The surveyability of long proofs

Edwin Coleman
University of Melbourne

Abstract. The specific characteristics of mathematical argumentation all depend on the centrality that writing has in the practice of mathematics, but blindness to this fact is near universal. What follows concerns just one of those characteristics, justification by proof. There is a prevalent view that long proofs pose a problem for the thesis that mathematical knowledge is justified by proof. I argue that there is no such problem: in fact, virtually all the justifications of mathematical knowledge are 'long proofs', but because these real justifications are distributed in the written archive of mathematics, proofs remain surveyable, hence good.

Keywords: mathematical practice; mathematical proof; surveyability; written discourse

1. Introduction: a twofold need

Both philosophy of mathematics and the study of argumentation seem to me to be in rather poor shape.[1] One way they can both be improved is by recognition of the role of media in communication. My aim here

[1] This is not the place to argue for these large claims, but here is an indication of why I hold them. Philosophy of mathematics is in poor shape for, as Putnam (1979) put it long ago, 'in philosophy of mathematics, nothing works.' Apart from some welcome attention to practice, not much has changed. This is because the only positions taken seriously are variations of those associated with foundational studies — logicism, intuitionism, formalism — all of which have mortal wounds. There's lots of activity in the field, but almost all of it takes for granted the claim of formal logic to authority over the argumentation in at least the proofs of pure mathematics (even the occasional deviant). Correlative to that is an obsession with the 'objects' of mathematics and their metaphysical salvation from skepticism, or subjection to it. The real nature of mathematical knowledge is hardly discussed, the assumption being that if we get the nature of the objects straight, then what knowing them consists in will be a simple corollary. Many people deny that there are any mathematical objects ('really') because if there were we couldn't know anything about them. This leads to much tortuous nonsense.

The study of argumentation might seem to be just the source for a better understanding of mathematical knowledge, since it is concerned with real argumentation rather than the idealised simulacrum that formal logic indicates. There are two reasons why I doubt this. The first is that the quite proper concern of the study of argumentation with non-deductive argumentation might lead the discussion toward philosophically peripheral aspects of mathematical practice. This is evident in uses which some writers in the mathematical education community have made of the so-called 'Toulmin model' (for example Inglis et al., 2007). My point is not that I want to insist that these elements of practice can be ignored for all serious purposes — on the is to discuss the nature of the core argumentation in mathematics — proofs — and to show why there is no threat from the existence of 'long proofs' to the claim that MATHEMATICAL KNOWLEDGE IS JUSTIFIED BY PROOFS. It is rarely acknowledged in either field that the kinds of argument most of significance in modern culture are all heavily dependent on **writing** — they are practices pursued *essentially* through the development, accumulation, preservation, transmission and use of **archives**. None is more so than mathematics. The next section will provide a simple example.

2. Hobbes' delight, modern disquiet

Recent philosophy of mathematics has obsessed about the alleged mysteriousness of mathematical objects, but I think the primary phenomenon to understand is the peculiarity of our mathematical knowledge. That peculiarity is well-expressed by a story John Aubrey (1958) tells about Thomas Hobbes:

> He was (vide his life) 40 yeares old before he looked on geometry; which happened accidentally. Being in a gentleman's library, Euclid's Elements lay open, and 'twas the 47 El. libri I.1. He read the proposition. 'By G–' sayd he, 'this is impossible!' So he reads the demonstration of it, which referred him back to such a proposition; which proposition he read. That referred him back to another, which he also read. Et sic deinceps, that at last he was demonstratively convinced of that trueth. This made him in love with geometry.

What was it made Hobbes 'in love with geometry'? Aubrey's 'this' is ambiguous. I take it to refer primarily to the process of justification through which Hobbes moved back from the statement of Pythagoras' Theorem ['47 El. libri I.1']:

THEOREM 1. *The square on the hypotenuse of a right-angled triangle equals the squares on the other two sides*

to the propositions on which it depends, to the propositions on which *they* depend, and so on, to the axioms, common notions and definitions with which Euclid begins. (Students of argumentation will of course

contrary — but that they distract us from trying to give a better treatment of those elements traditionally and *rightly* seen as central for the philosophy of mathematics, such as proofs as given by Euclid. My other reason for doubting the usefulness of the extant study of argumentation is that there is no real *theory* of argumentation to rival formal logic, just a congeries of interesting and useful but weak 'approaches' to the study of argumentation (cf. the survey book of van Eemeren et al 1996).

recognise the similarity between Hobbes' convincing himself and the common dialogic process of asking for and giving reasons.) But I also take Aubrey's 'this' to refer to the initial incredulity which that process dispelled. The peculiarity which struck Hobbes so forcefully has two aspects: the theorem of Pythagoras is unexpected, and yet it is demonstrable. We will need to keep this duality in sight below.

It's commonly thought that proof is what distinguishes mathematical knowledge. Euclid's proof is paradigmatic. But several achievements of modern mathematics have beeen seen as posing problems for the notion that proof is the central vehicle of mathematical knowledge making and learning. Tymoczko questions whether the Appel-Haken proof (Appel et al., 1977) of the four color theorem is really a proof, or if it is, then whether the nature of proof has changed. What the theorem says is simple — four colors suffice to color any proper planar map — but the proof required some hundreds of hours of computer calculations which have not been printed, and even if they were, no-one could read them in a lifetime.

There has been less discussion about the proof of *The Enormous Theorem* — the classification of the finite simple groups —

THEOREM 2. *Every finite simple group belongs (up to isomorphism) to at least one of three infinite families (cyclic groups with prime order; alternating groups of degree at least 5; simple groups of Lie type), or is one of 26 sporadic simple groups.*

The proof is 'the union of about 500 papers totalling about 10000 pages' (Gorenstein, 1985). Very few people have even read all this material, yet alone verified it as correct. Some group theorists have privately expressed doubts whether anyone has.

A third example is Andrew Wiles' proof (Wiles, 1995) of Fermat's Last Theorem, which says

THEOREM 3. *If $x^p + y^p = z^p$ with x, y, z, p natural numbers and $p > 2$, then $xyz = 0$.*

That proof is not so long in terms of pages — merely a couple of hundred, including a patch from Wiles and Taylor — but those pages are very demanding pages — they require acquaintance with a large range of highly technical material. No serious doubts about the correctness of this proof have been expressed recently, but there is some disquiet about how properly we can speak of a proof which is not available to most mathematicians, let alone most people. (I recently read a mathematican blogging that he intends to spend *two years full-time* getting up to speed on the theories required, in order to read the proof.)

These proofs among others are thought to be problematic because they lack a quality found in Euclid's: they are said not to be 'surveyable'.

3. Requirements for proofs

There's a good deal of agreement that proofs need to be surveyable, though this agreement disguises considerable vagueness as to just what surveyability is. It is not the difference between a correct proof and one which is mistaken, for example. Nor should the proof be called unsurveyable that Fermat claimed to have in his famous marginal comment, rather it is (so far) simply unseen. Before we consider what surveyability in proofs might be, and why it might be valuable, we need first to distinguish several related but different things that may be referred to as 'proofs'. Conditions for proof fall into three groups, roughly corresponding to the traditional tripartite division of metaphysics, epistemology and ethics.

The first point is that 'proof' is a success noun: when I say without qualification that P is a proof of T, I am endorsing P as really proving T, demonstrating T, establishing T as true. Sometimes we call proofs what should more accurately be called candidate, putative or attempted proofs, or even failed proof attempts. Students offer candidate proofs which are frequently, alas, not really proofs. Mathematicians do the same, but less frequently. For example, Lamé put forward in 1847 a supposed proof of Fermat's Last Theorem which was immediately questioned by Liouville, and indeed had previously in effect been shown by Kummer to be wrong, for the reason Liouville suggested. A genuine proof is at least a *sound argument* to the theorem's claim.

The second point is that a proof must be given in an *actual written instance*. Although proofs, or more often their leading ideas, are frequently conveyed by word of mouth, only written proofs count as establishing a theorem. (More argument for this claim is given below.) Moreover, describing a possible written instance does not count as giving a proof because our grasp of possibilities is too weak. For example, it would be possible to use a mixture of Chinese characters and astrological symbols as the labels for points in Euclid's proof of I.47; I don't suppose anyone has done that, but there are clearly various possible such proofs. These would be mere notational variants of Euclid's, but none of them is actually a proof because none of them is actual and so cannot be examined for correctness.

Those two conditions might be called ontological. The next conditions are epistemological. The third point is one familiar in argumentation theory: a putative proof is only genuinely a proof if it can be known to be a proof, because the point of proof is to establish knowledge. A correct deduction is not necessarily a sound argument, for just this reason. If the premises involve concepts we cannot grasp, or inferences we cannot see, their correctness is not enough to make a sound argument from them. This gap between correct deduction and sound argument is not confined to mathematical proofs.

The fourth point is that knowability requires comprehensibility. No doubt there can be 'theories' with hundreds of primitive operations and hundreds of axioms about them and programs can be written to correctly derive consequences of these axioms. But these will (mostly) not be proofs since the claims derived will be incomprehensible, and hence unknowable, to agents like us. Such purported proofs might not be long, but they would be in another way too complex to be comprehensible. We might want to rule them out for the same reason some long 'proofs' might need to be ruled out.

Only items satisfying all four of the preceding conditions are genuinely proofs; those that do not satisfy the first are incorrect, the second not actual, the third not cogent, the fourth incomprehensible.

The third kind of condition will probably be more contentious, since they concern the value of a proof. So I will say that these are the conditions which must be met for a proof to be *good*, allowing that there can be genuine proofs — satisfying the preceding conditions — which are not good ones.

The fifth point is that a proof P of T can only be good if T is interesting. From a formal point of view, there is a proof in Euclid of a theorem obtained by stopping one step short of the real conclusion of I.47. Formally, theorems, lemmas and corollaries are all on a par. Indeed a calculation showing that $77 \times 180 = 13860$ is formally a proof. But all this shows is one of the numerous limitations of the formal point of view for understanding mathematics. Theorems have to be discoveries, they ought to elicit the Hobbes response 'By G–' etc. Showing that 77 triangles have angle sum 13860 degrees is not a proof in Euclidean geometry.

Finally, P is only a good proof of T if P is interesting — if the ideas on which it depends are valuable. A real proof justifies asserting the theorem; but a *good* proof must also give some kind of insight. As Lord Rayleigh is reported (for example, by Huntley and by Kline, though I have been unable to locate the original source) to have said: 'Some proofs command assent. Others woo and charm the intellect. They evoke delight and an overpowering desire to say, 'Amen, Amen'.'

Actually there are many dimensions along which (real) proofs might be compared for goodness. One proof is better than another because it is more elegant, or informative, or shorter, or easier to follow, or less conceptually demanding, or more easily generalisable, or with less fussy detail, or more direct, or more constructive, or even more novel... These various values might usefully be classified and compared, but not here.[2] For our purposes it is enough to stick to the essence of Hobbes' delight: MATHEMATICAL KNOWLEDGE IS JUSTIFIED BY GOOD PROOFS.

Now let's consider how we might fit in to this picture the claim that proofs need to surveyable. (In a recent survey [!] of the literature, Bassler (2006) distinguishes local (each step) and global (overall) surveyability and says that both are needed for a real proof. I'll discuss that distinction in a subsequent section.)

4. 'A mathematical proof must be surveyable'

Doubts about Hobbes' delight surface from time to time when the question is raised as to whether long proofs are real proofs. Why might we doubt that long proofs really give us sufficient reason to claim knowledge of the 'theorems' they supposedly prove? Azzouni (1994) says that in 'the old days' one accepted that

> The mathematician proves truths ... proofs were (more or less) detailed arguments; these arguments were (more or less) valid ... and these arguments were surveyable, provided one had the training.

Azzouni suggests that more surveyability makes for better proofs *because they are easier to understand* (1994, p. 125). Later he rebuts an objection to one of his arguments by saying that certain proofs are not unsurveyable, though long. He recognises that in reality individual mathematicians have not actually surveyed many of the proofs which justify their knowledge, but accept them on the authority of others who have — a point to which I will return in section 10. But he evidently thinks that this 'in principle' surveyability, because it is shown by actual survey *distributed* in the community, is actual surveyability. By contrast proofs such as Appel-Haken which

> cannot even be checked by the mathematical community as a whole by examining the proof the computer has generated ... are really not proofs at all.

Unfortunately his argument for this claim consists in a thought experiment, a technique which is very rarely convincing to anyone who disagrees. But anyway, Azzouni requires that a real proof be given at least *distributed* expert survey, and he explicitly rules out the four color theorem and by implication some other long proofs, as not being

[2] For a start, cf. Tao (2007) on good mathematics.

real proofs. So, IF REAL PROOF REQUIRES AZZOUNI-WARRANTABILITY, THEN LONG PROOFS ARE NOT REAL PROOFS. Azzouni thinks the four colour theorem put an end to the old days; but concerns about long proofs lacking surveyability go back at least to Descartes.

Descartes was concerned that such proofs could not be held whole before the mind's eye and so not reach the cognitive standard of intuitions. He writes (Cottingham, 1985) in his seventh rule for directing the mind

> In order to make our knowledge complete, every single thing relating to our undertaking must be surveyed in a continuous and wholly uninterrupted sweep of thought, and be included in a sufficient and well-ordered enumeration.

In order to explain why the 'sweep of thought' is needed, Descartes makes it plain that knowledge is best if it is an intuition, second best if it is immediately deduced from such, and that third best is something achieved by a repeatedly surveyed deduction. Without it, mistakes are likely to creep in because our memory of links in the deductive chain is unreliable. (What 'enumeration' means in his statement is rather obscure, even confused, and I will ignore it as far as possible; it seems to be connected to making sure no deductive links are omitted.) So the problem is with memory: 'our undertaking' will only be knowledge if all of it is *present before the mind's eye together*, which any reliance on memory prevents. It is certainly true that many long proofs don't satisfy Descartes-surveyability — nobody can sweep through the calculations in the four color theorem in the requisite way. Nor can anyone do it with either of the proofs of the Enormous Theorem or Fermat's Last Theorem. In fact *most* proofs in real mathematics fail this test. IF REAL PROOFS REQUIRE THIS CARTESIAN GAZE, THEN LONG PROOFS ARE NOT REAL PROOFS.

Kitcher (1983) discusses the 'worry from long proofs' from a rather Cartesian point of view, though his target is the claim that mathematical knowledge is apriori rather than an elucidation of surveyability. Since I regard 'apriori' as just a historical label for that *evident* peculiar quality of mathematical knowledge which Hobbes expresses and we want to understand, I have no sympathy with Kitcher's project or his main argument that any doubtfulness somehow proves mathematical knowledge to be empirical.

Wittgenstein (1978) writes

> 'A mathematical proof must be perspicuous' [übersehbar].[3] Only a structure whose reproduction is an easy task is called a 'proof'. It must be possible to decide with certainty whether we really have the same proof twice over, or not. The proof must be a configuration whose exact reproduction can be certain.

Wittgenstein's interest in the concept of surveyability for proofs seems to have been transferred from some idea about 'perspicous representations' which he got from Frege. His transfer is not perhaps a happy one. His example of such a perspicuous representation is a color octagon which gives us an overview of the 'grammar' of color, which we can take in at a glance, and his first remarks about surveyability relate to finding such perspicuous representations for other parts of *language*. All that is part of the 'perfectly logical language' wild goose chase. It is later, when he is discussing mathematics, that he makes the remarks quoted. His concern is not about what Bassler calls local surveyability, which roughly means I must be able to check each step of the proof myself. Rather, he thought some 'proofs' might not be easily, exactly iterable. Given what else Wittgenstein says about proof, that a theorem is a new rule for the terms in it, that the proof gives the theorem its content, and so on, I think we should take him to be concerned with the subsequent use of the proof in the subsequent use of the theorem. You can't use the theorem, use its novel sense, unless you grasp the proof properly. (An example in section 7 below will elucidate this.) So Wittgenstein's argument is that proofs need to be surveyable because good proofs are usable and only surveyable proofs are usable. IF REAL PROOFS REQUIRE WITTGENSTEIN-SURVEYABILITY, THEN LONG PROOFS ARE NOT REAL PROOFS.

Tymoczko thinks long proofs may require the introduction of empirical elements, and hence fallibility, into mathematics. He writes (1998, p. 247)

> A proof is a construction that can be looked over, reviewed, verified by a rational agent. We often say that a proof must be perspicuous, or capable of being checked by hand. It is an exhibition, a derivation of the concluson, and it needs nothing outside itself to be convincing. The mathematician surveys the proof in its entirety and thereby comes to know the conclusion.... to say that [proofs] can be surveyed is to say that they can be definitively checked by members of the mathematical community.

[3] The word 'übersehbar' translated sometimes 'perspicuous' and sometimes 'surveyable' is actually commented on by the editors of Wittgenstein as difficult to render.

Because their computer calculations are not able to be looked over by mathematicians, Tymoczko questions whether the Appel-Haken proof (Appel et al., 1977) of the four color theorem is really a proof, or if it is, then whether the nature of proof has changed. IF REAL PROOFS NEED TYMOCZKO-SURVEYABILITY, LONG PROOFS ARE NOT REAL PROOFS.

Bassler says that the Appel-Haken proof lacks local surveyability — no-one can check every step of the calculations — so on his view too, long proofs lack surveyability. All these writers connect their concern about long proofs with surveyability or the lack of it. Their common argument is this: LONG PROOFS ARE NOT SURVEYABLE, BUT REAL PROOFS MUST BE SURVEYABLE, SO LONG PROOFS ARE NOT REAL PROOFS. Their conceptions of surveyability are rather varied though. Before I criticise them and propose a better, here's a further motivation.

5. Short proofs are really long!

Formalists regard the proofs given in ordinary mathematics as 'informal': they think the real proofs that anchor mathematical truths are their formal counterparts in some more polished descendent of *Principia Mathematica* or *Grundgesetze*. But a more convincing suggestion for the 'real' proofs is that they are fully expanded informal proofs. Unfortunately virtually all such proofs are long.

Euclid's proof of Pythagoras' Theorem (I.47) is a short proof. Moreover, it's easy to get the idea of the proof — you divide the big square with a line from the right-angle vertex parallel to its sides, and then use theorems about triangles between parallels to show that the two parts are equal to the two smaller squares (cf. Figure 1). Of course the proof as given explicitly calls on earlier theorems, and so invokes the recursive tree of justifications referred to by Aubrey's description of Hobbes' enlightenment.

Consider this suggestion: Euclid's justification of Pythagoras' Theorem is not simply the proof as given in theorem I.47, because it takes for granted the truth of propositions 4, 14, 41 and 46 as well as axiom 6, all of which are called on in the proof. They in turn call on other theorems, axioms, postulates and common notions. If we recursively eliminate these dependencies, by replacing them with suitablly relabelled versions of what they cite, we will obtain the *fully written proof* of I.47. That is the real proof of Pythagoras' theorem, because it is what justifies Hobbes' knowledge and that of myriads of students of geometry.

How long is this 'real proof' of Pythagoras' Theorem? Well, I.1 for instance is used *over thirty times* in the development leading to Pythagoras' Theorem, so the real proof will include over thirty in-

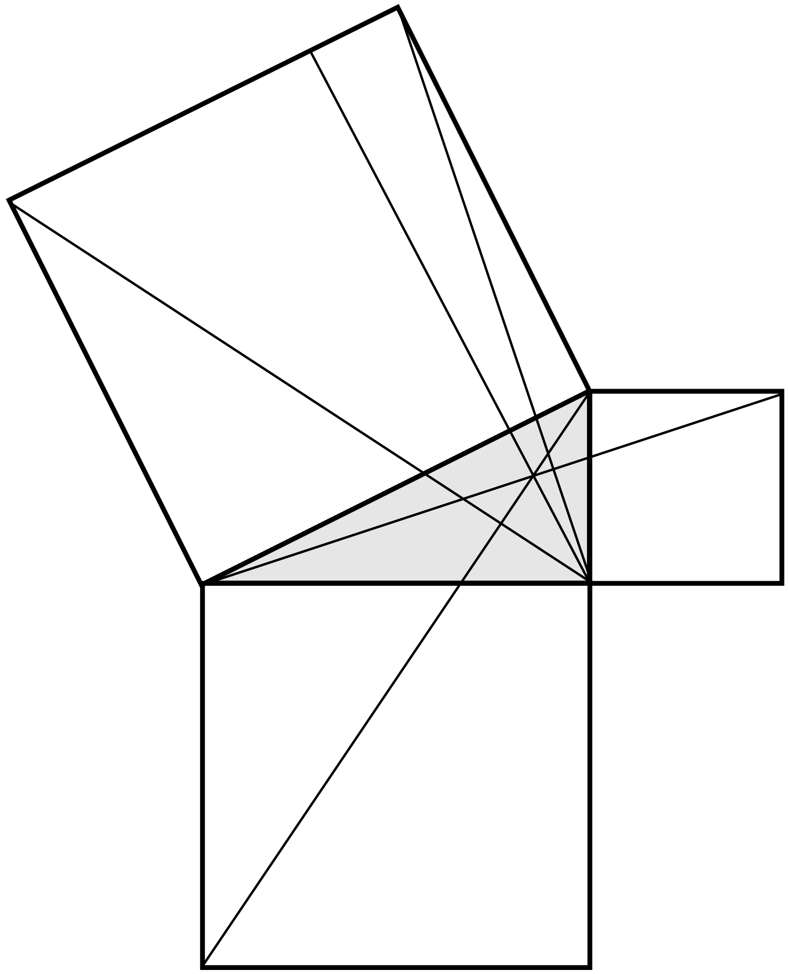

Figure 1. Euclid's proof of Pythagoras' Theorem. (Diagram by Scott Sutherland, Stony Brook University. Reproduced with permission.)

stances of the reasoning in that proposition, with minor variations in the labelling of points. To get the real proof requires these replacements and many others for other theorems used in the development. The fully written proof occupies at least 50 pages (or one extremely large piece of paper — so large that if you can read individual words then you can't see all the text). Is this real proof surveyable? Pace Descartes, reading it over and over won't provide a way for the mind's eye to do what the real eye can't. Pace Wittgenstein, this will make it impossibly harder to use Pythagoras' Theorem in a way that reflects the specific idea of the I.47 proof. Pace Azzouni, if no-one had actually carried this out, we would not actually have (Euclid's) apriori warrant for Pythagoras' Theorem, so we would not know it (of course, there are other proofs but the same considerations apply to them).

Euclid's proof is actually short and simple — mathematics is replete with much more complicated and moderately lengthy good proofs which rest on far more prior mathematics than does Pythagoras' Theorem, for example the proofs that e and π are transcendental. Their fully expanded counterparts would be enormous — every one would be like the 'proof' of 1+1=2 in *Principia Mathematica*, which rests on a prodigious prior development of some 200 dense pages.

The consequence is that the 'real proof' of Pythagoras' Theorem, and virtually every other proof in actual mathematics, is very long, so long as to be unsurveyable, and hence unperspicuous. But if so, then none of the theorems which these proofs 'prove' are actually known by any of us. This conclusion is so silly that something must be wrong in the reasoning which leads to it. What is wrong of course is the claim that the proofs we have, that is, the proofs in the literature, are not the real proofs. If so, does that mean that the proofs we rely on are (mostly) surveyable — but not really justifications ? And if so, what is the value in surveyability?

6. What surveyability could not be

Neither what surveyability consists in, nor its value, may be simple; different writers clearly have in mind somewhat different things. Before making my proposal, let's see what's wrong with others.

Surveyability in the sense of Descartes — that every detail of the proof may be present to consciousness together — is not really a requirement of knowledge, more a rationalist prejudice. Descartes is just plain wrong about the value of the mind's eye — the real eye trumps it every time. Euclid I.47 *in writing* can be held before the real eye altogether, but although one cannot attend simultaneously to every part of it, one can move one's attention from part to part *knowing that it does not change*. This is just not true about a mental version. Real proofs are written not intuited. Most of them cannot even be viewed altogether anyway, even with the real eye, (though possible versions of them on very large pieces of paper could)!

Do real proofs need surveyability in the sense of Wittgenstein? Well, can you *easily exactly* reproduce this number?

93326215443944152681699238856266700490715968264381621468592
96389521759999322991560894146397615651828625369792082722375825
1185210916864

Copying it mindlessly does not count, any more than parroting a proof shows that you have grasped it! But in fact it is easy to reproduce it, once I tell you that it is 100! with the trailing zeros omitted. Wittgenstein (1978, p. 143) says:

> I want to say: if you have a proof-pattern that cannot be taken in, and by a change in notation you turn it into one that can, then you are producing a proof, where there was none before.

This apparently implies that that in saying '100! with the trailing zeros omitted' I have made a different, surveyable representation of that number. But no: it is *the original numeral* that is thereby shown to be surveyable after all! This observation will be enlarged upon below.

Why does Wittgenstein want surveyability? He seems to think that to be capable of the specific use of a theorem which a new proof makes possible we must be able to *reproduce* its proof. This is just false, indeed perversely so — without understanding anything about Wiles' proof of Fermat's Last Theorem you can use it to rule out the truth of $a^{17}+b^{17} = c^{17}$ where a, b and c are any three integers, even hundreds of digits long — for example I know that $123456789^{17} + 112233445566778899^{17}$ *can't be* equal to $12345678901234567890^{17}$ *without needing to calculate any of the three powers.* Wittgenstein's whole effort in (1978) is to understand the force of 'can't' in that sentence, without invoking any kind of 'ultraphysics' of numbers. But he goes too far in identifying the sense of a theorem with its proof. Moreover, it's not really true that someone who does grasp a proof as a whole can reproduce it exactly. You might not be able to remember all the variables or labels used, and so produce an unimportantly different proof, for example. (Gowers (2007) has an interesting discussion of the question when two proofs are 'essentially' the same.)

Surveyability for Azzouni is not threatened by 'short proofs', but he seems to recognise that in practice one could not survey the real proofs (the fully written proofs) of all of one's mathematical knowledge — any single theorem perhaps, but not all of them. His solution is to widen the scope of mathematical justification to make it a community, not necessarily individual, possession. I agree about that, but in two ways his account seems wrong: it rules out the Appel-Haken proof (and perhaps also those of The Enormous Theorem and Fermat's Last Theorem), which conflicts with mathematical practice; and he does not ever explain *how* surveyability is an 'epistemic virtue' — *how* it contributes to understanding.

As for Bassler's distinction, I question the usefulness of his claim that local surveyability consists in the possibility of 'surveying of each of the individual steps of a proof in some order' (p. 101) unless we are independently told what *surveying a step* is. He *seems* to mean verifying each inference as valid. But in what sense does verifying an inference require 'survey'? Perhaps Bassler here has half-recognised that checking a proof is a matter of *reading a written text*. As for global surveyability, Bassler writes (p. 102):

> At the minimum end, I want to assert that even if we have locally surveyed the individual steps of a proof this is not conceptually sufficient to account for the theorem having been proved: in addition to the surveying of each of the individual steps in a proof in some order, to see that the theorem is proved we require a recognition that the inference steps are sufficient to establish the theorem. Here the point is that the collective force of the proof steps requires a further conceptual acknowledgment in addition to a recognition of the validity of each of the steps and their respective positioning within the proof. Such a conceptual acknowledgement, that the proof steps fit together in such a way that they establish the claim, is a minimal requirement for global surveyability.

Here he seems to vacillate between meaning the comprehensibility of a proof as a whole, and one (last) instance of local surveyability, in the sense just adumbrated. (Other writers share this vacillation.) These ideas need keeping apart, because (fully) acknowledging the last step of a proof just *is* counting the proof as sufficient to establish the theorem. To be able to certify that, one must indeed have an overall grasp of the proof as a whole. But the grasp and the acknowledgement are distinct. I will now try to show why they are.

7. Uses of proofs presuppose *written* practice

What are proofs *really*, and what must surveyability be in order to contribute to them? To answer these questions we need to understand that REAL PROOFS ARE WRITTEN TEXTS. It is a bad mistake to think that 'in principle' proofs are just the contents of proof tokens, and can be just as good however embodied, in speech or just in pure thought (as Descartes imagined). Thoughts can't be used as proofs must be. There are two main uses of proofs: not only the justification of claims to knowledge, but also application in further mathematical development.

What Azzouni wants in a proof requires that it be written. The use of a proof to justify mathematical knowledge requires it be written if one is to check it properly oneself. In some cases, like Pythagoras' Theorem or The Enormous Theorem, it's not really possible even to make clear what the theorem *says* without writing. Try stating in speech the orders of the 26 sporadic simple groups perspicuously! (They are 7920; 95040; 175560; 443520; 604800; 10200960; 44352000; 50232960;

244823040; 898128000; 4030387200; 145926144000; 448345497600; 460815505920; 495766656000; 42305421312000; 64561751654400; 273030912000000; 51765179004000000; 90745943887872000; 4089470473293004800; 4157776806543360000; 86775571046077562880; 1255205709190661721292800; 4154781481226426191177580544000000; and 808017424794512875886459904961710757005754368000000000.)

But in every case, even the simple cases of short and apparently obvious proofs like Euclid's proof of the infinity of the primes, a written proof is needed (exercise: try to give a purely spoken explanation; you will find that referring clearly to the several numbers involved is extremely difficult in speech). The reason is the same as the reason why calculations written down are so much more reliable than those done in the head: they can be checked and rechecked, and by several persons. A written proof remains there unchanged while the attention shifts from part to part and person to person. Verbally presented or merely imagined proofs (Fermat!) are too prone to unseen errors to be taken as certifying — though they are important in communicating (we shall see shortly something about how this works).

The use of a proof to make another proof requires writing. Here is an example: to convert Euclid's proof of the irrationality of $\sqrt{2}$ into a proof of the irrationality of $\sqrt{5}$ we can simple replace 2 by 5 everywhere in it — but we need the precise text so we can see that 'the proof still goes through' (as they say):

Suppose $\sqrt{2} = \frac{a}{b}$, where $(a, b) = 1$.
Then $a^2 = 2b^2$.
So $2|a^2$, so $2|a$, say $a = 2k$.
Then $a^2 = 4k^2$, so $2b^2 = 4k^2$ and $b^2 = 2k^2$, so $2|b^2$ so $2|b$.
So $2|(a, b)$, contradiction.

becomes

Suppose $\sqrt{5} = \frac{a}{b}$, where $(a, b) = 1$.
Then $a^5 = 5b^5$.
So $5|a^5$, so $5|a$, say $a = 5k$.
Then $a^5 = 4k^5$, so $5b^5 = 4k^5$ and $b^5 = 5k^5$, so $5|b^5$ so $5|b$.
So $5|(a, b)$, contradiction.

But *actually this is wrong* — because we should not have replaced the '2's that indicate squaring, whereas we should have replaced the '4's with '25's. We need

Suppose $\sqrt{5} = \frac{a}{b}$, where $(a, b) = 1$.
Then $a^2 = 5b^2$.
So $5|a^2$, so $5|a$, say $a = 5k$.
Then $a^2 = 25k^2$, so $5b^2 = 25k^2$ and $b^2 = 5k^2$, so $5|b^2$ so $5|b$.
So $5|(a, b)$, contradiction.

I think this illustrates what Wittgenstein might be getting at when he demands exact easy reproduction — we have to understand how *every single character* contributes to the proof's working, in order to make use of it. But this character by character scrutiny requires we use a written proof not a spoken one.

None of this is refuted by observing the common practice of mathematicians in telling one another new theorems verbally or with back of the envelope sketches — they still insist on written documentation for the ideas that are conveyed like that, precisely because too many great proof-ideas turn out to be wrong. *Remember Fermat!*

8. Surveying

Before I try to say explicitly in what way surveyability is a requisite for good proofs, let's consider surveying generally. Some people will want to claim that calling proofs 'surveyable' appeals to the simple, literal meaning of 'survey'. What is that? Originally perhaps the idea was simply overview, as etymology suggests. (These days in dictionaries the sense in 'opinion survey' usage has taken first place, but I don't think this is relevant to the surveyability of proofs.) But an overview generally neglects some detail, there is no reason why a survey must include a view of every included point. The canonical survey of Argumentation Theory (van Eemeren et al., 1996) obviously does not mention everything in Perelman's writings, for example, important to the field though they are. 'Overview' implies view, clearly, and adding 'over' suggests to me either vantage point or considerable inclusiveness. But note also this non-technical definition from another dictionary: 'to look carefully at the whole of something, especially in order to get a *general impression* of it' (my emphasis).

Some sources suggest that the technical use of 'survey' — 'take linear measurements of a tract of ground' — is as old in English as the simpler 'view in detail'. Because surveying is one of the oldest applications of mathematics, we should take notice of the technical idea in attributing surveyability to proofs. If not, all we are saying about them is that they can be looked at carefully and a general impression formed. This is not going to distinguish among proofs at all — any uneducated lout can examine Euclid with the greatest care and form the general impression that it's a kind of gobbledegook.

What do surveyors do? At bottom, what surveyors do is this: using instruments to measure lengths and angles, and calculations based on

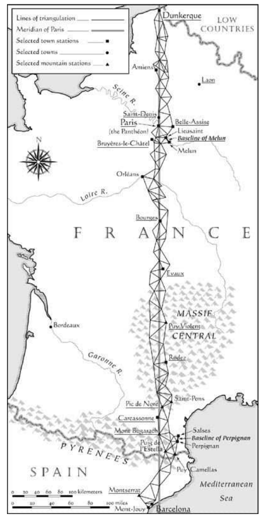

Figure 2. Cassini's triangulation of the meridian of Paris. (Map by Chris Robinson. Reproduced from Alder (2002) with permission.)

geometry and trigonometry, they determine as accurately as possible the positions in space of all the points of some region, relative to a fixed baseline (cf. Irvine, 1988). Actually, what surveyors do can be described a bit more precisely, with useful application to proofs. Frequently they work by successive triangulation, as in the Great Triangulation Survey of India (see Keay, 2000), or the triangulation of the meridian of Paris (cf. Alder, 2002). These two books provide much illuminating detail about those examples. Figure 2 is an old diagram used in the latter.

From a baseline with known ends A and B we can fix C if we measure the length of AB and the angles ABC and ACB, and calculate AC and BC. Then from AC we can progress to D, from DC to E and so on. This work is repeated until enough of the terrain has been covered with a net of measured triangles. *But this is not all*. Surveyors also locate the points they determine with respect to salient features of the terrain, sometimes installing their own salient features like pegs and suchlike, and they *record* their measurements. This record permits the construction of maps, and calculation of volumes and areas. In this way during several decades of the nineteenth century a huge grid was cast over India from which maps were made (and the conquest properly carried out), and a little earlier the standard meter was established using the measurements from the survey of France by Delambre and Méchain. Thus modern dictionaries have definitions like 'to examine and measure an area of land and record the details on a map', 'to measure and record the features of an area of land, for example in order to make a map or in preparation for building'.

9. What surveyability is for proofs

How then should we characterise surveyability for proofs, having regard for this information? I think there are two requirements of technical surveyability: [S1] surveyability of P requires that P can be effectively labelled and measured in sufficient detail, and [S2] surveyability of P requires that the labels and measurements of P can be recorded and made into a map. The *point* of survey is lost without S2: consider the paradigm cases of the surveys of India and France — what a waste of effort if the measurements had not been written down! The process of survey has its preconditions (don't try it on the Sun), but the *product* is the point. Carrying this over to proofs should retain two main aspects — that point-to-point scrutiny of a written proof is possible, and that supplementary perspicuous representation of the proof-idea can be made from the scrutiny. I take the phrase 'perspicuous representation' from Wittgenstein's translators, and bearing in mind his color octagon example, we will regard many kinds of 'map' of a proof as filling this role. An idea somewhat similar to this is the use of structure diagrams in informal logic to represent the argumentation of a text. (Also, a referee drew my attention to Lamport (1995), in which there is some discussion of proof-sketches and proof structure which can be usefully compared with the present work. However Lamport's concern is with the correctness of proofs, mine with their comprehensibility. That these are quite distinct will be argued below.)

To show that a proof is surveyable it suffices to give a perspicuous representation of it. Some simple examples have already been given:

- the idea of Euclid's proof of Pythagoras' Theorem is that you divide the big square with a line from the right-angle vertex parallel to its sides, and then use theorems about triangles between parallels to show that the two parts are equal to the two smaller squares;

- the idea of the proof of the irrationality of $\sqrt{2}$ is to argue by contradiction that no candidate fraction is in lowest terms;

- the idea of the classical proof (Euclid IX.20) of the infinity of the primes is to construct a bigger one from any candidate biggest by forming its factorial plus 1.

A more complicated example is the Appel-Haken proof: its leading idea is to show that a certain set of reducible configurations is unavoidable. (This needs some explanation to make it perspicuous!)

Here's a considerably more complex example. Duke and Hopkins (2005) gives a proof of a generalisation of quadratic reciprocity (a famous theorem of number theory) to finite groups. They explain in their introduction that 'we will combine this classic technique [Gauss' sixth proof of classical quadratic reciprocity, based on Gauss sums using cyclotomic fields and the Frobenius automorphism] with another technique of Frobenius, the character table, to prove a law of reciprocity for the quadratic symbol for any finite group G.' In their proof they first do some matrix calculations using the character table, then they say 'To prove the first statement of Theorem 1 we apply an argument used by Schur to prove Stickelberger's theorem about the discriminant of a number field', and later on 'That [their quadratic symbol] is nontrivial if d is not a square follows easily from Dirichlet's theorem on primes in arithmetic progressions'. These remarks are their way of enabling the reader to survey their proof.

Here is an example of a proof of Pythagoras' theorem which is *not* surveyable, despite being extremely short.

> *The axioms, postulates and common notions in Euclid entail that the square on the hypotenuse of a right-angled triangle equals the squares on the other two sides; so, the square on the hypotenuse of a right-angled triangle equals the squares on the other two sides.*

There is nothing wrong with this as an *argument* justifying Pythagoras' Theorem, but it is not a good *proof* because it is impossible to glean *how* Pythagoras' Theorem follows from the axioms etc; no *proof-idea* is made explicit *and nor could one be*, since there are various ways one might actually do it, none of which belongs intrinsically to this argument. There are no salient features to label! But a perspicuous representation needs to lable salient features of the proof and connect them.

We can now understand Bassler's mistake in distinguishing local and global surveyability in a proof. There are **two** cognitive instruments in play, not one: one is the written proof text and the other is the proof-idea, general impression or perspicuous representation, which may be written too (Duke and Hopkins) or remain only in the minds or speech of proof readers. A real proof must be **verifiable** step by step, and this is indeed a required property of the written proof. This is an objective property of the written proof itself. A good proof must also be surveyable, which is also a property of the written proof but not the same kind of property. It is a matter of degree, and a given proof is surveyable to some degree for a specific community of mathematicians.

It's a perfectly objective fact about the proof and the community, but a relative one. Also, there are more or less fleshed-out perspicuous representations of the proof, since all real proofs involve a combination of ideas. There is often one main idea — if two, the proof could be split (and this often happens). But this does not preclude there being other non-routine ideas involved. What is most needed is a guide to the structure of the proof; for example, a proof that all of a set of statements are equivalent often starts with an indication of which implications will be shown in what order. Complex proofs frequently have a number of helpful hints scattered about and within them to assist the reader's survey, as in Duke and Hopkins.

But although a good proof must be both verifiable and surveyable, these are *independent* requirements. The Appel-Haken proof is not very good because it is hard to verify and in fact it cannot be 'hand-checked' by humans. But that does not make it uncheckable *or* unsurveyable — its survey is easy. It is surveyable but not humanly checkable. But it is checkable. The idiotic proof of Pythagoras' theorem given two paragraphs ago, on the other hand, is checkable but not surveyable. (In the remarks of Tymoczko quoted above he silently conflates perspicuity and checkability by hand, and explicitly equates checkability and surveyability.)

A proof must be verifiable, and that requires checkability and correctness. But that is not what surveyability requires. In fact, a surveyable proof need not be verifiable, because it need not be correct! My first attempt to convert the proof of irrationality for $\sqrt{2}$ into a proof of the irrationality of $\sqrt{5}$ is surveyable for just the same reason as are the original and the final version (their proof-ideas are the same), but it is not correct. SURVEYABILITY IS THE REQUIREMENT THAT THE PROOF BE CAPABLE OF SUPPORTING THE CONSTRUCTION OF A PERSPICUOUS REPRESENTATION OF THE PROOF-IDEA. But this perspicuous representation need not be, indeed almost certainly won't be, constructed using all the details. *The product of a survey is the map not the territory.*

Once we separate verifiability and surveyability it becomes apparent that putative ideal agents need not be any better off than we are as far as *understanding* proofs are concerned. They could check proofs like Appel-Haken by 'hand' perhaps, but it is not our 'mere medical limitations' (as Russell put it) that determines whether we can survey a proof. No-one, no matter how ideal, could survey the bad proof of Pythagoras because *it* is simply unsurveyable. Equally, we do not need to be concerned with second and third generation proofs of the four color theorem since while they may have improved verifiability, Appel-Haken is already surveyable. Automating checking may well become more useful in the future, as the Lamport paper and work by Gonthier and others suggests, but this is not relevant to surveyability.

10. Distributed justification: the archives and the community

Azzouni is half-right about the involvement of the mathematical community is one's knowledge of mathematics. A lot of one's knowledge of mathematical truths is not mathematical knowledge, or not purely so, because it depends on the authority of others in the mathematical community. For example, we all know now that $123456789^{17} + 112233445566778899^{17}$ *can't be* equal to $12345678901234567890^{17}$, thanks to Andrew Wiles. This is knowledge we have of a mathematical fact but not in the strictest sense mathematical knowledge because we don't understand Wiles' proof. But all the real *mathematical* knowledge someone has is fully justified by mathematical support, which is mostly proof. But this justification is itself distributed over the mathematical archives, not located in an integral superproof for each thing known.

Like Hobbes, I know Pythagoras' Theorem through Euclid's proof, because I know all the proofs that put together justify it. Knowing the theorem of Duke and Hopkins similarly requires a lot of knowledge which adds up to justification for it, but has probably never been written down in one sequence like Euclid, and different mathematicians who know it will have a wide variety of differing justifications (involving the way they know theorems of group theory, number theory, matrix theory etc), though they converge on the Duke and Hopkins proof.

The moral of the story is that we need to distinguish between proof and justification. A proof is a kind of conditional justification. The full justification of virtually any item of mathematical knowledge is a long and unsurveyable 'proof', but this is not a problem because the archives are so organised that each chunk (proof) is surveyable. Chunks vary in size — chunking devices are used to make surveyable justifications which we call proofs. Many mathematics texts provide several devices to enable their being surveyed, including the traditional table of contents and index, maybe a list of notation or glossary, but also a diagram of the dependencies among the chapters.

Indexes are one of many devices used to organise written materials for just the same reason that we want proofs to be surveyable — you *can't* take everything in at once, what you need are ways to navigate to the bits currently in question. Euclid's Elements is actually paradigmatic in this regard too, not only as an example of mathematical knowledge, but as a surveyable document. Its surveyability is achieved by division into books, by breaking the deductions up into propositions, by labelling them and using citations to tie them to those they use, etc. Many devices used in Heath's version of Euclid are not specific to mathematical documentation but to documentation in general, of course — mathematics is only one among several important writing-based practices. Indeed, it is not only proofs that need to be surveyable in the sense that we require for them. Any complicated discourse presents the same kinds of problems of understanding. The peculiarity of mathematical knowledge, which I think is the focus that philosophy of mathematics should have, has not been explained by what I have pointed out about surveyability. That's because I have said nothing here about how the written mathematical sign-system, involving notation and diagrams as well as words, enables the construction of *correct deductions* about mathematical objects. But that is a different topic.

11. Conclusion: mathematics *is* long proofs

Long proofs are not a problem, and nor is the understanding of long proofs. What has been a problem is the understanding of the understanding of long proofs. The key to that, is to recognise that mathematics is a written practice which depends on the accumulation and deployment of an archive.[4] Because of that, one does not need to have all of a proof in one's head because it is all on record. Understanding it means you have the main idea of it in your head, and knowing it requires both that, and the facility with the archive to get at the details if wanted. The 'fully written out proofs' are there, in a sense — the distributed sense in which the mathematically educated agent can access them, and so the theorems of mathematics are indeed justified by proofs, though the true justifications are not the proofs labelled as such. Really, mathematics is long proofs.

Author's Vitae

Dr Edwin Coleman is currently a Senior Research Fellow in Philosophy at the University of Melbourne. He holds graduate degrees in mathematics, philosophy and computing from ANU, Adelaide and RMIT. He has written extensively about both the philosophy of mathematics and informal logic; most recently he published yet another paper about fallacies in the procedings of the sixth ISSA conference.

[4] Many other 'problems' about mathematics have their solution here too.

References

Alder, K.: 2002, *The Measure of the World*. New York: Free Press.

Appel, K., W. Haken, and J. Koch: 1977, 'Every Planar Map is Four Colorable'. *Illinois Journal of Mathematics* **21**, 439–567.

Aubrey, J.: 1958, *Brief Lives*. London: Secker and Warburg.

Azzouni, J.: 1994, *Metaphysical Myths, Mathematical Practice*. Cambridge: Cambridge University Press.

Bassler, O. B.: 2006, 'The Surveyability of Mathematical Proof'. *Synthese* **148**(1), 99–133.

Cottingham, J. (ed.): 1985, *Philosophical Writings of Descartes*, Vol. 1. Cambridge: Cambridge.

Duke, W. and K. Hopkins: 2005, 'Quadratic Reciprocity in a Finite Group'. *American Mathematical Monthly* **112**(3), 251–256.

Gorenstein, D.: 1985, 'The Enormous Theorem'. *Scientific American* **253**, 104–115.

Gowers, T.: 2007, 'When are two proofs essentially the same?'. http://gowers.wordpress.com/2007/10/04/when-are-two-proofs-essentially-the-same/.

Inglis, M., J. P. Mejia-Ramos, and A. Simpson: 2007, 'Modelling Mathematical Argumentation: The Importance of Qualification'. *Educational Studies in Mathematics* **66**(1), 3–21.

Irvine, W.: 1988, *Surveying for Construction*. Maidenhead: McGraw-Hill.

Keay, J.: 2000, *The Great Arc*. London: HarperCollins.

Kitcher, P.: 1983, *The Nature of Mathematical Knowledge*. Oxford: Oxford University Press.

Lamport, L.: 1995, 'How to Write a Proof'. *American Mathematical Monthly* **102**(7), 600–608.

Putnam, H.: 1979, 'Philosophy of Mathematics: A Report'. In: P. D. Asquith and J. Henry E. Kyburg (eds.): *Current Research in Philosophy of Science: Proceedings of the P.S.A. Critical Research Problems Conference*. East Lansing, MI, Philosophy of Science Association.

Tao, T.: 2007, 'What is Good Mathematics?'. *Bulletin of the American Mathematical Society* **44**(4), 623–634.

Tymoczko, T. (ed.): 1998, *New Directions in the Philosophy of Mathematics*. Princeton, NJ: Princeton University Press.

van Eemeren, F., R. Grootendorst, F. S. Henkemans, J. A. Blair, R. H. Johnson, E. C. W. Krabbe, C. Plantin, D. N. Walton, C. A. Willard, J. A. Woods, and D. F. Zarefsky: 1996, *Fundamentals of Argumentation Theory: A Handbook of Historical Backgrounds and Contemporary Developments*. Mahwah, NJ: Lawrence Erlbaum Associates.

Wiles, A.: 1995, 'Modular Elliptic Curves and Fermat's Last Theorem'. *Annals of Mathematics* **141**(3), 443–551.

Wittgenstein, L.: 1978, *Remarks on the Foundations of Mathematics*. Cambridge, MA: MIT Press. Edited by G. H. von Wright, R. Rhees, and G. E. M. Anscombe.