THE NATURE OF MATHEMATICAL PROOF

David Tall

Introduction

There is a legendary story of the sage who posed the question: 'A normal elephant has four legs; if an elephant's trunk is called a leg, how many legs does it have?' He asked a mathematician, who continued to stare at a pile of paper on which he was scribbling as he muttered: 'four and one make five'. Next to him a philosopher mused enigmatically and puffed for a few moments on his pipe before observing: 'The fact that it is *called* a leg, doesn't change the fact that it is *not* a leg, so the answer is *four*'. 'Excuse me,' said a passing zoologist, 'if a trunk is classified as a leg, clearly this will also apply to the tail, so it has *six* legs, and it's an insect'. A logician joined the conversation: 'A *normal* elephant has four legs, but you did not actually say that *this* elephant is normal, so there is insufficient evidence...'

Continuing to seek enlightenment, the sage in his wisdom passed the query on to a statistician who returned the following day asserting 'the mean is 0.33'. 'Might I ask how you came by this information?' queried the sage, concealing his innermost thoughts behind an inscrutible smile. 'The best way to solve such a question is to obtain empirical information,' replied the statistician, 'so I went to the local zoo and got the answer from the horse's mouth, so to speak. Two elephants refused to respond and the third blew his own trumpet just once.'

Still bemused the sage went along to the local school which was deeply embroiled in GCSE investigations and once again stated his problem. 'That's a very interesting question,' said the teacher.

The moral of this story is that, as Humpty Dumpty once said, 'when I use a word, it means just what I want it to mean, and nothing else'. The term 'proof' is just such a word. In different contexts it means very different things. To a judge and jury it means something established by evidence 'beyond a reasonable doubt'. To a statistician it means something occurring with a probability calculated from assumptions about the likelihood of certain events happening randomly. To a scientist it means something that can be tested - the proof that water boils at 100° C is to carry out an experiment. A mathematician wants more - simply predicting and testing is not enough - for there may be hidden assumptions (that the water boiling is always carried out at normal atmospheric pressure and not, say, on the top of Mount Everest).

The Problem of Proof in School Mathematics

Mathematics students live in a world in which the term 'proof' means many different things and so their interpretation of the meaning may be different from that of the teacher, just as one teacher's interpretation may differ from another's. I well remember as an A-level student many years ago being drilled in the different nuances of the words 'show', 'demonstrate', 'verify', 'prove' and 'prove from first principles'. Proof very much meant reproducing a

Published in *Mathematics Teaching*, **127**, 28–32 (1989).

sequence of deductions to establish an important result. Rarely did it grow from a genuine problem.

Proof has nominally been a major ingredient of senior secondary mathematics for many years, historically through the medium of Euclidean geometry. Euclid vanished from the British curriculum many years ago (though it is still a staple diet in other countries, including the USA). In the nineteenth century I understand that it was the practice to insist that Euclid was learned verbatim so that, if a pupil reproduced the exact proof but used different letters, it was considered wrong.

So what do we mean by proof in school mathematics today? It is my belief that an adequate concept of mathematical proof is rarely, if ever, satisfactorily considered in school. In response to what will surely be a chorus of indignation, it is necessary to look at the practice of what we call 'proof' in school mathematics. In A-level studies proof appears in such things as 'proving that the derivative of cosx is -sinx', which means going through a sequence of symbolic manipulations that many students find hard to follow, only to arrive at a result which they are quite prepared to accept. Why is it necessary to *prove* something that is *known* to be true?

Using a computer to sketch the gradient of sinx gives a graph which looks like 'sinx upside down', can give a strong sense of understanding why the derivative is *minus* sinx. So why, a student may ask, is a proof necessary? An amusing consequence of such an approach occurred when a group of students who used a graphical approach to guess the gradients of x^2 and x^3 generalized this to assert that the derivative of x^n was nx^{n-1}. When asked if there was a need to *prove* this they replied in the negative. A follow-up question querying whether the computer had actually 'proved' the result for all n, for instance, for n=7 or n=–1 or $n=\frac{1}{2}$, elicited the rejoinder: 'tell us the value of n and we will use the computer to verify the result'.

The student response to proof here seems closer to the scientist than the mathematician. Yet is it? Think for a moment how mathematicians define continuity: 'give me an epsilon and I'll find you a delta such that ...'. Is this, on the surface, much different from 'give me a value of n and I'll find a good enough numerical approximation to the gradient ...'?

In A-level, proof often occurs in the form 'show that if something occurs then something else happens' - for instance a mechanics problem might ask to show that '*if* the block slides, *then* the coefficient of friction is less than a certain value'. For over a quarter of a century it was my privilege to mark A-level examination papers in which, invariably, many pupils answered the question by showing that 'if the coefficient of friction was less than a certain value, then the block slides'. In other words pupils were not able to distinguish between the statements 'if P then Q' and 'if Q then P'. In examiner's meetings we invariably agreed not to penalize this mistake, largely because in virtually every question under consideration the two conditions P and Q were equivalent and so *both* statements were true simultaneously. Under these

circumstances the concentration on deduction in one way only is a fine distinction.

Of course, we teachers would never make such a mistake, would we? The truth is otherwise. Every year virtually everyone teaching A-level maths commits just this error. We assert that two indefinite integrals differ by an arbitrary constant, that is we say '*if* f′(x)=g′(x) *then* f(x)=g(x)+c'. We deduce this statement by reference to the true statement

'*If* f(x)–g(x)=c *then* f'(x)–g'(x)=0'

and commit the cardinal sin of simply turning it the other way round, perpetrating the very error I remarked on in mechanics examinations. Only in this case the 'theorem' is false.

A counter-example may be given using the signum function

$$\operatorname{sgn}(x)=\begin{cases}-1 & x<0\\ 0 & x=0\\ +1 & x>0\end{cases}.$$

The functions f(x)=sgn(x)+$\frac{1}{x}$ ad g(x)=$\frac{1}{x}$ have the same derivatives everywhere (except x=0 where they are not defined), but they do not differ by a constant[1]. Oh, you may say, that's cheating, we don't normally meet functions like that in the calculus...

No we don't. Nor do we normally have the personal experience of boiling water on the top of Mount Everest, which would prove that water doesn't always boil at 100˚ C.

If we insist on claiming a general result is true just because we find it to be true in most of the cases we meet, how can we pretend to honour the ideal of mathematical proof?

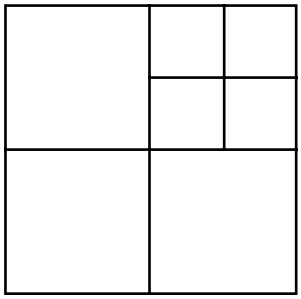
This example reminds me of another well-known story of the experimental physicist who claimed to prove that 60 is divisible by every other number. He came to this conclusion by considering a sequence of cases to establish the pattern: 1,2,3,4,5,6 and then moved on to a few others at random to test out the theory : 10,12, 20, 30, and concluded that his result was experimentally verified. He was surpassed in this endeavour by an engineer who noticed that all odd numbers seemed to be prime... One - well that's an oddity, but we'll include it in - three, five, seven, good, we're getting somewhere - nine ? Oh, nine... Let's leave that a moment - eleven, thirteen - fine. The exceptional case of nine must have been an experimental error.

If we are truly to address the notion of mathematical proof in the A-level curriculum, we must begin to show students the difference between asserting something is true on empirical evidence and proving it true by logical deduction from known facts.

[1] The true theorem states "if f'(x)=g'(x) on a *connected* domain, then f(x)=g(x)+c". The counter-example does not violate this theorem because the domain of f(x)=1/x is in two separate connected pieces, x<0 and x>0, and on each of these there is a single arbitrary constant.

Problem Solving and Convincing Arguments

When a problem is encountered, the question of providing a convincing argument to explain the solution often arises. It has been a joyful experience over the last few years to teach an annual problem-solving course based on 'Thinking Mathematically' by John Mason, Leone Burton and Kay Stacey. In this book are lots of diverting problems and puzzles which give mathematical students of all abilities the chance to develop problem solving strategies. Consider the problem 'into how many squares can you cut a square?'. Initially students often say 'as many as you like', or 'infinity'. Then they begin to realize that they could cut a square into 4, 9 or 16 by dividing it into pieces of equal size, and suddenly they see that any of these squares could itself be cut into four smaller squares. Aha! A square could be cut into four quarters and one quarter cut into four again, losing the quarter as a counted square but gaining four smaller ones - so a square can be cut into *seven* squares.



Now the problem begins to strike home - cutting any of the seven squares in four gives a new total of 10, then 13, then 16, and so on. Soon the student is on the track of identifying all the possible number of squares that might arise. I won't spoil the fun by telling you the answer. But in all the years I have studied this problem with some of the brightest mathematical undergraduates around, virtually *all* of them see sequences of possible numbers increasing by three such as 4,7,10,13,.. and 9,12,15,... and almost *none* of them enunciate the obvious general result 'IF I can cut it into n squares THEN I can cut it into n+3'. After three years of university mathematics they still prefer to *assert* the truth of a statement using specific numbers rather than *deduce* the truth of one general statement from another.

After an hour or so on this problem almost all the students have not only found the numbers which can be done, including one or two surprising ones they had not initially thought about, they are concentrating on the few numbers that do not seem to be possible. Some students were able to offer some arguments as to *why* the particular numbers would not work. But it was not until I had run the course three times, with a total of well over a hundred students passing through my hands, that a student produced a pleasing proof showing precisely which numbers could not be done.

I was so concerned about this lack of formalities that the following year I offered the prize of a bottle of vintage wine to anyone who produced what I termed a 'sweet proof' of the result. I had to award the prize to a student who

neatly word-processed his answer on a Macintosh computer and coated the paper with sugar!

To help the student focus on the various stages of putting up a convincing argument, 'Thinking Mathematically' suggests three stages:

convince yourself,
convince a friend,
convince an enemy.

The idea is first to get a good idea how and why the result works, sufficient to believe its truth. Convincing oneself is, regrettably, all too easy. So pleased is the average mortal when the 'Aha!' strikes that, even if shouting 'Eureka' and running down the street in a bath towel is de rigeur, it is very difficult to believe that the blinding stroke of insight might be wrong. So the next stage is to convince a friend - another student, perhaps - which has the advantage that, to explain something to someone else at least makes one sort out the ideas into some kind of coherent argument. The final stage, according to 'Thinking Mathematically' is to convince an enemy - a mythical arbiter of good logic who subjects every stage of an argument with a fine toothcomb to seek out weak links.

But absent from 'Thinking Mathematically' is the *formal* notion of mathematical proof, not because the authors do not believe in it, but because, in my experience too, the nature of a formal mathematical proof is very difficult for students (and others) to comprehend.

Mathematical Proof

Mathematical proof differs from convincing a friend or enemy in that it must be based on two important ideas. One is that it requires *clearly formulated definitions and statements*, and the other is that it requires *agreed procedures to deduce the truth of one statement from another*.

The reason why mathematical proof is so difficult to introduce successfully in the sixth form is that *neither* of these concepts is found in the sixth form curriculum. We do not propose definitions of mathematical concepts in the sixth form in the way that is done (with limited success) in university courses. By this I mean definitions such as 'a group is a set G and a binary operation * such that ...'. Attempts at introducing such ideas in the 'New Maths' of the sixties were doomed to failure. As examiner of an A-level paper including group theory I found that virtually all students avoided the questions involving the simplest of proofs, preferring instead to go for a predictable but horribly messy question on integration. It is much easier to carry out a routine, but nasty, calculation rather than make deductions from an abstract definition, however mathematically trivial the deduction may be.

Without adequate definitions of concepts we cannot be absolutely certain what we are talking about and so we cannot be sure that we have a sound formal proof. Experience shows that the introduction of formal definitions is not a suitable way of teaching mathematics at this level. The French tried it for years, in best Bourbaki style, and such an approach still persists in some coun-

tries (Greece for one). But the French are now leading activities which take into account the cognitive development of pupils, based on the premise that informal experience of *using* ideas must precede the logical analysis of them. It is only by building on experience that we can hope to show students the subtleties of different forms of proof: that 'IF P THEN Q' is the same as 'IF Q is false THEN P is false', but is not the same as 'IF Q THEN P'.

The problem of proof has been with us since time immemorial and, if we do not address the problem seriously, it could get worse in new curricula with an emphasis on informal methods of enquiry through mathematical investigations. My friends in Computer Science at university often express to me their worries about students entering their course without any concept of proof. For how can they teach students about the necessity for exact logical thinking in the development of software if the students have no concept of deduction? They are seriously worried about the need to produce computer scientists and programmers who write provably correct software that does not contain horrendous bugs. The development of a sensible concept of proof is surely an important part of A-level.

Three men were going by train to a conference in the nether region of the United Kingdom. The engineer looked out of the window and said, 'Look, all the sheep in Scotland are black'. The theoretical physicist thought for a moment and said, 'No, there exists a field in Scotland in which all the sheep are black'. There was silence from the logician who mused for some time in the corner of the compartment before declaring 'No, there exists a field in Scotland in which all the sheep are at least half black...'

The Beginnings of Proof in School

By asserting that formal mathematical proof is unattainable in a course which lacks formal definitions and formal methods of deduction, have I forced myself into a Catch 22 situation where the concept of proof cannot be part of A-level? Will the difficulties not grow worse in new A-level courses which must follow on from GCSE and cater for a wider range of abilities, including students who only obtain a C-grade at 16+?

No. By giving students the opportunity to start along the path of producing convincing arguments in practical situations, we can move towards the idea of making logical deductions in more general situations.

First it is imperative for students to be involved in making deductions of the form 'if I know something, then I know something else', in cases where the second may hold when the first does not. Surprisingly such ideas are around us all the time and we make them every day in the classroom:

If $x+1=3$ then $(x+1)^2=9$.

The converse is false: If $(x+1)^2=9$, then it does not follow that $x+1=3$, because it might equal -3.

It may be helpful to look at general statements such as

If x>5 then x>3

which is likely to meet with universal agreement as a true statement. Yet if P is the statement ‘x>5’ and Q is ‘x>3’, this leads to some interesting possibilities. Certainly we see that whenever P is true, Q is also true. But there are also cases, such as x=4 when P is false and Q is true, and others, such as x=2 when P is false and Q is false. In general, this suggests that the statement ‘If P then Q’ means ‘If P is true then Q also must be true, but if P is false then it doesn’t matter whether Q is true or false’. My experience in trying to explain this idea shows that students find it very hard to think about implications of the form ‘If P then Q’ when P is false. They don’t understand why we waste our time talking about it.

My own personal preference is to go for meaningful examples of ‘IF P THEN Q’ where the converse doesn’t hold. There are a number of such cases that arise in the sixth form if only we were to be brave enough to confront them properly. One of my pet hates is the way that students invariably look for a maximum or minimum by seeking zeros of f'(x). They cannot cope with f(x)=|x| which has a minimum at the origin, but no derivative. If they have graphical experience which reveals that a curve only has a derivative where it magnifies to look straight, then they are quite happy to see that f(x)=|x| has no derivative at the origin. The correct theorem is ‘IF a function has a derivative at a local maximum or minimum, THEN the derivative is zero’. The proof is that if a function has a positive derivative, under high magnification it looks locally straight and is increasing, whilst if it has a negative derivative, locally it is decreasing. In either case it cannot have a maximum or minimum, so the only possible thing that can happen at a maximum or minimum is that the derivative is zero.

The power of proof can be emphasized nicely by showing how a general algebraic statement covers a far wider number of cases than a specific numerical calculation. A nice example from a recent book on ‘Introduction to Proof in Mathematics’ by James Franklin and Albert Daoud is the proof that

$$\frac{1}{1\,000} - \frac{1}{1\,001} < \frac{1}{1\,000\,000}$$

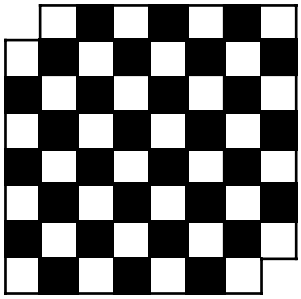
Although this can be done by simple arithmetic, it is just as easy, and more powerful, to show that

$$\frac{1}{n} - \frac{1}{n+1} < \frac{1}{n^2}\,.$$

This now applies not just to n=1000, but to any value of n, say n=100000000002.

Another kind of proof that we should emphasize is the proof that certain things are not possible - for instance, is it possible to cut two opposite corner squares off a chess board and to cover the remaining squares with dominoes, each covering two adjoining squares. After a time one begins to realize that perhaps it cannot be done. One strategy is to look not at 8 by 8 boards, but something more manageable, say 2 by 2, 3 by 3 or 4 by 4. The clue for the 8 by 8 board eventually comes from looking at the *colours* of the chessboard

squares. A domino covers one white and one black square. But what are the colours of the squares in the opposite corners? Does this help us show that the board with the missing corners cannot be covered by dominoes?



Proof means being precise about arguments and getting things right. But it must be placed in the broader context of its power and generality, not just the finnicky nitty gritty of dotting all the i's and crossing the t's. It should take its proper place as the last and culminating stage of the process of mathematical enquiry when all the threads are drawn together in an orderly fashion with a clear statement of assumptions and a clear sequence of deduction.

There is an unfair story, with a grain of truth in it, that is told about mathematicians which is a warning against excessive reliance on proof without practicality.

'How can you tell a mathematician from an engineer or physicist?' The answer is, you set fire to his wastepaper basket. The engineer will make a cursory calculation and swamp the basket with enough water to put out the fire and more. The physicist will sit down, calculate exactly how much water is needed and pour the exact quantity on the fire. The mathematician? The mathematician will sit down and calculate exactly how much water is needed.

What is needed in A-level mathematics are experiences that encourage students to make convincing arguments in meaningful situations. What we must do is to introduce these experiences in a way that is both an end in itself for the vast majority of students who will go on to study other disciplines, but also provides the cognitive foundations of formal proof for the tiny minority of mathematics specialists who will later make logical deductions from precise definitions.