Thoughts on the Essay Question

T. W. Körner

August 27, 2005

Small print The opinions expressed in this note are the author's own. Even the best advice (and there is no reason to suppose that the advice here is the best advice) does not apply to all people and in all circumstances[1]. The advice given is intended for students taking 1A, 1B and, particularly, Part II of the Cambridge mathematics tripos. It does not apply to Part III which is a very different course with different objectives.

I should **very much** appreciate being told of any corrections or possible improvements. This document is written in LaTeX2e and stored in the file labelled `~twk/FTP/Excess.tex` on emu in (I hope) read permitted form. It also available via my web home page. My e-mail address is `twk@dpmms`.

1 What is an essay question?

What students dislike most about essay questions is that they have to make choices for themselves about what to include and how to treat it. Since this dislike is often shared by examiners (giving students genuine choices makes marking much more difficult) many essay questions are simply disguised bookwork. Consider the following essay question.

Q. 1. *Write an essay on the Steinitz exchange lemma and its use in establishing the notion of dimension.*

A little thought shows that it could be rewritten as follows.

Q. 1′. *Define the terms 'spanning set', 'linearly independent set' and 'basis'.*
State and prove the Steinitz exchange lemma.
Show that any vector space V with a finite spanning set has a basis. Use the Steinitz exchange lemma to show that all bases of V have the same number of elements and so we can define the dimension of V.

[1] 'Two days wrong!' sighed the Hatter 'I told you butter wouldn't suit the works!' he added, looking angrily at the March Hare.
'It was the *best* butter,' the March Hare meekly replied.
'Yes, but some crumbs must have got in as well,' the Hatter grumbled: 'you shouldn't have put it in with the breadknife.' [Alice in Wonderland]

Now consider a possible essay question on numerical analysis.

Q. 2. *Suppose that we wish to find*

$$\int_{-1}^{1} f(x)\,\mathrm{d}x$$

but that it is expensive to obtain values of f. Discuss the use of orthogonal polynomials in evaluating the integral.

In principle, this is a more open question than Question 1 but in practice it is limited by the fact that the essay must be on a subject that you have covered. If the only method of integration discussed in the course is Gaussian quadrature then the question can only be on Gaussian quadrature and can be restated as follows.

Q. 2′. *Explain the method of Gaussian quadrature.*

A little reflection converts this into a bookwork question.

Q. 2″. *If $x_1,\ x_2,\ \dots,\ x_n$ are distinct points of $[-1,1]$ show that we can find $\lambda_1,\ \lambda_2,\ \dots,\ \lambda_n$ such the formula*

$$\int_{-1}^{1} f(x)\,\mathrm{d}x = \sum_{j=1}^{n} \lambda_j f(x_j) \qquad \bigstar$$

is exact for all polynomials of degree $n-1$ or less.

If the $x_1,\ x_2,\ \dots,\ x_n$ are chosen to be the roots of the Legendre polynomial of degree n show that the formula $\bigstar$ is exact for all polynomials of degree $2n-1$ or less. Conclude that the method of approximate integration given by $\bigstar$ is likely to be improved by this choice of x_j.

There is substantially more that can be said on Gaussian quadrature but (supposing the course to say no more) the rules of the game say that you are not expected to know more than the syllabus demands.

The next example of an essay question is, in my opinion, rather more demanding.

Q. 3. *Assuming any results about* exp *and* log *that you need, define x^α for $x > 0$ and $\alpha > 0$ and establish the basic results concerning the function $x \mapsto x^\alpha$.*

Here the problem is to decide which are the basic properties. Perhaps the reader should make her own list before looking at mine.

Q. 3′. *Set*

$$x^\alpha = \exp(\alpha \log x).$$

Stating carefully any results about exp *and* log *that you need prove the following results. (Here $x, y > 0$, α and β are real and n is strictly positive integer.)*

(i) $(xy)^\alpha = x^\alpha y^\alpha$.
(ii) $x^{(\alpha+\beta)} = x^\alpha x^\beta$.
(iii) $x^{\alpha\beta} = (x^\alpha)^\beta$.
(iv) $\dfrac{dx^\alpha}{dx} = \alpha x^{\alpha-1}$.
(v) $x^n = \overbrace{xx\ldots x}^{n}$.

Of course there are other ways to define x^α and the examiner would accept these[2]. If I was marking Question 3, I think that I would give full marks to an essay which included a satisfactory treatment of (v) (it is essential that our new definition of x^n should coincide with the old) together with three out of the four points (i), (ii), (iii) and (iv).

Perhaps you disagree with my list (remember that I only ask for three out of the first four points). Do you disagree that points (i) to (v) are basic or do you feel that that I should have included more points? Some people might feel that I should have included the fact that $x^0 = 1$ (though this follows from (ii)) or shown explicitly that our new and old definitions of x^α coincide when α is rational (although this follows from (v), I would certainly do this in any lecture course). Others might wish to include the fact that $x^\alpha \to \infty$ as $x \to \infty$ if $\alpha > 0$. On the other hand, most people would agree that the Taylor series for $(1+x)^\alpha$ is not a basic result in this context. There is room for disagreement but I hope that, after reflection, most mathematicians will agree that my list was reasonable (particularly if my marking scheme was flexible).

My final example is a fully fledged essay.

Q. 4. *Write an essay on Möbius maps.*

I shall use this as my typical example and I shall assume only the material contained in the first year course 'Algebra and Geometry'.

Question 4 is a full fledged essay because it involves organising material in a different way from the course (results on Möbius maps are scattered throughout the lectures), there is too much material to cover in detail in a short essay so we must select which topics to treat and, even after making a choice of topics, it is not possible to give all the proofs in full so we must choose what to prove, what to sketch and what to state.

[2] However, the other ways that I can think of involve quite a lot of hard work.

2 How is the essay marked?

If examiners have to mark a large number of mathematics essays quickly and reasonably fairly, it seems to me that they have have very little choice but to prepare a mark scheme in advance. It will not be quite as crude as 'give 2 marks for each fact mentioned' but will consist of instructions like 'give 2 marks for mention of either A or B and 3 marks if both mentioned'. (Although the examiner will try to cover all possibilities in advance the mark scheme will almost inevitably need to be modified to cover unexpected paths taken by examinees. If the changes are substantial the examiner may need to run through all the scripts a second time to maintain consistency).

If the essay allows genuine choice then the total marks available will exceed the maximum permitted mark for a question and the examiner will need to apply a scaling formula. Often the scaling formula is the crude

$$\text{final mark} = \min(\text{total marks}, \text{maximum mark}).$$

It is generally accepted that the distribution of marks for essay questions looks very different from the distribution of marks for ordinary 'problem' or 'bookwork' questions[3]. Since you can either do a piece mathematics or not, marks for 'problem' questions tend to be either very high or very low and the same is true to a lesser extent for 'bookwork' questions. However, it is harder to produce a very good or very bad essay and so it is harder to get very high or very low marks for an essay question.

Most directors of study would therefor advise you do do a non-essay question in preference to an essay question *if you can.* The sting is, of course, in the words *if you can.* A glance around the room two hours into an examination establishes that very few students can find enough questions that they can do to occupy the full examination period. You are certainly better off writing an essay than staring at the ceiling. I therefore tell my students that they should normally expect to tackle any appropriate essay.

Here is some further specific advice.

- *Tackle non-essay questions first.* If you get stuck there is longer for your subconscious to work at the problem. Return to the question at intervals to see if your subconscious has come up with anything.

- *Do the essay questions last.* The difficulty with essay questions is getting the organisation right rather than solving mathematical problems.

[3] Of course, any distribution can be made to resemble any other by 'grading to the curve' but, in my view, such drastic rescaling would be ridiculous. More to the point, no such rescaling is attempted in Cambridge mathematics exams.

Visit the question from time to time to make notes and to let your subconscious consider the matter. When you have definitely finished with the non-essay questions look at your notes, pick up your pen and write like hell.

- *Allocate roughly equal time to all the essay questions you tackle.* If your essay is good it will already have gained most of the marks available and extra work can not gain many marks. If your essay is poor any improvement is likely to gain marks.

3 Planning an essay for supervision

Since most of the marks for an essay are for specific points rather than for presentation and since any examination question must be done in a hurry you neither can nor should aim very high in the actual examination.

There are, however, various reasons for trying to do substantially better when preparing the same essay for supervision.

(1) If a similar essay turns up in the exam you will have a good model to guide you.

(2) If your supervision standards are high you can relax them in the exam. If they are low you will find it hard to improve you.

(3) Writing a good essay is a good form of revision. It helps you to see how different parts of the course hang together.

(4) Remember, however, that in a few years you may be defending a PhD, proposing unpopular changes to top management, giving a talk to a selection board or even lecturing to your own students. In such circumstances, it matters not merely what you say but how you say it. You will need to marshal your arguments and present them attractively. Treat the essay as training for such circumstances.

If an essay forms part of a week's work for a supervision you should start thinking about it at the beginning of the week, jotting down ideas as they occur to them. It is a good idea to run through the essay in your head several times during the week before putting pen to paper. (Compare my advice for doing the essay in the exam.)

Now imagine that you have to give a talk to other mathematicians who know nothing about the subject. Part of your job is to do some hard mathematics but you must do more than this. You must also try to give some idea of where your topic comes from and where it is going. You also wish to convince your audience that the topic is interesting.

One problem that faces you at once is that there is not time to prove all the results that you wish to discuss. This should not be treated as an invitation to prove nothing. Instead you must be selective. Your imaginary audience wants to see the important things proved properly but is happy to take the less important things on trust. Thus your first business is to choose the central theorem of your essay which will be proved fully. Next you need to consider the lemmas you will need in your proof. Some of these you may decide to prove, for some you may decide to give a sketch proof, others you will decide to merely state. In the same way, your central theorem may have associated corollaries, and counter examples and you will need to decide how much detail to give for each.

Here is a possible check list. Not all the items will be appropriate to each essay and those that do appear may occur in a different order. (For example, the preliminary definitions may mixed up with the preliminary lemmas.

(1) History. [If you decide to include some history, keep it down to a couple of sentences. This is a mathematics exam.]

(2) Preliminary definitions. [You may decide to give some illustrative examples.]

(3) Preliminary lemmas. [Proved, sketched or stated.]

(4) Main theorem. [Proved]

(5) Corollaries. [Proved, sketched or stated.]

(6) Counter-examples showing that the results can not be extended. [I would tend to sketch rather than prove or state.]

(7) In what directions does the result lead? Why is it useful?

You do not have to write your essay in linear order. Since most people tend to spend too much time on the preliminaries it may be a good idea to start by writing out the main theorem and then add the preliminaries. If you find that your essay is too short you can expand it by adding further proofs at the end with brief directions towards the proofs in the text. It is a good idea to write on alternate lines to leave room for such alterations.

Every supervisor knows that students can be divided into a large class who write too little for supervisions and a small class who write too much. If you belong to the class which writes too much, remember that condensation and selection are key skills which the essay is intended to encourage. If you can not prevent yourself from producing fifteen pages of beautifully neat handwriting covering every aspect of the subject in complete detail, write your masterpiece and then set yourself the task of producing a three page summary[4].

[4]You may on other hand, be one of those undergraduates who spent the first part of the week rehearsing for the Boat Club's production of Aida, but had genuinely left lots of time but your pet hamster died and then you lost your room key on a ten mile hike in aid of charity and then you had to take tea with the Master and then there was so much to

All this advice may have a contrary effect to that intended and make the task of writing an essay seem dauntingly complicated. It is not. Try it and see. Your first essay will be hard to write but not as hard as you thought. Your second essay will be easier to write and read better than the first and your third essay will be easier to write than your second. Thereafter essay questions will be simply part of the day's work.

Up to now I have stressed the similarities between an essay and a talk. Let me finish this section by pointing out one important difference. At best, your essay will be a sort of condensed talk which would need to be diluted by commentary and by the jokes, anecdotes and other devices which provide pauses for the audience to catch up. It is instructive to look at a TV science programme or to read a popular science essay and then list the number of ideas that the author tries to get across. Usually there are only one or two. Such a style is unsuitable for mathematics but it should, nonetheless, be observed that mathematicians are prone to confuse high information content with high information transmission.

4 An essay with commentary

[We write the commentary in sanserif and the essay in ordinary type. Recall the question

Q. 4. *Write an essay on Möbius maps.*

We start by jotting down all the topics we can remember concerned with Möbius maps.

(1) They take circles and straight line to circles and straight lines.
(2) They form a group.
(3) They can be broken up into simpler maps.
(4) They can be used to map any three points into any three points.
(5) They are connected with $SL(\mathbb{C}, 2)$.
(6) We have to deal with ∞.

We can not remember anything about (5) so we decide not to include it. Point (1) seems the most important so we decide to make it the centre of our essay. We also take the key decision to make as much use of (3) as we can. We shall treat the points in order (3), (2), (1), (4). Since we can not think of any interesting historical titbit we go straight to the definitions.]

---

do that you were too depressed to do anything, particularly because of your cold. In such a case remember that producing a skeleton plan of your essay is substantially better than producing nothing.

We work on the extended complex plane $\mathbb{C}^* = \mathbb{C} \cup \infty$. Suppose that a, b, c, $d \in \mathbb{C}$ with $ad - bc \neq 0$. If $c = 0$ we define

$$\begin{aligned} f(z) &= (az+b)/d \text{ if } z \in \mathbb{C} \\ f(\infty) &= \infty \end{aligned}$$

If $c \neq 0$ we define

$$\begin{aligned} f(z) &= (az+b)/(cz+d) \text{ if } z \in \mathbb{C} \text{ and } z \neq -d/c \\ f(-d/c) &= \infty \\ f(\infty) &= a/c. \end{aligned}$$

We write

$$f(z) = \frac{az+b}{cz+d}$$

and call f a Möbius map. In what follows we shall often leave the treatment of special cases like $c = 0$ or $z = \infty$ to the reader. [This is the kind of fiddly but easy stuff which can be omitted from an essay.]

We write $\mathcal{M}$ for the set of Möbius maps. We are particularly interested in the following elements of $\mathcal{M}$ which we call 'elementary maps' (here $\lambda \neq 0$)

$$\begin{aligned} T_a(z) &= z + a \\ D_\lambda &= \lambda z \\ J(z) &= 1/z. \end{aligned}$$

We observe that T_a is a translation and that if $\lambda = r \exp i\theta$ with $r > 0$, θ real then D_λ is a rotation through θ about 0 followed by a dilation. We observe also that $T_a T_{-a} = I$, $D_\lambda D_{\lambda^{-1}} = I$ and $J^2 = I$ where I is the identity on $\mathbb{C}^*$.

Lemma 4.1. *(i) If S is an elementary Möbius map and $T \in \mathcal{M}$ then $TS \in \mathcal{M}$.*

(ii) Every Möbius map is the composition of elementary maps.

(iii) $\mathcal{M}$ is precisely the set of compositions of elementary maps.

[I would be inclined to leave this completely unproved. If we decide to prove anything it should be the least obvious thing which in my opinion is (ii).]

Combining our observations on the nature of the elementary maps with part (iii) of Lemma 4.1 we obtain the following result.

Theorem 4.2. *$\mathcal{M}$ is a set of bijections $\mathbb{C}^* \to \mathbb{C}^*$ forming a group under composition.*

[The proof of Lemma 4.2 given here could hardly be presented to beginning students without a great deal of discussion. However the presumed audience for an essay consists of good mathematicians who merely know nothing about the particular subject of the essay[5].]

We now study the effect of Möbius maps on circles and straight lines. We begin by finding a suitable characterisation of these objects.

The equation of a circle is

$$|z-a|^2 = r^2$$

with $a \in \mathbb{C}$ and $r > 0$. Expanding, using the relation $|w|^2 = ww^*$ gives

$$zz^* - a^*z - az^* - (r^2 - |a|^2) = 0.$$

Finally, multiplying by a real non-zero constant A gives

$$Azz^* + Bz^* + B^*z + C = 0$$

with A, C real, $A \neq 0$ and $|B|^2 > AC$. We note that the working is reversible so we have a complete characterisation of a circle.

The equation of a real line is

$$\alpha x + \beta y = \gamma$$

with α, β, γ real and α, β not both zero. This may be rewritten

$$\alpha(z + z^*) - i\beta(z - z^*) = 2\gamma$$

so that, collecting terms we have

$$(\alpha - i\beta)z + (\alpha + i\beta)z^* - 2\gamma = 0$$

that is

$$B^*z + Bz^* + C = 0$$

with C real and $B \neq 0$ or equivalently $B^2 > 0C$. Once again the working is reversible.

Combining these facts with some further simple observations we obtain our key theorem.

[5] As Szilard, is claimed to have said 'You may assume zero knowledge and infinite intelligence'.

Theorem 4.3. *The general equation of a circle or line is*

$$Azz^* + Bz^* + B^*z + C = 0$$

with A, C real and $|B|^2 > AC$.

We have a line if and only if $A = 0$. The circle or line passes through 0 if and only if $C = 0$.

Since Theorems 4.3 and 4.4 form the centre of our essay we give more detail in the proof of Theorem 4.3 than elsewhere in the essay.

Theorem 4.4. *(i) The elementary Möbius maps translation $z \mapsto z + a$ and dilatation and rotation $z \mapsto \lambda z$ $[\lambda \neq 0]$ takes circles to circles and straight lines to straight lines.*

(ii) The elementary Möbius map $z \mapsto 1/z$ takes circles and straight lines passing through 0 to straight lines and all other circles and straight lines to circles. It takes straight lines to circles and straight lines passing through 0.

(iii) Möbius maps take circles and straight lines to circles and straight lines.

Proof. (i) Obvious.

(ii) By Theorem 4.3 the general equation of a circle or line is

$$Azz^* + Bz^* + B^*z + C = 0. \qquad (*)$$

with A, C real and $|B|^2 > AC$.

The mapping $z \mapsto 1/z$ takes the curve described by equation $(*)$ to the curve described by

$$Czz^* + B^*z^* + Bz + A = 0. \qquad (**)$$

with C, A real and $|B^*|^2 > CA$, that is another circle or straight line. The remaining statements follow from the second paragraph of Theorem 4.3.

(iii) Every Möbius map is the composition of elementary Möbius maps. □

It can be shown that Möbius maps preserve the angle between curves. [An essay can contain a certain number of unproved results though a long sequence of such assertions makes rather dull reading.]

Here is an example of the use of Theorem 4.4. [Since this example is not in the syllabus, it should be possible to get full marks without including it. However, it makes the essay more interesting and we would certainly get credit for it.]

Theorem 4.5 (Steiner's porism). *Suppose C and C' are non intersecting circles with C' inside C. If C_1, C_2, ...form a sequence of circles arranged in clockwise order with C_1 tangent to C and C' and C_{j+1} tangent to C and C' and C_j $[j \geq 1]$ we say that they form a chain of circles. If $C_n = C_1$ we say that the chain is closed. Steiner's porism states that if one choice of C_1 gives a closed chain then all choices do.*

[This should be illustrated by a diagram. Use diagrams freely throughout your essay. Remember, a picture is worth a thousand words.]

Sketch proof. We can find a Möbius transform which takes C and C' to concentric circles. The transform takes chains of circles to chains. Since Steiner's porism is trivial for concentric circle we are done. □

[This really is a sketch. It is not obvious without some further argument that there is a Möbius transform which takes C and C' to concentric circles. We have not shown explicitly that tangent circles are taken to tangent circles.]

Just as a simple translation allows us to move any point in $\mathbb{R}^n$ to the origin so Möbius maps allow us to move any three points in $\mathbb{C}^*$ to any other. [Note that we are trying to show that Möbius maps are useful. A skeptical reader would not be fully convinced without the inclusion of material from later courses.] Here is a particular example.

Lemma 4.6. *(i) If $w \neq \infty$, the Möbius map $z \mapsto 1/(z-w)$ takes w to ∞.*

(ii) If $w \neq \infty$ the Möbius map $z \mapsto z - w$ takes w to 0 and fixes ∞.

(iii) If $w \neq 0,\ \infty$ the Möbius map $z \mapsto z/w$ takes w to 1 and fixes 0 and ∞.

(iv) If w_1, w_2, w_3 are distinct then there exists a Möbius map T with $Tw_1 = 0$, $Tw_2 = 1$ and $Tw_3 = \infty$

Using Lemma 4.6 it can be shown that the following general result holds. [We are not obliged to prove everything, but it helps the reader if we make it clear when we are not proving something.]

Theorem 4.7. *If w_1, w_2, w_3 are distinct and z_1, z_2, z_3 are distinct then there exists a Möbius map T with $Tw_j = z_j$*

The map T of Theorem 4.7 is unique. To show this we introduce the idea of the cross ratio $[z_1, z_2, z_3, z_4]$. [Generally speaking it is bad exposition to introduce a definition that is used only once but this is an exam and we wish to display our knowledge.]

Definition 4.8. *We set*

$$[z_1, z_2, z_3, z_4] = \frac{(z_1 - z_2)(z_3 - z_4)}{(z_1 - z_4)(z_3 - z_2)}$$

treating ∞ as as in the definition of the Möbius map.

[This is more of sketch of a definition than a definition]

Lemma 4.9. *If T is a Möbius map*

$$[Tz_1, Tz_2, Tz_3, Tz_4] = [z_1, z_2, z_3, z_4].$$

Sketch proof. The result is true for elementary Möbius maps by direct calculation and so true for all Möbius maps by Lemma 4.1 (iii). □

Lemma 4.10. *The Möbius T of Theorem 4.7 is unique.*

Sketch proof. We shall show that the result is true for the case $w_1 = z_1 = 0$, $w_2 = z_2 = 1$, $w_3 = z_3 = \infty$ since then

$$z = [z, 0, 1, \infty] = [Tz, T0, T1, T\infty] = [Tz, 0, 1, \infty] = Tz.$$

The general result may be obtained from this. □

This shows that, in general there does not exist a Möbius map allowing us to move any four points in $\mathbb{C}^*$ to any other.

[The syllabus also includes *discussion of relations between eigenvalues of a matrix and fixed points of a Möbius map*. This idea did not occur in the list of jottings with which we started. My feeling is that the essay is complete without it but of course the examiner could disagree. One could write a very nice essay ending with a discussion of the behaviour of the sequence $T^n z$ for different Möbius maps T. Many students will have seen a discussion of the homomorphism

$$\theta : SL(\mathbb{C}, 2) \to \mathcal{M}$$

given by

$$\theta(A)(z) = \frac{az + b}{cz + d} \text{ where } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

This could also have been included in the essay but I have chosen not to.

Not only could we have chosen different topics, but we could have chosen much the same topics but dealt with them very differently.]

Exercise 5. *Write an essay on the Möbius transform based on the following outline.*

(i) Define the cross ratio.

(ii) Show that the set of maps $T : \mathbb{C}^* \to \mathbb{C}^*$ *which leave the cross product of every four points unchanged is a group of bijections of* $\mathbb{C}^*$. *We call this group* $\mathcal{M}$.

(iii) Show that $\mathcal{M}$ *consists precisely of the maps*

$$z \mapsto \frac{az + b}{cz + d}$$

with $ad - bc \neq 0$.

(iv) Show by elementary geometry (recall that angles on the same chord of a circle are equal or complementary) that four points lie on a circle or straight line if and only if their cross ratio is real.

(v) Prove that the Möbius maps take circles and straight lines to circles and straight lines.

5 What can you expect from your supervisor?

When your supervisor looks at your essay he or she will, like the examiner, have a series of points in mind. Have you included results A and B? If you used method C did you take care to define D? If, on the other hand, you used method E how did you deal with the rather tricky argument to establish F?

If your arguments are fallacious or you leave out an important point your supervisor will tell you but often he or she will be less prescriptive merely saying that you could have included this or excluded that but that it was very reasonable to proceed as you did. Some students find this very upsetting but it reflects the fact that there are many different ways of writing a good essay. Your supervisor is more concerned with getting you to think about the general idea of essay writing than polishing a particular essay.

An ordinary supervision runs more effectively if students bring along a list of the problems they can not resolve and the points they wish to discuss. In the same way the discussion of an essay question is likely to be more productive if you make a list of the decisions you made in writing your essay and, in particular, any decisions you are worried about. Your supervisor may not be able to resolve all your worries (there is, I repeat, no unique recipe for writing an essay) but the discussion will usually be both helpful and reassuring.

If you have no specific problems it may be instructive to discuss the following kinds of question.

(1) If one extra result was to be added to the essay to make it longer which should it be and why?

(2) If one result was to be removed from the essay to make it shorter which should it be and why?

(3) If one proof was to be expanded which should it be and why?

(4) If one proof was to be shortened which should it be and why?

It may be noted that you do not need the presence of a supervisor to ask yourself these questions.

Exercise 6. *What would your replies to questions (1) to (4) if they were asked about the essay in Section 4.*

Good luck with your supervisions and your exams.