# Mathematical Simplicity

Jeremy Avigad

Department of Philosophy and Department of Mathematical Sciences
Carnegie Mellon University
(currently visiting the INRIA-MSR Joint Research Centre in Orsay)

Jnanuary, 2010

## The new epistemology of mathematics

Since Plato, the philosophy of mathematics has been concerned with:

- the nature of mathematical objects, and
- the appropriate justification for mathematical knowledge claims.

But we employ other normative judgments as well:

- some theorems are interesting
- some questions are natural
- some concepts are fruitful, or powerful
- some proofs provide better explanations than others
- some historical developments are important
- some observations are insightful

. . . and so on.

## Simplicity

One such value is simplicity: mathematical developments are often valued for providing shorter proofs, easier calculations, or streamlined solutions to problems.

Note that, in this sense, mathematical simplicity need not coincide with notions of simplicity in the philosophy of science.

It seems to have more to do with streamlining our thought processes and modes of expression than simplifying models and reducing the number of parameters.

## Outline

- Why a theory of mathematical simplicity is important.
- Some case studies from number theory
- What a theory of mathematical simplicity should look like.

## Worries about the new epistemology

Tappenden:

> *Judgements of "naturalness" and the like are* reasoned*. It is not just that some brute aesthetic response or sudden, irrational "aha!" reaction that brings about that judgement that — for example — "the scheme is the more natural setting for many geometric arguments" . . . Quite the contrary: elaborate reasons can be and are given for and against these choices. One job facing the methodologist of mathematics is to better understand the variety of reasons that can be given, and how such reasons inform mathematical practice.*
>
> *The factual observation should be beyond controversy: reasoned judgements about the "proper" way to represent and prove a theorem inform mathematical practice. I have found that more contention is generated by the disciplinary classification of the study of these judgements and the principles informing them: is this* philosophy*, or something else, like cognitive psychology?*

## Worries about the new epistemology

Arana:

> *But why should we expect that the right definitions will be uniform, rather than having lots of case distinctions? It would be nice if that were so, but wishing doesn't make it so, unless what makes a definition right is that it's the one we want. Correct definition as wish-fulfillment: if this were Dedekind's view, he would have been a relativist.*

Also:

> *This is surely an important labor-saving technique. But why should we think that this technique leads to the right definitions for a domain? There would have to be something "inevitable" about these properties if the technique were to avoid being another type of relativism.*

## Worries about the new epistemology

A bundle of anxieties:

- Merely psychological?
- Merely pragmatic?
- Subjective?
- Historical relativism.

Senses of "getting a definition right":

- presuppositions (existence assumptions) justified
- captures intuitions
- good model of empirical phenomena
- useful, *appropriate* to the theory

## A way out

Distinguish two sorts of foundational concerns:

- Ontological / metaphysical concerns: do the objects referred to really exist, and have the properties we think they have? Or: are we justified in referring to such objects, and attributing to them the properties that we do?
- Methodological concerns: are our mathematical definitions fruitful, useful, and well-suited to our mathematical goals?

The philosophy of mathematics has traditionally focused on the first (and has run out of steam).

In fact, there is a lot of progress that can be made on the second question; and that provides the best means for addressing the first.

## A way out

Two senses of "objective":

- Criteria are described in such a way that we can engage in rational debate and come to agreement over whether they are met.
- Criteria are forced upon any rational being (or, at least, any rational being with cognitive capacities roughly like ours).

The second seems to be Tappanden's and Arana's primary concern.

A theory of mathematical ease / difficulty or complexity / simplicity would help address the second.

## Case studies

Let's look at three brief examples:

- Sums of squares.
- Euler's theorem.
- Composition of quadratic forms.

## Sums of squares

**Theorem.** Suppose $x$ and $y$ can each be written as a sum of two integer squares. Then so can $xy$.

**Proof.** Write $x = a^2 + b^2$ and $y = c^2 + d^2$. Then

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

This seems to have been known to Diophantus.

## Sums of squares

Another proof: if $\alpha = a + bi$ is a Gaussian integer, write $\overline{\alpha} = a - bi$ for the complex conjugate of $\alpha$, and write $N(\alpha) = \alpha\overline{\alpha} = a^2 + b^2$.

Conjugation is an automorphism, so

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

**Proof 2.** If $x = N(\alpha)$ and $y = N(\beta)$, then $xy = N(\alpha\beta)$.

Is this simpler? Let's hold off on that discussion...

## Euler's theorem

Let $n \geq 2$, and let $\phi(n) = |\{i \mid 0 \leq i < n, i \text{ is coprime to } n\}|$.

**Theorem.** If $a$ is coprime to $n$, then $a^{\phi(n)} \equiv 1 \bmod n$.

For example, $\phi(14) = |\{1, 3, 5, 9, 11, 13\}| = 6$ and the theorem predicts $3^6 \equiv 1 \bmod 14$.

**Proof.** $\mathbb{Z}/n\mathbb{Z} = \{\overline{0}, \ldots, \overline{n-1}\}$ is a ring.

The set $\{\overline{i} \mid 0 \leq i < n, i \text{ is coprime to } n\}$ are exactly the units of that ring, which form a group.

Lagrange's theorem implies that if $G$ is any finite group and $g \in G$, then $g^{|G|} = 1$.

## Euler's theorem

This is "essentially" Euler's proof. (Full theorem: E271 on Euler archive; special case where $n$ is prime is E262, and has been translated.)

Differences:

- Euler doesn't have the concept of a ring, or the units of a ring.
- He doesn't have the concept of a group.
- He doesn't use notation for congruence.
- In proving (an instance of) Lagrange's theorem, he doesn't have use notion of a "coset" and operations on cosets.
- Rather than say "by induction," he repeats the inductive step twice and says "and so on."

## Composition of quadratic forms

Back to representations by binary quadratic forms. . .

Recall that the product of two integers of the form $x^2 + y^2$ is again of that form, and an odd prime is represented by that form if and only if $p$ is congruent to 1 modulo 4.

With arbitrary binary quadratic forms $ax^2 + bxy + cy^2$, the situation is much more complicated.

For example, Fermat conjectured that the product of two primes each congruent to 3 or 7 modulo 20 is of the form $x^2 + 5y^2$.

Legendre showed that the product of two numbers of the form $2x^2 + 2xy + 3y^2$ is of the form $x^2 + 5y^2$.

## Composition of quadratic forms

In *Disquisitiones Arithmeticae*, Gauss defined an equivalence relation on forms, and a notion of "composition" of forms. He showed, among other things, that composition is associative.

Goldman: "Gauss' proof . . . is difficult to follow."

Stillwell: "The proof is monstrous."

Edwards: "In a *tour de force* of algebraic manipulation. . . "

In the middle of the 19th century, Dirichlet simplified the analysis by working in terms of the roots of the associated quadratic polynomial.

By the end of the 19th century, Dedekind had solidified the relationship between quadratic forms and the class group of an associated number field.

## Observations

- Simplicity judgments are often contextual: a simple proof or calculation can rely on an elaborate background theory.
- Often the general utility of the background theory means that we are not required to charge the complexity against the individual application. (In accounting terms: these are capital expenditures, rather than operating expenses.)
- But theoretical infrastructure can even be useful for a single application, in that it helpfully manages the data in front of us and reduces the amount of detail we need to pay attention to at any given point.
- Sometimes theoretical expansions are ontologically or methodologically dubious; in that case, ontological or methodological qualms are balanced against the gains.

Regarding the last: in our examples, consider the uses of complex numbers, and the treatments of cosets as objects.

## Observations

Suppose we come up with objective criteria of simplicity that are plausibly correspond our cognitive capacities.

In methodologically uncontroversial situations, aren't we justified in formulating definitions that simplify the our tasks?

## Observations

In methodologically controversial cases, the story is more elaborate:

- Expansions are met with caution and concern.
- Sometimes the expansions can be explained away, in terms of the more conservative theory (e.g. complex numbers as ordered pairs, or translating algebraic proofs back to geometric proofs).
- Otherwise, sometimes the expansions can at least be explained away in particular instances (e.g. nonconstructive proofs can be constructivized), and, moreover, come with clear rules of use.
- Over time, the expansions become more than useful shorthands, and the nonconservative aspects don't seem to cause problems.

## Observations

(Cf.: the introduction of negative numbers, algebraic methods in geometry, infinitesimals in the calculus, points at infinity, abstract algebraic arguments, ideals, cosets, equivalence classes, nonconstructive definitions, infinitary objects, and so on.)

Question (for Andy, Jamie): as far as justification, what more do you want?

## A unifying theme

Math is hard.

We look for concepts and methods that make it *easier* to do what we want to do: solve problems, prove theorems, build theories, . . .

Good mathematics makes it easier for us to pursue our mathematical goals.

This doesn't dispell concerns about contextualism and relativism: evaluations are relative to goals.

But it does provide a perspective that can help unify and explain methodological criteria.

## Measuring complexity

Some measures on offer:

- Computer science: algorithmic complexity
- Logic: quantifier complexity, length of proof, etc.
- Cognitive science, psychology: timing tasks, etc.

## Length of proof

**Conservative extensions:** Let $T_2$ be an extension of $T_1$. $T_2$ is *conservative over* $T_1$ (with respect to $\Gamma$) if whenver $T_2$ proves $\varphi$ then $T_1$ also proves $\varphi$ (for $\varphi$ in $\Gamma$).

**Speedup phenomena:** It is often the case that $T_2$ has polynomial-size proofs of a sequence of statements $\varphi_1, \varphi_2, \varphi_3, \ldots$ for which the shortest proofs in $T_1$ are much longer (say, iterated exponential length).

Examples (Solovay): $ACA_0$ over $PA$, Gödel-Bernays set theory over Zermelo-Fraenkel.

## Length of proof

**Example 1:** Let $\varphi_i$ say "$T_1$ can't prove me with less than $f(i)$ symbols," and let $T_2$ be $T_1 + \{\varphi_i\}$.

**Example 2:** Let $\varphi_i$ say "a big chunk of $T_1$ is consistent." (E.g. $ACA_0$ over $PA$).

**Example 3:** Let $\varphi_i$ be a combinatorial statement encoding the above. (E.g. instances of the Paris-Harrington statement, Kruskal's theorem. Striking bounds due to Friedman.)

## Length of proof

This isn't quite the right sort of thing:

- They are asymptotic comparisons.
- They apply to formal axiomatic derivations, which don't model higher-level aspects of proof and background knowledge.
- They rely on extreme (cooked) cases.
- They are overly dramatic.

Similarly, computational complexity and psychological models are not quite right, either.

But that does not mean that they are entirely wrong: clearly length of proofs, complexity of tasks, and our cognitive abilities have something to do with it.

These provide a good starting point.

## Summary

We want to understand a sense in which common normative judgments have an "objective" component.

A more dynamic understanding of "mathematical knowledge," and the way that various methodological strategies simplify the tasks before us, can help us make sense of common normative assessments.

Suitable measures of complexity / simplicity then provide objective normative criteria.

Then challenge is to characterize mathematical goals and our cognitive constraints at the right level of idealization, and develop robust and compelling accounts of mathematical simplicity.