Reliability of mathematical inference

Jeremy Avigad

Department of Philosophy and
Department of Mathematical Sciences
Carnegie Mellon University

January 2020

Formal logic and mathematical proof

An important mathematical goal is to get the answers right:

- Our calculations are supposed to be correct.
- Our proofs are supposed to be correct.

Mathematical logic offers an idealized account of correctness, namely, formal derivability.

Informal proof is viewed as an approximation to the ideal.

- A mathematician can be called on to expand definitions and inferences.
- The process has to terminate with fundamental notions, assumptions, and inferences.

Formal logic and mathematical proof

Two objections:

- Few mathematicians can state formal axioms.
- There are various formal foundations on offer.

Slight elaboration:

- Ordinary mathematics relies on an informal foundation: numbers, tuples, sets, functions, relations, structures, ...
- Formal logic accounts for those (and any of a number of systems suffice).

Formal logic and mathematical proof

What about intuitionstic logic, or large cardinal axioms?

Most mathematics today is classical, and does not require strong assumptions.

But even in those cases, the assumptions can be make explicit and formal.

Formal logic and mathematical proof

So formal derivability provides a standard of correctness.

Azzouni writes:

> *The first point to observe is that formalized proofs have become the norms of mathematical practice. And that is to say: should it become clear that the implications (of assumptions to conclusion) of an informal proof cannot be replicated by a formal analogue, the status of that informal proof as a successful proof will be rejected.*

Formal verification, a branch of computer science, provides corroboration: computational proof assistants make formalization routine (though still tedious).

Formal logic and mathematical proof

The fact that some arguments draw on visual and spatial intuitions leads some to challenge claims of formalizability.

But the history and contemporary practice corroborate the story:

- Space filling curves and other monsters led to rigorization.
- We have topological, metric, geometric, and analytic language to spell out our intuitions.
- Diagrams play a limited role in mathematics journals.
- Intuitively obvious theorems like the Jordan Curve Theorem, have been proved.

Formal logic and mathematical proof

Formal derivability has been held to be the standard of correctness implicitly or explicitly throughout twentieth century philosophy of mathematics.

Hamami finds clear expressions of the view in Mac Lane and Bourbaki.

Detlefsen calls it the *common view*, and Antonutti Marfori and Hamami call it the *standard view*.

Formal logic and mathematical proof

Rav has criticized the standard view on a number of grounds.

Azzouni has responded to some of them.

Others have leveled criticisms or raised concerns, including Antonutti Marfori, Detlefsen, Tanswell, Larvor, and Pelc.

Yacin Hamami has recently provided a precise formulation and spirited defense.

My goals here:

- to raise a strong objection to the standard view
- to respond to it

I believe the standard view is essentially correct, but understanding the sense in which it is correct can tell us a lot about mathematics.

Outline

- The standard view
- The problem
- General strategies
- Specific strategies
- Conclusions

The problem

According to the standard view, when a mathematical referee certifies a mathematical result, the correctness of the judgment stands or falls with the existence of a formal derivation.

How can our mathematical judgments possibly warrant the existence of such a thing?

The problem

Formal derivations are fragile objects.

A formal derivation can require thousands of inferences. If even one is incorrect, it is not a formal derivation.

With 1% error, the odds of correctly assessing 100 inferences is about 37%. The odds drop exponentially.

Informal proofs provide less information. A priori, this only makes it harder.

The fact that mathematical results rely on prior results in the literature makes matters even worse.

The problem

Mathematical texts abound in terms such as "it follows from . . . that," "given that . . . it is clear that" and the like; the antecedents are held to be true, from which the truth of the consequent is taken by necessity to follow. The issue is that beyond the verbal phrase "it follows from . . . that so and so is the case" (equals "is true"), a mathematical proof in general only says that it follows, not why by logical necessity it has to follow. Hence the frequent need to interpolate further and further intermediate links as "bridges," leading from claimed antecedents to the asserted conclusion. Why the consequent follows from the antecedents has to be figured out by the reader of a proof, based on the reader's understanding of the meaning of the terms in the antecedent and consequent and requiring the reader's familiarity with the underlying theory to which the proof is intended to be a contribution. None of these can be formally captured. (Rav 2007)

The problem

Detlefsen and Tanswell also dissociate epistemic justification of mathematics from its formalizability.

Detlefsen writes:

> *Mathematical proofs are not commonly formalized, either at the time they're presented or afterwards. Neither are they generally presented in a way that makes their formalizations either apparent or routine. This notwithstanding, they are commonly presented in a way that does make their rigor clear—if not at the start, then at least by the time they're widely circulated among peers and/or students. There are thus indications that rigor and formalization are independent concerns.*

The problem

Azzouni credits formal norms with the stability of mathematics as a shared practice, but he denies that they are *epistemic* norms; a mathematician need not produce such a derivation to know or justify such a proof.

According to Azzouni, mathematicians rely on "inference packages" and, more recently, "algorithmic languages," that circumvent formal detail.

Hamami is most optimistic about the prospects of a formal account.

- Mathematical experts have a stock of theorems, learned rules.
- When verifying a proof, they invoke procedures that reduce inferences to those.

The problem

Rav, Azzouni, and Hamami offer valuable insights:

- Background knowledge and expertise is important.
- These can be used to supply additional detail.

But the core concern remains: formal derivations are long and complex, and there is no room for error.

We need a more compelling account of how we bridge the gap between the informal and the formal.

Is it philosophy?

Two questions:

- Why isn't this psychology?
- Why isn't this just pragmatic?

Is it philosophy?

It is useful to keep a methodological distinction:

- philosophy: normative account of grounds for knowledge, based on what we see in the mathematical literature.
- psychology: descriptive account of human behavior, based on empirical studies

Understanding mathematical inference requires both.

Here my focus is on the mathematics.

Is it philosophy?

Philosophy of mathematics is supposed to provide a normative ideal.

Conventional accounts assume mathematical agents have finite capacities.

The problem raised here becomes pressing when we recognize that we have only bounded resources.

Cognitive efficiency is a legitimate norm, and we can still idealize away.

Our core concern remains the same as that of Plato and Aristotle: understanding how we can, should, and do come to know mathematics.

Outline

- The standard view
- The problem
- General strategies
- Specific strategies
- Conclusions

General strategies

A attempted formal derivation with a single error is not a formal derivation.

Not so with informal proof: an informal proof can have minor errors and still be "essentially correct."

The real question: how can informal mathematical texts warrant mathematical truth while being *robust* with respect to error.

This is an engineering problem.

General strategies

Informal proofs have much more structure. Good ones

- are *modular* (Avigad, to appear),
- are *motivated* (Morris, 2015),
- and show evidence of a *rational plan* (Hamami and Morris, in preparation).

Metaphors:

- modularity, reliability, and robustness in engineering
- grasping the plot of a novel
- planning a trip

This is not progress *per se*, but it gives us something to work with.

General strategies

Broad strategies:

- Isolate and minimize critical information.
- Maximize exposure to error detection.
- Leverage redundancy.

I will describe these in metaphorical terms, and then get more specific.

Isolate and minimize critical information

View informal proof as a form of data compression, like:

- signal coding
- version control (send the “diff”)
- image compression
- making a coarse plan

Maximize exposure to error detection

When a proof is wrong, we want it to be obviously wrong, or as clearly wrong as possible.

They should be falsifiable, not by empirical data, but by mathematical reasoning.

Slogan: maximize probability by maximizing probe-ability

Analogies:

- complete proof search
- mental models vs. formal rules in cognitive science

Leverage redundancy

Ensure there are multiple ways to succeed. Analogies:

- multiple paths to a goal
- backup plans
- extra support and backup components in engineering
- error correction in coding schemes

In this case, exponential decay works in our favor.

Outline

- The standard view
- The problem
- General strategies
- Specific strategies
- Conclusions

General strategies

Broad strategies:

- Isolate and minimize critical information.
- Maximize exposure to error detection.
- Leverage redundancy.

Specific strategies

Some specific strategies:

1. Reason by analogy.
2. Modularize.
3. Generalize.
4. Use algebraic abstraction.
5. Collect examples.
6. Classify.
7. Develop complementary approaches.
8. Visualize.

I will try to convince you that

- these occur in mathematics, and
- they support reliability and robustness.

Reason by analogy

If a proof I have in mind is similar to one you have already seen, I only need to explain the differences.

Example: unique factorization of Gaussian integers, writing $a = qb + r$ with $0 \leq \|r\| < \|b\|$

Error checking can focus on the points of difference.

Another example: various ways of completing a space (metric completions, ideal completions) look the same.

Modularize

Modular systems can be decomposed into components with limited interactions between them, modulated by interfaces.

- One can use theorems and lemmas without knowing how they are proved.
- Proofs are decomposed into smaller lemmas.
- Structures are defined in terms of simpler ones.

Positive effects:

- Modularity minimizes information (between proofs, within proofs).
- Components can be checked and repaired independently.
- Interfaces support multiple realizations.

Generalize

Mathematical lemmas are often of the form $\forall x\, (A(x) \rightarrow B(x))$.

Making A as weak as possible and B as strong as possible maximizes exposure to error.

Indeed, mathematical practice supports breaking out lemmas and stating them as strongly as possible.

General theories are analogous to special cases. For example, optimization in univariate calculus is generalized to multiple dimensions, infinite dimensions, nonsmooth functions, multivalued functions.

Similarly, discrete probability generalizes to continuous probability measures.

Use algebraic abstraction

Algebra involves characterizing classes of structures axiomatically.

- It is a prototypical means of generalization.
- Algebraic results guided by concrete cases.
- Supports transferring and reusing results.
- Supports counterexamples.

Examples:

- Groups from permutations (substitutions), geometry, number theory
- Integers and Gaussian integers are Euclidean domains, hence principle ideal domains, hence unique factorization domains and Dedekind domains.
- Topological spaces, metric spaces, inner products spaces, and normed spaces all generalize the Euclidean plane.

Collect examples

When assessing the correctness of a general proof, it is often helpful to think about a particular instance.

If the statements holds of the instance, we try to understand why. If not, we have found an error.

Having an abundance of examples is a virtue.

Textbooks often provide standard examples, as well as standard counterexamples. (Cf. *Counterexamples in Topology* and *Counterexamples in Analysis*.)

Classify

It is helpful not only to have plenty of examples, but also to have them categorized and sorted in useful ways.

Groups can be finite, abelian, nilpotent, solvable, discrete, and so on. Trained mathematicians know where hypotheses are typically useful, and an absence can raise a red flag.

Classification theorems describe families in terms of invariants and give concrete representations.

These give alternative means of proof, and also parameters that can be varied to find counterexamples.

Develop complementary approaches

Leveraging redundancy means fostering multiple ways of carrying out inferences.

There are more than 200 published proofs of the law of quadratic reciprocity.

Papers often say things like "X follows from Theorem Y in paper Z, but, for completeness, we provide a more direct proof here."

The algebraic method provides various perspectives: e.g. one can view a structure in topological, metric, or geometric terms.

Visualize

Examples of visual intuition in mathematics are compelling, and have received the most philosophical attention.

Larkin and Simon famously argued that diagrammatic representations leverage our spatial cognitive capacities.

Features in a diagram suggest properties from the constraints, and the ability to vary the diagram produces counterexamples.

A number of authors have explored ways this plays out in mathematical reasoning, and Hamami, Mumma, and Amalric have done experimental work.

Terms like “space,” “continuous transformation,” “distance,” “interior,” “endpoint,” and “smooth curve” invoke spatial intuitions even in textual proofs.

Outline

- The standard view
- The problem
- General strategies
- Specific strategies
- Conclusions

General strategies

Broad strategies:

- Isolate and minimize critical information.
- Maximize exposure to error detection.
- Leverage redundancy.

Specific strategies

Some specific strategies:

1. Reason by analogy.
2. Modularize.
3. Generalize.
4. Use algebraic abstraction.
5. Collect examples.
6. Classify.
7. Develop complementary approaches.
8. Visualize.

Conclusions

Original question: how can an informal proof reliably indicate the existence of a formal derivation?

Modified question: what general cognitive mechanisms and design principles support reliable assessement of proof?

We have begun to explore some of them.

Experimental psychology might help, but there is a lot of data in the mathematical literature, and we can begin to form models.

Conclusions

The account incorporates insights from Rav, Azzouni, and Hamami:

- Background knowledge is important.
- Expertise is important.
- We use procedural knowledge to fill in information.

Additional features:

- There is no simple answer. There are lots of mechanisms at play.
- The heuristics are reliable but fallible.
- They rely on metacognitive reflection.

But we can still try to better understand how they work.

Conclusions

This approach preserves the standard view that informal proofs work by indicating formal derivations.

It also extends the normative account: in addition to being correct, we want our proofs to support robust and reliable assessment.

The ability of mathematics to bridge the gap between the informal and the formal is one of the most important and interesting aspects of the practice.

Mathematics requires us not only to be correct, but to be correct about complex things.

Conclusions

Rav has emphasized that we get a lot more from a proof than a certificate of correctness, and that we value proofs for reasons that go beyond the ability to construct a formal derivation.

I agree wholeheartedly.

But the standard view is fully compatible with the desire to develop a broader epistemology of mathematics.

Conclusions

In fact, the approach I have taken here points to a reconciliation.

The cognitive mechanisms and design principles we have explored have as much to do with discovery and understanding as they have to do with correctness and justification.

So even if one is primarily concerned with grounds for knowledge and standards of correctness, one *has* to come to terms with the mechanisms that support mathematical understanding to make sense of how mathematics makes it possible to meet those standards.

Both verification and discovery are then subsumed under the more general umbrella of mathematical understanding.

Summary

To sum up, I have defended two complementary claims:

1. The gap between informal proof and formal derivation is not a good reason to reject the latter as a normative standard of correctness.
2. Accepting this standard is not at odds with the task of making sense of higher-level epistemic features of mathematical reasoning. Rather, the latter is an essential prerequisite to understanding how the normative standard can and should be met.

I have taken initial steps towards developing a positive account of the features that make reliable and robust assessment possible.

Summary

Based on 1, supporters of the standard view may conclude: "see, mathematics is formalizable!"

Based on 2, opponents may conclude: "see, formalization doesn't say anything about the really important stuff!"

But the two claims together give us a clearer understanding of what the standard account does, and does not, accomplish.

So we can lay the argument to rest, and focus on understanding how mathematics works as well as it does.

Extras

If you now give me an equation that you have chosen at your pleasure, and if you want to know if it is or is not solvable by radicals, I could do no more than to indicate to you the means of answering your question, without wanting to give myself or anyone else the task of doing it. In a word, the calculations are impracticable.

From that, it would seem that there is no fruit to derive from the solution that we propose. Indeed, it would be so if the question usually arose from this point of view. But, most of the time, in the applications of the Algebraic Analysis, one is led to equations of which one knows beforehand all the properties: properties by means of which it will always be easy to answer the question by the rules we are going to explain. ...All that makes this theory beautiful and at the same time difficult, is that one has always to indicate the course of analysis and to foresee its results without ever being able to perform [the calculations].

(Galois 1830)

Extras

. . . to seek proofs based immediately on fundamental characteristics, rather than on calculation, and indeed to construct the theory in such a way that it is able to predict the results of calculation. . .

(Dedekind 1877)