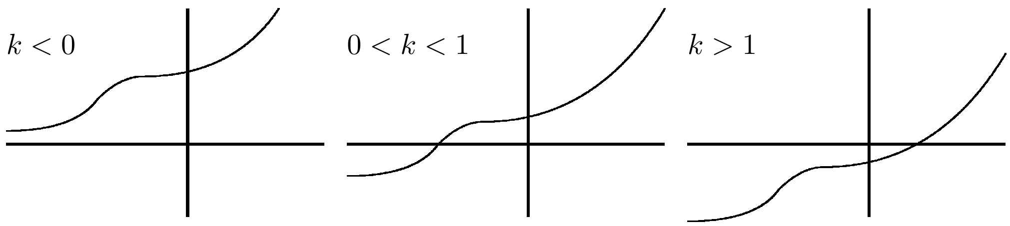
Solution to question 30

(i) Differentiating gives $f'(x) = 2xe^x + (1+x^2)e^x = (1+x)^2e^x$ which is non-negative (because the square is non-negative and the exponential is positive.

There is one stationary point, at $x = -1$, which is a point of inflection, since the gradient on either side is positive.

Also, $f(x) \to -k$ as $x \to -\infty$, $f(0) = 1 - k$ and $f(x) \to \infty$ as $x \to +\infty$. The graph is essentially exponential, with a hiccup at $x = -1$. The differences between the three cases are the positions of the horizontal asymptote ($x \to -\infty$) and the place where the graph cuts the y axis (above or below the x axis?).

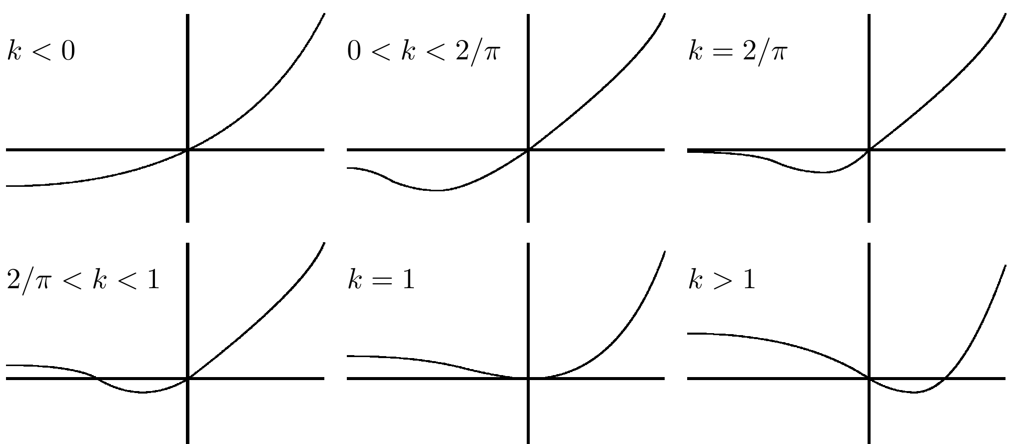
Since the graph is increasing and $0 < f(x) < \infty$, the given equation has one real root if $k > 0$ and no real roots if $k \leqslant 0$.



(ii) First we collect up information required to sketch the graph, as in part (i). We have

$f(0) = 0$, $\quad f(x) \to k\pi/2 - 1$ as $x \to -\infty$, $\quad f(x) \to \infty$ as $x \to \infty$, $\quad f'(x) = e^x - k(1+x^2)^{-1}$ (which we know from the first part can only be zero only if $k > 0$), $\quad f'(0) = 1 - k$.

The critical values of k are 0, $2/\pi$ and 1 so we will have to consider four cases.



As you see, the equation $(e^x - 1) - k\tan^{-1} x = 0$ has one root if $k \leqslant \pi/2$ and two roots otherwise. The case $k = 1$ can be thought of as having one root at $x = 0$ or as having two roots, both at $x = 0$ (it is a double root).

Question 31$(**)$

For each positive integer n, let

$$a_n = \frac{1}{n+1} + \frac{1}{(n+1)(n+2)} + \frac{1}{(n+1)(n+2)(n+3)} + \cdots ;$$
$$b_n = \frac{1}{n+1} + \frac{1}{(n+1)^2} + \frac{1}{(n+1)^3} + \cdots .$$

(i) Show that $b_n = 1/n$.

(ii) Deduce that $0 < a_n < 1/n$.

(iii) Show that $a_n = n!\mathrm{e} - [n!\mathrm{e}]$, where $[x]$ is the integer part of x.

(iv) Hence show that e is irrational.

Comments

Each part of this looks horrendously difficult, but it doesn't take much thought to see what is going on. If you haven't come across the concept of *integer part* some examples should make it clear: $[21.25] = 21$, $[\pi] = 3$, $[1.999999] = 1$ and $[2] = 2$.

For part (iii), you need to know the series for the number e, namely $\sum_{n=0}^{\infty} 1/n!$.

If you are stumped by the last part, just remember the definition of the word *irrational*: x is rational if and only if it can be expressed in the form p/q where p and q are integers; if x is not rational, it is irrational. Then look for a proof by contradiction ('Suppose that x is rational ...').

The proof that e is irrational was first given by Euler (1736). He also named the number e, though he didn't 'invent' it; e is the basis for natural logarithms and as such was used implicitly by John Napier in 1614. In general, it is not easy to show that numbers are irrational. Johann Lambert showed that π is irrational in 1760 but it is not known if $\pi + \mathrm{e}$ is irrational.

Solution to question 31

(i) Although the series for b_n does not at first sight look tractable, it is in fact just a geometric progression: the first term is $1/n+1$ and the common ratio is also $1/n+1$. Thus

$$b_n = \frac{1}{n+1}\left(\frac{1}{1-1/(n+1)}\right) = \frac{1}{n}\,.$$

(ii) Each term (after the first) of a_n is less than the corresponding term in b_n, so $a_n < b_n = 1/n$.

(iii) Mulitiplying the series for e by $n!$ gives

$$n!\mathrm{e} = n! + n! + n!/2 + \cdots + 1 + a_n$$

and the result follows because $a_n < 1$ and all the other terms on the right hand side of the above equation are integers.

(iv) We use proof by contradiction. Suppose that there exist integers k and m such that $\mathrm{e} = k/m$. Then $m!\mathrm{e}$ is certainly an integer. But if $m!\mathrm{e}$ is an integer then $[m!\mathrm{e}] = m!\mathrm{e}$, which contadicts the result of part (iii) since we know that $a_m \neq 0$ (it is obvious from the definition that $a_m > 0$).

Question 32$(***)$

To nine decimal places, $\log_{10} 2 = 0.301029996$ and $\log_{10} 3 = 0.477121255$.

(i) Calculate $\log_{10} 5$ and $\log_{10} 6$ to three decimal places. By taking logs, or otherwise, show that

$$5 \times 10^{47} < 3^{100} < 6 \times 10^{47}\,.$$

Hence write down the first digit of 3^{100}.

(ii) Find the first digit of each of the following numbers: 2^{1000}; $2^{10\,000}$; and $2^{100\,000}$.

Comments

This nice little question shows why it is a good idea to ban calculators from some mathematics examinations.

When I was at school, before calculators were invented, we had to spend quite a lot of time in year 8 (I think) doing extremely tedious calculations by logarithms. We were provided with a book of tables of four-figure logarithms and the book also had tables of values of trigonometric functions. When it came to antilogging, to get the answer, we had to use the tables backwards, since there were no tables of inverse logarithms (exponentials).

A logarithm consists of two parts: the *characteristic* which is the number before the decimal point and the *mantissa* which is the number after the decimal point. It is the mantissa that gives the significant figures of the number that has been logged; the characteristic tells you where to put the decimal point. This question is all about calculating mantissas.

The characteristic of a number greater than 1 is non-negative but the characteristic of a number less than 1 is negative. The rules for what to do in the case of a negative characteristic were rather complicated: you couldn't do ordinary arithmetic because the logarithm consisted of a negative characteristic and a positive mantissa. In ordinary arithmetic, the number -3.4 means $-3 - 0.4$ whereas the corresponding situation in logarithms, normally written $\bar{3}.4$, means $-3 + 0.4$. Instead of explaining this, the teacher gave a complicated set of rules, which just had to be learned — not the right way to do mathematics.

There are no negative characteristics in this question, I'm happy to say.

Solution to question 32

(i) For the very first part, we have

$\log 2 + \log 5 = \log 10 = 1$ so $\log 5 = 1 - 0.301029996 = 0.699$ to 3 d.p.

$\log 6 = \log 2 + \log 3 = 0.778$ to 3 d.p.

Taking logs preserves the inequalities, so we need to show that

$$47 + \log 5 < 100 \log 3 < 47 + \log 6$$

i.e. that

$$47 + 0.699 < 47.7121 < 47 + 0.778$$

which is true. The first digit of 3^{100} is therefore 5.

(ii) To find the first digit of these numbers, we use the method of part (i).

We have $\log 2^{1000} = 1000 \log 2 = 301.030 = 301 + 0.030$ (to 3 d.p.). But $\log 1 < 0.030 < \log 2$, so the first digit of $2^{1\,000}$ is 1.

Similarly, $\log 2^{10\,000} = 10\,000 \log 2 = 3010.29996$. But $\log 1 < 0.29996 < \log 2$, so the first digit of $2^{10\,000}$ is 1.

Finally, $\log 2^{100\,000} = 30102 + 0.9996$ (4 d.p.) and $\log 9 = 2 \log 3 = 0.95$ (2 d.p.), so the first digit of $2^{100\,000}$ is 9.

Postmortem

Although the ideas in this question are really quite elementary, you needed to understand them deeply. You should feel pleased with yourself it you got this one out.

Question 33($**$)

For any number x, the largest integer less than or equal to x is denoted by $[x]$. For example, $[3.7] = 3$ and $[4] = 4$.

(i) Sketch the graph of $y = [x]$ for $0 \leqslant x < 5$ and evaluate

$$\int_0^5 [x] \, \mathrm{d}x \, .$$

(ii) Sketch the graph of $y = [\mathrm{e}^x]$ for $0 \leqslant x < \ln n$, where n is an integer, and show that

$$\int_0^{\ln n} [\mathrm{e}^x] \, \mathrm{d}x = n \ln n - \ln(n!) \, .$$

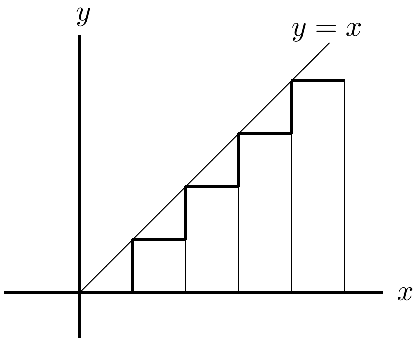
Hence show that $n! \geqslant n^n \mathrm{e}^{1-n}$.

Comments

Again, I have had to add a bit to the original question because it was all dressed up with nowhere to go. The question is clearly about estimating $n!$ so I added in the last line (which makes the question a bit longer but not much more difficult). I could have added another part, but I thought that you would probably add it yourself: you will easily spot, once you have drawn the graphs, that a similar result for $n!$, with the inequality reversed, can be obtained by considering rectangles the tops of which are above the graph of e^x instead of below it. You therefore end up with a nice sandwich inequality for $n!$.

Solution to question 33

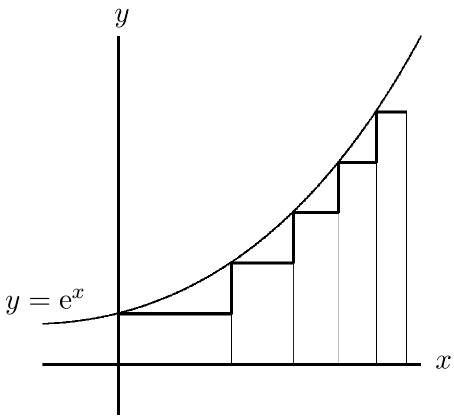
(i)



The graph of $[x]$ is an ascending staircase with 5 stairs, the lowest at height 0, each of width 1 unit and rising 1 unit.
The integral is the sum of the areas of the rectangles shown in the figure:

$$\text{area} = 0 + 1 + 2 + 3 + 4 = 10\,.$$

(ii)



The graph of $[\mathrm{e}^x]$ is a also staircase: the height of each stair is 1 unit and the width decreases as x increases (because the gradient of e^x increases). It starts at height 1 and ends at height $(n-1)$.
The value of $[\mathrm{e}^x]$ changes from k to $k+1$ when $\mathrm{e}^x = k$ i.e. when $x = \ln(k+1)$. Thus $[\mathrm{e}^x] = k$ when $\ln k \leqslant x < \ln(k+1)$ and the area of the corresponding rectangle is $k(\ln(k+1) - \ln k)$.
The total area under the curve is therefore

$$1(\ln 2 - \ln 1) + 2(\ln 3 - \ln 2) + \cdots + (n-1)(\ln n - \ln(n-1))$$

i.e. $n \ln n - \ln n - \ln(n-1) - \ln(n-2) - \cdots - \ln 1$ which sums to $n\ln n - \ln(n!)$ as required.

For the last part, note that $[\mathrm{e}^x] \leqslant \mathrm{e}^x$ (this is clear from the definition) so

$$\int_0^{\ln n} [\mathrm{e}^x]\,\mathrm{d}x \leqslant \int_0^{\ln n} \mathrm{e}^x\,\mathrm{d}x = n - 1 \qquad \text{i.e.} \qquad n\ln n - \ln(n!) \leqslant n - 1\,.$$

Taking exponentials gives the required result.

Postmortem

Considering $[\mathrm{e}^x + 1]$ gives rectangles above the graph of $y = \mathrm{e}^x$ rather than below. The calculations are roughly the same, so you should easily arrive at $n! \leqslant n^{n+1}\mathrm{e}^{1-n}$. We have therefore proved that

$$n^n \mathrm{e}^{1-n} \leqslant n! \leqslant n^{n+1}\mathrm{e}^{1-n}\,.$$

This is a pretty good (given the rather elementary method at our disposal) approximation for $n!$ for large n. Stirling (1692 – 1770) proved in 1730 that $n! \approx \sqrt{2\pi}\, n^{n+\frac{1}{2}}\mathrm{e}^{-n}$ for large n. This was a brilliant result — even reading his book, it is hard to see where he got $\sqrt{2\pi}$ from (especially as he wrote in Latin) — but he went on to obtain the approximation in terms of an infinite series. The expression above is just the first term.

Question 34$(***)$

Given that $\tan\frac{\pi}{4} = 1$ show that $\tan\frac{\pi}{8} = \sqrt{2} - 1$.

Let

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

Show, by using the change of variable $x = \sin 4t$, that

$$I = \int_{0}^{\pi/8} \frac{2\cos 4t}{\cos^2 t}\, \mathrm{d}t\ .$$

Hence show that

$$I = 4\sqrt{2} - \pi - 2\ .$$

Comments

This tests trigonometric manipulation and integration skills. You will certainly need $\tan 2\theta$ in terms of $\tan\theta$, $\cos 2\theta$ in terms of $\cos\theta$, and maybe other formulae. A good idea is required in the first part to deal with $\sqrt{1 \pm \sin 4t}$ — look away if you want to find it for yourself — try writing 1 as $\cos^2 2t + \sin^2 2t$.

Both parts of the question are what is called *multistep*: there are half a dozen consecutive steps, each different in nature, with no guidance. This sort of thing is unusual at A-level but normal in university course mathematics. In this question, the most steps require trigonometric manipulation similar to that appearing in earlier parts of the question.

I checked the answer on `www.integrals.com`, which turned out to be very good indeed at doing this sort of thing. I asked it to do the indefinite integral and, in less than a second, it came up with

$$\sqrt{1-x}\left[-1 - (\sqrt{x+1}+1)^{-1}\right] - 2\arcsin\sqrt{(x+1)/2} + [\sqrt{x+1}+1]^{-1}.$$

Not a pretty sight and not in its neatest form by a long way (the arcsin term reduces, after a bit of algebra, to $-\pi/2 + \arcsin x$(!)). I also asked it to do the same integral with the 2 replaced by a which it did in less than four seconds. The answer was about 20 times longer than the above result. It seemed to enjoy the problem, as far as I could tell.

I think that it does it by multiplying top and bottom of the fraction by suitable expressions (maybe $\sqrt{(1+x)} - \sqrt{(1-x)} + 2$ to start with) until the square roots have been removed from the denominator.

Solution to question 34

Let $t = \tan\frac{\pi}{8}$. Then

$$\frac{2t}{1-t^2} = 1 \Rightarrow t^2 + 2t - 1 = 0 \Rightarrow t = \frac{-2 \pm \sqrt{8}}{2} = -1 \pm \sqrt{2}\,.$$

We take the root with the + sign since we know that t is positive.

Now the integral. Substituting $x = \sin 4t$ gives:

$$I = \int_{-\pi/8}^{\pi/8} \frac{4\cos 4t}{\sqrt{1+\sin 4t} + \sqrt{1-\sin 4t} + 2}\,\mathrm{d}t = 2\int_0^{\pi/8} \frac{4\cos 4t}{(\cos 2t + \sin 2t) + (\cos 2t - \sin 2t) + 2}\,\mathrm{d}t\,,$$

using

$$1 \pm \sin 4t = 1 \pm 2\sin 2t\cos 2t = \cos^2 2t + \sin^2 2t \pm 2\sin(t/2)\cos 2t = (\cos 2t \pm \sin 2t)^2$$

and remembering to take the positive square root ($\cos 2t \geqslant \sin 2t$ for values of t in the range of integration). One more trig. manipulation gives the required result.

Now

$$\cos 4t = 2\cos^2 t - 1 = 2(2\cos^2 t - 1)^2 + 1 = 8\cos^4 t - 8\cos^2 t + 1$$

so

$$I = 2\int_0^{\pi/8} \big(8\cos^2 t - 4 + \sec^2 t\big)\mathrm{d}t = \Big[4\sin 2t - 8t + \tan t\Big]_0^{\pi/8},$$

which gives the required result.

Postmortem

Manipulating the integral after the change of variable was really quite demanding. If you didn't think of the idea referred to in the comments for dealing with $\sqrt{1 \pm \sin 4t}$ you could of course have verified that $\sqrt{1+\sin 4t} + \sqrt{1 \pm \sin 4t} + 2 = 4\cos^2 t$, though even this would require some thought: you would have to subtract 2 from both sides before you squared them, and then you would have to worry about signs, since square-rooting the squares would lead to a $\pm$.

It occurs to me that there ought to be a better way of doing the integral, without at any stage spoiling the $x \to -x$ symmetry of the integrand; maybe I'll think again about this one day.

Question 35(∗∗)

(i) Show that, for $0 \leqslant x \leqslant 1$, the largest value of $\dfrac{x^6}{(x^2+1)^4}$ is $\dfrac{1}{16}$.

What is the smallest value?

(ii) Find constants A, B, C and D such that, for all x,

$$\frac{1}{(x^2+1)^4} \equiv \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{Ax^5+Bx^3+Cx}{(x^2+1)^3}\right) + \frac{Dx^6}{(x^2+1)^4}\,.$$

(iii) Hence, or otherwise, prove that

$$\frac{11}{24} \leqslant \int_0^1 \frac{1}{(x^2+1)^4}\,\mathrm{d}x \leqslant \frac{11}{24} + \frac{1}{16}\,.$$

Comments

You should think about part (i) graphically, though it is not necessary to draw the graph: just set about it as if you were going to (starting point, finishing point, turning points, etc).

The equivalence sign in part (ii) indicates an equality that holds for all x — you are not being asked to solve the equation for x. After doing the differentiation, you obtain four linear equations for A, B, C and D by equating coefficients of powers of x. You could instead put four carefully chosen values of x into the equation; one good choice would be $x = i$ to eliminate terms with factors of x^2+1 and setting $x = 0$ would be useful; after that, it becomes more difficult.

Note that there are only odd powers of x inside the derivative. You could include even powers as well, but you would find that their coefficients would be zero: the derivative has to be an even function (even powers of x) so the function being differentiated must be odd.

For part (iii), you need to know that inequalities can be integrated: this is 'obvious' if you think about integration in terms of area, though a formal proof requires a formal definition of integration and this is the sort of thing you would do in a first university course in mathematical analysis.

This question is about estimating an integral. Although the idea is good, the final result is a bit feeble. It only gives the value of the integral to an accuracy of about (1/16)(11/24) which is about 15%. The actual value of the integral can be found fairly easily using the substitution $x = \tan t$ and is $11/48 + 5\pi/64$ so the inequalities give (not particularly good) estimates for π namely ($2.933 \leqslant \pi \leqslant 3.733$).

Solution to question 35

(i) The derivative of $\dfrac{x^6}{(x^2+1)^4}$ is

$$\frac{6x^5}{(x^2+1)^4} - \frac{8x^7}{(x^2+1)^5} = \frac{2(3-x^2)x^5}{(x^2+1)^5}$$

which is positive for $0 < x < 1$. The function therefore increases in value from 0 to 1. The largest value occurs when $x = 1$ and is $1/16$ and the smallest value occurs at $x = 0$ and is 0.

(ii) Doing the differentiation gives

$$\frac{1}{(x^2+1)^4} \equiv \frac{5Ax^4+3Bx^2+C}{(x^2+1)^3} - \frac{6x(Ax^5+Bx^3+Cx)}{(x^2+1)^4} + \frac{Dx^6}{(x^2+1)^4}$$

and multiplying by $(x^2+1)^4$ gives the following identity:

$$1 \equiv (5Ax^4+3Bx^2+C)(x^2+1) - 6x(Ax^5+Bx^3+Cx) + Dx^6$$

i.e.

$$1 \equiv (D-A)x^6 + (5A-3B)x^4 + (3B-5C)x^2 + C\ .$$

This holds for all values of x so we can equate coefficients of the different powers of x on each side of the equivalence sign. This gives

$1 = C$, $0 = 3B - 5C$, $0 = 5A - 3B$, $0 = D - A$,

so $A = 1$, $B = 5/3$, $C = 1$ and $D = 1$.

With these values of A, B, C and D the identity in the question holds (i.e. the these values are not merely necessary, they are also sufficient).

(iii) Using the results of parts (i) and (ii), we see that for $0 \leqslant x \leqslant 1$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{x^5+(5/3)x^3+x}{(x^2+1)^3}\right) \leqslant \frac{1}{(x^2+1)^4} \leqslant \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{x^5+(5/3)x^3+x}{(x^2+1)^3}\right) + \frac{1}{16}\ .$$

Inequalities can be integrated, so

$$\int_0^1 \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{x^5+(5/3)x^3+x}{(x^2+1)^3}\right)\mathrm{d}x \leqslant \int_0^1 \frac{1}{(x^2+1)^4}\mathrm{d}x \leqslant \int_0^1 \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{x^5+(5/3)x^3+x}{(x^2+1)^3}\right)\mathrm{d}x + \int_0^1 \frac{1}{16}\,\mathrm{d}x$$

i.e.

$$\left[\frac{x^5+(5/3)x^3+x}{(x^2+1)^3}\right]_0^1 \leqslant \int_0^1 \frac{1}{(x^2+1)^4}\mathrm{d}x \leqslant \left[\frac{x^5+(5/3)x^3+x}{(x^2+1)^3}\right]_0^1 + \left[\frac{x}{16}\right]_0^1$$

from which the required result follows immediately.

Question 36$(**)$

Show that
$$\int_{-1}^{1} |\, x\mathrm{e}^{x} \,|\mathrm{d}x = -\int_{-1}^{0} x\mathrm{e}^{x}\mathrm{d}x + \int_{0}^{1} x\mathrm{e}^{x}\mathrm{d}x$$
and hence evaluate the integral.

Evaluate the following integrals:

(i) $$\int_{0}^{4} |\, x^3 - 2x^2 - x + 2 \,|\, \mathrm{d}x\,;$$

(ii) $$\int_{-\pi}^{\pi} |\, \sin x + \cos x \,|\, \mathrm{d}x\,.$$

Comments

The very first part shows you how to do this sort of integral (with mod signs) by splitting up the range of integration at the places where the integrand changes sign. In parts (i) and (ii) you have to use the technique on different examples. This learning process and the nice mixture of ideas make it an ideal STEP question.

Solution to question 36

For the first part, note that $|xe^x| = -xe^x$ if $x < 0$. Then integrate by parts to evaluate the integrals:

$$-\int_{-1}^{0} xe^x \mathrm{d}x + \int_0^1 xe^x \mathrm{d}x = -\left(\Big[xe^x\Big]_{-1}^{0} - \int_{-1}^{0} e^x \mathrm{d}x\right) + \left(\Big[xe^x\Big]_0^1 - \int_0^1 e^x \mathrm{d}x\right) = 2 - 2e^{-1}\,.$$

For the next parts, we have to find out where the integrand is positive and where it is negative.

(i) $x^3 - 2x^2 - x + 2 = (x-1)(x+1)(x-2)$ (spotting the factors), so the integrand is positive for $0 \leqslant x < 1$, negative for $1 < x < 2$ and positive for $2 < x < 4$. Splitting the range of integration into these ranges and integrating gives

$$\begin{aligned}
&\int_0^4 \left| x^3 - 2x^2 - x + 2 \right| \mathrm{d}x \\
&\qquad = \int_0^1 (x^3 - 2x^2 - x + 2)\,\mathrm{d}x - \int_1^2 (x^3 - 2x^2 - x + 2)\,\mathrm{d}x + \int_2^4 (x^3 - 2x^2 - x + 2)\,\mathrm{d}x \\
&\qquad = \left[\tfrac14 x^4 - \tfrac23 x^3 - \tfrac12 x^2 + 2x\right]_0^1 - \left[\tfrac14 x^4 - \tfrac23 x^3 - \tfrac12 x^2 + 2x\right]_1^2 + \left[\tfrac14 x^4 - \tfrac23 x^3 - \tfrac12 x^2 + 2x\right]_2^4 \\
&\qquad = 22\tfrac16\,.
\end{aligned}$$

(ii) $\sin x + \cos x$ changes sign when $\tan x = -1$, i.e. when $x = -\pi/4$ and $x = 3\pi/4$. Splitting the range of integration into these ranges and integrating gives

$$\begin{aligned}
\int_{-\pi}^{\pi} \left| \sin x + \cos x \right| \mathrm{d}x &= -\int_{-\pi}^{-\pi/4} (\sin x + \cos x)\,\mathrm{d}x + \int_{-\pi/4}^{3\pi/4} (\sin x + \cos x)\,\mathrm{d}x - \int_{3\pi/4}^{\pi} (\sin x + \cos x)\,\mathrm{d}x \\
&= -\Big[-\cos x + \sin x\Big]_{-\pi}^{-\pi/4} + \Big[-\cos x + \sin x\Big]_{-\pi/4}^{3\pi/4} - \Big[-\cos x + \sin x\Big]_{3\pi/4}^{\pi} \\
&= 4\sqrt{2}
\end{aligned}$$

Alternatively, start by writing $\sin x + \cos x = \sqrt{2}\cos(x - \pi/4)$ which makes the changes of sign easier to spot and the integrals easier to do.

Postmortem

If you give a bit more thought to part (ii), you will see easier ways of doing it. Since the trigonometric functions are periodic with period 2π the integrand is also periodic. Write the integrand in the form $\sqrt{2}\left|\cos(x - \pi/4)\right|$. Integrating this over any 2π interval gives the same result. Indeed, we may as well integrate $\sqrt{2}\left|\sin x\right|$; from 0 to π and double the answer.

Question 37($*$)

A number of the form $1/N$, where N is an integer greater than 1, is called a *unit fraction.*

Noting that

$$\frac{1}{2} = \frac{1}{3} + \frac{1}{6} \quad \text{and} \quad \frac{1}{3} = \frac{1}{4} + \frac{1}{12},$$

guess a general result of the form

$$\frac{1}{N} = \frac{1}{a} + \frac{1}{b} \tag{$*$}$$

and hence prove that any unit fraction can be expressed as the sum of two distinct unit fractions.

By writing ($*$) in the form

$$(a - N)(b - N) = N^2$$

and by considering the factors of N^2, show that if N is prime, then there is only one way of expressing $1/N$ as the sum of two distinct unit fractions.

Prove similarly that any fraction of the form $2/N$, where N is prime number greater than 2, can be expressed uniquely as the sum of two distinct unit fractions.

Comments

Fractions written as the sum of unit fractions are called *Egyptian fractions*: they were used by Egyptians. The earliest record of such use is 1900BC. The Rhind papyrus in the British Museum gives a table of representations of fractions of the form $2/n$ as sums of unit fractions for all odd integers n between 5 and 101 — a superb achievement when you consider that algebra was 3.500 years away.

It is not clear why Egyptians represented fractions this way; maybe it just seemed a good idea at the time. Certainly the notation they used, in which for example $1/n$ was denoted by n with an oval on top, does not lend itself to generalisation to fractions that are not unit. One of the rules was that all the unit fractions in an Egyptian fraction should be distinct, so $\dfrac{2}{7}$ had to be expressed as $\dfrac{1}{4} + \dfrac{1}{28}$, which seems pretty daft.

It is not obvious that every fraction can be express in Egyptian form; this was proved by Fibonacci in 1202. There are however still many unsolved problems relating to Egyptian fractions.

Egyptian fractions have been called a wrong turn in the history of mathematics; if so, it was a wrong turn that favoured style over utility which is no bad thing, in my opinion.

Solution to question 37

A good guess would be that the first term of the decomposition of $\dfrac{1}{N}$ is $\dfrac{1}{N+1}$. In that case, the other term is $\dfrac{1}{N} - \dfrac{1}{N+1}$ i.e. $\dfrac{1}{N(N+1)}$. That proves the result that every unit fraction can be expressed as the sum of two unit fractions.

The only factors of N^2 (since N is prime) are 1, N and N^2. The possible factorisations of N^2 are therefore $N^2 = 1 \times N^2$ or $N^2 = N \times N$ so, since $a \neq b$, $a - N = 1$ and $b - N = N^2$ (or the other way round). Thus $N + 1$ and $N^2 + N$ are the only possible values for a and b and the decomposition is unique.

For the second half, set

$$\frac{2}{N} = \frac{1}{a} + \frac{1}{b}$$

where $a \neq b$. Proceeding as before, we have $ab - \dfrac{(a+b)N}{2} = 0$ which we write as

$$\left(a - \frac{N}{2}\right)\left(b - \frac{N}{2}\right) = \frac{N^2}{4} \qquad \text{i.e.} \qquad (2a - N)\,(2b - N) = N^2\,.$$

Thus $2a - N = N^2$ and $2b - N = 1$ (or the other way round). The decomposition is therefore unique and given by

$$\frac{2}{N} = \frac{1}{(N^2+N)/2} + \frac{1}{(N+1)/2}\,.$$

The only possible fly in the ointment is the $1/2$ in the denominators: a and b are supposed to be integers. However, all prime numbers greater than 2 are odd, so $N + 1$ and $N^2 + N$ are both even and the denominators are indeed integers.

Postmortem

Another way of getting the last part (less systematically) would have been to notice that

$$\frac{1}{N} = \frac{1}{(N+1)} + \frac{1}{N(N+1)} \Longrightarrow \frac{2}{N} = \frac{1}{(N+1)/2} + \frac{1}{N(N+1)/2}$$

which gives the result immediately. How might you have noticed this? Well, I noticed by trying to work out some examples, starting with what is given at the very beginning of the question. If $\dfrac{1}{3} = \dfrac{1}{4} + \dfrac{1}{12}$ then $\dfrac{2}{3} = \dfrac{2}{4} + \dfrac{2}{12}$ which works.

Of course, the advantage of the systematic approach is that it allows generalisation: what happens if N is odd but not prime; what happens if the numerator is 3 not 2?

Question 38(∗)

Prove that if $(x-a)^2$ is a factor of the polynomial $\mathrm{p}(x)$, then $\mathrm{p}'(a)=0$. Prove a corresponding result if $(x-a)^4$ is a factor of $\mathrm{p}(x)$.

Given that the polynomial

$$x^6+4x^5-5x^4-40x^3-40x^2+kx+m$$

has a factor of the form $(x-a)^4$, find k and m.

Comments

The very first part has an 'obvious' feel to it but the difficulty lies in setting out a clear mathematical proof. The key is the factor theorem: if $(x-a)$ is a factor of a polynomial, then it can be written as the product of polynomial and $(x-a)$.

Please take some care over the presentation of next part.

There are a few numerical calculations to do. Since calculators are neither permitted nor needed in STEP, it is probably best if you don't use yours.

In the original question, k was given (it is 32). But this seemed rather unsatisfactory because it meant that not all the information available is required to solve the problem. Even my new improved version is not perfect in this respect because $\mathrm{p}''(a)=0$ is used only to distinguish between three possibilities for a.

Solution to question 38

If $(x-a)^2$ is a factor of $\mathrm{p}(x)$, we can write $\mathrm{p}(x) = (x-a)^2\mathrm{q}(x)$, where $\mathrm{q}(x)$ is a polynomial. Differentiating gives

$$\mathrm{p}'(x) = 2(x-a)\mathrm{q}(x) + (x-a)^2\mathrm{q}'(x)$$

and $\mathrm{p}'(a) = 0$.

Similarly, if $(x-a)^4$ is a factor, then $\mathrm{p}(x) = (x-a)^4\mathrm{q}(x)$ for some polynomial $\mathrm{q}(x)$. Following the above argument, we have

$$\mathrm{p}'(x) = 4(x-a)^3\mathrm{q}(x) + (x-a)^4\mathrm{q}'(x)$$

so $\mathrm{p}'(x)$ has a factor $(x-a)^3$ and can be expressed in the form $\mathrm{p}'(x) = (x-a)^3\mathrm{r}(x)$ for some polynomial. Similarly, $\mathrm{p}''(x)$ has a factor $(x-a)^2$ and $\mathrm{p}'''(x)$ has a factor $(x-a)$. A necessary condition for $\mathrm{p}(x)$ to have a factor $(x-a)^4$ is therefore $\mathrm{p}(a) = \mathrm{p}'(a) = \mathrm{p}''(a) = \mathrm{p}'''(a) = 0$.

For the given polynomial,

$$\mathrm{p}'''(a) = 0 \Rightarrow 120a^3 + 240a^2 - 120a - 240 = 0 \Rightarrow a^3 + 2a^2 - a - 2 = 0 \Rightarrow (a-1)(a+1)(a+2) = 0$$

(spotting factors), so the possible values of a are 1, -1 and -2. But $\mathrm{p}''(1) \neq 0$ and $\mathrm{p}''(-1) \neq 0$, so $a = -2$.

Now we have two remaining equations.

$$\mathrm{p}'(-2) = 0 \Rightarrow 6(-2)^5 + 20(-2)^4 - 20(-2)^3 - 120(-2)^2 - 80(-2) + k = 0$$

so $k = 32$. Similarly

$$\mathrm{p}(-2) = 0 \Rightarrow m = 48\ .$$

Postmortem

Another way of doing the question would be simply to write

$$x^6 + 4x^5 - 5x^4 - 40x^3 - 40x^2 + kx + m \equiv (x-a)^4(bx^2 + cx + d)$$

and equation coefficients of x^n on either side of the equation. That gives 7 equations for 6 unknowns. In general, the system would be overdetermined (too many equations for the number of unknowns) but here we know that a solution exists.

Question 39$(***)$

Show that
$$\sin\theta = \frac{2t}{1+t^2}\ , \qquad \cos\theta = \frac{1-t^2}{1+t^2}\ , \qquad \frac{1+\cos\theta}{\sin\theta} = \tan(\pi/2 - \theta/2)\ ,$$
where $t = \tan(\theta/2)$.

Let
$$I = \int_0^{\frac{\pi}{2}} \frac{1}{1+\cos\alpha\sin\theta}\, \mathrm{d}\theta\ .$$
Use the substitution $t = \tan(\theta/2)$ to show that, for $0 < \alpha < \pi/2$,
$$I = \int_0^1 \frac{2}{(t+\cos\alpha)^2 + \sin^2\alpha}\, \mathrm{d}t\ .$$
By means of the further substitution $t + \cos\alpha = \sin\alpha\tan u$ show that
$$I = \frac{\alpha}{\sin\alpha}\ .$$

Deduce a similar result for
$$\int_0^{\frac{\pi}{2}} \frac{1}{1+\sin\alpha\cos\phi}\, \mathrm{d}\phi\ .$$

Comments

The first of the two substitutions is familiarly known as the 't substitution'. It would have been very standard fare 30 years ago, but seems to have gone out of fashion now. The second is the normal substitution for integrals with quadratic denominators and is only just over the edge of the core A-level syllabus.

For the last part, 'deduce' implies that you don't have to do any further integration. This can only really mean a change of variable to make the integrand into the one you want; you have to hope that the limits then transform to the one given. Note that the change of variable is signalled by the use of the variable ϕ in the last integral; since it is a definite integral, it doesn't matter what the variable is called, and could have been called θ as I; the use of a different variable was just a kindness on the part of the examiner (me). Of course, you can change α without any such complications.

The 'similar result' result should include the conditions under which it is true. It is worth thinking about why the condition $0 < \alpha \leqslant \pi/2$ is required – or indeed, if it is required.

Solution to question 39

The three identities just require use of the half-angle formulae $\cos\theta = \cos^2(\theta/2) - \sin^2(\theta/2)$ and $\sin\theta = 2\sin(\theta/2)\cos(\theta/2)$. Remember, for the last one, that $\cot x = \tan(\pi/2 - x)$.

For the first change of variable, we have $\mathrm{d}t = \frac{1}{2}\sec^2(\theta/2)\,\mathrm{d}\theta = \frac{1}{2}(1+t^2)\mathrm{d}\theta$ and the new limits are 0 and 1, so

$$\begin{aligned} I = \int_0^{\frac{\pi}{2}} \frac{1}{1+\cos\alpha\sin\theta}\,\mathrm{d}\theta &= \int_0^1 \frac{1}{1+(\cos\alpha)2t/(1+t^2)}\,\frac{2}{1+t^2}\,\mathrm{d}t \\ &= \int_0^1 \frac{2}{1+2t\cos\alpha+t^2}\,\mathrm{d}t = \int_0^1 \frac{2}{(t+\cos\alpha)^2+\sin^2\alpha}\,\mathrm{d}t\,. \end{aligned}$$

For the second change of variable, we have $\mathrm{d}t = \sin\alpha\sec^2 u\,\mathrm{d}u$. When $t=0$, $\tan u = \cot\alpha$ so $u = \pi/2-\alpha$. When $t=1$, $\sin\alpha\tan u = 1+\cos\alpha$ so $u = \pi/2 - \alpha/2$. Thus

$$\begin{aligned} I &= \int_{\pi/2-\alpha}^{\pi/2-\alpha/2} \frac{2}{\sin^2\alpha(1+\tan^2 u)}\,\sin\alpha\sec^2 u\mathrm{d}u \\ &= \int_{\pi/2-\alpha}^{\pi/2-\alpha/2} \frac{2}{\sin\alpha}\,\mathrm{d}u = \frac{2}{\sin\alpha}\Big[\,u\,\Big]_{\pi/2-\alpha}^{\pi/2-\alpha/2} = \frac{\alpha}{\sin\alpha}\,. \end{aligned}$$

For the last part, we want to make a change of variable the changes the cosine in the denominator to a sine. One possibility is to set $\phi = \pi/2 - \theta$. This will swaps the limits but also and introduces a minus sign since $\mathrm{d}\theta = -\mathrm{d}\phi$. Thus

$$\frac{\alpha}{\sin\alpha} = I = -\int_{\frac{\pi}{2}}^{0} \frac{1}{1+\cos\alpha\cos\phi}\,\mathrm{d}\phi = \int_0^{\frac{\pi}{2}} \frac{1}{1+\cos\alpha\cos\phi}\,\mathrm{d}\phi\,.$$

This is almost the integral we want: we still need to replace $\cos\alpha$ in the denominator by $\sin\alpha$. Remembering what we did a couple of lines back, we just replace α by $\pi/2-\alpha$ in the integral and in the answer, giving

$$\int_0^{\frac{\pi}{2}} \frac{1}{1+\cos\alpha\cos\phi}\,\mathrm{d}\phi = \frac{\pi/2-\alpha}{\cos\alpha}\,.$$

Postmortem

The restriction on α in the integral might prevent the denominator being zero for some value of θ; a zero in the denominator usually means that the integral is undefined. However, the only value of α for which $\cos\alpha\sin\theta$ could possibly be as small as -1 (for $0\leqslant\theta\leqslant\pi/2$) is π (and of course 3π, etc). From the point of view of the integral, the only restriction should be $\alpha\neq\pi$ (etc). There is a slight awkwardness in the answer when $\alpha=0$ but this can be overcome by taking limits:

$$\lim_{\alpha\to 0}\frac{\alpha}{\sin\alpha} = 1$$

which you can easily verify is the correct answer in this case.

You may by now have noticed a curious thing: increasing α by 2π does not change I but does change the answer. If you work through the solution with this in mind you see that the only step where it can make a difference is in the line where you work out $\tan^{-1} u$, which by definition lies in the range $-\pi/2$ to $\pi/2$. It is this that determines the given restriction. (Note that negative values of α are not required because setting $\alpha\to-\alpha$ changes neither the integral nor the given solution.)

Question 40$(**)$

The line ℓ has vector equation $\mathbf{r} = \lambda\mathbf{s}$, where

$$\mathbf{s} = (\cos\theta + \sqrt{3}\,)\,\mathbf{i} + (\sqrt{2}\,\sin\theta)\,\mathbf{j} + (\cos\theta - \sqrt{3}\,)\,\mathbf{k}$$

and λ is a scalar parameter. Find an expression for the angle between ℓ and the line $\mathbf{r} = \mu(a\,\mathbf{i} + b\,\mathbf{j} + c\,\mathbf{k})$. Show that there is a line m through the origin such that, whatever the value of θ, the acute angle between ℓ and m is $\pi/6$.

A plane has equation $x - z = 4\sqrt{3}$. The line ℓ meets this plane at P. Show that, as θ varies, P describes a circle, with its centre on m. Find the radius of this circle.

Comments

It is not easy to set vector questions at this level: they tend to become merely complicated (rather than difficult in an interesting way). Usually, there is some underlying geometry and it pays to try to understand what this is. Here, the question is about the geometrical object traced out by ℓ as θ varies.

You will need to know about scalar products of vectors for this question, but otherwise it is really just coordinate geometry.

Vectors form an extremely important part of almost every branch of mathematics (maybe *every* branch of mathematics) and will probably be one of the first topics you tackle on your university course.

Solution to question 40

Both ℓ and the line $\mathbf{r} = \mu(a\,\mathbf{i} + b\,\mathbf{j} + c\,\mathbf{k})$ pass through the origin, so the angle α between the lines is given by the scalar product of the unit vectors, i.e. the scalar product between the given vectors divided by the product of the lengths of the two vectors:

$$\begin{aligned}
\cos\alpha &= \frac{a(\cos\theta+\sqrt{3}) + \sqrt{2}b\sin\theta + (\cos\theta-\sqrt{3})c}{\sqrt{[(\cos\theta+\sqrt{3})^2 + 2\sin^2\theta + (\cos\theta-\sqrt{3})^2]}\ \sqrt{[a^2+b^2+c^2]}} \\
&= \frac{(a+c)\cos\theta + \sqrt{2}\,b\sin\theta + (a-c)\sqrt{3}}{2\sqrt{2}\ \sqrt{[a^2+b^2+c^2]}}.
\end{aligned}$$

Now we want to show that there is some choice of a, b and c such that $\cos\alpha$ does not depend on the value of θ. By inspection, this requires $a=-c$ and $b=0$. For these values, the terms depending on θ disappear from the expression for $\cos\alpha$ leaving $\cos\alpha = \pm\frac{1}{2}\sqrt{3}$ according to the whether a is positive or negative. If we choose $a=1$ then $\cos\alpha = +\frac{1}{2}\sqrt{3}$ and $\alpha = \pi/6$ as required. The line m is given by

$$\mathbf{r} = \mu(\mathbf{i} - \mathbf{k})\ .$$

Note that as θ varies, the line ℓ generates a cone with axis m; this is what the question is about.

The coordinates of a general point on the line ℓ are

$$x = \lambda(\cos\theta+\sqrt{3}) \qquad y = \lambda\sqrt{2}\sin\theta \qquad z = \lambda(\cos\theta-\sqrt{3})\ .$$

For a point which is also on the plane $x - z = 4\sqrt{3}$ we have

$$\lambda(\cos\theta+\sqrt{3}) - \lambda(\cos\theta-\sqrt{3}) = 4\sqrt{3}$$

so $\lambda = 2$. The point P on the intersection between the line and the plane has coordinates

$$(2\cos\theta + 2\sqrt{3}\,,\ 2\sqrt{2}\sin\theta\ ,\ 2\cos\theta - 2\sqrt{3})\ .$$

A general point on the line m has coordinates $x=\mu$, $y=0$, $z=-\mu$. It meets the plane $x-z=4\sqrt{3}$ when $\mu = 2\sqrt{3}$ i.e. at the point O with coordinates $(2\sqrt{3},\,0\,,-2\sqrt{3})$.

We need to check that, as θ varies, P moves round a circle with centre O, i.e. that the length OP is independent of θ. We have

$$OP^2 = (2\cos\theta)^2 + (2\sqrt{2}\sin\theta)^2 + (2\cos\theta)^2 = 8\ ,$$

which is indeed independent of θ as required. The radius of the circle is therefore $2\sqrt{2}$.

Question 41(∗∗)

Given that

$$x^4 + px^2 + qx + r \equiv (x^2 - ax + b)(x^2 + ax + c) \ ,$$

express p , q and r in terms of a , b and c .

Show that a^2 is a root of the cubic equation

$$u^3 + 2pu^2 + (p^2 - 4r)u - q^2 = 0 \ .$$

Verify that $u = 9$ is a root in the case $p = -1$, $q = -6$, $r = 15$.

Hence, or otherwise, solve the equation

$$y^4 - 8y^3 + 23y^2 - 34y + 39 \ .$$

Comments

The long-sought solution of the general cubic was found, in 1535 by Niccolò Tartaglia (c. 1500-57). He was persuaded to divulge his secret (in the form of a poem) by Girolamo Cardano (1501–76), who promised not to publish it before Cardano did. However, Cardano discovered that it had previously been discovered by del Ferro (1465–1525/6) before 1515 so he published it himself in his algebra book *The Great Art*. There followed an acrimonious dispute between Tartaglia and Cardano, in which the latter was championed by his student Ferrari (1522–1565). The dispute culminated in a public mathematical duel between Farrari and Tartaglia held in the church of Santa Maria in Milan in 1548, in which they attempted to solve each others' cubics. The duel ended in a shouting match with Tartaglia storming off. It seems Ferrari was the winner. Tartaglia was sacked from his job as lecturer and Ferrari made his fortune as a tax assessor before becoming a professor of mathematics at Bologna. He was poisoned by his sister, with arsenic, in 1565.

Ferarri found the a way of reducing quartic equations to cubic equations; his method (roughly) is used in this question to solve a quartic which could probably be solved easier 'otherwise'. But it is the method that is interesting, not the solution.

The first step of the Ferrari method is to reduce the general quartic to a quartic equation with the cubic term missing by means of a linear transformation of the form $x \to x - a$. Then this reduced quartic is factorised (the first displayed equation in this question). The factorisation can be found by solving a cubic equation (the second displayed equation above) that must be satisfied by one of the coefficients in the factorised form.

The solutions of the quartic are all complex, but don't worry if you haven't come across complex numbers: you will be able to do everything except perhaps write down the last line.

You will no doubt be full of admiration for this clever method of solving quartic equations. One thing you are bound to ask yourself is how the other two roots of the cubic equation fit into the picture. In fact, the cubic equation gives (in general) 6 distinct values of a so there is quite a lot of explaining to do, given that the quartic has at most 4 distinct roots.

Solution to question 41

We have

$$\begin{aligned}(x^2-ax+b)(x^2+ax+c) &= (x^2-ax)(x^2+ax)+b(x^2+ax)+c(x^2-ax)+bc\\ &= x^4+(b+c-a^2)x^2+a(b-c)x+bc\end{aligned}$$

so

$$p=b+c-a^2, \qquad q=a(b-c)\,, \qquad r=bc\,. \tag{$*$}$$

To obtain an equation for a in terms of p, q and r we eliminate b and c from $(*)$ using the identity $(b+c)^2=(b-c)^2+4bc$. This gives $(p+a^2)^2=(q/a)^2+4r$ which simplifies easily to the given cubic with u replaced by a^2. For the purposes of this equation, all that is required is to verify that setting $u=a^2$ in the given cubic equation works (on use of $(*)$), but this is probably no easier than constructing the equation by elimination, as above.

We can easily verify that $u=9$ satisfies the given cubic by direct substitution.

To solve the quartic equation, the first task is to reduce it to the form x^4+px^2+qx+r, which has no term in x^3. This is done by means of a translation. Noting that $(y-a)^4=y^4-4ay^3+\cdots$, we set $x=y-2$. This gives

$$(x+2)^4-8(x+2)^3+23(x+2)^2-34(x+2)+39\,.$$

which (not surprisingly) boils down to $x^4-x^2-6x+15$, so that $p=-1$, $q=-6$ and $r=15$. We have already shown that one root of the cubic corresponding to these values is $u=9$. Thus we can achieve the factorisation of the quartic into two quadratic factors by setting $a=3$ (or $a=-3$ — it doesn't matter which we use) and

$$\begin{aligned}b&=\frac{1}{2}\left(p+a^2+\frac{q}{a}\right)=3\\ c&=\frac{1}{2}\left(p+a^2-\frac{q}{a}\right)=5\,.\end{aligned}$$

Thus

$$x^4-x^2-6x+15=(x^2-3x+3)(x^2+3x+5)\,.$$

Setting each of the two quadratic factors equal to zero gives

$$x=\frac{3\pm i\sqrt{3}}{2} \qquad \text{and} \qquad x=\frac{-3\pm i\sqrt{11}}{2}\,,$$

so

$$y=\frac{7\pm i\sqrt{3}}{2} \qquad \text{or} \qquad y=\frac{1\pm i\sqrt{11}}{2}\,.$$

Postmortem

Did you work out the relation between the six possible values of a (corresponding to given values of p, q and r) and the roots of the quartic? The point is that the quartic can be written as the product of four linear factors (by the fundamental theorem of algebra) and there are six ways of grouping the linear factors into two quadratic factors. Each way corresponds to a value of a.

Question 42($***$)

Show that $\tan 3\theta = \dfrac{3\tan\theta - \tan^3\theta}{1 - 3\tan^2\theta}$.

Given that $\theta = \cos^{-1}(2/\sqrt{5})$ and $0 < \theta < \pi/2$, show that $\tan 3\theta = \dfrac{11}{2}$.

Hence, or otherwise, find all solutions of the equations

(i) $\tan(3\cos^{-1} x) = \dfrac{11}{2}$,

(ii) $\cos\left(\dfrac{1}{3}\tan^{-1} y\right) = \dfrac{2}{\sqrt{5}}$.

Comments

Perhaps three stars is a bit over the top for this question, but I thought it was tricky. Very few candidates got more than half marks in the examination (Paper I, 2001).

It is not hard to see what you have to do, but great care is required when writing out the argument. The important thing to remember is that if (say) $\tan A = \tan B$ then you cannot say that $A = B$; rather, you must say $A = B + n\pi$.

For the very first part, you might wonder where to start. This did not trouble the candidates in the examination: they all quoted the formula for $\tan(\alpha + \beta)$ and worked from there.

Solution to question 42

First part:

$$\tan 3\theta = \frac{\tan 2\theta + \tan\theta}{1 - \tan 2\theta \tan\theta} = \frac{2t/(1-t^2)+t}{1-2t^2/(1-t^2)} = \frac{3t-t^3}{1-3t^2}\,,$$

where $t = \tan\theta$.

For θ in the given range, we can use Pythagoras's theorem (think of a triangle with sides 1, 2 and $\surd 5$) to show that

$$\arccos\big(\frac{2}{\sqrt5}\big) = \arcsin\big(\frac{1}{\sqrt5}\big) = \arctan\big(\frac{1}{2}\big)$$

and then set $t = 1/2$ in the above formula. Alternatively, we can obtain $\tan\theta$ from $\tan^2\theta = \sec^2\theta - 1$ taking the positive square root because $0 < \theta < \pi/2$.

(i) We have

$$\tan(3\arccos x) = \frac{11}{2} = \tan 3\arccos\big(\frac{2}{\sqrt5}\big)\,,$$

so $3\arccos x = 3\arccos\big(\frac{2}{\sqrt5}\big) + n\pi$. Thus

$$x = \cos\Big(\arccos(\frac{2}{\sqrt5}) + n\pi/3\Big) = \frac{2}{\sqrt5}\cos\big(\frac{n\pi}{3}\big) - \frac{1}{\sqrt5}\sin\big(\frac{n\pi}{3}\big)\,.$$

This gives six possibilities:

$$\begin{array}{lcll}
x = \dfrac{2}{\sqrt5} & = & \dfrac{2}{\sqrt5} & (n=0)\\[2ex]
x = (\dfrac{2}{\sqrt5})(-1) & = & -\dfrac{2}{\sqrt5} & (n=3)\\[2ex]
x = (\dfrac{2}{\sqrt5})(\dfrac12) - (\dfrac{1}{\sqrt5})(\dfrac{\sqrt3}{2}) & = & \dfrac{2-\sqrt3}{2\sqrt5} & (n=1)\\[2ex]
x = (\dfrac{2}{\sqrt5})(\dfrac12) - (\dfrac{1}{\sqrt5})(-\dfrac{\sqrt3}{2}) & = & \dfrac{2+\sqrt3}{2\sqrt5} & (n=-1)\\[2ex]
x = (\dfrac{2}{\sqrt5})(-\dfrac12) - (\dfrac{1}{\sqrt5})(\dfrac{\sqrt3}{2}) & = & \dfrac{-2-\sqrt3}{2\sqrt5} & (n=2)\\[2ex]
x = (\dfrac{2}{\sqrt5})(-\dfrac12) - (\dfrac{1}{\sqrt5})(-\dfrac{\sqrt3}{2}) & = & \dfrac{-2+\sqrt3}{2\sqrt5} & (n=-2)
\end{array}$$

(ii) Similarly

$$\begin{aligned}
\cos(\frac13\arctan y) = \cos\Big(\frac13\arctan(\frac{11}{2})\Big) \quad &\Rightarrow\quad \frac13\arctan y = \pm\frac13\arctan(\frac{11}{2}) + 2n\pi\\
&\Rightarrow\quad \arctan y = \pm\arctan(\frac{11}{2}) + 6n\pi\\
&\Rightarrow\quad y = \tan\Big(\pm\arctan(\frac{11}{2}) + 6n\pi\Big)\\
&\Rightarrow\quad y = \tan\Big(\pm\arctan(\frac{11}{2})\Big) = \pm\frac{11}{2}\,.
\end{aligned}$$

Question 43$(***)$

The curve C_1 passes through the origin in the x–y plane and its gradient is given by

$$\frac{\mathrm{d}y}{\mathrm{d}x} = x(1-x^2)e^{-x^2}.$$

Show that C_1 has a minimum point at the origin and a maximum point at $\left(1, \frac{1}{2}\,e^{-1}\right)$. Find the coordinates of the other stationary point. Give a rough sketch of C_1.

The curve C_2 passes through the origin and its gradient is given by

$$\frac{\mathrm{d}y}{\mathrm{d}x} = x(1-x^2)e^{-x^3}.$$

Show that C_2 has a minimum point at the origin and a maximum point at $(1,k)$, where $k > \frac{1}{2}\,e^{-1}$. (You need not find k.)

Comments

No work is required to find the x coordinate of the stationary points, but you have to integrate the differential equation to find the y coordinate. For the second part, you cannot integrate the equation — other than numerically, or in terms of rather obscure special functions that you almost certainly haven't come across, such as the incomplete gamma function, defined by

$$\Gamma(x,a) = \int_0^a t^{x-1}\mathrm{e}^{-t}\mathrm{d}t\,.$$

However, you can obtain an estimate, which is all that is required, by comparing the gradients of C_1 with C_2 and thinking of the graphs for $-1 \leqslant x \leqslant 1$. This is perhaps a bit tricky; an idea that you may well not alight on under examination conditions.

Solution to question 43

C_1 has stationary points when $\dfrac{\mathrm{d}y}{\mathrm{d}x} = 0$, i.e. when $x = 0$, $x = +1$ or $x = -1$. To find the y coordinates of the stationary points, we integrate the differential equation, using integration by parts:

$$\begin{aligned} y &= \int (1 - x^2) x \mathrm{e}^{-x^2} \mathrm{d}x \\ &= (1-x^2)(-\tfrac{1}{2}\mathrm{e}^{-x^2}) - \int x\mathrm{e}^{-x^2}\mathrm{d}x \\ &= -\tfrac{1}{2}(1-x^2)\mathrm{e}^{-x^2} + \tfrac{1}{2}\mathrm{e}^{-x^2} + \text{const} \\ &= \tfrac{1}{2}x^2\mathrm{e}^{-x^2}\,, \end{aligned}$$

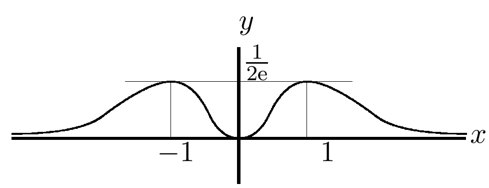
where we have used the fact that C_1 passes through the origin to evaluate the constant of integration. The coordinates of the stationary points are therefore $(1, \frac{1}{2}\,\mathrm{e}^{-1})$, $(0,0)$ and $(-1, \frac{1}{2}\,\mathrm{e}^{-1})$.

One way of classifying the stationary points is to look at the second derivative:

$$\frac{\mathrm{d}y}{\mathrm{d}x} = (x - x^3)e^{-x^2} \Longrightarrow \frac{\mathrm{d}^2y}{\mathrm{d}x^2} = (1-3x^2)\mathrm{e}^{-x^2} - 2x^2(1-x^2)\mathrm{e}^{-x^2} = (1 - 5x^2 + 2x^4)\mathrm{e}^{-x^2}\,,$$

which is positive when $x = 0$ (indicating a minimum) and negative when $x = 1$ and $x = -1$ (indicating maxima).

Since $y \to 0$ as $x \to \pm\infty$, the sketch is:



For C_2, the stationary points are again at $x = 0$ and $x = \pm 1$. To classify the stationary points, we calculate the second derivative:

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = (1 - 3x^2 - 3x^3 + 3x^5)\mathrm{e}^{-x^3}\,,$$

which is positive at $x = 0$ (a minimum) and negative at $x = 1$ (a maximum).

Since we cannot integrate this differential equation explicitly, we must compare C_2 with C_1 to see whether the value of y at the maxima is indeed greater for C_2 than for C_1.

For $0 < x < 1$, $x^3 < x^2$, so $\mathrm{e}^{x^3} < \mathrm{e}^{x^2}$ and $\mathrm{e}^{-x^3} > \mathrm{e}^{-x^2}$. Thus the gradient of C_2 is greater than the gradient of C_1. Since both curves pass through the origin, we deduce that C_2 lies above C_1 for $0 < x \leqslant 1$ and therefore the maximum point $(1, k)$ on C_2 has $k > \frac{1}{2}\mathrm{e}^{-1}$.

Postmortem

You

Question 44($* * *$)

A spherical loaf of bread is cut into parallel slices of equal thickness. Show that, after any number of the slices have been eaten, the area of crust remaining is proportional to the number of slices remaining.

A European ruling decrees that a parallel-sliced spherical loaf can only be referred to as 'crusty' if the ratio of volume V (in cubic metres) of bread remaining to area A (in square metres) of crust remaining after any number of slices have been eaten satisfies $V/A < 1$. Show that the radius of a crusty parallel-sliced spherical loaf must be less than $2\frac{2}{3}$ metres.

[*The area A and volume V formed by rotating a curve in the x–y plane round the x-axis from $x = -a$ to $x = -a + t$ are given by*

$$A = 2\pi \int_{-a}^{-a+t} y \left(1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2\right)^{\frac{1}{2}} \mathrm{d}x \,, \qquad V = \pi \int_{-a}^{-a+t} y^2 \mathrm{d}x \,. \;]$$

Comments

The first result came as a bit of a surprise to me — though no doubt it is well known. I wondered if it was the only surface of revolution with this property. You might like to think about this.

For the last part, you will need to minimise a ratio as a function of t (the 'length' of loaf remaining — which is of course proportional to the number of slices remaining if the slices are all the same thickness); to find the ratio you have to do a couple of integrals. It is this 'multi-stepping' that makes the problem difficult (and very different from typical A-level questions) rather than any individual step.

Solution to question 44

The first thing we need is an equation for the surface of a spherical loaf. The obvious choice, especially given the hint at the bottom of the question, is $x^2 + y^2 = a^2$ (and $z = 0$), rotated about the x-axis.

If the loaf is cut at a distance t from the end $x = a$, then the area remaining is

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

This is proportional to the length t of remaining loaf, so proportional to the number of slices remaining (if the loaf is evenly sliced).

The remaining volume is

$$\begin{aligned} \pi \int_{-a}^{-a+t} y^2\mathrm{d}x &= \pi \int_{-a}^{-a+t} (a^2-x^2)\mathrm{d}x = \pi(a^2x - x^3/3)\Big|_{-a}^{-a+t} \\ &= \pi[(a^2(-a+t) - (-a+t)^3/3] - \pi[(a^2(-a) - (-a)^3/3] = \pi(at^2 - t^3/3). \end{aligned}$$

As a quick check on the algebra, notice that this is zero when $t = 0$ and $4\pi a^3/3$ when $t = 2a$.

Thus $V/A = (3at - t^2)/(6a)$. This is a quadratic curve with zeros at $t = 0$ and $t = 3a$, so it has a maximum at $t = 3a/2$ (by differentiating or otherwise), and where $V/A = 3a/8$. Since we require this ratio to be less than one, we must have $a < 8/3$ metres.

Question 45$(**)$

A tortoise and a hare have a race to the vegetable patch, a distance X kilometres from the starting post, and back. The tortoise sets off immediately, at a steady v kilometers per hour. The hare goes to sleep for half an hour and then sets off at a steady speed V kilometres per hour. The hare overtakes the tortoise half a kilometre from the starting post, and continues on to the vegetable patch, where she has another half an hour's sleep before setting off for the return journey at her previous pace. One and quarter kilometres from the vegetable patch, she passes the tortoise, still plodding gallantly and steadily towards the vegetable patch. Show that

$$V = \frac{10}{4X - 9}$$

and find v in terms of X.

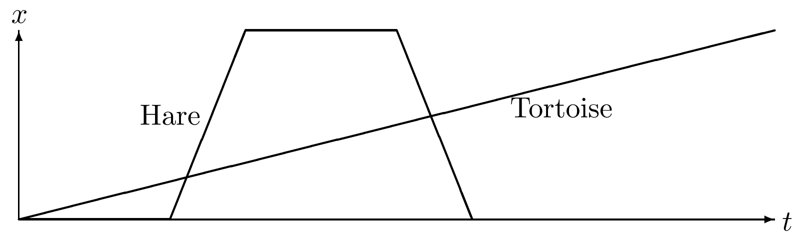
Find X if the hare arrives back at the starting post one and a half hours after the start of the race.

How long does it take the tortoise to reach the vegetable patch?

Comments

You might find it helpful to draw a distance–time diagram. You might also find it useful to let the times at which the two animals meet be T_1 and T_2.

Solution to question 45



Let the times of the first and second meetings be T_1 and T_2. Then

$$vT_1 = \tfrac{1}{2}, \quad V(T_1 - \tfrac{1}{2}) = \tfrac{1}{2}, \quad vT_2 = (X - \tfrac{5}{4}), \quad V(T_2 - 1) = X + \tfrac{5}{4}.$$

The first pair and the second pair of equations give, respectively,

$$\frac{1}{v} - \frac{1}{V} = 1, \qquad \frac{X - 5/4}{v} = \frac{X + 5/4}{V} + 1,$$

and hence (first eliminating v):

$$V = \frac{10}{4X - 9}, \qquad v = \frac{10}{4X + 1}.$$

The total distance travelled by the hare in 1/2 hour is $2X$, so

$$2X = \frac{10}{4X - 9} \times \frac{1}{2}, \quad \text{i.e.} \quad 8X^2 - 18X - 5 = 0.$$

Factorizing the quadratic gives two roots $-1/4$ and $5/2$, and we obviously need the positive root. The speed of the tortoise is given by

$$v = \frac{10}{4X + 1} = \frac{10}{11}$$

so the time taken to travel 5/2 kilometres is 11/4 hours.

Question 46(∗)

A single stream of cars, each of width a and exactly in line, is passing along a straight road of breadth b with speed V. The distance between successive cars (i.e. the distance between back of one car and the front of the following car) is c.

A chicken crosses the road in safety at a constant speed u in a straight line making an angle θ with the direction of traffic. Show that

$$u \geqslant \frac{Va}{c\sin\theta + a\cos\theta}. \qquad (*)$$

Show also that if the chicken chooses θ and u so that it crosses the road at the least possible (constant) speed, it crosses in time

$$\frac{b}{V}\left(\frac{c}{a}+\frac{a}{c}\right).$$

Comments

I like this question because it relates to (an idealized version) of a situation we have probably all thought about. Once you have visualized it, there are no great difficulties. As usual, you have to be careful with the inequalities, though it turns out here that there is no danger of dividing by a negative quantity.

When I arrived at (∗), I noticed that there is another inequality that might be relevant (did you?); but when I looked at the rest of the question I realized I had worried unnecessarily.

Is the time given in the final part (which corresponds to the least possible speed) the longest time that the chicken can spend crossing the road?

Solution to question 46

Let t be the time taken to cross the distance a in which the chicken is at risk. Then $a = ut\sin\theta$.

For safety, the chicken must choose $ut\cos\theta \geqslant Vt - c$: this is the case where the chicken starts at the near-side rear of one car and just avoids being hit by the far-side front of the next car.

Eliminating t from these two equations gives the required inequality:

$$\begin{aligned}
ut\cos\theta &\geqslant Vt - c \\
\implies (u\cos\theta - V)t &\geqslant -c \\
\implies (u\cos\theta - V)\frac{a}{u\sin\theta} &\geqslant -c \\
\implies au\cos\theta - aV &\geqslant -cu\sin\theta \\
\implies u(c\sin\theta + a\cos\theta) &\geqslant aV
\end{aligned}$$

which is the required result

For a given value of θ, the minimum speed satisfies

$$u(c\sin\theta + a\cos\theta) = aV.$$

The smallest value of this u is therefore obtained when $c\sin\theta + a\cos\theta$ is largest. This can be found by calculus (regard it as a function of θ and differentiate: the maximum occurs when $\tan\theta = c/a$) or by trigonometry:

$$c\sin\theta + a\cos\theta = \sqrt{a^2 + c^2}\cos\left(\theta - \arctan(c/a)\right)$$

so the maximum value is $\sqrt{a^2+c^2}$ and it occurs when $\tan\theta = c/a$.

The time of crossing is

$$\frac{b}{u\sin\theta} = \frac{b(c\sin\theta + a\cos\theta)}{Va\sin\theta} = \frac{b(c + a\cot\theta)}{Va} = \frac{b(c + a^2/c)}{Va}.$$

Postmortem

There is another inequality besides $(*)$ that you should have noticed:

$$u\cos\theta t \leqslant Vt + c$$

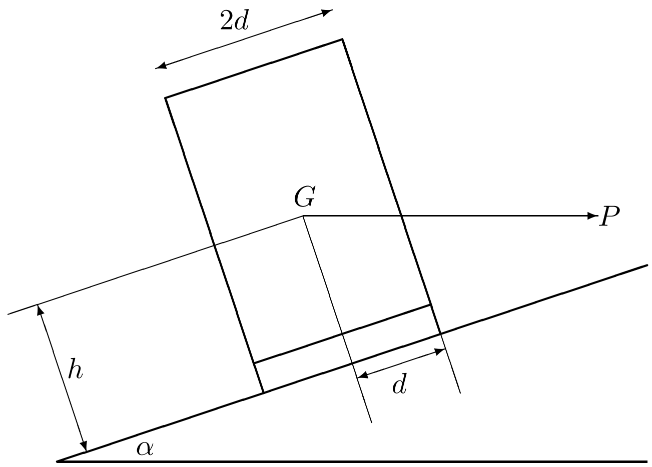
otherwise the chicken will run into the back of the car ahead. This leads to (changing the sign of c and reversing the inequality in $(*)$)

$$u(-c\sin\theta + a\cos\theta) \leqslant aV\ .$$

This looks potentially as if it might also give a lower bound on u (if $(-c\sin\theta + a\cos\theta) < 0$). But it turns out that this is not relevant.

Of course, the chicken can take as long as it wants to cross the road just by going at an extremely acute angle at a speed marginally greater than V. (For maximum time, the minimum value of $u\sin\theta$, not of u) is required.)

Question 47(*)



A lorry of weight W stands on a plane inclined at an angle α to the horizontal. Its wheels are a distance $2d$ apart, and its centre of gravity G is at a distance h from the plane, and halfway between the sides of the lorry. A horizontal force P acts on the lorry through G, as shown.

(i) If the normal reactions on the lower and higher wheels of the lorry are equal, show that the sum of the frictional forces between the wheels and the ground is zero.

(ii) If P is such that the lorry does not tip over (but the normal reactions on the lower and higher wheels of the lorry need not be equal), show that

$$W\tan(\alpha - \beta) \leqslant P \leqslant W\tan(\alpha + \beta) \,,$$

where $\tan\beta = d/h$.

Comments

There is not much more to this than just resolving forces and taking moments about a suitable point.

You might think it a bit odd to have a force that acts horizontally through the centre of gravity of the lorry: it is supposed to be centrifugal. The first draft of the question was intended as a model of a lorry going round a bend on a cambered road. The idea was to relate the speed of the lorry to the angle of camber: the speed should be chosen so that there is be no tendency to skid. However, the wording was rather difficult, because motion in a circle is not included in the STEP syllabus, and the modelling part of the question was eventually abandoned.

You can do the second part algebraically (using equations derived from resolving forces and taking moment), but there is a more direct approach.

There was a slight mistake in the wording of the question, which I have not corrected here. Maybe you will spot it in the course of your solution. The diagram is correct.

Solution to question 47

Let the normal reactions at the lower and upper wheels be N_1 and N_2, respectively, and let the frictional forces at the lower and upper wheels be F_1 and F_2, respectively, (both up the plane).

(i) Taking moments about G, we have

$$(N_1 - N_2)d = (F_1 + F_2)h \qquad (*)$$

so if $N_1 = N_2$ the sum of the frictional forces is zero.

(ii) The condition for the lorry not tip over down the plane is $N_2 \geqslant 0$, which is the same as saying that the total moment of W and P about the point of contact between the lower wheel and the plane in the clockwise sense is positive. This gives one of the inequalities directly (after a bit of geometry to find the distance between the lie of action of the forces and the points of contact of the wheels), and the other follows similarly.

Alternatively, resolving forces perpendicular to the plane and parallel to the plane gives

$$\begin{aligned} N_1 + N_2 &= W\cos\alpha + P\sin\alpha \\ F_1 + F_2 &= W\sin\alpha - P\cos\alpha \end{aligned}$$

Multiplying the first equation by d and the second by h and adding (using $(*)$ gives

$$2dN_1 = W(h\sin\alpha + d\cos\alpha) + P(-h\cos\alpha + d\sin\alpha) \geqslant 0\,,$$

i.e.

$$W\sin(\alpha+\beta) \geqslant P\cos(\alpha+\beta)\,. \qquad (*)$$

Similarly, the first equation by d and the second by h and subtracting (using $(*)$ gives

$$2dN_2 = W(-h\sin\alpha + d\cos\alpha) + P(h\cos\alpha + d\sin\alpha) \geqslant 0\,,$$

i.e

$$P\cos(\alpha-\beta) \geqslant W\sin(\alpha-\beta)\,. \qquad (**)$$

Postmortem

To obtain the required inequalities from $(*)$ and $(**)$ above, you have to divide by a cosine. As always with inequalities, you have to worry about the sign of anything you divide by. In this case, $\cos(\alpha-\beta) > 0$, since both angles are acute. But what about $\cos(\alpha+\beta)$? If this is negative, then $(**)$ is satisfied with no constraint and it would be wrong to divide through the inequality without changing the direction of the inequality. Can $\cos(\alpha+\beta)$ be negative? The answer, which was overlooked by the setter (me), is that it is negative if $\alpha+\beta > \pi/2$. In this case, the line of action of P passes between the wheels of the lorry (not, as shown in the diagram, higher than the higher wheel). Thus increasing P will not cause to the lorry to tip up the slope and there is, in this case, no upper limit to P.

An alternative way of obtaining the inequalities of the second part is to work with the resultant of P and W. The direction of this resultant force (given by $\tan^{-1}(W/P)$) must be such that the line of action of the force passes between the wheels. The line of action makes an angle of $\tan^{-1}(P/W)$ with the vertical through G. The lines through G to points of contact of the upper wheel and the lower wheel with the plane make angles of $\beta+\alpha$ and $\alpha-\beta$, respectively, with the vertical through G. Hence the inequalities in part (ii).

Question 48$(**)$

In a game of cricket, a fielder is perfectly placed to catch a ball. She watches the ball in flight takes the catch just in front of her eye. The angle between the horizontal and her line of sight at a time t after the ball is struck is θ. Show that $\dfrac{\mathrm{d}}{\mathrm{dt}}(\tan\theta)$ is constant during the flight.

The next ball is also struck in the direction of the fielder but at a different velocity. In order to be perfectly placed to catch the ball, the fielder runs at constant speed towards the batsman. Assuming that the ground is horizontal, show that again $\dfrac{\mathrm{d}}{\mathrm{dt}}(\tan\theta)$ is constant during the flight.

Comments

I came across this while looking through some forty year old Entrance Examination papers. Candidates sat these papers in December and were called for interview if they did well enough. This sounds more efficient in terms of everyone's time than the current system in which all applicants are interviewed, but the disadvantage of implementing it now would be that students would have to sit the examination after only four terms in the sixth form. Forty years ago the situation was different because nearly all applicants to Oxford and Cambridge were post A-level.

As with all the very best questions, 9/10 of this question is submerged below the surface. It uses the deepest properties of Newtonian dynamics, and a good understanding of the subject makes the question completely transparent. However, it can still be done without too much trouble by a straightforward approach: take the height above the fielder's eye-level at which the ball is struck to be h, and later eliminate h by writing it in terms of the flight time, the speed and angle of elevation at which the ball is hit.

The cleverness of this question lies in its use of two fundamental invariances of Newton's law

$$\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} = \mathbf{g}. \tag{\dagger}$$

The first is invariance under what is called *time reflection symmetry*. The equation (†) is invariant under the transformation $t \to -t$, because replacing t by $-t$ does not affect the equation. This means that any given solution can be replaced by one where the projectile goes back along the trajectory.

The second is invariance under what are called *Galilean transformations*. Equation (†) is also invariant under the transformation $\mathbf{x} \to \mathbf{x} + \mathbf{v}t$, where $\mathbf{v}$ is an arbitrary constant velocity. This means that we can solve the equation in a frame that moves with constant speed.

Solution to question 48

We take a straightforward approach. Let the height above the fielder's eye-level at which the ball is struck be h. Let the speed at which the ball is struck be u and the angle which the ball makes with the horizontal be α.

Then, taking x to be the horizontal distance of the ball at time t from the point at which the ball was struck and y to be the height at of the ball at time t above the fielder's eye-level, we have

$$x = (u\cos\alpha)t\ , \quad y = h + (u\sin\alpha)t - \tfrac{1}{2}gt^2\ .$$

Let d be the horizontal distance of the fielder from the point at which the ball is struck, and let T be the time of flight of the ball. Then

$$d = (u\cos\alpha)T\ , \quad 0 = h + (u\sin\alpha)T - \tfrac{1}{2}gT^2 \qquad (*)$$

and

$$\begin{aligned}
\tan\theta &= \frac{y}{d-x} \\
&= \frac{h + (u\sin\alpha)t - \frac{1}{2}gt^2}{d - (u\cos\alpha)t} \\
&= \frac{-(u\sin\alpha)T + \frac{1}{2}gT^2 + (u\sin\alpha)t - \frac{1}{2}gt^2}{(u\cos\alpha)T - (u\cos\alpha)t} && \text{(using $*$)} \\
&= \frac{-(u\sin\alpha)(T-t) + \frac{1}{2}g(T^2-t^2)}{(u\cos\alpha)(T-t)} \\
&= \frac{-u\sin\alpha + \frac{1}{2}g(T+t)}{u\cos\alpha}. && \text{(cancelling the factor $(T-t)$)}
\end{aligned}$$

This last expression is a polynomial of degree one in t, so its derivative is constant, as required.

For the second part, let l be the distance from the fielder's original position to the point at which she catches the ball. Then $l = vT$ and

$$\tan\theta = \frac{y}{(l-vt)+d-x} = \frac{y}{v(T-t)+d-x} = \frac{-u\sin\alpha + \frac{1}{2}g(T+t)}{v + u\cos\alpha}$$

cancelling the factor of $(T-t)$ as before. This again has constant derivative.

Postmortem

The invariance mentioned above can be used to answer the question almost without calculation.

Using time reflection symmetry to reverse the trajectory shows that the batsman is completely irrelevant: it only matters that the fielder caught a ball. We just think of the ball being projected from the fielder's hands (the time-reverse of a catch). Taking her hands as the origin of coordinates, and using u to denote the final speed of the ball and α to be the final value of θ, we have $y = (u\sin\alpha)t - \frac{1}{2}gt^2$, $x = (u\cos\alpha)t$ and $\tan\theta = (u\sin\alpha - \frac{1}{2}gt)/u\cos\alpha$. The first derivative of this expression is constant, as before.

We use the Galilean transformation for the second part of the question. Instead of thinking of the fielder running with constant speed v towards the batsman, we can think of the fielder being stationary and the ball having an additional horizontal speed of v. The situation is therefore not changed from that of the first part of the question, except that $u\cos\theta$ should be replaced by $v + u\cos\theta$.

Question 49$(**)$

A rigid straight rod AB has length l and weight W. Its weight per unit length at a distance x from B is $\alpha W l^{-1}(x/l)^{\alpha-1}$, where α is a constant greater than 1. Show that the centre of mass of the rod is at a distance $\alpha l/(\alpha+1)$ from B.

The rod is placed with the end A on a rough horizontal floor and the end B resting against a rough vertical wall. The rod is in a vertical plane at right angles to the plane of the wall and makes an angle of θ with the floor. The coefficient of friction between the floor and the rod is μ and the coefficient of friction between the wall and the rod is also μ. Show that, if the equilibrium is limiting at both A and B, then

$$\tan\theta = \frac{1-\alpha\mu^2}{(1+\alpha)\mu} \,.$$

Given that $\alpha = 3/2$ and given also that the rod slides for any $\theta < \pi/4$ find the greatest possible value of μ.

Comments

This is a pretty standard situation: a rod leaning against a wall, prevented from slipping by friction at both ends. The only slight variation is that the rod is not uniform; the only effect of this is to alter the position of the centre of gravity through which the weight of the rod acts.

The question is not general, since the coefficient of friction is the same at both ends of the rod. I think that it is a good idea (at least, if you are not working under examination conditions) to set out the general equations (with μ_A and μ_B), and at first no limiting friction (with frictional forces F_A and F_B). This will help understand the structure of the equations and highlight the symmetry between the ends of the rod.

You might wonder why the weight per unit length is given in such a complicated way; why not simply kx^α? The reason is to keep the dimensions honest. The factor of W/l appears so that the dimension is clearly weight divided by length. Then x appears divided by l to make it dimensionless. Finally, the factor of α is required for the total weight of the rod is W

Postmortem

In the original version, the question spoke about limiting friction without saying that it was limiting at both ends. We then wondered whether the friction could be limiting at one end only. The answer is yes. The three equations that govern the system do not determine uniquely the four unknown forces (normal reaction and frictional forces at each end of the rod); extra information is required. If the angle is also to be determined, two extra pieces of information are required; in this case the information is that the friction is limiting at both ends.

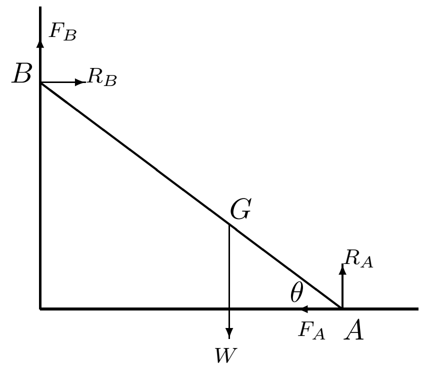
The same lack of determinism occurs in many statics problems. For example, the forces on the legs of a four-legged table cannot all be determined from the three available equations (vertical forces, moment about the x-axis and moments about the y-axis) — only three legs are required for the table to stand (at least,in the case of a three-legged table with an added fourth leg), so extra information is needed to determine all the forces. This extra information might come from considerations of the internal structure of the table; otherwise, the way the table arrived at its current state might matter; for example, whether all four legs were placed on the floor at the same time.

Solution to question 49

The distance, $\bar{x}$, of the centre of mass from the end B of the rod is given by

$$\bar{x} = \frac{\int_0^l \alpha W l^{-\alpha} x^{\alpha}\,\mathrm{d}x}{\int_0^l \alpha W l^{-\alpha} x^{\alpha-1}\,\mathrm{d}x} = \frac{\int_0^l x^{\alpha}\,\mathrm{d}x}{\int_0^l x^{\alpha-1}\,\mathrm{d}x} = \frac{(\alpha+1)^{-1} l^{\alpha+1}}{(\alpha)^{-1} l^{\alpha}} = \frac{\alpha l}{\alpha+1}.$$

The response to the rest of this question should be preceded by an annotated diagram showing all relevant forces on the rod and with G nearer to A than B. This will help to clarify ideas.



The equations that determine the equilibrium are

$$\begin{aligned}
F_B + R_A &= W && \text{(vertical forces)}\\
F_A - R_B &= 0 && \text{(horizontal forces)}\\
R_B\bar{x}\sin\theta + F_B\bar{x}\cos\theta &= R_A(l-\bar{x})\cos\theta - F_A(l-\bar{x})\sin\theta && \text{(clockwise moments about } G\text{)}
\end{aligned}$$

If the rod is about to slip then the frictional and normal forces at A can be specified as $F_A = \mu R_A$ and $F_B = \mu R_B$. Substituting in first two of the above questions gives

$$R_A = \frac{W}{\mu^2+1}, \quad R_B = \frac{\mu W}{\mu^2+1}, \quad F_A = \frac{\mu W}{\mu^2+1}, \quad F_B = \frac{\mu^2 W}{\mu^2+1}.$$

Substituting into the third (moments) equation gives the required result.

For the final part, setting $\alpha = 3/2$ gives

$$\tan\theta = \frac{2-3\mu^2}{5\mu}$$

If the rod slips for any angle less than $\theta = \pi/4$, then the angle at which limiting friction occurs at both ends must be at least $\pi/4$. Therefore

$$\frac{2-3\mu^2}{5\mu} \geqslant 1$$

i.e. $3\mu^2 + 5\mu - 2 \leqslant 0$, or $(3\mu-1)(\mu+2) < 0$. The greatest possible value of μ is $1/3$.

Question 50($**$)

N particles P_1, P_2, P_3, …, P_N with masses m, qm, q^2m, … , $q^{N-1}m$, respectively, are at rest at distinct points along a straight line in gravity-free space. The particle P_1 is set in motion towards P_2 with velocity V and in every subsequent impact the coefficient of restitution is e, where $0 < e < 1$. Show that after the first impact the velocities of P_1 and P_2 are

$$\left(\frac{1-eq}{1+q}\right)V \quad \text{and} \quad \left(\frac{1+e}{1+q}\right)V,$$

respectively.

Show that if $q \leqslant e$, then there are exactly $N-1$ impacts and that if $q = e$, then the total loss of kinetic energy after all impacts have occurred is equal to

$$\frac{1}{2}me\big(1 - e^{N-1}\big)V^2.$$

Comments

This situation models the toy called 'Newton's Cradle' which consists of four or more heavy metal balls suspended from a frame so that they can swing. At rest, they are in contact in a line. When the first ball is raised and let swing, there follows a rather pleasing pattern of impacts. In this case, the coefficient of restitution is nearly 1 and the balls all have the same mass, so, as the first displayed formula shows, the impacting ball is reduced to rest by the impact. At the first swing of the ball, nothing happens except that the first ball is reduced to rest and the last ball swings away. Note that this is consistent with the balls being separated by a very small amount; what actually happens is that the ball undergoes a small elastic deformation at the impact, and the impulse takes a small amount of time to be transmitted across the ball to the next ball. You can order a Newton's Cradle (and lots of other interesting toys and games) online from `www.Hawkin.com`.

Solution to question 50

Using conservation of momentum and Newton's law of impact at the first collision results in the equations

$$mV = mV_1 + mqV_2, \qquad V_1 - V_2 = -eV.$$

Solving these for V_1 and V_2 gives for the speeds of P_1 and P_2 after the collision

$$V_1 = \left(\frac{1-eq}{1+q}\right)V \quad \text{and} \quad V_2 = \left(\frac{1+e}{1+q}\right)V.$$

Note that V_1 is positive since $eq \leqslant e^2 < 1$, and that $V_1 < V$, and that $V < V_2$ since $(1+e) > (1+q)$.

Similarly, the speeds of P_2 and P_3 after the next collision are

$$\left(\frac{1-eq}{1+q}\right)V_2 = \left(\frac{1-eq}{1+q}\right)\left(\frac{1+e}{1+q}\right)V \quad \text{and} \quad \left(\frac{1+e}{1+q}\right)V_2 = \left(\frac{1+e}{1+q}\right)^2 V.$$

Note that the speed of P_2 is positive and greater than V_1 (the speed of P_1) so there is no further collision between P_1 and P_2. Applying this argument at each collision shows that there are exactly $N-1$ collisions: P_1 with P_2; P_2 with P_3; ... , P_{N-1} with P_N.

The speed of P_k after it has hit P_{k+1} is similarly

$$\left(\frac{1-eq}{1+q}\right)\left(\frac{1+e}{1+q}\right)^{k-1} V,$$

and the speed of P_{k+1} after this collision and before it hits P_{k+2} is

$$\left(\frac{1+e}{1+q}\right)^{k} V.$$

If $q = e$,

$$V_1 = \frac{1-e^2}{1+e}V = (1-e)V$$

and similarly the final speeds of $P_2, \ldots, P_{N-1}$ are $(1-e)V$. The final speed of P_N is V. Thus the total kinetic energy is

$$\begin{aligned}\frac{1}{2}(m + me + \cdots + me^{N-2})(1-e)^2V^2 + \frac{1}{2}me^{N-1}V^2 &= \frac{mV^2}{2}\frac{(1-e^{N-1})}{1-e}(1-e)^2 + \frac{mV^2}{2}e^{N-1} \\ &= \frac{mV^2}{2}(1-e+e^N),\end{aligned}$$

(replacing all the q's with e's and summing the geometric progression). Thus the loss of kinetic energy is

$$\tfrac{1}{2}mV^2 - \frac{mV^2}{2}(1-e+e^N),$$

as required.

Question 51($**$)

An automated mobile dummy target for gunnery practice is moving anti-clockwise around the circumference of a large circle of radius R in a horizontal plane at a constant angular speed ω. A shell is fired from O, the centre of this circle, with initial speed V and angle of elevation α. Show that if $V^2 < gR$, then no matter what the value of α, or what vertical plane the shell is fired in, the shell cannot hit the target.

Assume now that $V^2 > gR$ and that the shell hits the target, and let β be the (positive) angle through which the target rotates between the time at which the shell is fired and the time of impact. Show that β satisfies the equation

$$g^2\beta^4 - 4\omega^2V^2\beta^2 + 4R^2\omega^4 = 0.$$

Deduce that there are exactly two possible values of β.

Let β_1 and β_2 be the possible values of β and let P_1 and P_2 be the corresponding points of impact. By considering the quantities $(\beta_1^2 + \beta_2^2)$ and $\beta_1^2\beta_2^2$, or otherwise, show that the linear distance between P_1 and P_2 is

$$2R\sin\left(\frac{\omega}{g}\sqrt{(V^2 - Rg)}\right).$$

Comments

The rotation of the target is irrelevant for the first part, which contravenes the setters' law of not introducing information before it is required. In this case, it seemed better to describe the set-up immediately – especially as you are asked for a familiar result.

Remember, when you are considering the roots of the quartic (which is really a quadratic in β^2) that $\beta > 0$.

The hint in the last paragraph ('by considering ...') is supposed to direct you towards the sum and product of the roots of the quadratic equation; otherwise, you get into some pretty heavy algebra.

Solution to question 51

If the shell hits the target, the range is R, so

$$R = (V\cos\alpha)t, \qquad 0 = (V\sin\alpha)t - \tfrac{1}{2}gt^2.$$

Eliminating t gives $Rg = V^2\sin(2\alpha)$ which can be satisfied, for some value of α, if and only if $V^2 \geqslant Rg$.

The time of flight is $R/(V\cos\alpha)$, and this must also equal the time for the target to rotate through β, i.e. β/ω. Thus $\cos\alpha = R\omega/(\beta V)$. Substituting this into the range equation $Rg = 2V^2\sin\alpha\cos\alpha$ gives

$$Rg = 2V^2\frac{R\omega}{\beta V}\left(1 - \frac{R^2\omega^2}{\beta^2 V^2}\right)^{\frac{1}{2}}.$$

Squaring both sides of the equation, and simplifying, leads to the given quadratic in β^2. Solving the quadratic using the quadratic formula gives

$$g^2\beta^2 = 2\omega^2\left(V^2 \pm \sqrt{V^4 - g^2R^2}\right).$$

Since $V^2 > gR$, both the values of β^2 roots are real and positive, but they lead to only two relevant values of β since we only want positive values.

The linear distance between the two points of impact is $2R\sin\frac{1}{2}(\beta_2 - \beta_1)$. Now $\beta_1^2 + \beta_2^2 = 4\omega^2V^2/g^2$ and $\beta_1^2\beta_2^2 = 4R^2\omega^4/g^2$ (sum and product of the roots of the quadratic), so

$$(\beta_2 - \beta_1)^2 = \beta_2^2 + \beta_1^2 - 2\beta_1\beta_2 = \frac{4\omega^2V^2}{g^2} - \frac{4R\omega^2}{g},$$

which leads to the given result.

Question 52$(***)$

A particle of unit mass is projected vertically upwards with speed u. At any height x, while the particle is moving upwards, it is found to experience a total force F, due to gravity and air resistance, given by $F = \alpha e^{-\beta x}$, where α and β are positive constants. Calculate the energy expended in reaching a given height z. Show that

$$F = \tfrac{1}{2}\beta v^2 + \alpha - \tfrac{1}{2}\beta u^2 \ ,$$

where v is the speed of the particle, and explain why $\alpha = \frac{1}{2}\beta u^2 + g$, where g is the acceleration due to gravity.

Determine an expression, in terms of y, g and β, for the air resistance experienced by the particle on its downward journey when it is at a distance y below its highest point.

Comments

This is rather ingenious. The usual approach to this sort of problem is to say at the outset the particle experiences the force due to gravity and also a force due to air resistance of the form kv^2. In this question, the total force is given and you have to deduce the constituent parts of the force, making good use of the basic laws of mechanics (change of energy equals work done, for example).

I remember the discussion of the phrasing of the next part of the question 'explain why $\alpha = \frac{1}{2}\beta u^2 + g$'. How best to convey the idea that no real calculation is required, while not suggesting that the answer is trivially obvious? (All that is required, since we are trying to evaluate a constant, is to take a particular value of v.) I'm not sure 'explain why' was right; but I can't think of anything better.

For the last part, you have to integrate the equation of motion, using the standard form of the acceleration used when there is y dependence instead of t dependence.

Solution to question 52

The energy expended in moving a small distance $\mathrm{d}x$ is $F\mathrm{d}x$ so the energy expended in reaching height z is

$$\int_0^z \alpha \mathrm{e}^{-\beta x} \mathrm{d}x\,, \qquad \text{i.e.} \quad \frac{\alpha(1-\mathrm{e}^{-\beta z})}{\beta}\,.$$

By conservation of energy, the loss of kinetic energy is equal to the work done against the force in reaching this height, so

$$\frac{1}{2}u^2 - \frac{1}{2}v^2 = \frac{\alpha(1-\mathrm{e}^{-\beta z})}{\beta}\,,$$

where v is the speed at height y and

$$\mathrm{e}^{-\beta z} = 1 - \frac{\beta(\frac{1}{2}u^2 - \frac{1}{2}v^2)}{\alpha}\,.$$

The force experienced at this height is therefore given by

$$F \equiv \alpha e^{-\beta z} = \alpha - \beta(\frac{1}{2}u^2 - \frac{1}{2}v^2)$$

as required.

To obtain an expression for the constant α we just have to evaluate the above equation at a value of v for which F is known. The obvious choice is $v=0$: at maximum height the speed is zero so there is no air resistance, the only force on the particle being the force due to gravity. Thus $F=-g$ when $v=0$, which gives the required result.

Substituting for α in F gives

$$F = -g + \tfrac{1}{2}\beta v^2\,,$$

so the air resistance to the particle when it is moving at speed v (up or down) is equal to $\frac{1}{2}\beta v^2$.

For the last part, we just have to express $\frac{1}{2}\beta v^2$ in terms of y, β and g.

For the downward journey, let $v = \dfrac{\mathrm{d}y}{\mathrm{d}t}$. Then

$$v\frac{\mathrm{d}v}{\mathrm{d}y} = g - \tfrac{1}{2}\beta v^2 \qquad \text{i.e.} \qquad \int \frac{v}{g - \frac{1}{2}\beta v^2}\mathrm{d}v = \int \mathrm{d}y.$$

Integrating, and using the initial condition that $v=0$ when $y=0$, gives

$$y = -\frac{1}{\beta}\ln(g - \tfrac{1}{2}\beta v^2) + \frac{1}{\beta}\ln(g) = -\frac{1}{\beta}\ln\left(\frac{g - \frac{1}{2}\beta v^2}{g}\right)\,.$$

Thus $\frac{1}{2}\beta v^2 = g\left(1 - \mathrm{e}^{-\beta y}\right)$, which is therefore the air resistance.

Question 53(∗∗)

Hank's Gold Mine has a very long vertical shaft of height l. A light chain of length l passes over a small smooth light fixed pulley at the top of the shaft. To one end of the chain is attached a bucket A of negligible mass and to the other a bucket B of mass m.

The system is used to raise ore from the mine as follows. When bucket A is at the top it is filled with mass $2m$ of water and bucket B is filled with mass λm of ore, where $0 < \lambda < 1$. The buckets are then released, so that bucket A descends and bucket B ascends. When bucket B reaches the top both buckets are emptied and released, so that bucket B descends and bucket A ascends. The time to fill and empty the buckets is negligible. Find the time taken from the moment bucket A is released at the top until the first time it reaches the top again.

This process goes on for a very long time. Show that, if the greatest amount of ore is to be raised in that time, then λ must satisfy the condition $\mathrm{f}'(\lambda) = 0$ where

$$\mathrm{f}(\lambda) = \frac{\lambda(1-\lambda)^{1/2}}{(1-\lambda)^{1/2} + (3+\lambda)^{1/2}}.$$

Comments

One way of working out the acceleration of a system of two masses connected by a light string passing over a pulley is to write down the equation of motion of each mass, bearing in mind that the force due to tension will be the same for each mass (it cannot vary along the string, because then the acceleration of some portion of the *massless* string would be infinite). Then you eliminate the tension.

Alternatively, you can use the equation of motion of the system of two joined masses. The system has inertial mass equal to the sum of the masses (because both masses must accelerate equally) but gravitational mass equal to the difference of the masses (because the gravitational force on one mass cancels, partially, the gravitational force on the other), so the equation of motion is just (Newton's law of motion)

$$(m_1 + m_2)a = (m_1 - m_2)g.$$

In the last part, the question says that the process goes on for a very long time. The reason for this is just so that you don't have to worry about whether it is advantageous to stop in mid-cycle: the effect of doing so after completing many cycles is insignificant. It is therefore sufficient to choose λ so as to maximise the rate at which ore is raised in just one cycle.

Solution to question 53

For bucket A's downward journey, at acceleration a, the equations of motion for bucket A and bucket B, respectively, are

$$-T + 2mg = 2ma, \qquad T - (1+\lambda)mg = (1+\lambda)ma,$$

where T is the tension in the rope. Eliminating T gives so $a = g(1-\lambda)/(3+\lambda)$.

When bucket A ascends, the acceleration is g.

The time of descent (using $l = \frac{1}{2}at^2$) is $\sqrt{2l/a}$ and the time of ascent is $\sqrt{2l/g}$. The total time required for one complete cycle is therefore

$$\sqrt{\frac{2l}{g}}\left(1+\sqrt{\frac{3+\lambda}{1-\lambda}}\right).$$

Call this t. The number of round trips in a long time $t_{\rm long}$ is $t_{\rm long}/t$ so the amount of ore lifted in time $t_{\rm long}$ is $\lambda m t_{\rm long}/t$.

To maximise this, we have to maximise λ/t with respect to λ, and λ/t is exactly the $\mathrm{f}(\lambda)$ given.

Postmortem

If you are feeling exceptionally energetic, you could try to find the value of λ for which $\mathrm{f}'(\lambda) = 0$. This will eventually lead you to the rather discouraging quartic equation

$$\lambda^4 + 4\lambda^3 + 2\lambda^2 - 8\lambda + 3 = 0.$$

A quick look at the quartic equation solver `http://www.1728.com/quartic.htm` reveals, mysteriously, two real solutions: 0.7037579200493 and 0.5280411216973; 'mysteriously' because we were expecting, from the wording of the question, only one solution. In any case, there should be an odd number of turning points between $\lambda = 0$ and $\lambda = 1$, because $\mathrm{f}(\lambda) = 0$ when $\lambda = 0$ and $\lambda = 1$ and $\mathrm{f}(\lambda) > 0$ for $0 < \lambda < 1$.

It turns out that in order to remove surds, it was necessary to square an expression and this introduced an extra fictitious solution. The larger of the two solutions quoted above is in fact the correct one, as can be ascertained by substituting both back into the original equation.

Question 54($*$)

A smooth cylinder with circular cross-section of radius a is held with its axis horizontal. A light elastic band of unstretched length $2\pi a$ and modulus of elasticity λ is wrapped round the circumference of the cylinder, so that it forms a circle in a plane perpendicular to the axis of the cylinder. A particle of mass m is then attached to the rubber band at its lowest point and released from rest.

(i) Given that the particle falls to a distance $2a$ below the below the axis of the cylinder, but no further, show that

$$\lambda = \frac{9\pi mg}{(3\sqrt{3} - \pi)^2} \ .$$

(ii) Given instead that the particle reaches its maximum speed at a distance $2a$ below the axis of the cylinder, find a similar expression for λ.

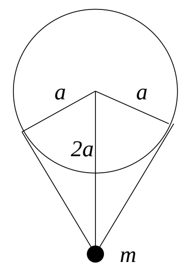
Comments

This question uses the most basic ideas in mechanics, such as conservation of energy. I included it in this selection of STEP questions without realising that the properties of stretched strings are not in the syllabus (which is the syllabus for STEP I and II given in the appendix). However, I didn't throw it out: it is a nice question and the only two things you need to know about stretched strings are, for a stretched string of natural (i.e. unstretched) length l and extended length $l + x$ with modulus of elasticity λ:

(i) The potential energy stored in the stretched string is $\dfrac{\lambda x^2}{2l}$

(ii) The tension in the stretched string is $\dfrac{\lambda x}{l}$.

Solution to question 54



The length of the extended band is

$$2\pi a - 2a\arccos(1/2) + 2\sqrt{3}\,a$$

so the extension is

$$-2a\arccos(1/2) + 2\sqrt{3}\,a\ , \qquad \text{i.e.} \qquad 2a(\sqrt{3} - \pi/3)\,.$$

(i) By conservation of energy (taking the initial potential energy to be zero), when the particle has fallen a distance d and has speed v,

$$0 = \frac{1}{2}mv^2 + \frac{\lambda}{2}\frac{[2a(\sqrt{3} - \pi/3)]^2}{2\pi a} - mgd\,.$$

At the lowest point, $v = 0$ and $d = a$, which gives the required answer.

(ii) The (vertical) acceleration ($\ddot{y}$) of the particle when it has fallen a distance d is given by

$$m\ddot{y} = 2T\cos\theta - mg$$

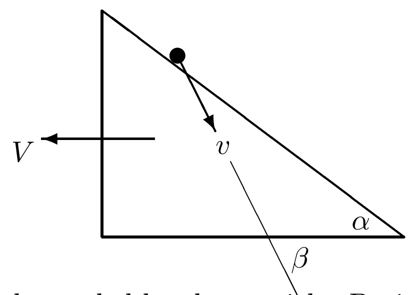
where $\sin\theta = a/(a + d)$ and T is the tension. The maximum speed occurs when $\ddot{y} = 0$ and $d = a$ as before, i.e. when

$$2\frac{2a(\sqrt{3} - \pi/3)\lambda}{2\pi a}\frac{\sqrt{3}}{2} - mg = 0$$

so that in this case $\lambda = \sqrt{3}\pi mg/(3\sqrt{3} - \pi)\,.$

Question 55(∗∗)

A wedge of mass M rests on a smooth horizontal surface. The face of the wedge is a smooth plane inclined at an angle α to the horizontal. A particle of mass kM slides down the face of the wedge, starting from rest. At a later time t, the speed V of the wedge, the speed v of the particle and the angle β of the velocity of the particle below the horizontal are as shown in the diagram.



Let y be the vertical distance descended by the particle. Derive the following results, stating in (ii) and (iii) the mechanical principles you use:

(i) $V \sin\alpha = v \sin(\beta - \alpha)$;

(ii) $\tan\beta = (1+k)\tan\alpha$;

(iii) $2gy = v^2(1 + k\cos^2\beta)$.

Write down a differential equation for y and hence show that

$$y = \tfrac{1}{2}g't^2$$

where

$$g' = \frac{g\sin^2\beta}{(1+k\cos^2\beta)} = \frac{(1+k)\sin^2\alpha}{k+\sin^2\alpha}.$$

Comments

I was surprised how difficult this problem is, compared with a particle on a fixed wedge. I don't think that there is an easier method: the numbered parts of the question (i) – (iii) seem unavoidable, since they embody the basic principles necessary to solve the problem. For a fixed wedge, the horizontal component of momentum is not conserved, because of the force required to hold the wedge; but (i) and (iii) are just what you would use in the fixed case.

Part (ii) is used only (apart from the very last result) to show that the angle β is constant. This is a bit surprising: it means that the particle moves in a straight line. To put it another way, the net force on the particle is parallel to the direction of motion. Why is this?

Most of the results of the important intermediate steps are given to you, but it is still very good discipline to check that they hold in special cases: for example, $k = 0$ corresponding to a massless particle or equivalently a fixed wedge; and $\alpha = 0$ or $\alpha = \pi/2$ corresponding to a horizontal or vertical wedge face. You should check that you understand what should happen in these special cases and that your understanding is consistent with the formulae.

Solution to question 55

(i) Particle remains in contact with the plane so components of the velocities of the particle and the wedge perpendicular to the face must be equal.

(ii) Horizontal momentum is conserved, so $MV = mv\cos\beta$. Substitution for V/v using (i) gives

$$k\cos\beta = \frac{V}{v} = \frac{\sin(\beta-\alpha)}{\sin\alpha} = \frac{\sin\beta\cos\alpha - \sin\alpha\cos\beta}{\sin\alpha} = \frac{\sin\beta}{\tan\alpha} - \cos\beta$$

which leads immediately to the required result. Note that (rather surprisingly) the angle β remains constant in the motion.

(iii) Conservation of energy gives

$$mgy = \frac{1}{2}MV^2 + \frac{1}{2}mv^2, \qquad \text{i.e.} \qquad kgy = \frac{1}{2}(kv\cos\beta)^2 + \frac{1}{2}kv^2.$$

Thus $2gy = (k\cos^2\beta + 1)v^2$.

The next step is to write v in terms of y to obtain the differential equation. The vertical component of velocity is $v\sin\beta$, so

$$v\sin\beta = \frac{\mathrm{d}y}{\mathrm{d}t}\ .$$

Now using the result of (iii) gives

$$\frac{\mathrm{d}y}{\mathrm{d}t} = v\sin\beta = \sqrt{\frac{2gy}{1+k\cos^2\beta}}\,\sin\beta = (2g')^{\frac{1}{2}}y^{\frac{1}{2}}.$$

Now, at last, we use the hard-won result that β, and hence g', is constant:

$$y^{-\frac{1}{2}}\frac{\mathrm{d}y}{\mathrm{d}t} = (2g')^{\frac{1}{2}} \Rightarrow 2y^{\frac{1}{2}} = (2g')^{\frac{1}{2}}t + \text{constant}.$$

The constant is zero, since y is the vertical distance descended by the particle. Squaring gives the required answer.

There is one final trigonometric manipulation to perform, using (ii) to convert β (the unknown angle) to α (the given angle). The methodical way (not necessarily the quickest) is to write all the trigonometric functions in terms of tan or cot:

$$\begin{aligned} g' &\equiv \frac{\sin^2\beta}{1+k\cos^2\beta} = \frac{\tan^2\beta}{\sec^2\beta + k} = \frac{\tan^2\beta}{\tan^2\beta + k + 1} \\ &= \frac{(1+k)^2\tan^2\alpha}{(1+k)^2\tan^2\alpha + k + 1} = \frac{(1+k)\tan^2\alpha}{(1+k)\tan^2\alpha + 1} = \frac{(1+k)\sin^2\alpha}{(1+k)\sin^2\alpha + \cos^2\alpha} = \frac{(1+k)\sin^2\alpha}{k\sin^2\alpha + 1} \end{aligned}$$

Postmortem

Why is β constant? There are two forces acting on the particle: gravity and the reaction of the wedge on the particle. One compenent of the reaction must exactly balance one component of gravity so as to keep the particle on the wedge; the only remaining components of both forces must be perpendicular to this direction and therefore in direction of motion of the particle. You can appreciate this argument best be assuming that the particle and wedge are instantaneously at rest, which is allowable because the argument is about forces acting at a given instant and the forces are unchanged by addition of a constant velocity ('Galilean invariance').

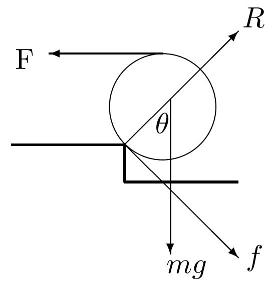
Question 56($*$)

A uniform solid sphere of radius a and mass m is drawn very slowly and without slipping from horizontal ground onto a step of height $a/2$ by a horizontal force of magnitude F which is always applied to the highest point of the sphere and is always perpendicular to the vertical plane which forms the face of the step. Find the maximum value of F in the motion, and prove that the coefficient of friction between the sphere and the edge of the step must exceed $1/\sqrt{3}$.

Comments

This is quite straightforward once you have realised that *very slowly* means so slowly that the sphere can be considered to be static at each position. You just have to solve a statics problem with the usual tools: resolving forces and taking moments.

Solution to question 56



Let the angle between the radius to the point of contact with the step and the downward vertical be θ, as shown.

Taking moments about the point of contact with the step gives

$$Fa(1 + \cos\theta) = mga \sin\theta,$$

so

$$F = mg\frac{\sin\theta}{1 + \cos\theta} = mg \tan(\theta/2).$$

This takes its maximum value when θ is largest, i.e. when the sphere just touches the hoqizontal ground. At this position, $\cos\theta = 1/2$ and $\sin\theta = \sqrt{3}/2$ and $F_{\max} = mg/\sqrt{3}$.

At the point of contact, let the frictional force, which is tangent to the sphere, be f and the reaction R (along the radius of the sphere). Taking moments about the centre of the sphere gives $F = f$, and resolving forces parallel to the radius at the point of contact gives

$$R = F \sin\theta + mg \cos\theta.$$

Now

$$F \sin\theta = mg\frac{\sin^2\theta}{1 + \cos\theta} = mg\frac{1 - \cos^2\theta}{1 + \cos\theta} = mg(1 - \cos\theta)$$

so $R = mg$. We therefore need $\mu mg > F_{\max}$.

Postmortem

To be precise about the meaning of *slow* in this context, you have to compare the dynamic forces connected with the motion of the sphere with the static forces. The motion of the sphere is rotation about the fixed point of contact with the step. If we assume that the centre moves with constant speed, the only extra force due to the motion is an additional reaction at the point of contact with the step. (No tangential force is required to maintain a constant speed in circular motion.) This reaction is centrifugal in nature and so is of the form kmv^2/d where d is some number which might be hard to calculate. The force calculated above is roughly mg, so the static approximation ('very slowly') is valid if $v^2/d \ll g$.

Question 57(∗∗)

Harry the Calculating Horse will do any mathematical problem I set him, providing the answer is 1, 2, 3 or 4. When I set him a problem, he places a hoof on a large grid consisting of unit squares and his answer is the number of squares partly covered by his hoof. Harry has circular hoofs, of radius 1/4 unit.

After many years of collaboration, I suspect that Harry no longer bothers to do the calculations, instead merely placing his hoof on the grid completely at random. I often ask him to divide 4 by 4, but only about 1/4 of his answers are right; I often ask him to add 2 and 2, but disappointingly only about $\pi/16$ of his answers are right. Is this consistent with my suspicions?

I decide to investigate further by setting Harry many problems, the answers to which are 1, 2, 3, or 4 with equal frequency. If Harry is placing his hoof at random, find the expected value of his answers. The average of Harry's answers turns out to be 2. Should I get a new horse?

Comments

Hans von Osten, a horse, lived in Berlin around the turn of the last century. He was known far and wide for his ability to solve complex arithmetical problems. Distinguished scientists travelled to Berlin to examine Hans and test his marvellous ability. They would write an equation on a chalkboard and Hans would respond by pawing the ground with his hoof. When Hans reached the answer he would stop. Though he sometimes made errors, his success rate was far higher than would be expected if his answers were random. The accepted verdict was that Hans could do arithmetic.

Hans's reputation as a calculating horse nosedived when an astute scientist simply made sure that neither the person asking the questions nor the audience knew the answers. Hans became an instant failure. His success was based on his ability to sense any change in the audience: a lifted eyebrow, a sigh, a nodding head or the tensing of muscles was enough to stop him from pawing the ground. Anyone who knew the answer was likely to give almost imperceptible clues to the horse. But we shouldn't overlook Hans's talents: at least he had terrific examination technique.

We are investigating the situation when Harry places his hoof at random, so that the probability of the centre of his hoof lands in any given region is proportional to the area of the region. We therefore move swiftly from a question about probability to a question about areas. You just have to divide a given square into regions. each determined by the number of squares that will be partially covered by Harry's hoof if its centre lands in the region under consideration. Remember that the areas must add to one, so the most difficult area calculation can be left until last and deduced from the others.

To answer the last part properly, you really need to set out a hypothesis testing argument: you will accept the null hypothesis (random hoof placing) if, using the distribution implied by the null hypothesis, the probability of obtaining the given result is greater than some predecided figure. Obviously, nothing so elaborate was intended here, since it is just the last line of an already long question: just one line would do.

Solution to question 57

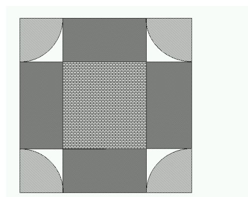
Harry's hoof will be completely within exactly one square if he places the centre of his hoof in a square of side $1/2$ centred on the centre of any square. The area of any such smaller square is $1/4$ square units.

His hoof will partially cover exactly four squares if he places the centre in a circle of radius $1/4$ centred on any intersection of grid lines. The total area of any one such circle is $\pi/16$ square units, which may be thought of as four quarter circles, one in each corner of any given square.

His hoof will partially cover exactly two squares if he places the centre in any one of four $1/2$ by $1/4$ rectangles, of total area $4 \times \dfrac{1}{8}$ in any given square.

Otherwise, his hoof will partially cover three squares; the area of this remaining region is

$$1 - \frac{1}{4} - \frac{\pi}{16} - \frac{1}{2} \,.$$



If Harry placed his hoof at random, the probabilities of the different outcomes would be equal to the corresponding area calculated above (divided by the total area of the square, which is 1). Thus the data given in the question are consistent with random placement.

The expected value given random placements is

$$1 \times \frac{1}{4} + 2 \times \frac{1}{2} + 3 \times \left(1 - \frac{1}{4} - \frac{1}{2} - \frac{\pi}{16}\right) + 4 \times \frac{\pi}{16} = 2 + \frac{\pi}{16}$$

The expected value (given that Harry gets all questions right) is $(1+2+3+4)/4 = 5/2$. Harry has a less accurate expected value even than the random expected value. He is clearly hopeless and should go.

Question 58$(*)$

In order to get money from a cash dispenser I have to punch in an identification number. I have forgotten my identification number, but I do know that it is equally likely to be any one of the integers 1, 2, $\ldots$, n. I plan to punch in integers in order until I get the right one. I can do this at the rate of r integers per minute. As soon as I punch in the first wrong number, the police will be alerted. The probability that they will arrive within a time t minutes is $1 - e^{-\lambda t}$, where λ is a positive constant. If I follow my plan, show that the probability of the police arriving before I get my money is

$$\sum_{k=1}^{n} \frac{1 - e^{-\lambda(k-1)/r}}{n} \, .$$

Simplify the sum.

On past experience, I know that I will be so flustered that I will just punch in possible integers at random, without noticing which I have already tried. Show that the probability of the police arriving before I get my money is

$$1 - \frac{1}{n - (n-1)e^{-\lambda/r}} \, .$$

Comments

This was originally about getting money from cash dispenser using a stolen card, but it was decided that STEP questions should not be immoral. Hence the rather more improbable scenario.

The *exponential distribution*, here governing the police arrival time is often used for failure of equipment (light-bulbs, etc). It has the useful property (not used here) that the reliability for a light bulb (here the probability of the police not coming within a certain time period of duration t doesn't depend on which time period you choose; i.e. given that the light bulb has survived to time T, the probability of it surviving until time $T + t$ is independent of T (which doesn't seem very suitable for light bulbs). This is referred to as the *memoryless property*.

To simplify the sums, you have to recognise a geometric series where the common ratio of terms is an exponential; not difficult, but easy to miss the first time you see it.

Solution to question 58

The probability that I get the right number on the kth go is $1/n$. The time taken to key in k integers is $(k-1)/r$. The probability that the police arrive before this is $1 - e^{-\lambda(k-1)/r}$ so the total probability that the police arrive before I get my money is

$$\sum_{k=1}^{n} \frac{1}{n} \times \left(1 - e^{-\lambda(k-1)/r}\right) .$$

We have

$$\sum_{k=1}^{n} \frac{1}{n} \times \left(1 - e^{-\lambda(k-1)/r}\right) = 1 - \sum_{k=1}^{n} \frac{e^{-\lambda(k-1)/r}}{n} = 1 - \frac{1 - e^{-\lambda n/r}}{n(1 - e^{-\lambda/r})} ,$$

since the last part of the sum is a geometric progression with common ratio $e^{-\lambda/r}$.

This time, the probability of getting my money on the kth go is $1/n$ times the probability of not having punched in the correct integer in the preceding $k-1$ turns:

$$\mathrm{P}(\text{money on } k\text{th go}) = \frac{1}{n} \times \left(\frac{n-1}{n}\right)^{k-1} .$$

Thus the probability of the police arriving before I get my money is

$$\sum_{k=1}^{\infty} \frac{1}{n} \times \left(\frac{n-1}{n}\right)^{k-1} \times \left(1 - e^{-\lambda(k-1)/r}\right) = \frac{1}{n} \times \frac{1}{1-(n-1)/n} - \frac{1}{n} \times \frac{1}{1 - e^{-\lambda/r}(n-1)/n} .$$

Postmortem

For each part, you have to calculate the probablity of the first success occuring on the kth go. You have to remember to multiply the probability of success on this go by the probability of failure on the preceding $k-1$ goes — a very typical idea in this sort of probability question. After that, the question is really just algebra. STEP questions on this sort of material nearly always involve significant algebra or calculus.

Car batteries as well as light bulbs often crop up as examples of objects to which exponential reliability can be applied. Suppose your car battery has a guarantee of three years, and your car is completely destroyed after two years. Your insurance company offers to give you one third of the price of the battery. What do you think they would say to your counterclaim for the full price, based on the fact that, given it had lasted two years, the exponential model says that it would have had an expected further three years life in it?

Question 59(*)

Four students, Arthur, Bertha, Chandra and Delilah, exchange gossip. When Arthur hears a rumour, he tells it to one of the other three without saying who told it to him. He decides whom to tell by choosing at random amongst the other three, omitting the ones that he knows have already heard the rumour. When Bertha, Chandra or Delilah hear a rumour, they behave in exactly the same way (even if they have already heard it themselves). The rumour stops being passed round when it is heard by a student who knows that the other three already have aready heard it.

Arthur starts a rumour and tells it to Chandra. By means of a tree diagram, or otherwise, show that the probability that Arthur rehears it is $3/4$.

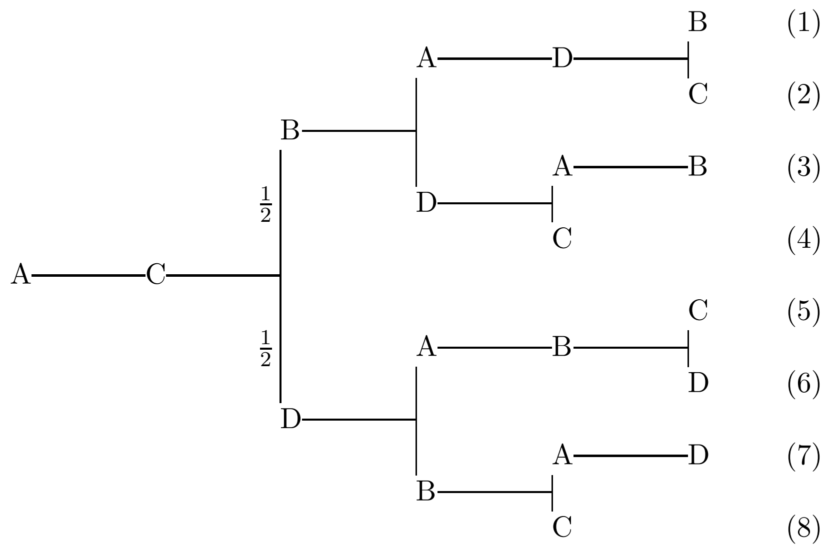
Find also the probability that Bertha hears it twice and the probability that Chandra hears it twice.

Comments

This is pretty straightforward, though you have to marshall the information carefully. Arthur, the hero of this question, is named after one of my sons, who was born in the year this was set. I can't remember who Bertha and co are. Next time this comes up, you can expect five students since I now have another son called Edward. But don't hold your breath: the tree diagram is so horrible to produce in LaTeX (the mathematical word-processing language in which these papers, and the solutions, are set) that such questions are rare.

Solution to question 59

The probability tree is



Each of the branches has probablility $\frac{1}{2}$, so the probability of each outcome is $\frac{1}{8}$.

A rehears the rumour on 1, 2, 3, 5, 6, and 7, so with probability $\frac{6}{8}$.

B hears the rumour twice on 1 and 3, so with probability $\frac{2}{8}$.

C hears the rumour twice on 2, 4, 5 and 8, so with probability $\frac{4}{8}$.

Postmortem

I've nothing to add here. The question has not much underlying interest. So here is the source code, mentioned in the comment on the previous page, for the tree diagram:

```
\addtolength{\tabcolsep}{-2.1 mm}
\newcommand{\lin}{\rule[1mm]{15mm}{0.2mm}}
\begin{tabular}{llllllllllllllll}
&&&&&&&&&&B&\hspace{1cm}(1) \\
&&&&&&A&\lin&D&\lin&\vline \\
&&&&&&\vline&&&&C &\hspace {1cm}(2)\\
&&&&B&\lin&\vline \\
&&&&\vline&&\vline&&A&\lin&B &\hspace{1cm}(3)\\\
&&&\mbox{\hspace{1.2cm}$\frac12$} &\vline&&D&\lin&\vline\\
&&&&\vline&&&&C&&&\hspace{1cm}(4)\\
A&\lin&C&\lin&\vline\\
&&&&\vline&&&&&&C&\hspace{1cm}(5)\\
&&&\mbox{\hspace{1.2cm}$\frac12$}&\vline&&A&\lin&B&\lin&\vline \\
&&&&\vline&&\vline&&&&D &\hspace{1cm}(6)\\
&&&&D&\lin&\vline&\\
&&&&&&\vline&&A&\lin&D&\hspace{1cm}(7)\\
&&&&&&B&\lin&\vline\\
&&&&&&&&C&&&\hspace{1cm}(8)\\
\end{tabular}
```

Question 60$(*)$

Four students, one of whom is a mathematician, take turns at washing up over a long period of time. The number of plates broken by any student in this time obeys a Poisson distribution, the probability of any given student breaking n plates being $e^{-\lambda}\lambda^n/n!$ for some fixed constant λ, independent of the number of breakages by other students. Given that five plates are broken, find the probability that three or more were broken by the mathematician.

Comments

The way this is set up, it is largely a counting exercise (but see postmortem). To start with, you work out the probability of five breakages, then follow that with the probability that the mathematician broke three of more plates. You need to calculate the number of ways that 5 plates can be shared amongst four students, for which you have to consider each *partition* of the number 5 and the number of different ways it can arise.

Solution to question 60

First we work out the probability of five breakages.

Let P(5, 0, 0, 0) denote the probability that student A breaks 5 plates and students B, C and D break no plates. Then

$$\mathrm{P}(5,0,0,0) = (\text{Prob. of breaking 5}) \times (\text{Prob of breaking none})^3 = \frac{\mathrm{e}^{-\lambda}\lambda^5}{5!}\left(\mathrm{e}^{-\lambda}\right)^3 = \frac{\lambda^5\mathrm{e}^{-4\lambda}}{5!}.$$

The probability that one student breaks all the plates is $\mathrm{P}(5,0,0,0)+\mathrm{P}(0,5,0,0)+\mathrm{P}(0,0,5,0)+\mathrm{P}(0,0,0,5)$, i.e. $(4/5!)\times\lambda^5\mathrm{e}^{-4\lambda}$ (the factor 4 because it could be any one of the four students).

Let P(4, 1, 0, 0) denote the probability that student A breaks 4 plates. student B breaks one plate and students C and D break no plate. Then

$$\begin{aligned}\mathrm{P}(4,1,0,0) &= (\text{Prob. of breaking 4}) \times (\text{Prob. of breaking one}) \times (\text{Prob of breaking none})^2 \\ &= \frac{\mathrm{e}^{-\lambda}\lambda^4}{4!}\,\frac{\mathrm{e}^{-\lambda}\lambda}{1!}\left(\mathrm{e}^{-\lambda}\right)^2 = \frac{\lambda^5\mathrm{e}^{-4\lambda}}{4!}\,.\end{aligned}$$

The probability that any one student breaks four plates and any other student breaks one plate is therefore $12\times\lambda^5\mathrm{e}^{-4\lambda}/4!$.

Considering P(3, 2, 0, 0), P(3, 1, 1, 0), P(2, 2, 1, 0) and P(2, 1, 1, 1) in turn, together with the number of different ways these probabilities can occur, shows that the probability of five breakages is

$$\lambda^5\mathrm{e}^{-4\lambda}\left(\frac{4}{5!}+\frac{12}{4!\,1!}+\frac{12}{3!\,2!}+\frac{12}{3!\,1!\,1!}+\frac{12}{2!\,2!\,1!}+\frac{4}{2!\,1!\,1!\,1!}\right) = \lambda^5\mathrm{e}^{-4\lambda}\left(\frac{1024}{5!}\right)\,.$$

The probability that the mathematician breaks three or more is $\mathrm{P}(5,0,0,0)+3\mathrm{P}(4,1,0,0)+3\mathrm{P}(3,2,0,0)+3\mathrm{P}(3,1,1,0)$, i.e.

$$\lambda^5\mathrm{e}^{-4\lambda}\left(\frac{4}{5!}+\frac{3}{4!\,1!}+\frac{3}{3!\,2!}+\frac{4}{3!\,1!\,1!}\right) = \lambda^5\mathrm{e}^{-4\lambda}\left(\frac{106}{5!}\right)\,.$$

The probability that the mathematician breaks three or more, given that five are broken is therefore 106/1024.

Postmortem

If you did the question by the method suggested above, you will probably be wondering about the answer: why is it independent of λ; and why is the denominator 4^5. You will quickly decide that the Poisson distribution was a red herring (though I promise you that it was not an intentional herring), since there is no trace of the distribution in the answer.

The denominator is pretty suggestive. A completely different approach is as follows, It is clear that the probability that the mathematician breaks any given plate is $\frac{1}{4}$. The probability that he or she breaks k plates is therefore binomial:

$$\binom{5}{k}\left(\frac{1}{4}\right)^k\left(\frac{3}{4}\right)^{5-k}$$

and adding the cases $k=3$, $k=4$ and $k=5$ gives rather rapidly

$$\frac{90+15+1}{1024}.$$

Question 61(∗∗)

The national lottery of Ruritania is based on the positive integers from 1 to N, where N is very large and fixed. Tickets cost £1 each. For each ticket purchased, the punter (i.e. the purchaser) chooses a number from 1 to N. The winning number is chosen at random, and the jackpot is shared equally amongst those punters who chose the winning number.

A syndicate decides to buy N tickets, choosing every number once to be sure of winning a share of the jackpot. The total number of tickets purchased in this draw is $3.8N$ and the jackpot is £W. Assuming that the non-syndicate punters choose their numbers independently and at random, find the most probable number of winning tickets and show that the expected net loss of the syndicate is approximately

$$N - \frac{5\big(1-e^{-2.8}\big)}{14}\,W\,.$$

Comments

This is a binomial distribution problem: the probability that n out of m punters choose the winning ticket is

$$\binom{m}{n}p^n q^{m-n}$$

where here $m = 2.8N$, $p = 1/N$ and $q = 1-p$. It is clear that an approximation to the binomial distribution is expected (for example, the question uses the word 'approximately' and you have to think about how an approximation might arise); and the presence of the exponential in the given result gives a pretty broad hint that it should be a Poisson distribution — the use of which has to be justified. One can expect the Poisson approximation to work when the number of trials (call it m) is large (e.g. $m > 150$) and when $np \approx npq$, i.e. the mean is roughly equal to the variance (since these are equal for the Poisson distribution) — so $q \approx 1$.

You can't find the most probable number of winning tickets by differentiation (unless you fancy differentiating the factorial $x!$); instead, you must look at the ratios of consecutive terms and see when these turn from being greater than one to less than one.

Solution to question 61

Let X be the random variable whose value is the number of winning tickets out of the $2.8N$ tickets purchased by the non-syndicate punters. Then $X \sim \text{Poisson}(2.8)$.

Let $p_j = \mathrm{P}(X = j)$. Then

$$p_j = \frac{(2.8)^j \mathrm{e}^{-2.8}}{j!} \Rightarrow \frac{p_{j+1}}{p_j} = \frac{2.8}{j+1}\,.$$

This fraction is greater than 1 if $j < 1.8$, so the most probable number of winning tickets by non-syndicate punters is 2. Overall, the most probable number of winning tickets is 3.

The expected winnings of the syndicate is

$$\begin{aligned}\left(Wp_0 + \frac{1}{2}Wp_1 + \frac{1}{3}Wp_2 + \cdots\right) &= \mathrm{e}^{-2.8}W\left(1 + \frac{1}{2}\times\frac{1}{1!}(2.8) + \frac{1}{3}\times\frac{1}{2!}(2.8)^2 + \cdots\right)\\ &= \mathrm{e}^{-2.8}W\frac{\mathrm{e}^{2.8}-1}{2.8}\,,\end{aligned}$$

so the expected loss is as given.

Postmortem

This is rather interesting. You might wonder whether it would be worth borrowing the money to buy every single lottery combination, in order to win a share of the jackpot. Clearly, the people who run the lottery have to think about this sort of thing.

Suppose the number of tickets sold, excluding the N that we plan to buy, is expected to be kN. Suppose also that a fraction α of the total is paid out in the jackpot. Then the above formula becomes

$$N\left(1 - \frac{1-e^{-k}}{k}\alpha(k+1)\right).$$

If k is very small (take $k = 0$), we lose $N(1-\alpha)$ (obviously). If k is large, $k = 3$ say, the exponential can be ignored and we lose $N(1-\alpha(k+1)/k)$. If $k \gg 1$, this becomes $N(1-\alpha)$ again. In between, there is a value of k that, for each fixed α, gives a minimum loss (which may be a gain if α is close to 1).

Note how informative it is to have k rather than 2.8; the numerical value was chosen in the question to model roughly that lottery system in the UK.

Having gone back to this solution after a break, I am now wondering about the use of the Poisson approximation. Of course, the set-up (large m, small p) begs us to approximate, but did we need to? Certainly not for the first result, since we can just as well look at the ratio of two terms of the Binomial distribution as at two terms of the Poisson distribution. Try it; the result is of course the same.

The second part is more difficult. What we want is the expectation of $1/(n+1)$, and this turns out to be a difficult sum using the Binomial distribution (in fact, it can only be expressed in terms of a hypergeometric function, which would then have to be approximated to get a less obscure answer).

It seems to me that the solution above is therefore a bit unsatisfactory. It would surely have been better to work with the exact distribution until it was necessary to approximate, even though one knows that the approximation is so good that the answers would be the same. We are, after all, mathematicians and not engineers.

Question 62(*)

I have k different keys on my key ring. When I come home at night I try one key after another until I find the key that fits my front door. What is the probability that I find the correct key, for the first time, on the nth attempt in each of the following three cases?

(i) At each attempt, I choose a key that I have not tried before, each choice being equally likely.

(ii) At each attempt, I choose a key from all my keys, each of the k choices being equally likely.

(iii) At the first attempt, I choose from all my keys, each of the k choices being equally likely. At each subsequent attempt, I choose from the keys that I did not try at the previous attempt, each of the $k-1$ choices being equally likely.

Comments

This is very easy, and would really be far too easy if the answers were given.

You should set out your argument clearly and concisely, because if you come up with the wrong answer and an inadequate explanation you will not get many marks. Even if you write down the correct answer you may not get the marks if your explanation is inadequate.

You should of course run the usual checks on your answers. Do they lie in the range $0 \leqslant p \leqslant 1$? If you sum over all outcomes, do you get 1? (The latter is a good and not difficult exercise; I insist that you try it.)

You will probably be struck by the simplicity of the answer to part (i), after simplification. Is there an easy way of thinking about it (or, if you did it the easy way, is there a hard way)?

Solution to question 62

(i)

$$\text{P(finding correct key, for the first time, on } n\text{th attempt)}$$

$$=\text{P(fail first, fail second, ... , fail } (n-1)\text{th, succeed } n\text{th)}$$

$$=\begin{cases} \dfrac{1}{k} & \text{for} \quad n=1; \\ \dfrac{k-1}{k}\times\dfrac{k-2}{k-1}\times\cdots\times\dfrac{k-(n-1)}{k-(n-2)}\times\dfrac{1}{k-(n-1)}=\dfrac{1}{k} & \text{for} \quad 2\leqslant n\leqslant k; \\ 0 & \text{for} \quad n>k. \end{cases}$$

(ii)

$$\text{P(find correct key, for the first time, on } n\text{th attempt)}$$

$$=\text{P(fail first, fail second, ... , succeed } n\text{th)} = \left(\frac{k-1}{k}\right)^{n-1}\frac{1}{k}.$$

(iii)

$$\text{P(find correct key, for the first time, on } n\text{th attempt)}$$

$$=\text{P(fail first, fail second, ... , succeed } n\text{th)}$$

$$=\begin{cases} \dfrac{1}{k} & \text{for} \quad n=1; \\ \left(\dfrac{k-2}{k-1}\right)^{n-2}\dfrac{1}{k} & \text{for} \quad n\geqslant 2. \end{cases}$$

Postmortem

The above solution is the 'hard way'. The easy way is to consider the keys all laid out in a row, instead of being picked sequentially. Exactly one of the keys is the correct one, which is equally likely to be any of the keys in the row, so has a probability of $1/k$ of being in any given position. Clearly, this corresponds exactly to picking them one by one.

I gave the hard solution above, because I think most people will be drawn into doing it this way by the phrasing of the question: keys are picked one by one and tried before going on to the next.

Question 63$(**)$

Tabulated values of $\Phi(\cdot)$, the cumulative distribution function of a standard normal variable, should not be used in this question.

Henry the commuter lives in Cambridge and his working day starts at his office in London at 0900. He catches the 0715 train to King's Cross with probability p, or the 0720 to Liverpool Street with probability $1-p$. Measured in minutes, journey times for the first train are $N(55, 25)$ and for the second are $N(65, 16)$. Journey times from King's Cross and Liverpool Street to his office are $N(30, 144)$ and $N(25, 9)$, respectively. Show that Henry is more likely to be late for work if he catches the first train.

Henry makes M journeys, where M is large. Writing A for $1 - \Phi(20/13)$ and B for $1 - \Phi(2)$, find, in terms of A, B, M and p, the expected number, L, of times that Henry will be late and show that for all possible values of p,

$$BM \leqslant L \leqslant AM.$$

Henry noted that in 3/5 of the occasions when he was late, he had caught the King's Cross train. Obtain an estimate of p in terms of A and B.

[**Note**: A random variable is said to be $N\left(\mu, \sigma^2\right)$ if it has a normal distribution with mean μ and variance σ^2.]

Comments

This is impossible unless you know the following result:

If X_1 and X_2 are independent and normally distributed according to $X_1 \sim N(\mu_1, \sigma_1^2$ and $X_2 \sim N(\mu_2, \sigma_2^2$, then $X_1 + X_2$ is also normally distributed and $X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Even if you didn't know this result, you do now and it should not be difficult to complete the first parts question.

For the last part, you need to know something about conditional probability, namely that

$$\mathrm{P}(A\big|B) = \frac{\mathrm{P}(B \cap A)}{\mathrm{P}(B)}$$

which is intuitively and can easily be understood in terms of Venn diagrams. The denominator is essentially a normalising constant. The formula may be taken as the definition of conditional probability on the left hand side.

Solution to question 63

Let T_1 be the random variable representing the total journey time via Kings Cross, so that

$$T_1 \sim N(55+30, 25+144) = N(85, 169),$$

using the theorem mentioned on the previous page, and and let T_2 be the random variable representing the total journey time via Liverpool Street, so that

$$T_2 \sim N(65+25, 16+9) = N(90, 25).$$

Then the probabilities of being late are, respectively, $\mathrm{P}(T_1 > 105)$ and $\mathrm{P}(T_2 > 100)$, i.e. $1-\Phi(20/13)$ and $1-\Phi(2)$. Note that $\Phi(20/13) < \Phi(2)$.

We have

$$L = [pA + (1-p)B]M = [B + (A-B)p]M.$$

L increases as p increases, since $A > B$, hence given inequalities corresponding to $p = 0$ and $p = 1$.

We have

$$\mathrm{P}(\text{Kings Cross given late}) = \frac{\mathrm{P}(\text{Late and Kings Cross})}{\mathrm{P}(\text{Late})}$$

An estimate for p (all it $\tilde{p}$) is therefore given by

$$\frac{3}{5} = \frac{A\tilde{p}}{A\tilde{p} + B(1-\tilde{p})},$$

so $\tilde{p} = 3B/(2A+3B)$.

Postmortem

The result mentioned in the comment on the previous page is just the sort of thing you should try to prove yourself rather than take on trust. Unfortunately, such results in probability tend to be pretty hard to prove. You can also prove it fairly easily using generating functions (which are not in the syllabus for STEP I and II). You can also prove it from first principles, but that would be difficult for you but not impossible because of the awkward integrals involved.

Question 64($*$)

A group of biologists attempts to estimate the magnitude, N, of an island population of voles (*Microtus agrestis*). Accordingly, the biologists capture a random sample of 200 voles, mark them and release them. A second random sample of 200 voles is then taken of which 11 are found to be marked. Show that the probability, p_N, of this occurrence is given by

$$p_N = k\frac{\big((N-200)!\big)^2}{N!(N-389)!},$$

where k is independent of N.

The biologists then estimate N by calculating the value of N for which p_N is a maximum. Find this estimate.

All unmarked voles in the second sample are marked and then the entire sample is released. Subsequently a third random sample of 200 voles is taken. Using your estimate for N, write down the probability that this sample contains exactly j marked voles, leaving your answer in terms of binomial coefficients.

Deduce that

$$\sum_{j=0}^{200}\begin{pmatrix}389\\ j\end{pmatrix}\begin{pmatrix}3247\\ 200-j\end{pmatrix} = \begin{pmatrix}3636\\ 200\end{pmatrix}.$$

Comments

This is really just an exercise in combinations. (Recall that a permutation is a reordering of a set of objects, and a combination is a selection of a subset from a set.) You assume that you are equally likely to choose any given subset of the same size, so that the probability of a set of specific composition is the number of ways of choosing a set of that composition divided by the total number of ways of choosing any set of the same size. Of course, you are assuming that the voles are indistinguishable, except for the marks made by the biologists.

Maximising a discrete (not a continuous) function of N came up on one of the previous questions: you have to compare adjacent terms.

Solution to question 64

200 out of N voles are marked, so p_N is just the number of ways of choosing 11 from 200 and 189 from $N-200$ divided by the number of ways of choosing 200 from N:

$$p_N \quad = \quad \frac{\dbinom{200}{11}\dbinom{N-200}{189}}{\dbinom{N}{200}} \quad = \quad \frac{\dfrac{200!}{11!\,189!}\;\dfrac{(N-200)!}{189!\,(N-389)!}}{\dfrac{N!}{200!\,(N-200)!}},$$

so

$$k = \frac{(200!)^2}{11!\,(189!)^2}.$$

At the maximum value, $p_N \approx p_{N-1}$, i.e.

$$\frac{(N-200)^2}{N(N-389)} \approx 1,$$

which gives $N \approx 200^2/11 \approx 3636$ (just divide 40000 by 11).

At the third sample, there are 389 marked voles and an estimated $3636 - 389 = 3247$ unmarked voles. Hence

$$\mathrm{P}(\text{exactly } j \text{ marked voles}) = \frac{\dbinom{389}{j}\dbinom{3247}{200-j}}{\dbinom{3636}{200}}.$$

The final part follows immediately, using

$$\sum_{j=0}^{200} \mathrm{P}(\text{exactly } j \text{ marked voles}) = 1.$$

Postmortem

Adding the Latin name of the species was a nice touch, I thought (not my idea); it adds an air of verisimilitude to the problem.

I suppose that this might be the basis of a method of estimating the population of voles — rather clever really. I don't know how well it works in practice though. The assumption mentioned earlier, that picking one set of voles of size 200 is just as likely as picking any other set, surely relies on perfect mixing of the marked voles, which would be rather difficult to achieve (especially as female voles can be highly territorial).

Question 65()

A stick is broken at a point, chosen at random, along its length. Find the probability that the ratio, R, of the length of the shorter piece to the length of the longer piece is less than r, where r is a given positive number.

Find the probability density function for R, and calculate the mean and variance of R.

Comments

Solution to question 65

Let X be the distance of the break from one end of the stick, and assume that $X \leqslant \ell/2$, where ℓ is the length of the stick (if this assumption does not hold, then we could measure from the other end of the stick). Then

$$\mathrm{P}(X < x) = \frac{x}{\ell/2}$$

Now $R = \dfrac{X}{\ell - X}$, by definition, so

$$X = \frac{\ell R}{1 + R}.$$

The cumulative distribution function for R is

$$\mathrm{P}(R < r) = \mathrm{P}\left(\frac{X}{l - X} < r\right) = \mathrm{P}\left(X < \frac{\ell r}{1 + r}\right) = \frac{\ell r}{(1 + r)(\ell/2)} = \frac{2r}{1 + r}.$$

[Check that this satisfies the conditions for a cumulative distribution function: it starts at 0, when $r = 0$; it ends at 1 when r takes its maximum value of 1.]

The probability density function is

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{2r}{1 + r}\right) = \frac{2}{(r + 1)^2}.$$

Integrating to find $\mathrm{E}(R)$ and $\mathrm{E}(R^2)$ gives

$$\mathrm{E}(R) = \int_0^1 \frac{2r}{(1 + r)^2}\mathrm{d}r = 2\ln 2 - 1; \quad \mathrm{E}(R^2) = 3 - 4\ln 2,$$

so $\mathrm{Var}\,(R) = 2 - 4(\ln 2)^2$.

Question 66()

An experiment produces a random number T uniformly distributed on $[0,1]$. Let X be the larger root of the equation

$$x^2 + 2x + T = 0.$$

What is the probability that $X > -1/3$? Find $E(X)$ and show that $\mathrm{Var}(X) = 1/18$.

The experiment is repeated independently 800 times generating the larger roots $X_1, X_2, \ldots, X_{800}$. If

$$Y = X_1 + X_2 + \cdots + X_{800}.$$

find an approximate value for K such that

$$\mathrm{P}(Y \leqslant K) = 0.08.$$

Comments

Solution to question 66

The larger root is given by $X = -1 + \sqrt{1-T}$, so $X > -1/3$ if $T < 5/9$. Thus $\mathrm{P}(X > -1/3) = 5/9$.

For the repeated experiment we use the normal distribution. First calcalate the expectation and variance:

$$\mathrm{E}(X) = \int_0^1 (-1 + \sqrt{1-t})\mathrm{d}t = -\frac{1}{3};$$

$$\mathrm{E}(X^2) = \int_0^1 (-1 + \sqrt{1-t})^2\mathrm{d}t = \frac{1}{6}; \quad \mathrm{Var}(X) = \frac{1}{6} - \frac{1}{9} = \frac{1}{18}\,.$$

Using standard results for the mean of a sum of random variables (always equal to the sum of the means) and the variance (equal to the sum of the variances if the variables are uncorrelated; independent implies uncorrelated), we have

$$\mu = \mathrm{E}(Y) = -\frac{800}{3} \quad \text{and} \quad \sigma^2 = \mathrm{Var}(Y) = \frac{800}{18}.$$

Now $P(Y \leqslant K) = \mathrm{P}(Z \leqslant (K-\mu)/\sigma)$, where Z is standard normal.

For a standard normal distribution, we have

$$P(Z \leqslant -1.40) = 0.08076$$

so we want K to satisfy

$$\frac{K + 800/3}{20/3} = -1.4\,.$$

Thus $K = -276$.

STEP Syllabus

General Introduction

There are two syllabuses, one for Mathematics I and Mathematics II, the other for Mathematics III. The syllabus for Mathematics I and Mathematics II is based on a single subject Mathematics Advanced GCE. Questions on Mathematics II are intended to be more challenging than questions on Mathematics I. The syllabus for Mathematics III is wider.

In designing the syllabuses, the specifications for all the UK Advanced GCE examinations, the Scottish Advanced Higher and the International Baccalauriat were consulted.

Each of Mathematics I, II and III will be a 3-hour paper divided into three sections as follows:

Section A (Pure Mathematics)	eight questions
Section B (Mechanics)	three questions
Section C (Probability and Statistics)	three questions

All questions will carry the same weight. Candidates will be assessed on the six questions best answered; no restriction will be placed on the number of questions that may be attempted from any section. Normally, a candidate who answers at least four questions well will be awarded a grade 1.

The marking scheme for each question will be designed to reward candidates who make good progress towards a complete solution.

Syllabuses

The syllabuses given below are for the guidance of both examiners and candidates. The following points should be noted.

1. Questions may test candidates' ability to apply mathematical knowledge in novel and unfamiliar ways.
2. Solutions will frequently require insight, ingenuity, persistence and the ability to work through substantial sequences of algebraic manipulation.
3. Some questions may be largely independent of particular syllabus topics, being designed to test general mathematical potential.
4. Questions will often require knowledge of several different syllabus topics.
5. Examiners will aim to set questions on a wide range of topics, but it is not guaranteed that every topic will be examined every year.
6. Questions may be set that require knowledge of elementary topics, such as basic geometry, mensuration (areas of triangles, volume of sphere, etc), sine and cosine formulae for triangles, rational and irrational numbers.

Formula booklets and calculators

Candidates may use an Advanced GCE booklet of standard formulae, definitions and statistical tables. The inter-Board agreed list of mathematical notation will be followed. Calculators are not permitted (or required).

MATHEMATICS I (9465) and MATHEMATICS II (9470)

Section A: Pure Mathematics

This section comprises the Advanced GCE common core (typically, modules C1 — C4 of an Advanced GCE specification) broadly interpreted, together with a few additional items ∗enclosed in asterisks∗.

Specification	Notes
General	
Mathematical vocabulary and notation	including: equivalent to; necessary and sufficient; if and only if; $\Rightarrow$; $\Leftrightarrow$; $\equiv$.
Methods of proof	including proof by contradiction and disproof by counterexample; ∗including, in simple cases, proof by induction∗.
Algebra	
Indices and surds	including rationalising denominators.
Quadratics	including proving positivity by completing a square.
The expansion for $(a+b)^n$	including knowledge of the general term; notation: $\binom{n}{r} = \dfrac{n!}{r!\,(n-r)!}$.
Algebraic operations on polynomials and rational functions.	including factorisation, the factor theorem, the remainder theorem; including understanding that, for example, if $x^3+bx^2+cx+d \equiv (x-\alpha)(x-\beta)(x-\gamma),$ then $d = -\alpha\beta\gamma$.
Partial fractions	including denominators with a repeated or quadratic factor.
Sequences and series	including use of, for example, $a_{n+1} = \mathrm{f}\,(a_n)$ or $a_{n+1} = \mathrm{f}\,(a_n, a_{n-1})$; including understanding of the terms convergent, divergent and periodic in simple cases; including use of $\sum_{k=1}^{n} k$ to obtain related sums.
The binomial series for $(1+x)^k$, where k is a rational number.	including understanding of the condition $\lvert x\rvert < 1$.
Arithmetic and geometric series	including sums to infinity and conditions for convergence, where appropriate.
∗Inequalities∗	including solution of, eg, $\frac{1}{a-x} > \frac{x}{x-b}$; including simple inequalities involving the modulus function; including the solution of simultaneous inequalities by graphical means.

Functions

Domain, range, composition, inverse	including use of functional notation such as $y = \mathrm{f}(ax+b)$, $x = \mathrm{f}^{-1}(y)$ and $z = \mathrm{f}\big(\mathrm{g}(x)\big)$.
Increasing and decreasing functions	both the common usages of the term 'increasing' (i.e. $x > y \Rightarrow \mathrm{f}(x) > \mathrm{f}(y)$ or $x > y \Rightarrow \mathrm{f}(x) \geqslant \mathrm{f}(y)$) will be acceptable.
Exponentials and logarithms	including $x = a^y \Leftrightarrow y = \log_a x$, $x = \mathrm{e}^y \Leftrightarrow y = \ln x$; ∗including the exponential series $\mathrm{e}^x = 1 + x + \cdots + x^n/n! + \cdots$. ∗
The effect of simple transformations	such as $y = a\mathrm{f}(bx+c)+d$.
The modulus function.	
Location of roots of $\mathrm{f}(x) = 0$ by considering changes of sign of $\mathrm{f}(x)$.	
Approximate solution of equations using simple iterative methods.	

Curve sketching*

General curve sketching	including use of symmetry, tranformations, behaviour as $x \to \pm\infty$, points or regions where the function is undefined, turning points, asymptotes parallel to the axes.

Trigonometry

Radian measure, arc length of a circle, area of a segment.	
Trigonometric functions	including knowledge of standard values, such as $\tan(\pi/4)$, $\sin 30°$; including identities such as $\sec^2\phi - \tan^2\phi = 1$; including application to geometric problems in two and three dimensions.
Double angle formulae	including their use in calculating, eg, $\tan(\pi/8)$.
Formulae for $\sin(A \pm B)$ and $\cos(A \pm B)$	including their use in solving equations such as $a\cos\theta + b\sin\theta = c.$
Inverse trigonometric functions	definitions including domains and ranges; notation: $\arctan\theta$, etc

Coordinate geometry

Straight lines in two-dimensions	including the equation of a line through two given points, or through a given point and parallel to a given line or through a given point and perpendicular to a given line; including finding a point which divides a segment in a given ratio.

Circles	using the general form $(x-a)^2+(y-b)^2=R^2$; including points of intersection of circles and lines.
Cartesian and parametric equations of curves and conversion between the two forms.	
Calculus	
Interpretation of a derivative as a limit and as a rate of change	including knowledge of both notations $\mathrm{f}'(x)$ and $\frac{\mathrm{d}y}{\mathrm{d}x}$.
Differentiation of standard functions	including algebraic expressions, trigonometric (but not inverse trigonometric) functions, exponential and log functions.
Differentiation of composite functions, products and quotients and functions defined implicitly.	
Higher derivatives	including knowledge of both notations $\mathrm{f}''(x)$ and $\frac{d^2y}{\mathrm{d}x^2}$; including knowledge of the notation $\frac{d^ny}{\mathrm{d}x^n}$.
Applications of differentiation to gradients, tangents and normals, stationary points, increasing and decreasing functions	including finding maxima and minima which are not stationary points; including classification of stationary points using the second derivative.
Integration as reverse of differentiation	
Integral as area under a curve	including area between two graphs; including approximation of integral by the rectangle and trapezium rules.
Volume within a surface of revolution	
Knowledge and use of standard integrals	including the forms $\int \mathrm{f}'(\mathrm{g}(x))\mathrm{g}'(x)\mathrm{d}x$ and $\int \mathrm{f}'(x)/\mathrm{f}(x)\mathrm{d}x$; including transformation of an integrand into standard (or some given) form; including use of partial fractions; not including knowledge of integrals involving inverse trigonometric functions.
Definite integrals	$*$including calculation, without justification, of simple improper integrals such as $\int_0^\infty \mathrm{e}^{-x}\mathrm{d}x$ and $\int_0^1 x^{-\frac{1}{2}}\mathrm{d}x$ (if required, information such as the behaviour of $x\mathrm{e}^{-x}$ as $x\to\infty$ or of $x\ln x$ as $x\to 0$ will be given).$*$
Integration by parts and by substitution	including understanding their relationship with differentiation of product and of a composite function; including application to $\int \ln x \mathrm{d}x$.

Formulation and solution of differential equations	formulation of first order equations; solution in the case of a separable equation or by some other method given in the question.

Vectors

Vectors in two and three dimensions	including use of column vector and **i**, **j**, **k** notation.
Magnitude of a vector	including the idea of a unit vector.
Vector addition and multiplication by scalars	including geometrical interpretations.
Position vectors	including application to geometrical problems.
The distance between two points.	
Vector equations of lines	including the finding the intersection of two lines; understanding the notion of skew lines (knowledge of shortest distance between skew lines is not required).
The scalar product	including its use for calculating the angle between two vectors.

Section B: Mechanics

This section is roughly equivalent to two GCE modules M1 and M2; candidates who have only studied one module of mechanics are likely to lack both the knowledge and the experience of mechanics required to attempt the questions. Questions may involve any of the material in the Pure Mathematics syllabus.

Force as a vector	including resultant of several forces acting at a point and the triangle or polygon of forces; including equilibrium of a particle; forces include weight, reaction, tension and friction.
Centre of mass	including obtaining the centre of mass of a system of particles, of a simple uniform rigid body (possible composite) and, in simple cases, of non-uniform body by integration.
Equilibrium of a rigid body or several rigid bodies in contact	including use of moment of a force; for example, a ladder leaning against a wall or on a movable cylinder; including investigation of whether equilibrium is broken by sliding, toppling or rolling; including use of Newton's third law; excluding questions involving frameworks.
Kinematics of a particle in a plane	including the case when velocity or acceleration depends on time (but excluding use of acceleration $= v\frac{\mathrm{d}v}{\mathrm{d}x}$) ; questions may involve the distance between two moving particles, but detailed knowledge of relative velocity is not required.
Energy (kinetic and potential), work and power	including application of the principle of conservation of energy.
Collisions of particles	including conservation of momentum, conservation of energy (when appropriate); coefficient of restitution, Newton's experimental law; including simple cases of oblique impact (on a plane, for example); including knowledge of the terms *perfectly elastic* ($e = 1$) and *inelastic* ($e = 0$); questions involving successive impacts may be set.
Newton's first and second laws of motion	including motion of a particle in two and three dimensions and motion of connected particles, such as trains, or particles connected by means of pulleys.

Motion of a projectile under gravity

including manipulation of the equation

$$y = x\tan\alpha - \frac{gx^2}{2V^2\cos^2\alpha},$$

viewed, possibly, as a quadratic in $\tan\alpha$; not including projectiles on inclined planes.

Section C: Probability and Statistics

The emphasis, in comparison with Advanced GCE and other comparable examinations is towards probability and formal proofs, and away from data analysis and use of standard statistical tests. Questions may involve use of any of the material in the Pure Mathematics syllabus.

Probability

Permutations, combinations and arrangements	including sampling with and without replacement.
Exclusive and complementary events	including understanding of $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$, though not necessarily in this form.
Conditional probability	informal applications, such as tree diagrams.

Distributions

Discrete and continuous probability distribution functions and cumulative distribution functions.	including calculation of mean, variance, median, mode and expectations by explicit summation or integration for a given (possibly unfamiliar) distribution (eg exponential or geometric or something similarly straightforward); notation: $\mathrm{f}(x) = \mathrm{F}'(x)$.
Binomial distribution	including explicit calculation of mean.
Poisson distribution	including explicit calculation of mean; including use as approximation to binomial distribution where appropriate.
Normal distribution	including conversion to the standard normal distribution by translation and scaling; including use as approximation to the binomial or Poisson distributions where appropriate; notation: $X \sim \mathrm{N}(\mu, \sigma^2)$.

Hypothesis testing

Basic concepts in the case of a simple null hypothesis and simple or compound alternative	including knowledge of the terminology *null hypothesis* and *alternative hypothesis*, *one* and *two tailed tests*.

MATHEMATICS III (9475)

In each section knowledge of the corresponding parts of the syllabus for Mathematics I and II is assumed, and may form the basis of some questions.

Section A: Pure Mathematics

Algebra

Summation of series	including use of method of differences, recognition of Maclaurin's series (including binomial series) and partial fractions; including investigation of convergence, by consideration of the sum to n terms.
Polynomial equations	including complex solutions (and the fact that they occur in conjugate pairs if the cooefficients are real); including the relationship between symmetric functions of the roots and the coefficients; including formation of a new equation whose roots are related to the roots of the old equation.
Induction	including application complicated situations such as divisibility problems.

Trigonometry

Formula for $\sin A \sin B$ (etc) and $\sin A + \sin B$ (etc)	

Coordinate geometry

Ellipse, parabola, hyperbola and rectangular hyperbola	including standard Cartesian forms and parametric forms; including geometrical understanding, but not detailed knowledge of, the terms *focus*, *directrix*, *eccentricity*.
Polar coordinates	including relation with Cartesian coordinates; including ability to sketch simple graphs; including integral formula for area.

Complex numbers

Algebra of complex numbers	including Cartesian and polar forms.
Argand diagram	including geometrical representation of sums and products; including identification of loci and regions (eg $\lvert z-a\rvert = k\lvert z-b\rvert$).

Euler's identity: $e^{i\theta} \equiv \cos\theta + i\sin\theta$	including the relationship between trigonometrical and exponential functions; including application to summation of trigonometric series.
de Moivre's theorem	including application to finding nth roots.

Hyperbolic functions

Definitions and properties of the six hyperbolic functions	including integrals and derivatives of cosh, sinh, tanh and coth.
Formulae corresponding to trigonometric formulae	such as $\cosh^2 x - \sinh^2 x = 1$ and $\sinh(x + y) = \cdots$.
Inverse functions	including ability to derive logarithmic forms and graphs; including derivatives.

Calculus

Integrals resulting in inverse trig. or hyperbolic functions	including integrands such as $(x^2 - ax)^{-\frac{1}{2}}$.
Maclaurin's series	including knowledge of expansions for $\sin x$, $\cos x$, $\cosh x$, $\sinh x$ and $\ln(1 + x)$.
Arc length of a curve expressed in Cartesian coordinates	including curves described parametrically.
First order linear differential equations	including the general form $$\frac{\mathrm{d}y}{\mathrm{d}x} + \mathrm{P}(x)y = \mathrm{Q}(x).$$
Second order linear differential equations with constant coefficients	including the inhomogeneous case $$a\frac{\mathrm{d}^2y}{\mathrm{d}x^2} + b\frac{\mathrm{d}y}{\mathrm{d}x} + cy = \mathrm{f}(x),$$ where $\mathrm{f}(x)$ is an exponential function, a sin/cos function or a polynomial (but excluding resonant cases); including equations that can be transformed into the above form; including initial or boundary conditions.

Section B: Mechanics

Motion of a particle subject to a variable force	including use of $v\frac{\mathrm{d}v}{\mathrm{d}x}$ for acceleration.
Elastic strings and springs	including Hooke's law and the potential energy; notation: $T = kx = \lambda x/l$, where k is stiffness, λ is modulus of elasticity.
Motion in a circle	including motion with non-uniform speed.
Simple harmonic motion	including obtaining from a given model an equation of the form $\ddot{x} + \omega^2 x = a$, where a is a constant, and quoting the solution.
Relative motion	

Moment of inertia	including calculations involving integration (such as a non-uniform rod, or a lamina); including parallel and perpendicular axis theorems.
Motion of a rigid body about a fixed axis	including understanding of angular momentum and rotational kinetic energy and their conservation; including the understanding of couple and the work done by a couple; including calculation of forces on the axis.

Section C: Probability and Statistics

Algebra of expectations	for example, $\mathrm{E}(aX + bY + c) = a\mathrm{E}(X) + b\mathrm{E}(Y) + c\,,$ $\mathrm{Var}\,(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2\,,$ $\mathrm{Var}\,(aX + b) = a^2\,\mathrm{Var}\,(X)\,.$
Further theory of distribution functions	including, for example, use of the cumulative distribution function to calculate the distribution of a related random variable (eg X^2).
Generating functions	including the use of generating functions to obtain the distribution of the sum of independent random variables.
Discrete bivariate distributions	including understanding of the idea of independence; including marginal and conditional distributions; including use of $\mathrm{Var}\,(aX + bY) =$ $a^2\,\mathrm{Var}\,(X) + b^2\,\mathrm{Var}\,(Y) + 2ab\mathrm{Cov}(X,Y)\,.$
Sampling	including sample mean and variance; including use of the central limit theorem.