Exploring strange new worlds of mahtenatics

Andrej Bauer
University of Ljubljana

Levi L. Conant Lecture Series
Worcester Polytechnic Institute
February 10, 2023

Thank you for the invitation. I am deeply honored to have the opportunity to deliver a Levi L. Conant talk.

BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 54, Number 3, July 2017, Pages 481–498
http://dx.doi.org/10.1090/bull/1556
Article electronically published on October 3, 2016

FIVE STAGES OF ACCEPTING CONSTRUCTIVE MATHEMATICS

ANDREJ BAUER

Abstract. On the odd day, a mathematician might wonder what constructive mathematics is all about. They may have heard arguments in favor of constructivism but are not at all convinced by them, and in any case they may care little about philosophy. A typical introductory text about constructivism spends a great deal of time explaining the principles and contains only trivial mathematics, while advanced constructive texts are impenetrable, like all unfamiliar mathematics. How then can a mathematician find out what constructive mathematics feels like? What new and relevant ideas does constructive mathematics have to offer, if any? I shall attempt to answer these questions.

The “5 stages” paper that got me where I am standing today was about mathematics done without the law of excluded middle – known as constructive mathematics. But the main point was not to present an alternative to classical mathematics.

We initially set out to understand the difference between the classical and the constructive world of mathematics, only to have discovered that there are not two but many worlds, some of which simply cannot be discounted as logicians' contrivances.

Excluded middle as the dividing line between the worlds is immaterial in comparison with having Cantor's paradise shattered into an unbearable plurality of mathematical universes.

The "depression stage" ends with the realization that things are much worse than there being just two alternatives, the classical and the intuitionistic. There is an entire mathematical multiverse. In today's talk I would like to tell you more about it. How does one accept the fact that there are many worlds of mathematics?

Talk outline

Worlds of mathematics

The real numbers

Conclusion

Here is the talk outline. I will first discuss what constitutes a “world of mathematics”, without dwelling into too many technical details. There are several alternatives, and it does not really matter which one we adopt in this talk.

Next, we shall look at the real numbers, a structure that is central to modern mathematics. Each world has its own version of the reals, and while these share many common features, they also differ quite a bit – also in ways that you might find a bit disoncerting.

In the end, we shall ask what good, if any, may come out of thee being many kinds of mathematics.

What is a world of mathematics?

Let us first address the question: what is a world of mathematics?

It is an environment, a system, a framework, which allows mathematicians to carry out the work that they do, using the language & techniques that they are used to. There may of course be some question on how large a chunk of mathematics can be incorporated in a particular case, but at the very least we expect there to be basic algebra, analysis, geometry and topology.

What is a world of mathematics?

A mathematical structure rich enough to support constructions and proofs that mathematicians carry out in their daily work.

Topos

- Ω object of truth values
  - $\bot \ \top \ \wedge \ \vee \ \Rightarrow \ \forall \ \exists \ \neg \ =$
- Constructions
  - $A \times B \quad A + B \quad A \to B \quad \{x \in A \mid \phi\} \quad A/{\approx}$
- $\mathbb{N}$ natural numbers object

For the purposes for this talk, we shall take the worlds of mathematics to be toposes.
These are categories with rich enough structure to interpret intuitionistic higher-order logic, as well as the common constructions used in mathematics: products, sums, function spaces, subsets, quotients, etc. And of course, there are the natural numbers, too.

Even though a category comprises objects, we shall refer to them as “sets” – out of habit, but also because toposes really can be seen as models of a certain kind of (intuitionistic) set theory.

The real numbers

To get a feel for how one works with toposes, let us review the construction of real numbers.

Integers & rational numbers

- $\mathbb{Z} = (\mathbb{N} \times \mathbb{N})/\sim$ where

  $(a,b) \sim (c,d) \Leftrightarrow a + d = c + b$

- $\mathbb{Q} = (\mathbb{Z} \times (\mathbb{N} \setminus \{0\}))/\approx$ where

  $(u,v) \approx (s,t) \Leftrightarrow u \cdot t = s \cdot v$

A topos postulates the natural numbers, from which we can construct the integers and the rationals in the usual way.

The integers are equivalence classes of pairs of natural numbers. Think of the pair (a,b) as representing the difference between a and b.

Similarly, the rational numbers are equivalence classes of pairs. These are just fractions, of course.

The Dedekind reals

$\mathbb{R} = \{ (L,U) \in P(\mathbb{Q}) \times P(\mathbb{Q}) \mid \mathrm{cut}(L,U) \}$

where cut(L,U) means:

- rounded & bounded:
  - $\forall q \in \mathbb{Q} . (q \in L \Leftrightarrow \exists r \in \mathbb{Q} . q < r \wedge r \in L)$
  - $\forall r \in \mathbb{Q} . (r \in U \Leftrightarrow \exists q \in \mathbb{Q} . q < r \wedge q \in U)$
- disjoint: $\forall q, r \in \mathbb{Q} . (q \in L \wedge r \in U \Rightarrow q < r)$
- located: $\forall q, r \in \mathbb{Q} . (q < r \Rightarrow q \in L \vee r \in U)$

There are several constructions of real numbers. Classically they are all equivalent, but not constructively. We shall take Dedekinds definition in terms of cuts.

The definition fits on a slide. We need not go through all the details, it’s better to show a picture. We are taking **two-sided** cuts. The lower and the upper cut are sets of rationals, think of them as all the lower and all the upper rational approximations of the real determined by L and U.

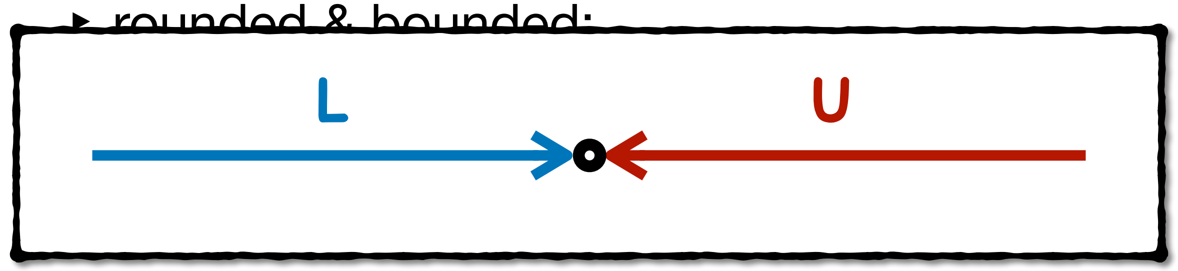
A note: sometimes the reals are constructed as one-sided cuts, but that creates unecessary asymmetry, and one has to take careful with the definition so that the other cut can be recovered without excluded middle.

The Dedekind reals

$\mathbb{R} = \{ (L,U) \in P(\mathbb{Q}) \times P(\mathbb{Q}) \mid \mathrm{cut}(L,U) \}$

where cut(L,U) means:

- rounded & bounded:



- disjoint: $\forall\, q, r \in \mathbb{Q}\ .\ (q \in L \wedge r \in U \Rightarrow q < r)$
- located: $\forall\, q, r \in \mathbb{Q}\ .\ (q < r \Rightarrow q \in L \vee r \in U)$

The Dedekind reals

$\mathbb{R} = \{ (L,U) \in P(\mathbb{Q}) \times P(\mathbb{Q}) \mid \mathrm{cut}(L,U) \}$

where cut(L,U) means:

- rounded & bounded:
  - $\forall q \in \mathbb{Q} . (q \in L \Leftrightarrow \exists r \in \mathbb{Q} . q < r \wedge r \in L)$
  - $\forall r \in \mathbb{Q} . (r \in U \Leftrightarrow \exists q \in \mathbb{Q} . q < r \wedge q \in U)$
- disjoint: $\forall q, r \in \mathbb{Q} . (q \in L \wedge r \in U \Rightarrow q < r)$
- located: $\forall q, r \in \mathbb{Q} . (q < r \Rightarrow q \in L \vee r \in U)$

The field of reals

- $\mathbb{R}$ is an archimedean complete ordered field.
- The following **cannot** be shown:
  - $\forall\, x, y \in \mathbb{R}\ .\ (x \leq y \lor y < x)$
  - "Every real has a digit expansion."
  - $\mathbb{R} \cong P(\mathbb{N})$ and $\mathbb{R} \cong \{0,1\}^{\mathbb{N}}$

In constructive mathematics the reals carry the usual algebraic structure of an archimedean ordered field.

However, there are also some surprises. For instance, the order is not decidable (but an approximate version that inserts ε-slack is).

We cannot show that every real has a digit expansion, but we can show that there are no reals without digit expansion (yes, it's confusing).

Also, you should stop thinking that the reals are "the same thing" as the powerset of $\mathbb{N}$, or the set of binary sequences.

The effective topos

- All objects have associated Gödel codes (realizers).
- Morphisms are the computable maps.
- A statement is valid if witnessed by a computable map.
- Example: a realizer for $x \in \mathbb{R}$ is (a number encoding) a Turing machine that computes approximations of x.
- Thus only computable reals exist in the effective topos.

We are ready to explore our first strange wolrd of mathematics – the effective topos.

Unfortunately, an hour's lecture is not enough to give a detailed description of the topos. At best I can try to convey rought ideas. In the effective topos every fact, construction and entity must be computed by a Turing machine. For example, a Dedekind real in this topos is a **computable** real: one which has a Turing machine computing abibtrarily good rational approximations.

Beware however that the slogan "everything is computable" can be misleading in some situations.

Axiom of Enumerability:
There are countably many countable subsets of $\mathbb{N}$.

The effective topos validates some strange statements. For instance, there are countably many countable subsets of $\mathbb{N}$.

This looks contradictory at first. Isn't every subset of $\mathbb{N}$ countable, so the statement claims that the powerset $P(\mathbb{N})$ is countable, which it isn't by Cantor's diagonalization argument (which is constructive)?

Not quite. The countable subsets of $\mathbb{N}$ in the effective topos are the **computably enumerable sets**, and those can be computably enumerated.

Specker sequence: There is a sequence in [0,1] without accummulation points.

How about the real numbers in the effective topos.

An old construction of Specker's shows that [0,1] is a locally non-compact space, because there is a bounded sequence of reals without an accummulation point.

Once such a sequence x_n is obtained, it is not hard to construct an unbounded continuous map: for each $n \in \mathbb{N}$ errect a "tent" centered at x_n of height n and sufficiently narrow to avoid previous tents.

And given an unbounded continuous map $f : [0,1] \to \mathbb{R}$, the sets $U_n = \{x \in \mathbb{R} \mid f(x) < n\}$ form an open cover without a finite subcover.

Specker sequence: There is a sequence in [0,1] without accummulation points.

There is an unbounded continuous map $[0,1] \to \mathbb{R}$.

Specker sequence: There is a sequence in [0,1] without accummulation points.

There is an unbounded continuous map [0,1] → $\mathbb{R}$.

There is an open cover of [0,1] without a finite subcover.

Function realizability

- All objects are coded by sequences of numbers.
- Morphisms are the continuous maps.
- A statement is valid if it is witnessed by a continuous map.

A variation of the effective topos is the so called function realizability.
This time we code objects with infinite sequences of numbers.
Morphisms are the continuous maps, with respect to the product topology on $\mathbb{N}^{\mathbb{N}}$.

Every open cover of [0,1] has a finite subcover.

Every map $[0,1] \to \mathbb{R}$ is uniformly continuous.

This topos could arguably be called “Brouwer’s paradise”.

It validates Brouwer’s principles: [0,1] is compact and all maps are (uniformly) continuous.

But also some strange things happen. For example, $\mathbb{R}$ cannot be decomposed into two non-trivial disjoint sets (of any kind, they need not be open).

Andrew Swan and I showed another curious fact about the topos: all metric spaces in it are separable. A consequence of this is: one cannot constructively define ℓ^∞ and other non-separable metric spaces.

Every open cover of $[0,1]$ has a finite subcover.

Every map $[0,1] \to \mathbb{R}$ is uniformly continuous.

If $\mathbb{R} = A \cup B$ and $A \cap B = \varnothing$ then $\mathbb{R} = A$ or $\mathbb{R} = B$.

Every open cover of $[0,1]$ has a finite subcover.

Every map $[0,1] \to \mathbb{R}$ is uniformly continuous.

If $\mathbb{R} = A \cup B$ and $A \cap B = \varnothing$ then $\mathbb{R} = A$ or $\mathbb{R} = B$.

Every metric space is separable.

Sheaves

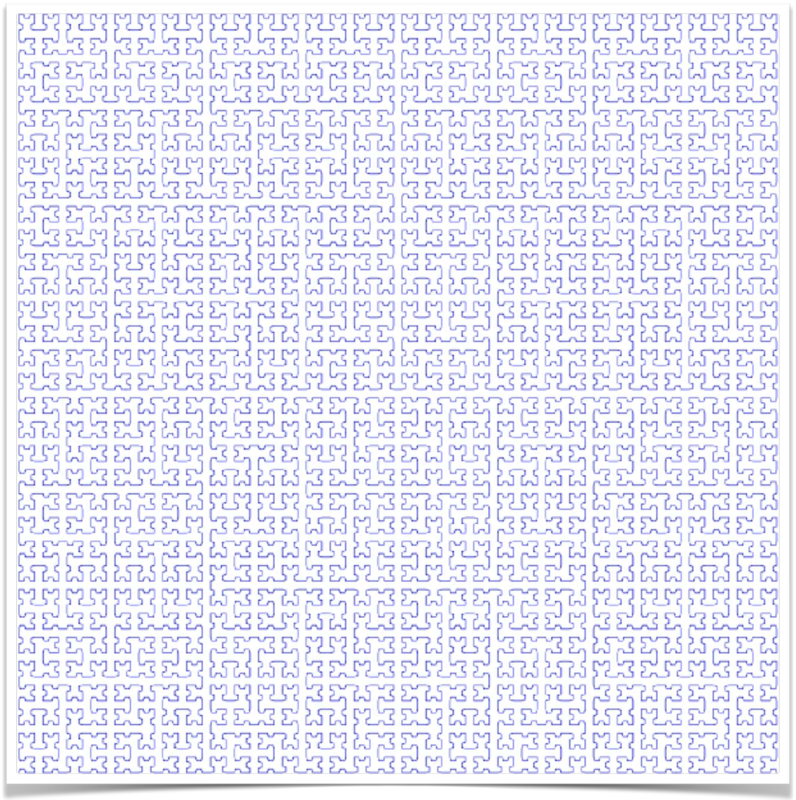
- Dubuc topos:
  Every map R → R is differentiable.
- Sh($[0,1]^2$):
  There are no space-filling curves.

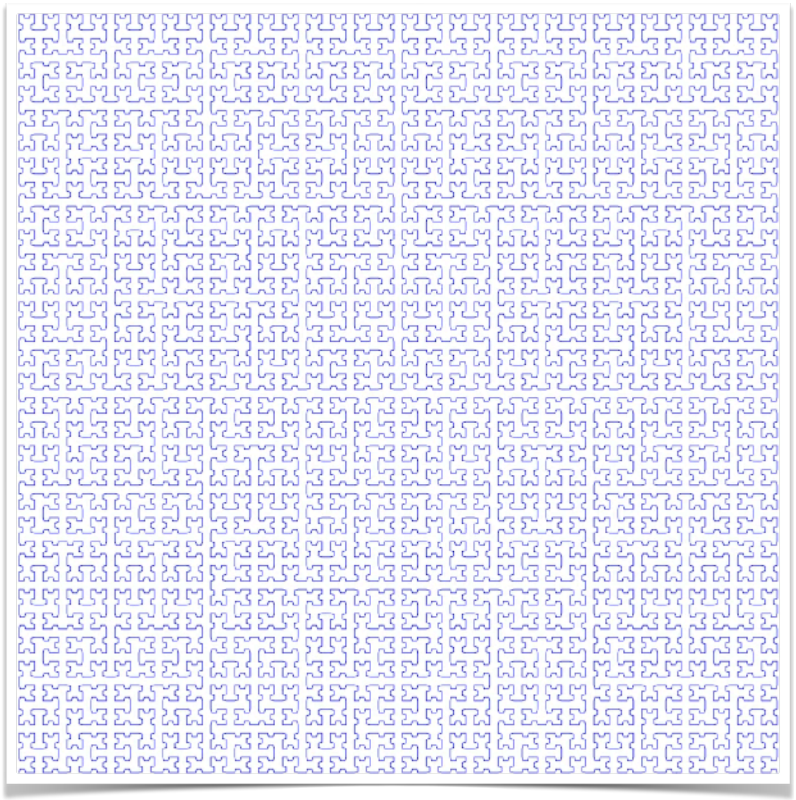
There are many other kinds of toposes, which we can visit only briefly.

An important class are toposes of sheaves. These were invented by Grothendieck for doing algebraic geometry, but have since found many uses. Let us mention two.

There are toposes with an object of **smooth reals** (not Dedekind reals) such that every map R → R is differentiable. In such toposes one can use 17th century nilpotent infinitesimals, just like physicists still do.

Were you upset when you first saw space filling curves? Well, now you can be upset again, because there is a topos in which there are no space filling curves. The usual construction produces a curve $[0,1] \to [0,1]^2$ whose complement is empty, but the curve is not surjective. (Yes, it's confusing.)





How large is $\mathbb{R}$?

Forcing the cardinality of $\mathbb{R}$.
Sub-countable reals.
Countable reals.

Making $\mathbb{R}$ large

- Cantor: there is no surjection $\mathbb{N} \to \mathbb{R}$
  (Uses excluded middle or countable choice)
- Cantor's hypothesis: $|\mathbb{R}| = \aleph_1$
- With Cohen's forcing we can make $|\mathbb{R}|$ arbitrarily large.

Set theorists have studied the cardinality of $\mathbb{R}$.

Cantor famously proved that the reals cannot be enumerated, and asked whether the cardinality of the reals is the next one after the cardinality of $\mathbb{N}$. We know today how to make classical worlds of mathematics in which the reals can have an abitrarily large cardinality.

Making ℝ small

Realizability topos over Joel Hamkin's infinite-time Turing machines:

- there is an injection $\mathbb{R} \to \mathbb{N}$
- there is no surjection $\mathbb{N} \to \mathbb{R}$

In constructive worlds we can go in the other direction, and try to make the reals small.

In the realizability topos over infinite-time Turing machine the reals embed into the natural numbers. Briefly, this is so because with infinite time it is possible to calculate a canonical code for a given infinite-time computable real. (This is not possible with ordinary Turing machines.)

Nevertheless, the reals are still uncountable: the topos validates countable choice, hence Cantor's diagonalization proves that there is no surjection from $\mathbb{N}$ onto $\mathbb{R}$.

The countable reals

- There is a topos in which $\mathbb{R}$ is countable (A. Bauer & J. Hanson)
- $\mathbb{R}$ has measure 0.
- Hilbert cube $[0,1]^\omega$ also turns out to be countable.

It has been an open question for a while whether Cantor's diagonalization can be carried out without excluded middle and without countable choice. Last year James Hanson and I gave an answer: it cannot.

We constructed a topos in which the Dedekind reals are countable. We used an idea of Joe Miller's who constructed certain sequences of reals, using a powerful fixed-point theorem from analysis, that collectively cannot be diagonalized against. To make everything work we formulated a new kind of realizability, where machines compute with oracles, but in a uniform way.

One quickly derives strange results, for example that $\mathbb{R}$ has measure zero: cover the n-th real with an interval of width $\varepsilon/2^n$.

It turns out that the Hilbert cube is also countable, from which Brouwer's fixed point theorem follows in a couple of lines, as a consequence of Lawvere's fixed point theorem.

Conclusion

We could go on for a while, but let us end here.

Denial:

“I shall forget this talk by lunchtime.”

Anger:

“So you have a PhD in logic?”

Bargaining:

“This is all very nice, but those worlds are not real.”

Depression:

“Mathematics is broken.”

Acceptance:

“How do we put the multiverse to good use?”

What significance do the strange worlds we’ve visited have for the working mathematician?

History teaches us that strange new ideas were never easily accepted. Imaginary numbers were considered, well, imaginary for a long time. Non-euclidean geometries were discovered but not presented to the public in fear of rejection. But ultimately new mathematics always finds acceptance and its place – it may just take a while.