THE HISTORY AND CONCEPT OF MATHEMATICAL PROOF

Steven G. Krantz
American Institute of Mathematics, Palo Alto, California 94306 U.S.A.

Keywords: Proof, axiom, postulate, definition, rigor, deduction, intuitionism, computer proof.

Contents

Summary

We consider the history and genesis of the concept of “proof” in the discipline of mathematics. A largely historical perspective is used to put the proof concept into context. Ideas about proofs are illustrated with many examples. The discussion segues into modern considerations of intuitionism, proof by computer, and other new methodologies in the subject. The chapter closes with a consideration of what will become of the role of proofs in mathematics in the future.

1. Introduction

A mathematician is a master of critical thinking, of analysis, and of deductive reasoning. These skills travel well, and can be applied in a large variety of situations—and in many different disciplines. Today, mathematical skills are being put to good use in medicine, physics, law, commerce, Internet design, engineering, chemistry, biological science, social science, anthropology, genetics, warfare, cryptography, plastic surgery, security analysis, data manipulation, computer science, and in many other disciplines and endeavors as well.

The unique feature that sets mathematics apart from other sciences, from philosophy, and indeed from all other forms of intellectual discourse, is the use of rigorous proof. It is the proof concept that makes the subject cohere, that gives it its timelessness, and that enables it to travel well. The purpose of this discussion is to describe proof, to put it in context, to give its history, and to explain its significance.

There is no other scientific or analytical discipline that uses proof as readily and routinely as does mathematics. This is the device that makes theoretical mathematics special: the tightly knit chain of reasoning, following strict logical rules, that leads inexorably to a particular conclusion. It is *proof* that is our device for establishing the absolute and irrevocable truth of statements in our subject. This is the reason that we can depend on mathematics that was done by Euclid (325 B.C.E.–265 B.C.E.) 2300 years ago as readily as we believe in the mathematics that is done today. No other discipline can make such an assertion.

2. The Concept of Proof

The tradition of mathematics is a long and glorious one. Along with philosophy, it is the oldest venue of human intellectual inquiry. It is in the nature of the human condition to want to understand the world around us, and mathematics is a natural vehicle for doing so. Mathematics is also a subject that is beautiful and worthwhile in its own right. A scholarly pursuit that had intrinsic merit and aesthetic appeal, mathematics is certainly worth studying for its own sake.

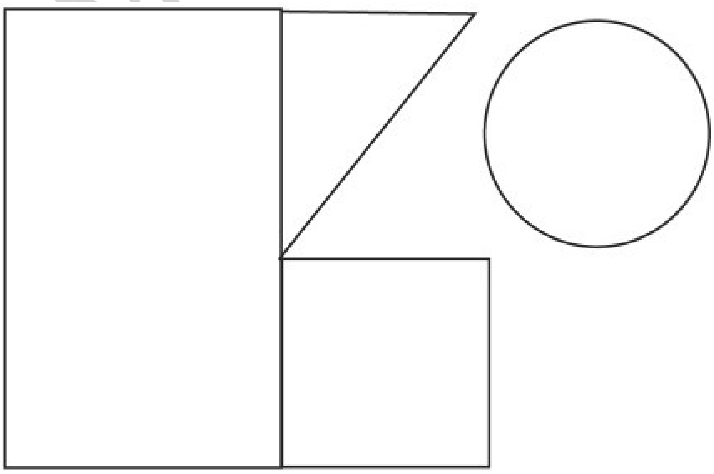

Figure 1. Mathematical constructions from surveying.

In its earliest days, mathematics was often bound up with practical questions. The Egyptians, as well as the Greeks, were concerned with surveying land. Refer to Figure 1. Thus it was natural to consider questions of geometry and trigonometry. Certainly triangles and rectangles came up in a natural way in this context, so early geometry concentrated on these constructs. Circles, too, were natural to consider—for the design of arenas and water tanks and other practical projects. So ancient geometry (and Euclid's axioms for geometry) discussed circles.

The earliest mathematics was phenomenological. If one could draw a plausible picture, or give a compelling description, then that was all the justification that was needed for a mathematical "fact". Sometimes one argued by analogy. Or by invoking the gods. The notion that mathematical statements could be *proved* was not yet an idea that had been developed. There was no standard for the concept of proof. The logical structure, the "rules of the game", had not yet been created.

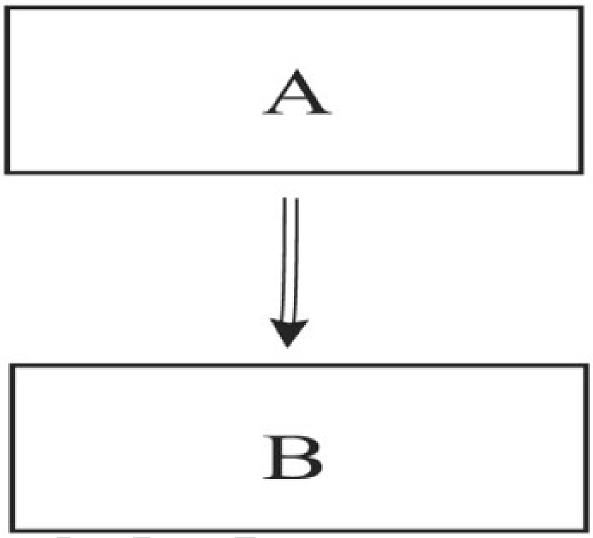

Figure 2. Logical derivation.

Thus we are led to ask: What is a proof? Heuristically, a proof is a rhetorical device for convincing someone else that a mathematical statement is true or valid. And how might one do this? A moment's thought suggests that a natural way to prove that something new (call it **B**) is true is to relate it to something old (call it **A**) that has already been accepted as true. Thus arises the concept of *deriving* a new result from an old result. See Figure 2. The next question then is, "How was the old result verified?" Applying this regimen repeatedly, we find ourselves considering a chain of reasoning as in Figure 3. But then one cannot help but ask: "Where does the chain begin?" And this is a fundamental issue.

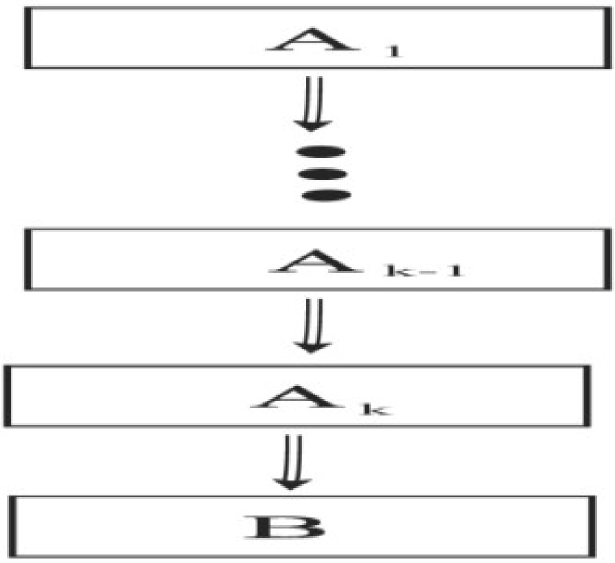

Figure 3. A chain of reasoning.

It will not do to say that the chain has no beginning: it extends infinitely far back into the fogs of time. Because if that were the case it would undercut our thinking of what a proof should be. We are endeavoring to justify new mathematical facts in terms of old mathematical facts. But if the reasoning regresses infinitely far back into the past, then we cannot in fact ever grasp a basis or initial justification for our reasoning. As we shall see below, the answer to these questions is that the mathematician puts into place definitions and axioms before beginning to explore the firmament, determine what is true, and then to prove it. Considerable discussion will be required to put this paradigm into context.

As a result of these questions, ancient mathematicians had to think hard about the nature of mathematical proof. Thales (640 B.C.E.–546 B.C.E.), Eudoxus (408 B.C.E.–355 B.C.E.), and Theaetetus of Athens (417 B.C.E.–369 B.C.E.) actually formulated theorems. Thales definitely proved some theorems in geometry (and these were later put into a broader context by Euclid). A theorem is the mathematician's formal enunciation of a fact or truth. Eudoxus is to be admired for the rigor of his work (his ideas about comparing incommensurable quantities—comparable to our modern method of cross-multiplication—are to be particularly noted), but he generally did not prove theorems. His work had a distinctly practical bent, and he was notably fond of calculations. Hippocrates of Chios (470 B.C.E.–410 B.C.E.) proved a theorem about ratios of areas of circles that he put to especially good use in calculating the areas of lunes (moon-shaped crescents).

It was Euclid of Alexandria who first formalized the way that we now think about mathematics. Euclid had definitions and axioms and then theorems—in that order. There is no gainsaying the assertion that Euclid set the paradigm by which we have been practicing mathematics for 2300 years. This was mathematics done right. Now, following Euclid, in order to address the issue of the infinitely regressing chain of reasoning, we begin our studies by putting into place a set of *Definitions* and a set of *Axioms*.

What is a definition? A definition explains the meaning of a piece of terminology. There are logical problems with even this simple idea, for consider the first definition that we are going to formulate. Suppose that we wish to define a *rectangle*. This will be the first piece of terminology in our mathematical system. What words can we use to define it? Suppose that we define rectangle in terms of points and lines and planes and right angles. That begs the questions: What is a point? What is a line? What is a plane? How do we define “angle”? What is a right angle?

Thus we see that our *first* definition(s) must be formulated in terms of commonly accepted words that require no further explanation. It was Aristotle (384 B.C.E.–322 B.C.E.) who insisted that a definition must describe the concept being defined in terms of other concepts already known. This is often quite difficult. As an example, Euclid defined a *point* to be that which has no part. Thus he is using words *outside of mathematics*, that are a commonly accepted part of everyday argot, to explain the precise mathematical notion of “point” (*It is quite common, among those who study the foundations of mathematics, to refer to terms that are defined in non-mathematical language—that is, which cannot be defined in terms of other mathematical terms—as undefined terms. The concept of “point” is an undefined term.*). Once “point” is defined, then one can use that term in later definitions—for example, to define “line”. And one will also use everyday language that does not require further explication. That is how we build up our system of definitions.

The definitions give us then a language for doing mathematics. We formulate our results, or *theorems*, by using the words that have been established in the definitions. But wait, we are not yet ready for theorems. Because we have to lay cornerstones upon which our reasoning can develop. That is the purpose of axioms.

What is an axiom? An axiom (or postulate) is a mathematical statement of fact, formulated using the terminology that has been defined in the definitions, that is taken to be self-evident. An axiom embodies a crisp, clean mathematical assertion. One does not *prove* an axiom. One takes the axiom to be given, and to be so obvious and plausible that no proof is required.

Generally speaking, in any subject area of mathematics, one begins with a brief list of definitions and a brief list of axioms. Once these are in place, and are accepted and understood, then one can begin proving theorems . And what is a proof? A proof is a rhetorical device for convincing another mathematician that a given statement (the theorem) is true. Thus a proof can take many different forms. The most traditional form of mathematical proof is that it is a tightly knit sequence of statements linked together by strict rules of logic. But the purpose of the present chapter is to discuss and consider the various forms that a proof might take. Today, a proof could (and often does) take the traditional form that goes back 2300 years to the time of Euclid. But it could also consist of a computer calculation. Or it could consist of constructing a physical model. Or it could consist of a computer *simulation* or *model*. Or it could consist of a computer algebra computation using `Mathematica` or `Maple` or `MatLab`. It could also consist of an agglomeration of these various techniques.

3. What Does a Proof Consist Of?

Most of the steps of a mathematical proof are applications of the elementary rules of logic. This is a slight oversimplification, as there are a great many proof techniques that have been developed over the past two centuries. These include proof by mathematical induction, proof by contradiction, proof by exhaustion, proof by enumeration, and many others. But they are all built on one simple rule: *modus ponendo ponens*. This rule of logic says that if we know that "**A** implies **B**", and if we know "**A**", then we may conclude **B**. Thus a proof is a sequence of steps linked together by *modus ponendo ponens* (*One of the most important proof techniques in mathematics is "proof by contradiction". With this methodology, one assumes in advance that the desired result is false and shows that that leads to an untenable position. But in fact proof by contradiction is nothing other than a reformulation of modus ponendo ponens.*).

It is really an elegant and powerful system. *Occam's Razor* is a logical principle posited in the fourteenth century (by William of Occam (1288 C.E.–1348 C.E.)) which advocates that your proof system should have the smallest possible set of axioms and logical rules. That way you minimize the possibility that there are internal contradictions built into the system, and also you make it easier to find the source of your ideas. Inspired both by Euclid's *Elements* and by Occam's Razor, mathematics has striven for all of modern time to keep the fundamentals of its subject as streamlined and elegant as possible. We want our list of definitions to be as short as possible, and we want our collection of axioms or postulates to be as concise and elegant as possible. If you open up a classic text on group theory—such as Marshall Hall's masterpiece [HAL], you will find that there are just three axioms on the first page. The entire 434-page book is built on just those three axioms (*In fact there has recently been found a way to enunciate the premises of group theory using just one axiom, and not using the word "and". References for this work are [KUN], [HIN], and [MCC].*). Or instead have a look at Walter Rudin's classic *Principles of Mathematical Analysis* [RUD]. There the subject of real variables is built on just twelve axioms. Or look at a foundational book on set theory like Suppes [SUP] or Hrbacek and Jech [HRJ]. There we see the entire subject built on eight axioms.

4. The Purpose of Proof

The experimental sciences (physics, biology, chemistry, for example) tend to use laboratory experiments or tests to check and verify assertions. The benchmark in these subjects is the *reproducible experiment with control*. In their published papers, these scientists will briefly describe what they have discovered, and how they carried out the steps of the corresponding experiment. They will describe the *control*, which is the standard against which the experimental results are compared. Those scientists who are interested can, on reading the chapter, then turn around and replicate the experiment in their own labs. The really classic, and fundamental and important, experiments become classroom material and are reproduced by students all over the world. Most experimental science is *not* derived from fundamental principles (like axioms). The intellectual process is more empirical, and the verification procedure is correspondingly practical and direct.

Mathematics is quite a different sort of intellectual enterprise. In mathematics we set our definitions and axioms in place *before* we do anything else. In particular, *before we* *endeavor to derive any results* we must engage in a certain amount of preparatory work. Then we give precise, elegant formulations of statements and we prove them. Any statement in mathematics which lacks a proof has no currency. Nobody will take it as valid. And nobody will use it in his/her own work. The proof is the final test of any new idea. And, once a proof is in place, that is the end of the discussion. Nobody will ever find a counterexample, nor ever gainsay that particular mathematical fact.

Another special feature of mathematics is its timelessness. The theorems that Euclid and Pythagoras proved 2500 years ago are still valid today; and we use them with confidence because we know that they are just as true today as they were when those great masters first discovered them. Other sciences are quite different. The medical or computer science literature of even three years ago is considered to be virtually useless. Because what people thought was correct a few years ago has already changed and migrated and transmogrified. Mathematics, by contrast, is here forever.

What is marvelous is that, in spite of the appearance of some artificiality in the mathematical process, mathematics provides beautiful models for nature (see the lovely essay [WIG], which discusses this point). Over and over again, and more with each passing year, mathematics has helped to explain how the world around us works. Just a few examples illustrate the point:

- Isaac Newton derived Kepler’s three laws of planetary motion from just his universal law of gravitation and calculus.
- There is a complete mathematical theory of the refraction of light (due to Isaac Newton, Willebrord Snell, and Pierre de Fermat).
- There is a mathematical theory of the propagation of heat.
- There is a mathematical theory of electromagnetic waves.
- All of classical field theory from physics is formulated in terms of mathematics.
- Einstein’s field equations are analyzed using mathematics.
- The motion of falling bodies and projectiles is completely analyzable with mathematics.
- The technology for locating distant submarines using radar and sonar waves is all founded in mathematics.
- The theory of image processing and image compression is all founded in mathematics.
- The design of music CDs is all based on Fourier analysis and coding theory, both branches of mathematics.

The list could go on and on.

The key point to be understood here is that *proof* is central to what modern mathematics is about, and what makes it reliable and reproducible. No other science depends on proof, and therefore no other science has the bulletproof solidity of mathematics. But mathematics is *applied* in a variety of ways, in a vast panorama of disciplines. And the applications are many and varied. Other disciplines often like to reduce their theories to mathematics—or at least explain them in mathematical terms—because it gives the subject a certain elegance and solidity. And it looks really sophisticated. Such efforts meet with varying success.

Many important periods of mathematical thought were important for the dynamic and the development and the continuity of mathematics, even though they did not stress proof. The medieval development of algebra by the Arabs, the construction of the Hindu-Arabic numerals, certainly the invention of calculus by Newton and Leibniz were all milestones in mathematical (and human) thought. They greatly enriched our understanding of mathematics, and were an important step in the history of the subject, but they did little to touch on the concept of proof.

5. The History of Mathematical Proof

In point of fact the history of the proof concept is rather inchoate. It is unclear just when mathematicians and philosophers conceived of the notion that mathematical assertions required justification. This was quite a new idea. Then it was another considerable leap to devise methods for *constructing* such a justification. In the present section we shall outline what little is known about the development of the proof concept.

Perhaps the first mathematical "proof" in recorded history is due to the Babylonians. They seem (along with the Chinese) to have been aware of the Pythagorean theorem (discussed in detail below) well before Pythagoras (*Although it must be stressed that they did not have Pythagoras's sense of the structure of mathematics, of the importance of rigor, or of the nature of formal proof. Some will dispute calling their treatment of the Pythagorean result a proof, but many sources use this terminology. The Babylonians certainly seem to have had algebraic calculations that pointed to the proof or verification of the result.*). The Babylonians had certain diagrams that indicate why the Pythagorean theorem is true, and tablets have been found to validate this fact (*We stress that the Babylonian effort was not a proof by modern standards. But it was at least an effort to provide logical justification for a mathematical fact.*). They also had methods for calculating Pythagorean triples—that is, triples of integers (or whole numbers) a, b, c that satisfy

$$a^2 + b^2 = c^2$$

as in the Pythagorean theorem.

-
-
-

Bibliography

E. Bishop, *Foundations of Constructive Analysis*, McGraw-Hill, New York, 1967. [*Errett Bishop's fundamental work on constructive methods in mathematical analysis]*.

E. Bishop and D. Bridges, *Constructive Analysis*, Springer-Verlag, New York, 1985. [*A modernized version of Errett Bishops's work on constructive methods in analysis. Published posthumously]*.

G. Chen, S. G. Krantz, D. Ma, C. E. Wayne, and H. H. West, The Euler-Bernoulli beam equation with boundary energy dissipation, in *Operator Methods for Optimal Control Problems* (Sung J. Lee, ed.), Marcel Dekker, New York, 1988, 67-96. [*A foundational work on the eigenvalue asymptotics of multiple clamped beams]*.

G. Chen, S. G. Krantz, C. E. Wayne, and H. H. West, Analysis, designs, and behavior of dissipative joints for coupled beams, *SIAM Jr. Appl. Math*., 49(1989), 1665-1693. [*A second foundational work on the eigenvalue asymptotics of multiple clamped beams]*.

S.-C. Chou and W. Schelter, Proving geometry theorems with rewrite rules, *Journal of Automated Reasoning* 2(1986), 253--273. *[A paper which demonstrates how new results in Euclidean geometry may be discovered with computers]*.

J. Gray, *The Hilbert Challenge*, Oxford University Press, New York, 2000. [*A historical work about the history of the Hilbert problems in mathematics]*.

Biographical Sketch

Steven G. Krantz received his B.A. degree from the University of California at Santa Cruz in 1971. He earned the Ph.D. from Princeton University in 1974. He has taught at UCLA, Princeton University, Penn State, and Washington University in St. Louis. Krantz is the holder of the UCLA Alumni Foundation Distinguished Teaching Award, the Chauvenet Prize, and the Beckenbach Book Prize. He is the author of 150 papers and 50 books. His research interests include complex analysis, real analysis, harmonic analysis, and partial differential equations. Krantz is currently the Deputy Director of the American Institute of Mathematics.