Mathematical Understanding

Jeremy Avigad

May 10, 2022

Carnegie Mellon University

Mathematical understanding

We often distinguish between *knowing that* something is true and *understanding why* something is true.

The topic is currently in vogue in epistemology and philosophy of science.

Looking at how understanding plays out in mathematics is a good place to start.

Mathematical understanding

For example, it is not unusual to say "I am convinced that the proof is correct, but I don't really understand what is going on."

Understanding involves something deeper and more satisfying.

Intuitions

Mathematics is hard.

Mathematical solutions, proofs, and calculations involve long sequences of steps, that have to be chosen and composed in precise ways.

To compound matters, there are too many options; among the many steps we may plausibly take, most will get us absolutely nowhere.

And we have limited cognitive capacities — we can only keep track of so much data, anticipate the result of a few small steps, remember so many background facts.

We rely on our understanding to help us and to guide us.

Mathematical understanding

Does understanding the demonstration of a theorem consist in examining each of the syllogisms of which it is composed in succession, and being convinced that it is correct and conforms to the rules of the game? In the same way, does understanding a definition consist simply in recognizing that the meaning of all the terms employed is already known, and being convinced that it involves no contradiction?

. . . Almost all are more exacting; they want to know not only whether all the syllogisms of a demonstration are correct, but why they are linked together in one order rather than in another. As long as they appear to them engendered by caprice, and not by an intelligence constantly conscious of the end to be attained, they do not think they have understood.

(Poincaré, *Science et méthod*, 1908)

Outline

- motivating questions and intuitions
- a dynamic account of understanding
- explanation, concepts, and representations
- methodological recommendations

The problem of multiple proofs

On the standard account, the value of a mathematical proof is that it warrants the truth of the resulting theorem.

Why, then, do we often value a new proof of a previous established theorem?

For example, Gauss published six proofs of the law of quadratic reciprocity in his lifetime, and left us two unpublished versions as well.

Franz Lemmermeyer has documented 246 proofs through 2013. The list is available online.

The problem of conceptual possibility

It is often said that some mathematical advance was “made possible” by a prior conceptual development.

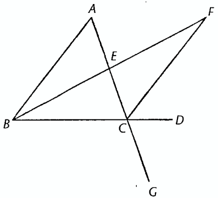
For example, Riemann’s introduction of the complex zeta function and the use of complex analysis made it possible for Hadamard and de la Vallée Poussin to prove the prime number theorem in 1896.

What is the sense of “possibility” here?

The nature of diagrammatic inference

Proposition 16

In any triangle, if one of the sides be produced, the exterior angle is greater than either of the interior and opposite angles.



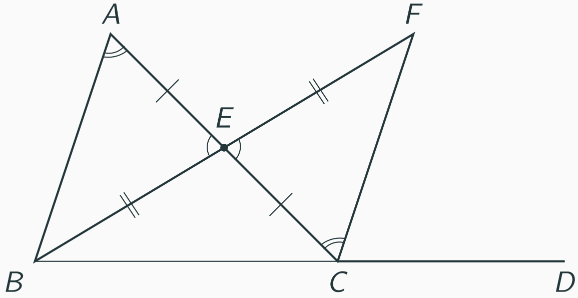
Let ABC be a triangle, and let one side of it BC be produced to D;
I say that the exterior angle ACD is greater than either of the interior and opposite angles CBA, BAC.

Let AC be bisected at E, **[I. 10]**
and let BE be joined and produced in a straight line to F;
let EF be made equal to BE, **[I. 3]**
let FC be joined, **[Post. 1]**
and let AC be drawn through to G. **[Post. 2]**

Then, since AE is equal to EC, and BE to EF,
the two sides AE, EB are equal to the two sides CE, EF respectively;
and the angle AEB is equal to the angle FEC, for they are vertical angles. **[I. 15]**

Therefore the base AB is equal to the base FC, and the triangle ABE is equal to the triangle CFE,
and the remaining angles are equal to the remaining angles respectively, namely those which the equal sides subtend; **[I. 4]**
therefore the angle BAE is equal to the angle ECF.

The nature of diagrammatic inference



By side-angle-side, $\triangle AEB \equiv \triangle CEF$. So $\angle BAC = \angle ACF$.

Clearly $\angle ACD > \angle ACF$. So $\angle ACD > \angle BAC$.

But why is it clear that $\angle ACD > \angle ACF$?

The problem of reliability

On a standard account, a proof is correct if each inference can be expanded to a formal derivation.

Such formal derivations can be extremely long. Even a single error renders one invalid.

How can ordinary mathematical proofs reliably warrant the existence of something so complex and fragile?

Why doesn't mathematics fall apart?

The role of abstraction

The value of algebraic reasoning is often attributed to its generality.

For example, the axiomatization of *groups* in the nineteenth century unified instances in Galois theory, number theory, and geometry.

But sometimes abstraction is valued even when there is only one instance.

In 1871, Dedekind introduced the notion of an *ideal* in a number ring. In 1882 he and Weber generalized it to rings of functions.

But Dedekind clearly thought the notion was useful, even before the 1882 generalization.

Why?

The use of computers in proofs

Kenneth Appel and Wolfgang Haken used extensive computation to prove the four-color theorem in 1976.

Thomas Hales announced a proof of the Kepler conjecture in 1998, again using extensive computation.

Propositional satisfiability solvers are being used to solve combinatorial problems, in some cases, producing proofs that are terabytes long.

Is this good mathematics?

Motivating questions

What the questions have in common:

- They have a general epistemological character.
- They raise normative questions. (What do we value? What makes for good mathematics?)
- They have to do with mathematical *understanding*.
- We have some intuitions.
- We care about the answers.

This doesn't guarantee that there is room for philosophy here.

But it should encourage us to take a look.

Motivating questions

How the questions relate to understanding:

- New proofs provide new understanding.
- Proving theorems requires a conceptual understanding.
- Reading a Euclidean proof requires a geometric understanding.
- We don't check proofs formally; we understand them.
- Abstraction can make a proof easier to understand.

Intuitions

Some tasks require understanding.

- reading a proof
- answering questions
- writing a proof
- discovering new theorems

Mathematical artifacts convey understanding

- proofs
- definitions
- concepts
- theories

Intuitions

Mathematical knowledge is static: definitions, theorems, and proofs.

Mathematical understanding is dynamic; it's the difference between knowing how and knowing that.

It's the capacity to think and reason mathematically.

Intuitions

Talking about mathematical understanding also means talking about:

- explanation
- concepts
- representations
- cognitive effort

Sums of squares

Let's consider an example.

In the *Arithmetic*, Diophantus notes that

- $5 = 2^2 + 1^2$
- $13 = 3^2 + 2^2$
- $5 \times 13 = 65 = 8^2 + 1^2 = 7^2 + 4^2$.

Theorem. If x and y can each be written as a sum of two integer squares, then so can xy.

Sums of squares

Proof #1. Suppose $x = a^2 + b^2$, and $y = c^2 + d^2$. Then

$$xy = (ac - bd)^2 + (ad + bc)^2,$$

a sum of two squares. $\square$

In more detail:

$$\begin{aligned}
&(ac - bd)^2 + (ad + bc)^2 \\
&\qquad = a^2c^2 - 2abcd + b^2d^2 + a^2d^2 + 2abcd + b^2d^2 \\
&\qquad = a^2c^2 + b^2d^2 + a^2d^2 + b^2d^2 \\
&\qquad = (a^2 + b^2)(c^2 + d^2)
\end{aligned}$$

Note: $(ac + bd)^2 + (ad - bc)^2$ works just as well.

Sums of squares

Define the *Gaussian integers*:

$$\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$$

If $\alpha = u + vi$, define the *conjugate*:

$$\overline{\alpha} = u - vi.$$

We have $\overline{\alpha\beta} = \overline{\alpha} \cdot \overline{\beta}$.

Define the norm:

$$N(\alpha) = \alpha\overline{\alpha} = (u + iv)(u - iv) = u^2 - i^2v^2 = u^2 + v^2.$$

Then

$$N(\alpha\beta) = \alpha\beta \cdot \overline{\alpha\beta} = \alpha \cdot \beta \cdot \overline{\alpha} \cdot \overline{\beta} = \alpha\overline{\alpha} \cdot \beta\overline{\beta} = N(\alpha)N(\beta).$$

Sums of squares

Proof #2. Suppose $x = N(\alpha)$ and $y = N(\beta)$ are sums of two squares. Then $xy = N(\alpha\beta)$, a sum of two squares. $\square$

Outline

- motivating questions and intuitions
- a dynamic account of understanding
- explanation, concepts, and representations
- methodological recommendations

Methods and abilities

Mathematical knowledge is often cast as *propositional* knowledge, like definitions and theorems.

But *understanding* seems to require something more dynamic, a kind of *procedural knowledge*.

Understanding guides thought.

One approach: talk about *methods*, i.e. heuristic, fallible, procedures for solving problems, searching for proofs, verifying inferences, etc.

Methods and abilities

Straightforward model:

- We face tasks (solving a problem, proving a theorem, verifying an inference, developing a theory, forming a conjecture).
- "Reasoning" involves passage though various epistemic states.
- "Understanding" (methods, techniques, procedures, protocols, tactics, strategies, . . .) makes this passage possible.

Talk of methods may be too fine-grained.

People multiply numbers in different ways. Sometimes we only care about the ability to do so.

Methods and abilities

Another approach: talk about *abilities*, or capacities for thought.

Understanding involves:

- Being able to recognize the nature of the objects and questions before us.
- Being able to marshall the relevant background knowledge and information.
- Being able to traverse space the of possibilities before us in a fruitful way.
- Being able to identify features of the context that help us cut down complexity.

Methods and abilities

The methodological thesis: for many purposes, we do not need anything more than an account of the abilities, or capacities, that we take to be constitutive of particular instances understanding.

To characterize a particular type of understanding, it suffices to characterize the abilities it confers.

If this explains the data (mathematical practice), we need not look any further.

See also Janet Folina, “Towards a better understanding of mathematical understanding.”

Methods and abilities

We can apply this point of view wherever talk of understanding arises:

- contemporary mathematics
- mathematical education
- history of mathematics
- automated reasoning and AI

To do that, we need better ways of talking about methods and abilities.

Methods and abilities

Challenges:

- Algorithms are overly specific; different methods may account for the same ability.
- Yet there is a compositional aspect to methods and abilities.
- Mathematical methods are heuristic and fallible.
- There are no clear criteria of identity.

Machine models, cognitive models, programming languages, psychological data, etc. seem to provide the wrong level of description.

We need a level of abstraction that is appropriate for talking about the interesting features of the *mathematics*.

Procedural aspects of proof

In *Mathematical Method and Proof*, I emphasized procedural language in proof:

> "…the first law may be proved by induction on n."
>
> "…by successive applications of the definition, the associative law, the induction assumption, and the definition again."
>
> "By choice of m, $P(k)$ will be true for all $k < m$."
>
> "Hence, by the well-ordering postulate…"
>
> "From this formula it is clear that…"
>
> "This reduction can be repeated on b and r_1…"
>
> "This can be done by expressing the successive remainders r_i in terms of a and b…"
>
> "By the definition of a prime…"

Procedural aspects of proof

"On multiplying through by b..."

"...by the second induction principle, we can assume $P(b)$ and $P(c)$ to be true..."

"Continue this process until no primes are left on one side of the resulting equation..."

"Collecting these occurrences, ..."

"By definition, the hypothesis states that..."

"...Theorem 10 allows us to conclude ..."

Birkhoff and Mac Lane, *A Survey of Modern Algebra*, Chapter 1.

Procedural aspects of proof

Fenner Tanswell has also written about imperative language in proofs. See:

- "Go Forth and Multiply: On Actions, Instructions and Imperatives in Mathematical Proofs"
- (with Matthew Inglis) "The Language of Proofs: A Philosophical Corpus Linguistics Study of Instructions and Imperatives in Mathematical Texts"

Sums of squares

Advantages of the more "conceptual" proof:

- The norm (and its square root, the modulus, or absolute value) are generally useful. For example, the Gaussian integers are a Euclidean domain.
- The proof is easy to remember and reconstruct.
- It avoids calculation.
- Generalizations to the quaternions and octonians give product rules for sums of 4 and 8 squares — and there are no others.
- Conjugates and norms lie at the heart of algebraic number theory.
- They provide, for example, a general theory of *quadratic forms* (expressions $ax^2 + bxy + cy^2$).

Aspects of these can be cast in terms of abilities.

Understanding and rationality

Rebecca Morris has argued that we expect proofs to be *motivated*:

- We want to understand why a proof step is reasonable, in the current state.
- We want to understanding how a proof step gets us closer to our goal.

See "Motivated proofs: What they are, why they matter and how to write them."

Yacin Hamami and Morris have argued that (one aspect of) understanding a proof involves the ability to cast the proof as an instance of a *rational plan*.

See "Plans and planning in mathematical proofs."

Understanding and rationality

Other aspects of understanding a proof:

- How is this hypothesis used?
- Why is this assumption necessary?
- Why is this step valid?
- Does ... provide an alternative proof?
- Can this be generalized to ... ?
- Can this be expressed in terms of ... ?
- How can the proof be varied?

Question: to what extent can these be cast in terms of grasping a proof plan?

Outline

- motivating questions and intuitions
- a dynamic account of understanding
- explanation, concepts, and representations
- methodological recommendations

Understanding vs. explanation

In *Mathematical Method and Proof* I expressed skepticism toward theories of explanation:

- the term is not so commonly used in mathematics
- judgments vary
- at best, it seems to refer to a basket of judgments.

I'd argue that the prospects are better for a theory of understanding, if only because it is more flexible and open-ended.

Understanding vs. explanation

Pease, Aberdein, and Martin have done empirical studies of explanation in collaborative discussions.

They distinguish between:

- exploration (mathematical discussion)
- publication (final mathematical presentation)

The argue that there explanations in mathematics, and that they are answers to *why* questions:

- trace explanations (reveal sequence of inferences)
- strategic explanations (place action in problem-solving context)
- deep explanations (relates question to user's knowledge base)

Concepts

In the psychological literature, concepts are sometimes thought of in terms of categorization (e.g. prototypes and exemplars).

From a logical perspective, a concept is given by a definition, in a suitable formal language.

These don't work so well for the philosophy of mathematics.

What does it mean to understand the concept of a *group*? Or the concept of a *function*? Or the concept of a *Riemannian manifold*?

Concepts

Mathematical concepts have some interesting features:

- Membership is often sharply defined.
- Mathematical concepts evolve over time.
- Understanding a concept admits degrees.
- Various things can “improve our understanding” of a concept.
- One can speak of implicit uses of a concept.

Concepts

One solution: think of a mathematical concept as a bundle of abilities.

For example, understanding the group concept includes:

- Knowing the definition of a group.
- Knowing common examples of groups, and being able to recognize implicit group structures when it is fruitful to do so.
- Knowing how to construct groups from other groups or other structures, in fruitful ways.
- Recognizing that there are different kinds of groups (abelian, nilponent, solvable, finite vs. infinite, continuous vs. discrete) and being able/prone to make these distinctions.
- Knowing various theorems about groups, and when and how to apply them.

Concepts

This renders “the group concept,” for example, vague and open-ended.

But the notion *is* vague and open-ended:

- We can talk about student understanding.
- We can talk about the role of the concept in contemporary mathematics.
- We can talk about the historical development.

The proposal suggests that we can make our talk more precise by being more precise about the abilities (or methods, or capacities) we have in mind.

Representations

In philosophy of mind, sometimes a concept is taken to be some sort of mental *representation*, maybe in a language of thought.

Understanding seems to have something to do with having the right representations.

In contemporary philosophy of mathematics, there has been a lot of interest in the nature of representations, especially diagrammatic representations.

Ken Manders has advocated using the word *artifacts*. Roy Wagner likes *presentations*.

What is important is not what they represent, but what we can do with them.

Cognitive effort

As cognitive agents, we have limited time, energy, memory, processing capacity.

We value developments that make things *easier*.

But how can we measure difficulty?

- Computer science: algorithmic complexity
- Logic: descriptive complexity, length of proof
- Experimental psychology: timing tasks

We can consider the number of pages in a proof, the number of symbols in an expression, or the number of steps in a calculation.

But we need better ways of talking about cognitive difficulty.

Outline

- motivating questions and intuitions
- a dynamic account of understanding
- explanation, concepts, and representations
- methodological recommendations

Goals

We are looking for a philosophical account that:

- is clear, precise, and internally coherent
- accords with our intuitions
- fits the data (what we see in mathematics)
- can inform (and can be informed by) other pursuits:
  - history of mathematics
  - interactive theorem proving and automated reasoning
  - psychology and cognitive science
  - mathematics education
  - mathematics itself

Toward a theory of mathematical understanding

General strategy: think globally, act locally.

Keep the big questions in mind, but address more focused questions:

- What kinds of things can be inferred from the diagrams in Euclid's *Elements*?
- How did Dedekind's introduction of *ideals*, or Dirichlet's introduction of *characters*, contribute to number theory?
- What mechanisms can be used to model algebraic hierarchies in interactive proof assistants?

Toward a theory of mathematical understanding

If we

- continue to make progress on specific questions and
- keep the general questions in mind,

a theory of mathematical understanding will eventually emerge.

What about the overarching question: Why do we do mathematics the way we do?

From formal methods to epistemology

View mathematics as a communal practice designed to meet fundamental constraints:

- scientific utility
- cognitive efficiency
- communicability
- reliability
- stability

The best justification for mathematics is that it serves its purposes well.

We need to better understand how and why.