Step II, Solutions
June 2007

Section A: Pure Mathematics

Q1 It is important to get off to a good start in any examination, especially so in STEPs, and Q1 is specifically designed to get as many candidates as possible off to such a start. Binomial series expansions are given in any of the permitted formulae books, and there is really no excuse for failing to pick up the marks on the introductory bit of the question. It is almost certainly to your advantage to simplify the terms of the expansion, but a little bit of care is in order here, else you are automatically losing accuracy marks later on in (a) and (b).

For part (a), you are told exactly what value of k to choose, and it is simply a case of using it **on both sides of the statement** – in the LHS to show that you can extract a sensible multiple of $\sqrt{3}$, and then in the RHS to see what you get as a decimal. Remember that working in powers of 10 makes the numerical working a lot simpler.

In (b), you have to choose a suitable value of k so that the LHS gives a multiple of $\sqrt{6}$. There is a small (but negative) integer value of k which will do this nicely. Many candidates, however, actually chose to work with $k = 50$ and, if you check, you will see that this *seems* to work equally well. However, the approximation gained is not nearly so accurate; this is because ….. ? Also, not a few candidates chose values of k greater than 100 in absolute value, and these are even worse, because ….. ?

In (ii), it is certainly possible to work back from the final answer in order to figure out what value of k to use here, but (again) you are looking for some (presumably) integer value that will this time yield a perfect cube multiple of 3 when $1+\frac{k}{1000}$ is written as an improper fraction.

For interest's sake, the original version of the question used the first *three* terms of the series expansion with $k = 24$ to find an approximation to $\sqrt[3]{2}$.

Answers: (i) $1+\frac{k}{200}-\frac{k^2}{80\,000}+\frac{k^3}{16\,000\,000}$; (a) 1.732 05; (b) 2.449 49.

Q2 It is fairly obvious that $x = p$ and $x = q$ are the two roots of the equation $\frac{dy}{dx} = 0$, which means that the derivative is a multiple of $(x-p)(x-q)$. Comparing the two then immediately gives b and c in terms of p and q. The sketch is a standard (positive) cubic, through the origin, with its two TPs in the first quadrant. Unintentionally, there are two possible candidates for the region $\boldsymbol{R}$, since the setters omitted to consider the one of them. Almost all candidates taking this paper identified the intended region, and this was because the question tries to get you to focus on the area around the point of inflection, which you are asked to mark on the diagram.

In (iii), m and n are simply the y-coordinates of the points corresponding to $x = p$ and q (respectively), and by this point you should know the curve's equation (in terms of p and q rather than b and c). Notice that $y(m)$ involves the extra qs and $y(n)$ involves extra ps, so the difference may just involve lots of $(q-p)$s, and the answer effectively tells you this much also. It may help in the working, both now and later, if you exploit this difference as much as possible.

Before embarking on the final part of the question, it would benefit you greatly to take a momentary pause and think about how the various bits of the question hang together. You were earlier asked to describe the symmetry of the cubic, and this was not just an idle bit of space-filler on the setter's part. Rather, it was an attempt to force you into recognising that the area of the region $\boldsymbol{R}$ can be found by means other than integration. Ignoring the coordinate axes on the diagram, and looking at the lines $x = p$, $x = q$, $y = n$ and $y = m$, you will see a nice rectangle appearing in the middle of the page. Because of the symmetry of the cubic, $\boldsymbol{R}$ is something to do with this rectangle, and this fact pretty much allows you to write the answer straight down, using the answer to (iii). On the other hand, if you want to do it by integration (as most candidates did) .....

And if you feel up to an algebraic challenge, see if you can work out, by integration, the area of the other possible region $\boldsymbol{R}$ – which also turns out – rather surprisingly, I felt – to be a rational multiple of $(q-p)^4$.

Answers: (i) $b = 3(p + q)$, $c = 6pq$; (ii) (two-fold) rotational symmetry about the P of I.

Q3 The first part of this question is a standard piece of bookwork, and requires only a modest ability to cope with substitution integration and a bit of trig. identity work. In (i) (a), you need to spot a suitable substitution for yourself – comparing the integrand with that in the introductory bit gives the game away, if you're stuck. In my day, the $t = \tan \frac{1}{2}x$ substitution was a very common bit of work, but you don't see it very often at A-level nowadays, so you could be forgiven for not being entirely familiar with it. Nonetheless, the principles of substitution still apply, and there may be the odd trig. identity to be employed, of course. The final two pieces of work here are greatly eased by the fact that they can be done in either direction. By that, I mean that one can eliminate all the *t*s in favour of *x*s, or vice versa. If you successfully complete part (i) (b), then (ii) is so much easier, since the only difference is that you must have $3\sin^2 x$ in the denominator to give $14t^2$ instead of $6t^2$. Another simple substitution then changes the form into a standard arctan integral and, with a little bit of care, the whole thing can be wrapped up quite smoothly.

Overall, I would suggest that this is a fairly routine question, with no great leaps of thought required for a good A-level candidate to be able to work their way through it. What *is* required, however, is a high level of thoroughness and familiarity with the basic techniques of the trig. and calculus involved therein. Such capabilities are an essential requirement if you are preparing for future STEPs.

Answers: (i) (a) $\frac{\pi}{4}$; (ii) $\frac{\pi}{6\sqrt{3}}$.

Q4 This was actually not a particularly popular, or well done, question, although I still maintain that it is quite an easy one when it comes down to it! To begin with, it is really, *really* obvious that you need to expand the given trig. expressions using the *Addition Formulae*. Then, in order to obtain tans throughout, rather than sines and cosines, you are going to have to divide by ..... (hint: note the introductory conditions at the very start of the question, which are given to enable you not to worry about dividing by .....). Wangling it into the given form and checking that the given condition holds is not much more than an algebraic exercise at this stage, and shouldn't prove too much of a burden. However, it is easy to overlook the fact that you are asked to prove an "if and only if" statement, which is two-directional. In point of fact, it is the case here that a clear line of reasoning from first equation to final one actually *is* entirely reversible, although it is best to (at least) point out that this is so, rather than ignore it.

For the next three parts, see how this result can now be used to solve each of the given equations, once the "*A*" and the "*B*" have been clearly identified. Also, don't forget to identify the α, β and γ (the same in each of the three parts) and verify that $\alpha^2 = \beta^2 + \gamma^2$. It is, of course, perfectly possible to start each bit from scratch, and the wording of the question doesn't actually prevent you from doing so, but it would seem a bit of a waste of time and effort to do so. Having said that, several candidates successfully did (iii) by collecting the two ($3x$) terms up together and collecting them up in an $R\sin(3x+\theta)$ form.

Incidentally, my favourite part of the question was (ii), in which I got to play a bit of a dirty trick – the statement looks like an equation, but is actually an ….. because …..

Answers: (i) $\frac{2\pi}{3}, \frac{5\pi}{3}$; (ii) all $x \in [0, 2\pi)$; (iii) $\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}$ and $\frac{\pi}{3}, \frac{4\pi}{3}$.

Q5 Part (i) is a standard opener using compositions of functions, and the algebra shouldn't prove too demanding if you're careful. Again, simplified answers at each stage are most helpful for successful further progress through a question like this. The sequence of powers of f turns out to be periodic with period 3, and so f^{2007} isn't quite the big ask that it might seem to be at first sight.

As you're told what to do in (ii), it is just a case of being careful in establishing the relationship. A grasp of the process of mathematical induction is an essential requirement for STEP II, even if it is no longer on single Maths syllabuses elsewhere, and this could be used in this case. An informal inductive proof was perfectly acceptable also, although it was equally acceptable to establish the cases for $n = 1$, 2 and 3 and then point out that the periodicity of the tan function guarantees the rest.

Now, part (iii) offers something a little more demanding. The simple approach involves spotting that the use of $t = \sin\theta$ gives $\sqrt{1-t^2} = \cos\theta$, and then a similar inductive argument to (ii)'s will lead to an admittedly unappealing but otherwise simple result for g^n in a $\sin(A+B)$ kind of way. However, if instead you note that $\sqrt{1-t^2}$ denotes the **positive** square-root of $1-t^2$, which may actually be $-\cos\theta$ for some values of θ (and hence t). Thus, in fact, g^2 *can* turn out to be just x again, so that the sequence $\{g, g^2, g^3, \ldots\}$ turns out to be oscillating (i.e. periodic with period 2). If you proceed further down this route, exploring which parts of g's domain give what "powers" of g, you get very interesting results which may be worth discussion, but were not expected under examination conditions here.

Answers: (i) $f^2(x) = \dfrac{x-\sqrt{3}}{1+\sqrt{3}\,x}$, $f^3(x) = x$; $f^{2007}(x) = x$.

(iii) Answer 1: $g^n(t) = \sin(\sin^{-1}t + \frac{n\pi}{6})$; Answer 2: $g^n(t) = \begin{cases} g(t) & n \text{ odd} \\ t & n \text{ even} \end{cases}$.

Q6 Once again, this starts off with a bit of very basic work that a realistic STEP candidate needs to be in a position to rattle off quickly and efficiently. The "Hence" at the start of line 2 of the question tells you that the answer to this integral is to be found in the two previous answers, without further calculus work being done. It is, therefore, very bad examination practice to ignore the "Hence" demand and go off on an "or otherwise" route that isn't actually needed. And there's a strong chance you may not get any marks at all for your alternative approach.

In (ii), there is no reason why you can't treat the given differential equation as a quadratic in $\frac{dy}{dx}$ and solve it to get two slightly different, and much simpler, differential equations than the original one. At this stage, if you have your wits about you, and you are **NOT** getting a $\sqrt{3+x^2}$ anywhere in sight, then you really ought to be a bit suspicious about why not! For the rest of it, it is a simple case of integrating using (i)'s result, and then applying the given initial condition to find the constants of integration in each of the two cases.

Answers: (i) $\dfrac{1}{\sqrt{3+x^2}}$, $\dfrac{2x^2+3}{\sqrt{3+x^2}}$; $\dfrac{1}{2}x\sqrt{3+x^2} + \dfrac{3}{2}\ln\left(x+\sqrt{3+x^2}\right)$ $(+\,C)$.

(ii) $y = \dfrac{1}{6}x\sqrt{3+x^2} + \dfrac{1}{2}\ln\left(x+\sqrt{3+x^2}\right) - \dfrac{1}{6}x^2 - \dfrac{1}{6} - \dfrac{1}{2}\ln 3$;

$y = -\dfrac{1}{6}x\sqrt{3+x^2} - \dfrac{1}{2}\ln\left(x+\sqrt{3+x^2}\right) - \dfrac{1}{6}x^2 + \dfrac{1}{2} + \dfrac{1}{2}\ln 3$.

Q7 I like this question, although I accept that lots of candidates were probably put off by a question that looks like something they've never seen before. However, it is often the case that questions of the "new and weird-looking kind" can actually turn out to be relatively easy IF you're prepared to be a bit adventurous.

The opening bit introduces you to a (possibly) new idea, and then gets you to practise this idea in a couple of cases in order that you get the hang of it. Then, in part (i), you actually get to use one of these ideas, and you're pretty much told exactly what to do, and which of the two initial functions to use to get the given result.

Next, in (ii), you're thrown in the deep-end rather more and left to decide what to do for yourself. Here, however, there is reference made to a mysterious "suitable function" to be used. Now, if you believe that the setter is out to trap you, trick you, and grind you into the ground then you probably think you're all on your own at this stage and have to find your own function. But you're wrong! The setters are actually trying to give you every opportunity to do some good mathematics, and every effort is made to point you in the right direction if is felt at all suitable to do so. In this case, you were initially asked to show that the sin and ln functions had the property being referred to. Then you used the sin function in (i). Perhaps, just *perhaps*, you are meant to be using the other one in (ii). If you can use the ln function to establish this next result (called the *Arithmetic Mean – Geometric Mean Inequality*), then parts (a) and (b) at the end simply use it twice; once with very little thought required, and one with a little more thought needed. Be brave! Give it a go.

Answers: (ii) -2.

Q8 If you don't know what is wanted in (i), then you really shouldn't be doing this question. It also really helps if you realise that if s and t are positive, then X is the point between B and C such that $BX : XC = t : s$. Once you have these ideas in place, this question involves nothing more than finding the points of intersection referred to, by equating two different line equations at a time. You will need to introduce a new pair of parameters each time, but if you keep each stage of working separate, then there is no reason not to use the same two symbols each time; and then solve pairs of simultaneous equations, gained by equating the **b**– and **c**–components of the two relevant line vector equations, for these two parameters in terms of β and γ. The result displayed is known as *Ceva's Theorem*.

Answers: (i) The straight line through B and C.

Section B: Mechanics

Q9 The greatest problem with marking mechanics questions on the STEPs is that candidates seem to be so unwilling, or unable, to mark up a decently labelled diagram with all relevant forces on them, or, in this case, relevant velocities and angles. On the face of it, this is just a collisions question dressed up a bit, and there really are only the two mechanical principles to be applied here: *Conservation of Linear Momentum* (CLM) and *Newton's Experimental Law of Restitution* (NEL or NLR). If you take a side-on view of the cone, then the collision – at the moment of impact – is effectively the same as would be given by a plan view of a particle striking a vertical wall: directly, in the first instance, and then obliquely in the second. Applying CLM parallel to this line of impact (which is very easy in the first case and, in fact, the reason why you were asked for an explanation to begin with so that you were pointed in the right direction) and NEL perpendicular to it are essential steps in both parts of the question. In order to prevent you worrying about how the cone might bounce off the plane, you are told that this does not happen. So there is no point considering CLM vertically for the particle-cone collision, but there is still the horizontal motion of the cone to consider.

In (ii), the collision is oblique to the line of the cone's side, so there are two angles involved, and a bit of trig. work might be needed to sort things out. Alternatively, rather than re-doing (i)'s working in this separate case, one could simply consider the components of the "incoming" velocity, and the second answer for w is exactly the same as the first, but with u replaced by ….. For the very final part of the question, a little calculus is in order.

Q10 The first thing to do here is to find the position of the centre of mass of the composite figure, and this is fairly easily done by taking moments about some suitable point. Most candidates who actually attempted this question then went very badly astray, largely due to lack of a discernible approach in their jottings. In slipping-tilting situations, the standard approach is to examine separately what happens at the instant when slipping occurs *assuming* that tilting hasn't, and then to examine what happens when tilting occurs *assuming* that the slipping hasn't. This then gives two sets of conditions on P which can be compared. Remember that P can be in either direction, hence the modulus sign in the answer, which needs to be explained somewhere along the line.

Q11 In a similar sort of way to Q9, this is just a reasonably standard projectiles question dressed up a bit, and the vector set-up should help you work in the third-dimension quite naturally. The given answer in (i) should help confirm that you're doing the right thing to begin with (or not!). Completing the square, or differentiating, will give the value of t when OP is a minimum, and this should then turn out to be the same instant/position as can be found in part (ii) by differentiating the vertical (**k** –) component of the displacement vector.

Part (iii) can be done in a couple of ways: one is very lengthy, pressing on with the vector formulation for as long as possible, but the intended approach is to work with distance and time as scalars on the assumption that the bullet moves in a straight line.

Answers: (i) $\underline{\mathbf{r}} = \left(50 - 5t\sqrt{5}\right)\mathbf{i} + \left(5t\sqrt{15}\right)\mathbf{j} + \left(5t\sqrt{5} - 5t^2\right)\mathbf{k}$;

$\underline{\mathbf{p}} = \frac{75}{2}\mathbf{i} + \frac{25\sqrt{3}}{2}\mathbf{j} + \frac{25}{4}\mathbf{k}$; (0)60°.

Section C: Probability and Statistics

Q12 With given answers, as here, it is important to make your method clear, since there is a lot of fiddling going on as candidates inevitably manage to wangle this answer somehow. Splitting the required event into a series of mutually exclusive events and recognising which of these events are independent, is crucial, and it helps both you and the examiner if there is an accompanying (brief!) explanation as to what you are doing. It seems to me that the first part can be approached in at least a couple of obvious ways. Firstly, one could work out the prob. that one die gives at least one 6 in the first r throws, $P(r)$, say, and then observing that the prob. that both dice have given a 6 **at** the r^{th} throw is $P(r) - P(r-1)$. Alternatively, one could write it as the sum of the probs. that {neither dice has recorded a 6 in the first $r - 1$ throws and then both give 6s} with the prob. that {one die gives a 6 before the r^{th} throw and then the 2^{nd} die first gives a 6 on the r^{th} throw}.

Finding the expected value of the number of throws is routine, in principle at least, and you are given a result to use to help you with this, if needed. In (ii), equating this expression (in terms of p only) to m and then re-arranging gives a ..... equation in p, which should now be very familiar territory.

Answers: (i) $p = \frac{1}{m}\left\{m + 1 - \sqrt{m^2 - m + 1}\right\}$.

Q13 The first couple of terms of the series expansion for e gives the opening result, which you are obviously intended to exploit later on in the question. Next, p(at least 1 matching pair) is best considered in the form 1 – p(no matching pairs), and you get (with a little imagination) a whole load of fractions of the form $\frac{n-r}{n}$ which can be approximated by the exponential result given at the outset. The laws of indices and a bit of summation of an AP then sort out the rest of the first problem.

The next two parts each involve working with an inequality, and the second requires another use of the initial exponential result. Each employs the remarkably accurate rational approximation to $\ln 2$ given in the question.

Answers: 23 ; 253.

Q14 The pdf sketch in (i) consists of three (actually, five – don't forget to indicate clearly the zero bits!) pieces. Then, equating expressions for the endpoints of these pieces, which are defined in two different ways, immediately gives a and b in terms of k. After this, equating the total area under this graph to 1 (total probability) then gives the exact value of k, and hence a and b also. This is the bulk of the question done, and most of it is really pure mathematical content.

The last part is similar in content, requiring – in statistical terms – only the observation that m is given by $\int_1^m \mathrm{f}(x) = \frac{1}{2}$. Now, it is not immediately clear which piece of the function that m lies in, so a little bit of justification needs to be given to explain the relevance of any subsequent working that you give. Some fairly simple approximations for e should enable you to show that m is **not** in the first piece but **is** in the second.

Answers: (ii) $a = 2\ln k$, $b = \frac{\ln k}{2k}$; $k = \mathrm{e}^{1/3}$, $a = \frac{2}{3}$, $b = \frac{1}{6}\mathrm{e}^{-1/3}$; (iii) $m = 3(\mathrm{e}^{1/3} - ½)$.