Category theory and its foundations: the role of diagrams and other "intuitive" material

Ralf Krömer

Arbeitsgruppe Didaktik und Geschichte der Mathematik
Bergische Universität Wuppertal

June 24, 2018

We're all former children

- My question in TO was not
  (1) "how to present something to people without expertise in a certain area?",

  but rather

  (2) "people *with* expertise in a certain area (the experts, for short) once started from "a few motivating examples"; which role do these still play in their mature thinking?"
- My only data: written utterings (and some interviews).
- More particularly, I am convinced that question (2) is related to the ways the experts justify the use of certain means, the consideration of certain objects, the relying on certain assumptions not provable to be consistent with something more generally accepted.
- The answers given to question (1) by authors writing introductory material in the field are relevant for a treatment of question (2).

Manders on Euclid

- [Man08a, Man08b]
- diagram vs. discursive text (equational information restricted to the discursive part)
- diagrams responsible for gaps etc. ("crisis of intuition"); diagram control
- "exact" and "co-exact" properties
- the problem of reductio proofs
- CT: diagrams display propositions, express equational information

Eilenberg-Steenrod on diagrams

"The diagrams incorporate a large amount of information. Their use provides extensive savings in space and in mental effort. In the case of many theorems, the setting up of the correct diagram is the major part of the proof. We therefore urge that the reader stop at the end of each theorem and attempt to construct for himself the relevant diagram before examining the one which is given in the text. Once this is done, the subsequent demonstration can be followed more readily; in fact, the reader can usually supply it himself." [ES52, xi]

"Intuition"

"Our intuition tells us that whenever two categories exist in our world, then so does the corresponding category of all natural transformations between the functors from the first category to the second" [Law66, 9].

- However, if one tries to reduce categorial constructions to set theory, one faces some serious problems in the case of a category of functors.
- Lawvere when relying on "intuition" stresses that those working with categorial concepts *despite* these problems have the feeling that the envisaged construction is clear, meaningful and legitimate.
- (And he is not exactly in the situation of Frege's full comprehension axiom. He just speaks about one particular application of comprehension—and no one denies, after all, that we can form *some* extensions of predicates, right?)

- In the tradition of philosophy, "intuition" means immediate, *i.e.*, not conceptually mediated cognition. In particular, the term is used in the context of validity (immediate insight in the truth of a proposition)
- And there is the use in the sensual context (geometrical intuition, the German *Anschauung*).
- Thus, I distinguish between *sensual intuition* and *validity intuition*.

Maddy: two types of justifications

Maddy distinguishes *intrinsic* and *extrinsic* justifications:

> *"The suggestion is that the axioms of ZFC follow directly from the concept of set, that they are somehow "intrinsic" to it (obvious, self-evident), while other axiom candidates are only supported by weaker, "extrinsic" (pragmatic, heuristic) justifications, stated in terms of their consequences, or intertheoretic connections, or explanatory power, for example." [Mad88a] p.482f*

Intrinsic justifications: the example of pairing

"[. . .] we acquire our most primitive physical and set-theoretic beliefs when we learn to perceive individual objects and sets of these. We come to believe, for example, that objects do not disappear when we are not looking at them, and that the number of objects in a set does not change when we move the objects around. [. . .] The simplest axioms of set theory, like Pairing, have their source in this sort of intuition. If they are not strictly part of the concept (whatever that comes to), they are acquired along with the concept. Given its origin in prelinguistic experience, the best indication of intuitiveness is when a claim strikes us as obvious." [Mad88b] p.758f

Extrinsic justifications

There are many types of extrinsic justification, according to [Mad88b] p.758f:

(1) confirmation by instances (the implication of known lower-level results)
(2) prediction (the implication of previously unknown lower level results)
(3) providing new proofs of old theorems
(4) unifying new results with old, so that the old results become special cases of the new
(5) extending patterns begun in weaker theories
(6) providing powerful new ways of solving old problems
(7) providing proofs of statements previously conjectured
(8) filling a gap in a previously conjectured "false, but natural proof"
(9) explanatory power
(10) intertheoretic connections

- example: axiom of choice (see Zermelo 1908)
- In recent work, she makes "the heretical suggestion that in fact intrinsic justifications are secondary to the extrinsic" [Mad11] p.134.
- She draws this conclusion from (actual or hypothetical) history of how new concepts have been introduced in group theory or set theory and subsumes:

"What's striking is that all these perfectly reasonable ways of proceeding are in fact grounded in their promise of leading to the realization of more of our mathematical goals, to the discovery of more fruitful concepts and theories, to the production of more deep mathematics. [. . .] intrinsic considerations are valuable, but only insofar as they correlate with these extrinsic payoffs. This suggests that the importance of intrinsic considerations is merely instrumental, that the justificatory force is all extrinsic" [Mad11] p.136.

Comparison Maddy-RK

- focus in the analysis of justification strategies: why does the fact that the axiom (or more generally, something to be justified) can play one of the roles (1)-(10) make it more convincing?
- To label a proposition as "obvious" or "self-evident" is done by relying on a capacity called intuition. Along with "our most primitive physical and set-theoretic beliefs" stemming from "prelinguistic experience" mentioned by Maddy, many much less primitive beliefs are labelled intuitive by experts (Lawvere).
- How do experts develop such an intuition? I think: it is developed during exactly such specific manipulations of theories like those yielding results of the types (1)-(10).
- I do not see why to privilege prelinguistic experience with respect to expert knowledge.
- Put differently, in my view it depends on the context in which a concept is used whether something counts as intrinsic or extrinsic to the concept.

Topos-theoretic foundations and Feferman's objection

Feferman quotes Mac Lane (personal communication):

> *"mathematicians are well known to have very different intuitions, and these may be strongly affected by training" [Fef77, 153]*

Feferman replies:

> *"I believe our experience demonstrates [the] psychological priority [of the general concepts of operation and collection with respect to structural notions such as 'group', 'category' etc.]. I realize that workers in category theory are so at home in their subject that they find it more natural to think in categorical rather than set-theoretical terms, but I would liken this to not needing to hear, once one has learned to compose music." [ibid.]*

The consistency of ZFC

- undecidable (Gödel)
- In particular, ZFC could be inconsistent
- but this could only be proved by finding one day a contradiction.
- Bourbaki: trust in the consistency of ZFC on “empirical” grounds (no contradiction discovered so far)
- Bourbaki seems to suggest a certain way in which contradictions are met with in the everyday work of mathematicians. Is this truly the usual way they have been met with historically? [Krö12]

"As far as sets occur and are necessary in mathematics (at least in the mathematics of today, including all of Cantor's set theory), they are sets of integers, or of rational numbers (. . .), or of real numbers (. . .), or of functions of real numbers (. . .), etc.; when theorems about all sets (or the existence of sets) in general are asserted, they can always be interpreted without any difficulty to mean that they hold for sets of integers as well as for sets of real numbers, etc. (. . .). This concept of set, however, according to which a set is anything obtainable from the integers (or some other well defined objects) by iterated application of the operation "set of", and not something obtained by dividing the totality of all existing things into two categories, has never led to any antinomy whatsoever; that is, the perfectly "naive" and uncritical working with this concept of set has so far proved completely self-consistent." [Göd47, p. 518f]

- CT proved to be extremely fruitful in various domains of mathematics
- (in Maddy's terms, this fact constitutes already an extrinsic justification of CT)
- in the same time, the problematic constructions don't occur like set-theoretic paradoxes but seem to be "intuitive"

"The restrictions employed [Grothendieck universes or NBG] seem mathematically unnatural and irrelevant. Though bordering on the territory of the paradoxes, it is felt that the notions and constructions [as the category of all structures of a given kind or the category of all functors between two categories] have evolved naturally from ordinary mathematics and do not have the contrived look of the paradoxes." [Fef77] p.155

"The well-known fact that some basic constructions applied to large categories take us out of the universe seems to me to indicate that the constructions are not yet properly presented. The discovery of proper presentations is too difficult, though, for all work on these constructions to wait for it." [Isb66] p.620

This quotation clearly shows that in this debate, justification by nondemonstrative arguments replaced temporarily the solution of a corresponding mathematical problem which had to be postponed, namely the problem "is there an axiomatic set theory allowing for unlimited category theory?"

Intended models vs. formalism

- My central focus: the relation between formal definitions and intended uses of mathematical concepts.
- In the case of non-standard models of formal systems, the formal definition is "overcomprehensive": it has more models than just the intended one(s).
- According to Kreisel, we have the capacity to cope with this situation:

"Many formal independence proofs consist in the construction of models which we recognize *to be different from the intended notion. It is a fact of experience that one can be honest about such matters! When we are shown a 'non-standard' model we can honestly say that it was not intended. [. . .] If it so happens that the intended notion is not formally definable this may be a useful thing to know about the notion, but it does not cast doubt on its objectivity [. . .]." [Kre70, 25]*

- In the case of category theory, the situation is complementary: usual formalizations are too restrictive to capture the naive theory.
- It seems at first glance that this situation provokes but another instance of the capacity stressed by Kreisel, see Feferman (we feel that something intended is not captured, so the formalization again is not adequate)
- (But we could very well be in the situation of Frege's full comprehension: something seems to be quite obvious, but then the counterexample comes around the corner!)

Unlimited category theory

Feferman in 1977 suggested the following requirements and in [Fef13] gave a foundational system (set-theoretical in nature and based on Quine's "New foundations") meeting them "nearly" (p.9):

(R1) Form the category of all structures of a given kind, e.g. the category $\mathbf{Grp}$ of all groups, $\mathbf{Top}$ of all topological spaces, and $\mathbf{Cat}$ of all categories.

(R2) Form the category B^A of all functors from A to B, where A, B are any two categories.

(R3) Establish the existence of the natural numbers N , and carry out familiar operations on objects $a, b, \ldots$ and collections $A, B, \ldots$, including the formation of $\{a, b\}$, (a, b), $A \cup B$, $A \cap B$, $A - B$, $A \times B$, B^A, $\bigcup A$, $\bigcap A$, $\prod B_x[x \in A]$, etc.

(R1)-(R3) spell out what workers in the field *intend* a foundation to provide

Ernst's result

- M. Ernst [Ern15] showed that "unlimited category theory" in the sense of [Fef13] is inconsistent, i.e. that presupposing (R1)-(R3) leads to a contradiction.
- Ernst's strategy: build a proof in analogy to the proof that there can be no set of all sets using Cantor's theorem.
- Cantor's theorem says that there is no surjection between any set and its powerset; established by a diagonal argument
- In Ernst's proof (relying on Lawvere's 1969 work on transporting diagonal arguments to the language of category theory) the set-theoretical notion of "surjection" is replaced by a notion of onto mapping in a category-theoretic sense, and "set of all sets" and "powerset" by corresponding objects in a certain category of graphs.

- The first important move: this category (like every category) can be considered as a graph and in this sense contains "itself" as an object
- Secondly, the category has exponentials, and with the help of the exponential of the object in question (and quite some technical lemmas), we get a contradiction (a monic arrow which exists and doesn't).
- (Note that this is not in parallel with the Russell antinomy, but with Cantor's theorem.)
- *"as soon as you satisfy these axioms you have a category. Such a claim is so obvious as to appear tautologous. If the axioms of category theory are satisfied then* of course *you have a category."* (p.308)
- Ernst's proof does not generalize to other candidates for contradiction (neither to **Set** not to **Cat**; p.319)
- The way Ernst actually arrived at the contradiction underlines my criticism concerning the Bourbaki viewpoint

- I ignore if there is already a final agreement of the community whether Ernst's proof is completely cogent. The details of the proof are quite tedious, and at least at some places I have personally the impression that the question might be begged altogether
- One should ask: on which grounds is Lawvere's theorem from [Law69] proved?
- If the proof is right
  - the discrepancy between what category theorists believe to need for their work and what set theory can furnish would stay put;
  - Ernst's result would enfeeble an extrinsic argument against ZFC: *"Finally, this removed some of the objections raised against various foundations, ZFC in particular. Since unlimited category theory cannot be founded, a failure to found it no longer represents a valid objection to ZFC or any other foundation similarly limited."* (p.320)
- if it is wrong
  - the search for a set-theoretic foundation of unlimited category theory would go on;
  - an analysis of flaws in the proof might lead to new extrinsic justifications of unlimited category theory.

A suggestion of mine from 2007 dismissed?

"there may be a shortcoming in the usual claims about the set-theoretical illegitimacy (or inconsistency) of **Cat***. For* **Cat** *contains "itself" not as the entire complicated building of points, arrows and labels (its inside) but as a single point connected to certain arrows (the functors between other categories and this category). Thus, to speak about "self-containing" here seems quite simplifying. This is naturally no proof for the claim that a category of all categories is consistent but a remedy to the usual arguments in favor of its illegitimacy." [Krö07] p.280*

- But Ernst takes this difference into account. The situation is the following:
  - there is a problem with self-containing pointed out by the standard proof that there can be no set of all sets using Cantor's theorem;
  - Ernst shows that this problem stays put when the category-theoretic equivalents of all set-theoretical notions intervening in this proof are used (when the category-theoretic equivalent of self-containing is used, so to say)
- (My idea was too naive with respect to the notion of cardinality, after all.)
- New idea: "world" vs. "language"

Is my epistemological investigation now pointless?

- Ernst's result at first glance seems to show that the debate on CT is ultimately only of historical interest: people when introducing the constructions felt justified to do so, but finally the feeling turned out to be misleading. (And one might wonder why they had it.)
- I maintain that the case is still of philosophical interest. I (still and even more) think that
  - thinking in terms of set membership is not the only form of thinking leading to mathematical knowledge
  - the debate on category theory and foundations should make us think about what we expect a foundation of mathematics to accomplish
  - Application of concepts to extensions, or of "languages" to "worlds", still constitutes a basic problem of phil of math.

Thank you!

[Ern15] Michael Ernst.
The prospects of unlimited category theory: doing what remains to be done.
The review of symbolic logic, 8(2):306–327, 2015.

[ES52] Samuel Eilenberg and Norman E. Steenrod.
Foundations of algebraic topology.
Princeton University Press, 1952.

[Fef77] Solomon Feferman.
Categorical foundations and foundations of category theory.
In *Butts, Robert E.; Hintikka, Jaakko (eds.): Logic, Foundations of Mathematics and Computability theory*, pages 149–169. Reidel, Dordrecht, 1977.

[Fef13] S. Feferman.
Foundations of unlimited category theory: what remains to be done.

The Review of symbolic logic, 6(1):6–15, 2013.

[Göd47] Kurt Gödel.
What is Cantor's Continuum Problem?
American Mathematical Monthly, 54:515–525, 1947.

[Isb66] J.R. Isbell.
Structure of categories.
Bull. Amer. Math. Soc., 72:619–655, 1966.
MR34 #5896.

[Kre70] Georg Kreisel.
The formalist–positivist doctrine of mathematical precision in the light of experience.
L'Age de la science, 3(1):17–46, 1970.

[Krö07] Ralf Krömer.
Tool and object. A history and philosophy of category theory, volume 32 of *Science Network Historical Studies*.

Birkhäuser, Basel, 2007.

[Krö12] Ralf Krömer.
Are we still babylonians? the structure of the foundations of mathematics from a wimsattian perspective.
In L. Soler, E. Trizio, Th. Nickles, and W. Wimsatt, editors, *Characterizing the Robustness of the Sciences After the Practical Turn of Philosophy of Science*, volume 292 of *Boston Studies in the Philosophy of Science*, pages 189–206. Springer, 2012.

[Law66] F. William Lawvere.
The category of categories as a foundation of mathematics.
In *Proc. Conf. Categorical Algebra (La Jolla, Calif., 1965)*, pages 1–21. Springer, 1966.
MR34 #7332.

[Law69] F. W. Lawvere.
Diagonal arguments and cartesian closed categories.
In P. J. Hilton, editor, *Category theory, Homology theory and their applications II*, volume 92 of *Lecture Notes in Mathematics*, pages 134–145. Springer, 1969.

[Mad88a] Penelope Maddy.
Believing the axioms. i.
The Journal of Symbolic Logic, 53(02):481–511, 1988.

[Mad88b] Penelope Maddy.
Believing the axioms. ii.
The Journal of Symbolic Logic, 53(02):736–764, 1988.

[Mad11] Penelope Maddy.

Defending the axioms. On the philosophical foundations of set theory.
Oxford University Press, 2011.

[Man08a] Kenneth L. Manders.
Diagram-based geometric practice.
In P. Mancosu, editor, *The Philosophy of Mathematical Practice*, chapter 3, pages 65–79.
Oxford Univ. Press, 2008.

[Man08b] Kenneth L. Manders.
The euclidean diagram (1995).
In P. Mancosu, editor, *The Philosophy of Mathematical Practice*, chapter 4, pages 80–133.
Oxford Univ. Press, 2008.