Hewitt, D. (Ed.) Proceedings of the British Society for Research into Learning Mathematics 25(2) June 2005

ASPECTS OF PROOF IN MATHEMATICS RESEARCH

Juan Pablo Mejía-Ramos

University of Warwick

Without having a clear definition of what proof is, mathematicians distinguish proofs from other types of argument. This has become increasingly difficult in the last thirty years, as mathematicians have been able to use ever more powerful computers to assist them in their research. An analysis of two types of proof (mathematical proof and formal proof) and two types of argument (mechanically-checked formal proof and computational experiment) reveals some aspects of proof in mathematics research. The emerging framework builds on the distinction between public and private aspects of proof, and revises the characterization of mathematical proof as being formal, convincing, and a source of understanding.

What is proof in mathematics research? Hersh (1997) differentiates between two meanings of "mathematical proof": what it is in practice and what it is in principle:

> Meaning number 1, the practical meaning, is informal, imprecise. Practical mathematical proof is what we do to make each other believe our theorems. It's argument that convinces the qualified, skeptical expert. It's done in Euclid and in The International Archive Journal of Absolutely Pure Homology. But what is it, exactly? No one can say. Meaning number 2, theoretical mathematical proof, is formal. Aristotle helped make it. So did Boole, Peirce, Frege, Russell, Hilbert, and Gödel. It's transformation of certain symbol sequences (formal sentences) according to certain rules of logic (modus ponens, etc). A sequence of steps, each a strict logical deduction, or readily expanded to a strict logical deduction. (p.49)

Often, when different terms are needed to differentiate these two versions of proof, "mathematical proof" is used to refer to the "practical" meaning, while "formal proof" is used for the theoretical version (The Flyspeck Project Fact Sheet, n.d.)

Distinguishing between mathematical proof and formal proof makes a lot of sense. While there are examples of ideal formal proofs (e.g. those in Whitehead and Russell's *Principia Mathematica*), actual mathematical proofs do not include all the logical steps of a formal proof, and they assume an enormous amount of implicit contextual knowledge (Stewart and Tall, 1977). This differentiation is particularly important today when formal proofs of more interesting theorems are in reach thanks to the growing popularity of mechanically-assisted proof-checking. However, mathematical proofs and formal proofs cannot be separated completely: first, mathematical proofs cannot be disassociated from logical deduction (i.e. mathematical proofs are, to a certain extent, formal proofs), and second, formal proofs help convince mathematicians (i.e. formal proofs are, to a certain extent, mathematical proofs).

FORMALISM

As noted above, the kind of proof done by mathematicians, although based on logical deduction, does not exactly resemble the ideal formal proof. Just how formal are mathematical proofs? Hersh (1997) and Ernest (1998) argue that while mathematicians know that mathematical proofs are not formal, some of them believe that all correct mathematical proofs "can, in principle, be translated into fully rigorous formal proofs" (Ernest, 1998, p.29), and that "a real mathematical proof is an abbreviation of a formal one" (Hersh, 1997, p.215). However, it is easy to see that the "potential formalizability" of mathematical proofs is a matter of faith.

The current production of mathematical proofs increases much faster than humanity's ability to formalize them. This makes the total formalization of (formalizable) mathematics seem practically impossible, at least without the help of computers. During the last thirty years some computer scientists and mathematicians have been working on the formalization of interesting mathematical proofs with the aid of computers. This is the case of Thomas Hales' proof of the Kepler Conjecture, and Georges Gonthier's proof of the Four Colour Theorem, two proofs that were originally so long and complicated (consisting of hundreds of pages and several gigabytes of computer code) that they were nearly impossible to verify. This led Hales and Gonthier to use a more transparent and reliable computer program, a *proof assistant*, to mechanically formalize and check the totality of the proof. While Gonthier has recently finished, Hales estimates it may take him as many as 20 work-years to complete this huge enterprise. (The Flyspeck Project Fact Sheet, n.d.)

Following this trend, a similar, but much more ambitious project is described as:

> QED is the very tentative title of a project to build a computer system that effectively represents all important mathematical knowledge and techniques. The QED system will conform to the highest standards of mathematical rigor, including the use of strict formality in the internal representation of knowledge and the use of mechanical methods to check proofs of the correctness of all entries in the system. (The QED Manifesto, Section 1, ¶ 1).

The long Manifesto does not indicate an estimated time of completion, but besides demanding what appears to be a huge amount of resources, this colossal project will have to withstand sceptics' remaining doubts about the reliability of proof assistants and computers' infrastructure in general. It is therefore fair to agree that, at the moment, the "potential formalizability" of mathematical proof cannot be fully justified and stands only as a belief of the idealist mathematician. In other words, mathematical proof is somewhat formal, but not necessarily formalizable.

A perhaps more important issue to consider is the effect of formal proofs on the development of mathematics. What will happen with Hales' proof after it is finally formalized? The result will most definitely be an ideal, formal proof, but it will still have to pass the test given to any other practical, mathematical proof: convincing the mathematical community.

CONVICTION

How convincing is formal proof? By definition, a formal proof of a statement removes all doubts about whether or not the statement *holds in a theory*. Therefore, *if a mathematician believes that a given formal proof is flawless*, then he/she will be convinced that the statement holds in that theory. Consequently, according to the definitions given above, every formal proof would be a mathematical proof. But human belief systems are more complicated than this. Conviction has internal and external aspects: *internal conviction* is achieved through intuition and personal belief systems, while *external conviction* follows external, socially established belief systems. Internal conviction removes an individual's doubts based on his/her own intuitions, knowledge, and beliefs; while external conviction is based on arguments given to remove other people's doubts. Of course these two aspects need not be exclusive, as external and internal belief systems may coincide. However, this distinction facilitates a better understanding of the role of proof in human conviction.

Therefore, even after the formal proof is mechanically checked, and after it is socially agreed that both the computer and the proof assistant can be trusted, there may remain important questions that prevent a mathematician from being internally convinced of the statement, e.g. what exactly do the thousands of lines of code *mean* mathematically speaking? To a lesser extent, this can also happen with mathematical proof. Although it is characterized as an argument that convinces the qualified expert, the role of mathematical proof is sometimes limited to external conviction. Even after a proof has been accepted by the mathematical community as complete (i.e. externally convincing), mathematicians might find it unintuitive, unbelievable, internally unconvincing. As noted by Hersh (1997), "it also happens that we don't believe, even in the presence of complete proof. […] There is a famous result of Banach and Tarski, which very few can believe, though all agree, it has been proved." (p.50) However, these examples are not representative of most mathematical proofs, which usually provide mathematicians with an internal sense of conviction. Therefore, although both mathematical and formal proof provide external conviction (the former through contextual arguments and the latter through logically exhaustive ones), mathematical proof also tends to be oriented at providing mathematicians with an internal sense of conviction.

On the other hand, mathematicians may be internally convinced of the truth of a statement through heuristic arguments that would not convince the "qualified, skeptical expert". As noted by Hersh himself (1997), "it commonly happens in mathematics that we believe something, even without possessing a complete proof" (p.50). Davis and Hersh (1981) exemplify this situation by quoting experimental evidence that supports the famous Riemann Hypothesis, which is "so strong that it carries conviction even without rigorous proof" (p.369).

This type of experimental methods plays an essential role in mathematical enquiry and has always been used by mathematicians (de Villiers, 2004; Epstein & Levy, 1995). However, the growing calculation and visualisation power of computers has dramatically increased the possibilities of experimentation, stimulating mathematicians to start sharing their results within the growing field of Experimental Mathematics. While these experiments may reveal paths leading to mathematical proofs (de Villiers, 2004; Epstein & Levy, 1995), this is not necessarily the case. On this subject, Jonathan Borwein, co-author of *Mathematics by Experiment: Plausible Reasoning in the 21st Century* and *Experimentation in Mathematics: Computational Paths to Discovery*, says: "one thing that's happening is you can discover many more things than you can explain." (cited in Klarreich, 2004, p.267) Similarly, regarding the Riemann Hypothesis, Davis and Hersh (1981) claim that mathematicians continue looking for a mathematical proof because they are not satisfied knowing *that* the hypothesis is true, they also want to know *why* it is true (p. 368).

UNDERSTANDING

Referring to his own research experience, Field's medallist William Thurston (1994, p.173) pointed out an important function of mathematical proof:

> When I started working on foliations, I had the conception that what people wanted was to know the answers. I thought that what they sought was a collection of powerful proven theorems that might be applied to answer further mathematical questions. But that's only one part of the story. More than the knowledge, people want personal understanding.

While conviction responds to the question "is the statement true?" personal understanding comes from a meaningful answer to the question "why is it true?" Accordingly, mathematics educators and mathematicians (e.g. de Villiers, 1999; Hersh, 1997) distinguish between these two functions of proof: verification/conviction (concerned with the truth of the statement), and explanation (providing insight into why the statement is true). Of course, a meaningful explanation that provides personal understanding will most likely also provide an internal sense of conviction. This suggests an additional differentiation of two types of internal conviction: internal conviction provided by heuristic, empirical arguments, and internal conviction that comes from a personal understanding of why a statement is true. As a result, three kinds of conviction are discernible:

> We have, then, three kinds of convictions. One is the formal extrinsic type of conviction indirectly imposed by a formal (sometimes a purely symbolical) argumentation. The second is the empirical inductive form of conviction derived from a multitude of practical findings which support the respective conclusions. The third is the intuitive intrinsic type of conviction, directly imposed by the structure of the situation itself. In the last case, the term "cognitive belief" seems to be very appropriate. (Fischbein, 1982, p.11)

Mathematicians' search for understanding helps explain the special status of mechanically-checked formal proofs and computational experiments in mathematics research. In the first case, the end product is a computer confirmation of the validity of a conjecture, and a massive amount of code in a syntax that is foreign to mathematics. Furthermore, as noted by Gonthier (n.d.), in order to be able to use existing proof assistants, all "visual reasoning" has to be replaced by "equational reasoning", which involves losing a major source of personal understanding in mathematical fields like graph theory. In the case of computational experiments, the end product is evidence that the conjecture *ought to be valid*, based on revealing visualisations, or the verification of a large number of cases. These experiments *may* become the source of ideas to explain why the statement is true, which makes them extremely valuable in the development of mathematics (de Villiers, 2004; Epstein & Levy, 1995), but they do not necessarily provide such insight.

DISCUSSION

The present theoretical perspective builds on Raman's (2002, 2003) aspectual framework, according to which proof involves both private and public arguments, which represent its essentially private and public aspects. According to Raman (2003), private arguments engender personal understanding, while public ones are "arguments with sufficient rigor for a particular mathematical community" (p.320). However Raman's perspective does not clearly state the role of private arguments (represented by heuristic ideas) in human conviction. While she claims that "heuristic ideas can lead to a private sense of conviction" (Raman, 2002, p.36), elsewhere she states that "a heuristic idea gives a sense of understanding, but not conviction" (Raman, 2003, p.322). The present framework (Figure 1) resolves this apparent inconsistency by expanding Raman's useful perspective to include two types of internal conviction, i.e. conviction that follows from an empirical argument, and conviction that responds to intuition and personal understanding.

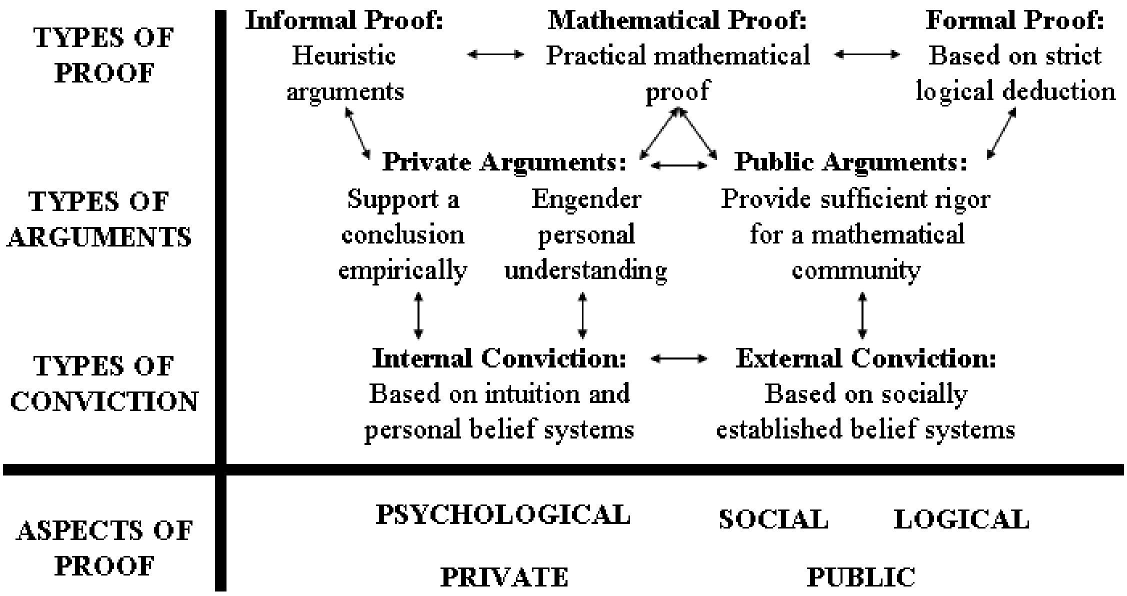

Figure 1. Different aspects of proof. Both mathematical and formal proof involve rigorous public arguments that promote the external conviction of mathematicians. However this only refers to the *social* and *logical* aspects of proof. The *psychological* aspect of proof (particularly important in mathematical proof) consists of private arguments that engender internal conviction and personal understanding. At its best, mathematical proof reaches equilibrium in these three aspects, but this is not always the case.

In summary, the distinction between mathematical and formal proof (established in terms of the level of logical rigor of a proof and its power to provide external conviction) concentrates on the public aspect of proof and disregards its private aspect, which involves arguments that engender understanding and/or internal conviction (see Figure 1). Also, mechanically-checked formal proofs fall towards the right end of this scheme, focusing on the logical facet of proof; while computational experiments fall towards its left end, focusing on the heuristic/empirical branch of the psychological facet of proof. It is mainly due to this imbalance that, by themselves, mechanically-checked formal proofs and computational experiments lie at the margin of practical mathematical proof.

REFERENCES

de Villiers, M. (1999). *The role and function of proof with sketchpad*. Retrieved April 28, 2005, from http://mzone.mweb.co.za/residents/profmd/proof.pdf.

de Villiers, M. (2004). The role and function of quasi-empirical methods in mathematics. *Canadian Journal of Sci., Math., and Tech. Education*, 4(3), 397-418.

Davis, P.J., & Hersh, R. (1981). *The mathematical experience*. Boston: Birkhäuser.

Epstein, D., & Levy, S. (1995). Experimentation and proof in mathematics. *Notices of the AMS*, 42(6), 670–674.

Ernest, P. (1998). *Social constructivism as a philosophy of mathematics*. Albany, NY: State University of New York.

Fischbein, E. (1982) Intuition and proof. *For the Learning of Mathematics*,3(2), 9-24

Gonthier, G. (n.d.). *A computer-checked proof of the Four Colour Theorem.* Retrieved from http://research.microsoft.com/~gonthier/4colproof.pdf

Hersh, R. (1997). *What is mathematics, really?* London: Jonathan Cape.

Klarreich, E. (2004, April 24). Math Lab: Computer experiments are transforming mathematics, *Science News*, 164, 266-268.

Raman, M. (2002). *Proof and justification in collegiate calculus*. Unpublished doctoral dissertation, University of California, Berkeley.

Raman, M. (2003). Key ideas: What are they and how can they help us understand how people view proof? Educational *Studies in Mathematics*, 52, 319-325.

Stewart, I., & Tall, D. O. (1977). *The foundations of mathematics*. Oxford: Oxford University Press.

The Flyspeck Project Fact Sheet. (n.d.) Retrieved from http://www.math.pitt.edu/%7Ethales/flyspeck/index.html

The QED Manifesto. (n.d.) Retrieved from http://www.cs.ru.nl/~freek/qed/qed.html

Thurston, W. (1994). On proof and progress in mathematics. *Bulletin of the American Mathematical Society*, 30(2), 161-177.