BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
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.

From a psychological point of view, learning constructive mathematics is agonizing, for it requires one to first *unlearn* certain deeply ingrained intuitions and habits acquired during classical mathematical training. In her book *On Death and Dying* [17] psychologist Elisabeth Kübler-Ross identified five stages through which people reach acceptance of life's traumatizing events: denial, anger, bargaining, depression, and acceptance. We shall follow her path.

1. Denial

I have met mathematicians who summarily dismiss constructive mathematics as nonsense because they misunderstand what constructivism says or they just make a simple logical error. Let us deal with these sorts of misconceptions first.

1.1. **Excluded middle.**

Constructive mathematics is mathematics done without the law of *excluded middle:*

For every proposition P, either P or not P.

This is not to say that we deny the law, but rather that we may use only those instances of it which we have proved. A good first exercise is to show from Peano axioms that any two natural numbers are equal or not equal.

Excluded middle is also known as the law of *excluded third*, or *tertium non datur* in Latin. One might think mistakenly that constructivism adopts the existence of a "third" possibility:

There is a proposition Q which is neither true nor false.

Received by the editors July 1, 2016.
2010 *Mathematics Subject Classification.* Primary 03F65.
This material is based upon work supported by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF under Award No. FA9550-14-1-0096.

©2016 American Mathematical Society

This statement is constructively false because it states both $\neg Q$ (Q is not true) and its negation $\neg\neg Q$ (Q is not false), which is impossible by the following law of *noncontradiction*:

> *There is no proposition P such that P and $\neg P$.*

We prove this as follows, using the fact that in constructive mathematics $\neg P$ is an abbreviation for $P \Rightarrow \bot$:

> Suppose there were a proposition P such that both P and $\neg P$.
> Since $\neg P$ is just $P \Rightarrow \bot$, we may use *modus ponens* to conclude $\bot$.
> We have reached a contradiction, therefore there is no such P. $\square$

Incidentally, this is *not* a proof by contradiction but rather a proof of negation, as we shall explain next.

1.2. **Proof by contradiction and proof of negation.** *Proof by contradiction*, or *reductio ad absurdum* in Latin, is the reasoning principle:

> *If a proposition P is not false, then it is true.*

In symbolic form it states that $\neg\neg P \Rightarrow P$ for all propositions P, and is equivalent to excluded middle. Confusingly, mathematicians call "proof by contradiction" any argument which derives a contradiction from a statement believed to be false, but there are *two* reasoning principles that have this form. One is indeed proof by contradiction, and it goes as

> *Suppose $\neg P$, ... (argument reaching contradiction) ..., therefore P.*

While the other is how a negation $\neg P$ is proved:

> *Suppose P, ... (argument reaching contradiction) ..., therefore $\neg P$.*

Because $\neg P$ abbreviates $P \Rightarrow \bot$, the rule for proving a negation is an instance of the rule for proving an implication $P \Rightarrow Q$: assume P and derive Q. Admittedly, the two arguments look and feel similar, but notice that in one case the conclusion has a negation removed and in the other added. Unless we already believe in $\neg\neg P \Leftrightarrow P$, we *cannot* get one from the other by exchanging P and $\neg P$. These really are different reasoning principles.

Famous theorems which use proofs of negations are irrationality of $\sqrt{2}$, Cantor's diagonal argument showing that there is no surjection from a set to its powerset, and Russell's antinomy showing there is no set of all sets. So these are all constructively valid theorems. Just to make sure, let us constructively prove Russell's theorem.

Theorem 1.1 (Russell)**.** *There is no set of all sets.*

Proof. Let us follow a classic text on set theory, such as that by Halmos [11, Sect. 2] or Jech [13, 1.11]. Suppose V were the set of all sets, and consider its subset

$$R = \{x \in V \mid x \notin x\}.$$

Observe that

(1) if $R \in R$, then by the definition of R also $R \notin R$, which is absurd;
(2) if $R \notin R$, then again by the definition of R also $R \in R$, which is absurd.

These two observations are both proofs of negation, the first one proving $\neg(R \in R)$ and the second $\neg(R \notin R)$. By the law of noncontradiction this cannot be the case, therefore there is no set of all sets. $\square$

This is indeed a proof of negation: we assume that the set of all sets exists, and we derive a contradiction. We *could* read the above splitting of the proof into two observations as an application of excluded middle (either $R \in R$ or $R \notin R$), but we do not have to!

The following proof is often presented as an exemplary application of excluded middle.

Theorem 1.2. *There exists irrational numbers a and b whose power a^b is rational.*

Classical proof. By excluded middle, the number $\sqrt{2}^{\sqrt{2}}$ is rational or not. Take $a = b = \sqrt{2}$ if it is, otherwise take $a = \sqrt{2}^{\sqrt{2}}$ and $b = \sqrt{2}$. □

As examples go, this one is poorly chosen because the splitting into two cases is inessential. The number $\sqrt{2}^{\sqrt{2}}$ is irrational, a fact not easily proved, and therefore we should have taken the second option right away. Here is a constructive proof:

Constructive proof of Theorem 1.2. $(\sqrt{2})^{\log_2 9} = \sqrt{2^{\log_2 9}} = \sqrt{9} = 3$. □

Of course, irrationality of $\log_2 9$ needs to be established, but the proof is no harder than the one of irrationality of $\sqrt{2}$.

1.3. **Axiom of choice.** In constructive mathematics we cannot afford the axiom of choice because it implies excluded middle. To prove this, we need a bit of set theory. Recall that the unordered pair $\{a, b\}$ is the set of those x's which are equal to a or to b:

$$\{a, b\} = \{x \mid x = a \vee x = b\}.$$

To show that $P(x)$ holds for all $x \in \{a, b\}$, it therefore suffices to check $P(a)$ and $P(b)$.

In constructive mathematics we need to take care with the formulation of the axiom of choice because there is a difference between a nonempty set, which is a set A such that $\neg(A = \emptyset)$, and an inhabited set, which is a set A such that there exists $x \in A$. The axiom of choice states that any family of *inhabited* sets has a choice function.

Theorem 1.3 (Diaconescu [8])**.** *The axiom of choice implies excluded middle.*

Proof. Consider an arbitrary proposition P. To decide P, define

$$A = \{x \in \{0, 1\} \mid P \vee (x = 0)\} \qquad \text{and} \qquad B = \{y \in \{0, 1\} \mid P \vee (y = 1)\}.$$

Every member of $\{A, B\}$ is inhabited because $0 \in A$ and $1 \in B$. By the axiom of choice there is a function $f : \{A, B\} \to A \cup B$ such that $f(A) \in A$ and $f(B) \in B$. Because $f(A) \in \{0, 1\}$, we have $f(A) = 0$ or $f(A) = 1$, and because $f(B) \in \{0, 1\}$, we have $f(B) = 0$ or $f(B) = 1$. We consider four cases organized as follows:

(1) if $f(A) = 1$, then $1 = f(A) \in A$, so $P \vee (1 = 0)$, which is equivalent to P;
(2) if $f(B) = 0$, then $0 = f(B) \in B$, so $P \vee (0 = 1)$, which is equivalent to P;
(3) if $f(A) = 0$ and $f(B) = 1$, then $\neg P$: if P were true, we would have $A = B = \{0, 1\}$, and so $0 = f(A) = f(B) = 1$, a contradiction.

In each case we decided whether P or $\neg P$ holds. □

There are different flavors of constructive mathematics, some of which adopt *countable* choice: a family of inhabited sets indexed by the natural numbers has a choice function. This form of choice suffices for many arguments that one encounters in analysis, so not all is lost. We shall talk about mathematics done without the axiom of choice later. For now let us show a typical inadvertent application of choice which is easily avoided.

Theorem 1.4 (Lebesgue covering theorem)**.** *For every open cover $\mathcal{A}$ of a compact metric space X, there exists $\epsilon > 0$ such that every open ball of radius ϵ is contained in a member of $\mathcal{A}$.*

Taken from Engelking [10, 4.3.31]. For every $x \in X$, take $\epsilon_x > 0$ such that the open ball $B(x, 2\epsilon_x)$ is contained in a member of $\mathcal{A}$. The open cover $\{B(x, \epsilon_x)\}_{x \in X}$ has a finite subcover, i.e., there exists a finite set $\{x_1, \ldots, x_k\} \subseteq X$ such that

$$X = B(x_1, \epsilon_{x_1}) \cup \cdots \cup B(x_k, \epsilon_{x_k}).$$

One can readily see that the number $\epsilon = \min(\epsilon_{x_1}, \ldots, \epsilon_{x_k})$ has the required property. □

Can you spot the application of the axiom of choice? It is in the very first sentence, when it says "for every $x \in X$, take $\epsilon_x > 0$". The mapping $x \mapsto \epsilon_x$ is a choice function. Such hidden choice happens whenever we introduce a new symbol and subscript it with another one to indicate a dependence, in our case ϵ_x. Had we started with "for every $x \in X$, there is $\epsilon > 0$", we would not have chosen anything, but we would also be stuck. The remedy in such situations is simple: do not choose!

Proof of Theorem 1.4 *without choice.* Consider the family of all open balls $B(x, \epsilon)$ where $x \in X$, $\epsilon > 0$, and $B(x, 2\epsilon)$ is contained in a member of $\mathcal{A}$. This is an open cover of X which therefore has a finite subcover, i.e., there is a finite set $\{(x_1, \epsilon_1), \ldots, (x_n, \epsilon_k)\}$ such that

$$X = B(x_1, \epsilon_1) \cup \cdots \cup B(x_k, \epsilon_k).$$

One can readily see that the number $\epsilon = \min(\epsilon_1, \ldots, \epsilon_k)$ has the required property. □

Before proceeding to the next stage, let us clear up one last misconception. Suppose that in a mathematical text we have the assumption that there exists x such that $\phi(x)$. We customarily say "choose an x satisfying $\phi(x)$" to give ourselves an x satisfying ϕ. This is *not* an application of the axiom of choice, but rather an elimination of an existential quantifier. Similarly, if we know that a set A is inhabited and we say "choose $x \in A$", it is not choice but existential quantifier elimination again.

It seems that constructive mathematics is not entirely unreasonable: constructivists do not deny excluded middle but are ambivalent about it; they agree that there is no proposition which is both true and false, or one that is neither true nor false; they are allowed to use negations and prove them by reaching a contradiction; and they find certain forms of choice acceptable. We also have to admit that with a bit of care some instances of excluded middle and choice can be removed, or just turn out to be illusions created by insufficient training in logic.

2. Anger

Constructivism may be self-consistent, but we still remember David Hilbert's words from 1927 [31]:

> *Taking the principle of excluded middle from the mathematician would be the same, say, as proscribing the telescope to the astronomer or to the boxer the use of his fists. To prohibit existence statements and the principle of excluded middle is tantamount to relinquishing the science of mathematics altogether. For compared with the immense expanse of modern mathematics, what would the wretched remnants mean, the few isolated results, incomplete and unrelated, that the intuitionists have obtained without the use of logical ϵ-axiom?*

Hilbert's attack was successful because at the time he was correct. The intuitionists had in fact produced only "the few isolated results", and had made things worse for themselves by talking about the unclear concept of "choice sequences" and by insisting that "all functions are continuous". The Russian constructivists, led by A. A. Markov, compounded the trouble by proving theorems which made a joke of mathematical analysis; for instance in their version of constructivism there was an unbounded continuous map $[0, 1] \to \mathbb{R}$. It is common knowledge that constructivism failed.

It is much less known in the wider mathematical community that things changed in 1967, a year after Brouwer's death, when Erret Bishop published a book on constructive analysis [5]. The importance of the work was best described by Michael Beeson [4]:

> *The thrust of Bishop's work was that both Hilbert and Brouwer had been wrong about an important point on which they had agreed. Namely both of them thought that if one took constructive mathematics seriously, it would be necessary to "give up" the most important parts of modern mathematics (such as, for example, measure theory or complex analysis). Bishop showed that this was simply false, and in addition that it is not necessary to introduce unusual assumptions that appear contradictory to the uninitiated. The perceived conflict between power and security was illusory! One only had to proceed with a certain grace, instead of with Hilbert's "boxer's fists".*

How can this be, when some amount of excluded middle and the axiom of choice is needed to prove such fundamental facts as the well-ordering of cardinals, existence of maximal ideals, existence of vector space bases, Hilbert's *Nullstellensatz*, compactness of the interval, the intermediate value theorem, Tychonoff's theorem, and many others? How did Erret Bishop do the impossible, and why is constructivism still alive?

Historical arguments are interesting, but a well-delivered mathematical punch carries real strength. No sane mathematician would reject the fact that a subset of a finite sets is finite, right?

Theorem 2.1. *Excluded middle is equivalent to the statement that subsets of finite sets are finite.*

Proof. Recall that a set A is *finite* when it can be put into bijective correspondence with $\{0, 1, \ldots, n\}$ for some $n \in \mathbb{N}$. That is, the elements of A are enumerated by a finite list $a_0, a_1, \ldots, a_{n-1}$ without repetition.

If excluded middle holds, we are in home territory: given a set A with an enumeration $a_0, \ldots, a_{n-1}$ and a subset $B \subseteq A$, each a_i either is or is not in B. Thus, to enumerate B, we just erase those a_i's which are not in B.

To prove the converse, assume that a subset of a finite set is always finite, and consider any proposition P. The set $A = \{0\}$ is finite, and $B = \{x \in A \mid P\}$ is a subset of it, and therefore finite. Let $b_0, \ldots, b_m$ be an enumeration of B. Either $m = 0$ or $m \neq 0$. In the first case $B = \emptyset$ and hence $\neg P$, while in the second case $0 \in B$, therefore P. We have decided P. □

Apparently, constructivists think that a subset of a finite set need not be finite. A cursory literature search reveals other bizarre statements considered in constructive mathematics: "$\mathbb{R}$ has measure zero" [28, §6.4.3], "there is a bounded increasing sequence without an accumulation point" [26], "ordinals form a set" [22], "there is an injection of $\mathbb{N}^{\mathbb{N}}$ into $\mathbb{N}$" [3], and so on.

A constructivist might point out that what counts as bizarre is subjective and remind us that once upon a time the discovery of non-Euclidean geometries was shelved in fear of rejection, that Weierstraß's continuous but nowhere differentiable function was and remains a curiosity, and that the Banach–Tarski theorem about conjuring two balls from one is even today called a "paradox". Still, we demand a mathematical explanation of constructivism, not history lessons and subjective opinions.

3. Bargaining

Classical mathematical training plants excluded middle so deeply into young students' minds that most mathematicians cannot even detect its presence in a proof. In order to gain some sort of understanding of the constructivist position, we should therefore provide a method for suspending belief in excluded middle.

If a geometer tried to disbelieve Euclid's fifth postulate, they would find helpful a model of non-Euclidean geometry—an artificial world of geometry whose altered meanings of the words "line" and "point" caused the parallel postulate to fail. Our situation is comparable, only more fundamental because we need to twist the meaning of "truth" itself. We cannot afford a full mathematical account of constructive worlds, but we still can distill their essence, as long as we remember that important technicalities have been omitted.

3.1. **Classical mathematics.** It is well worth pointing out that constructive mathematics is a *generalization* of classical mathematics, as was emphasized by Fred Richman [21], for a proof which avoids excluded middle and choice is still a classical proof. However, trying to learn constructive thinking in the classical world is like trying to learn noncommutative algebra by studying abelian groups.

3.2. **Realizability.** In our first honestly constructive world only that is true which can be computed. Let us imagine, as programmers do, that mathematical objects are represented on a computer as data, and that functions are programs operating on data. Furthermore, a logical statement is only considered valid when there is a program witnessing its truth. We call such programs *realizers*, and we say that

statements are *realized* by them. The *Brouwer–Heyting–Kolmogorov* rules explain when a program realizes a statement:

(1) falsehood $\bot$ is not realized by anything;
(2) truth $\top$ is realized by a chosen constant, say $\star$;
(3) $P \wedge Q$ is realized by a pair (p, q) such that p is a realizer of P and q of Q;
(4) $P \vee Q$ is realized either by $(0, p)$, where p realizes P, or by $(1, q)$, where q realizes Q;
(5) $P \Rightarrow Q$ is realized by a program which maps realizers of P to realizers of Q;
(6) $\forall x \in A \,.\, P(x)$ is realized by a program which maps (a representation of) any $a \in A$ to a realizer of $P(a)$;
(7) $\exists x \in A \,.\, P(x)$ is realized by a pair (p, q) such that p represents some $a \in A$ and q realizes $P(a)$;
(8) $a = b$ is realized by a p which represents both a and b.

The rules work for any reasonable notion of "program". Turing machines would do, but so would quantum computers and programs actually written by programmers in practice. To get used to realizers, let us unravel the computational meaning of:

For every natural number there is a prime larger than it.

This is a "for all" statement, so its realizer is a program p which takes as input a natural number n and outputs a realizer for "there is a prime larger than n", which is a pair (m, q) where m is again a number and q realizes "m is prime and $m > n$". If we forget about q, we see that p is essentially a program that computes arbitrarily large primes. Because such a program exists, there are arbitrarily large primes in the computable world.

A more interesting example is the statement

$$\forall x \in \mathbb{R} \,.\, x = 0 \vee x \neq 0. \tag{1}$$

If we define real numbers as the Cauchy completion of rational numbers, then a real number $x \in \mathbb{R}$ is represented by a program p which takes as input $k \in \mathbb{N}$ and outputs a rational number r_k such that $|x - r_k| \leq 2^{-k}$. Thus a realizer for (1) is a program q which accepts a representation p for any $x \in \mathbb{R}$ and outputs either $(0, s)$ where s realizes $x = 0$, or $(1, t)$ where t realizes $x \neq 0$. Intuitively speaking, such a q should not exist, for however good an approximation r_k of x the program q calculates, it may never be sure whether $x = 0$. To make a water-tight argument, we shall use q to construct the Halting oracle, which does not exist. (The usual proof of nonexistence of the Halting oracle is yet another example of a constructive proof of negation.) Given a Turing machine T and an input n, define the sequence $r_0, r_1, r_2, \ldots$ of rational numbers by

$$r_k = \begin{cases} 2^{-j} & \text{if } T(n) \text{ halts at step } j \text{ and } j \leq k, \\ 2^{-k} & \text{otherwise.} \end{cases}$$

This is a Cauchy sequence because $|r_k - r_m| \leq 2^{-\min(k,m)}$ for all $k, m \in \mathbb{N}$, and it is computable because the value of r_k may be calculated by a simulation of at most k steps of execution of $T(n)$. The limit $x = \lim_k r_k$ satisfies

- $x = 2^{-j} > 0$, if $T(n)$ halts at step j,
- $x = 0$, if $T(n)$ never halts.

The program p which outputs r_k on input k represents x because $|x - r_k| \leq 2^{-k}$ for all $k \in \mathbb{N}$. We may now decide whether $T(n)$ halts by running $q(p)$: if it outputs $(0, s)$, then $T(n)$ does not halt, and if it outputs $(1, t)$, then $T(n)$ halts.

As an exercise you should figure out where in the above argument you need the following (valid) instance of excluded middle: for every $k \in \mathbb{N}$, either $r_k = 2^{-k}$ or $r_k = 2^{-j}$ for some $j < k$.

Theorem 3.1. *Excluded middle is* not *realized.*

Proof. The statement (1) is an instance of excluded middle and it is not realized. □

The strategy to place constructivism inside a box is working! If one takes the limited view that everything must be computed by machines, then excluded middle fails because machines cannot compute everything. *Our* excluded middle is not affected because we are not machines.

We may use the computable world to explain Theorem 2.1. Let us show that the following instance of "subsets of finite sets are finite" is not realized:

(2) *All countable subsets of* $\{0, 1\}$ *are finite.*

In computable mathematics a finite set is represented by a finite list of its elements, and a countable set by a program which enumerates its elements, possibly with repetitions. The subsets $\{\}$, $\{0\}$, $\{1\}$ and $\{0, 1\}$ are all countable and finite, so (2) looks pretty true. Remember though that in the computable world "for all" means not "it holds for every instance" but rather "there is a program computing witnesses from instances". A realizer for (2) is a program q which takes as input a program p enumerating the elements of a subset of $\{0, 1\}$ and outputs a finite list of all the elements so enumerated. To see intuitively where the trouble lies, suppose p starts enumerating zeroes:

$$0, 0, 0, 0, 0, 0, \ldots.$$

The output list should contain 0, but should it contain 1? However long a prefix of the enumeration we investigate, if it is all zeroes, then we cannot be sure whether 1 will appear later. For an actual proof we use the same trick as before: with q in hand we could construct the Halting oracle. Given any Turing machine T and input n, consider the program p which works as follows:

$$p(k) = \begin{cases} 1 & \text{if } T(n) \text{ halts in fewer than } k \text{ steps,} \\ 0 & \text{otherwise.} \end{cases}$$

The subset $S \subseteq \{0, 1\}$ enumerated by p is constructed so that

- $1 \in S$ if $T(n)$ halts,
- $1 \notin S$ if $T(n)$ does not halt.

Now scan the finite list computed by $q(p)$: if it contains 1, then $T(n)$ holds, otherwise it does not.

3.3. **Sheaves.** Our second constructive model is the topos of sheaves on a topological space X. Recall that a sheaf of sets F on X is a collection of sets $F(U)$, one for each open subset $U \subseteq X$, together with an operation which to each $x \in F(U)$ assigns its *restriction* $x|_{V \subseteq U} \in F(V)$ to a smaller open set $V \subseteq U$, such that:

(1) restriction is functorial: $x|_{U \subseteq U} = x$ and $(x|_{V \subseteq U})|_{W \subseteq V} = x|_{W \subseteq U}$;

(2) the gluing property holds: if $(U_i)_{i\in I}$ is an open cover of $U \subseteq X$ and $x_i \in F(U_i)$ are compatible elements (which means that $x_i|_{U_i\cap U_j \subseteq U_i} = x_j|_{U_i\cap U_j\subseteq U_j}$ for all $i, j \in I$), then there is a unique $x \in F(U)$ such that $x|_{U_i\subseteq U} = x_i$ for all $i \in I$.

A map $\eta : F \to G$ between sheaves F and G is a collection of maps $\eta_U : F(U) \to G(U)$, one for each open $U \subseteq X$, which commute with restrictions, $\eta_U(x)|_{V\subseteq U} = \eta_V(x|_{V\subseteq U})$. Sheaves and maps between them form a category which is a topos, i.e., a model of intuitionistic mathematics [18].

Examples of sheaves may be more helpful than their definition. A familiar one is the sheaf of real-valued continuous maps $\mathcal{C}_\mathbb{R}$ which assigns the set

$$\mathcal{C}_\mathbb{R}(U) = \{f : U \to \mathbb{R} \mid f \text{ continuous}\}$$

to every open $U \subseteq X$. This is indeed a sheaf, for a continuous map remains continuous when restricted to a smaller domain, and it may be defined locally on the members of an open cover as long as the local definitions are compatible. A prototypical example of a sheaf is the sheaf of local sections S_p of a topological bundle $p : E \to X$,

$$S_p(U) = \{s : U \to E \mid s \text{ continuous and } p \circ s = 1_U\}.$$

The sheaf $\mathcal{C}_\mathbb{R}$ arises as the sheaf of sections of the trivial bundle $X \times \mathbb{R} \to X$, and in fact every sheaf of sets is isomorphic to a sheaf of sections of an étale map.

We shall think of the topos of sheaves on X as a mathematical universe in which objects *exist locally* and *vary continuously*. For example, a real number is not just a single quantity, but rather one varying continuously on an open $U \subseteq X$—it is an element of $\mathcal{C}_\mathbb{R}(U)$. The sheaf $\mathcal{C}_\mathbb{R}$ itself is a varying ring: for each open $U \subseteq X$, the ring operations on $\mathcal{C}_\mathbb{R}(U)$ are induced by those of $\mathbb{R}$, and restrictions $\mathcal{C}_\mathbb{R}(U) \to \mathcal{C}_\mathbb{R}(V)$ are ring homomorphisms. Other examples abound, as experts in algebraic geometry and algebraic topology will readily confirm.

Truth varies as well, so that a statement may be true on one open set and false on another. Restrictions and the gluing property of sheaves transfer to truth:

(1) if a statement is true on an open set $U \subseteq X$, then it is also true on a smaller open set $V \subseteq U$;
(2) if a statement is true on each member U_i of an open cover, then it is also true on the union $\bigcup_{i\in I} U_i$.

In the topos the truth values are the open subsets of X. The truth value of a statement is the largest open set on which it holds, and the logic is dictated by the topology of X:

- falsehood and truth are $\emptyset$ and X, the least and greatest open sets, respectively;
- conjunction $U \wedge V$ is $U \cap V$, the largest open set contained in U and V;
- disjunction $U \vee V$ is $U \cup V$, the least open set containing U and V;
- negation $\neg U$ is the topological exterior $\operatorname{ext}(U)$, the largest open set disjoint from U;
- implication $U \Rightarrow V$ is $\operatorname{ext}(U \setminus V)$, the largest open set whose intersection with U is contained in V.

Excluded middle amounts to saying that $U \cup \operatorname{ext}(U) = X$ for all open $U \subseteq X$, a condition equivalent to open and closed sets coinciding. Only a very special kind of

space X satisfies this condition, for as soon as it is a T_0-space (points are uniquely determined by their neighborhoods), it has to be discrete.

An attentive reader will wonder how it is possible to have as many truth values as there are open subsets of X, when in Section 1.1 we proved that there is no proposition which is neither false nor true. Is not an intermediate open set $\emptyset \subsetneq U \subsetneq X$ a truth value different from both $\emptyset$ and X? The mystery is resolved when we remember that we should work inside the topos and ask not whether it is different but *where* it is different. Saying "U is not true" is the same as negating U, and saying "U is not false" is the same as negating it twice, hence the truth value of

$$U \text{ is neither true nor false}$$

is the open set

$$\mathrm{ext}(U) \cap \mathrm{ext}(\mathrm{ext}(U)),$$

which is empty. The mathematicians who live in the sheaf topos see that *nowhere* is it the case that U is neither true nor false, and that is all they see. Only we, the external observers, can comprehend the *global* fact that U is an intermediate truth value.

To develop a bit of intuition about the world of sheaves, let us calculate the truth value of the Archimedean axiom:

(3) *For every real number x there is an integer k such that $x < k$.*

Of course, the words "real number" and "integer" must be understood inside the model. We already mentioned that the reals correspond to the sheaf $\mathcal{C}_{\mathbb{R}}$ of real-valued continuous maps. Similarly, the integers in the model are the sheaf of integer-valued continuous maps

$$\mathcal{C}_{\mathbb{Z}}(U) = \{f : U \to \mathbb{Z} \mid f \text{ continuous}\},$$

where $\mathbb{Z}$ is equipped with the discrete topology. To see whether (3) holds on an open set $U \subseteq X$, we consider a real $x : U \to \mathbb{R}$ and the statement

(4) *There is an integer k such that $x < k$.*

It looks like (4) can fail, for if U is connected, then a continuous map $k : U \to \mathbb{Z}$ must be constant, whereas $x : U \to \mathbb{R}$ could be unbounded. However, this is not the case. Because x is continuous, there is an open cover $U = \bigcup_{i \in I} U_i$ such that the restrictions $x|_{U_i} : U_i \to \mathbb{R}$ are bounded. On each U_i there is then a constant $k : U_i \to \mathbb{Z}$ which dominates $x|_{U_i}$ so that (4) holds on each U_i. By the gluing property it must hold on U, too. The Archimedean axiom holds in a sheaf topos by virtue of the local nature of truth.

By pursuing the Archimedean axiom a bit further, we can explain how the axiom of choice fails. We take the space X to be the open interval $(0, 1)$ with Euclidean topology and ask whether there exists a choice function which assigns to every real an integral upper bound. Remembering to think locally, we seek an open cover $(0, 1) = \bigcup_{i \in I} U_i$ such that each U_i has the desired choice function (a global choice function need not exist because the local ones need not be compatible with each other). If we had a choice function on an inhabited U_i, then its restriction to a small interval $(a, b) \subseteq U_i$ would constitute a choice function on (a, b), which it cannot because it would have to assign an upper bound $k \in \mathcal{C}_{\mathbb{Z}}(a, b)$ to the real number $x \in \mathcal{C}_{\mathbb{R}}(a, b)$ defined by $x(t) = 1/(t - a)$. The interval (a, b) is connected so

every $k \in \mathcal{C}_{\mathbb{Z}}(a, b)$ is constant and cannot bound x. Choice fails locally everywhere because of the topological properties of open intervals.

The method of building constructive models of mathematics gave us a solution: those who have contracted constructivism can be safely put away in one of the constructive worlds of mathematics where they will bother no one.

4. Depression

In Raymond Smullyan's story [24] there was a planet on which the concept of humor was unknown and laughter was treated as a disease. Those who laughed were sent to live in laugh-hospitals, away from normal people. As time passed ever more people contracted laughter and the laugh-hospitals grew into whole laugh-communities, until the few remaining normal people pretended to understand humor just so that they could join the rest. What if constructive models are like the laugh-hospitals? What if not understanding constructivism is like not having a sense of humor?

Apart from sheaves on a space there are many other toposes, each a model of constructive mathematics. They were invented by the great Alexander Grothendieck for the purposes of studying algebraic geometry, but have since proved generally useful in mathematics. The Dubuc topos [9] contains the 17th-century nilpotent infinitesimals, but without the 17th-century confusion and paradoxes. Joyal's theory of combinatorial species [16] is just a topos in disguise, and so are various kinds of graphs [19]. Simplicial sets, the home of homotopy theorists, form a presheaf topos. The realizability toposes [12] are computer scientists' Gardens of Eden in which everything is computable by design. Even such mundane topics as the syntax of programming languages get their own toposes [20].

Does anyone care about all these models of constructive mathematics? Well, if excluded middle is the only price for achieving rigor in infinitesimal calculus, our friends physicists just might be willing to pay it. After all they still use Newton's infinitesimals, despite our having lectured them about ϵ's and δ's since the time of Cauchy and Weierstraß. And how often does a physicist start a calculation by saying "suppose not"? The situation with computer scientists is worse, as some of them actually help spread constructive mathematics with slogans such as "propositions are types" [36]. The recently discovered homotopy-theoretic interpretation of Martin–Löf type theory [1, 34], a most extreme form of constructivism, has made some homotopy theorists and category theorists into allies of constructive mathematics. They even profess a new foundation of mathematics [30, 35] in which logic and sets are just two levels of an infinite hierarchy of homotopy types.

We turn to set theorists for advice. The axioms of Zermelo and Fraenkel stand as firm as ever, they assure us, and are the de facto foundation of today's mathematics. We are told that even the builders of toposes and modelers of homotopy types ultimately rely on set theory, and we need not renounce excluded middle to compute with infinitesimals. The relief however does not last long. Set theorists go on to explain that Grothendieck actually used set theory extended with universes, each of which is an entire model of classical mathematics. Ever since Cohen's work on the independence of Cantor's hypothesis [6, 7], set theorists have been exploring not one, but many worlds of classical mathematics. Would you like to have infinitesimals, or make all sets of reals measurable, or do you fancy $2^{\aleph_0} = \aleph_{42}$? A world of classical

mathematics is readily built to order for you. Or should we call it a classical laugh-hospital?

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.

5. Acceptance

As exciting as a multiverse of mathematics may be, the working mathematician has no time to spend their career wandering from one world to another. Nevertheless, they surely are curious about the newly discovered richness of mathematics, they welcome ideas coming from unfamiliar worlds, and they strive to make their own work widely applicable. They will find it easier to accomplish these goals if they speak the lingua franca of the mathematical multiverse—constructive mathematics.

Some aspects of constructive mathematics are just logical hygiene: avoid indirect proofs in favor of explicit constructions, detect and eliminate needless uses of the axiom of choice, know the difference between a proof of negation and a proof by contradiction. Of course, constructivism goes deeper than that. The stringent working conditions of constructive worlds require an economy of thought which is disheartening at first but eventually pays off with vistas of new mathematical landscapes that are proscribed by orthodox mathematics. We may at best offer a few glimpses.

5.1. Mathematics without excluded middle.

To demonstrate how one works without excluded middle, we adapt to the constructive setting a theorem which every student of mathematics learns.

Theorem 5.1 (Intermediate value theorem)**.** *If $f : [0,1] \to \mathbb{R}$ is a continuous map and $f(0) \leq 0 \leq f(1)$, then $f(x) = 0$ for some $x \in [0,1]$.*

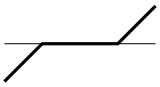
Classical proof. A common proof proceeds by bisection: start with the interval $[a_0, b_0] = [0,1]$, let $c_n = \frac{1}{2}(a_n + b_n)$, and set

$$[a_{n+1}, b_{n+1}] = \begin{cases} [c_n, b_n] & \text{if } f(c_n) < 0, \\ [a_n, c_n] & \text{otherwise.} \end{cases} \tag{5}$$

The intervals are defined so that $f(a_n) \leq 0 \leq f(b_n)$ for all $n \in \mathbb{N}$. The intersection of the descending chain $[a_0, b_0] \supset [a_1, b_1] \supset \cdots$ contains a single point x which by continuity of f must be its zero. Indeed, for every $\epsilon > 0$ there is $n \in \mathbb{N}$ such that $|f(x) - f(a_n)| < \epsilon$ and $|f(x) - f(b_n)| < \epsilon$, therefore $-\epsilon < f(x) - f(b_n) \leq f(x) \leq f(x) - f(a_n) < \epsilon$, which means that $f(x) = 0$. □

Unfortunately, the proof relies on excluded middle in the definition (5), as it presupposes that every real is either negative or nonnegative. To get a constructive version of the theorem we need to change something. We can try to eliminate the use of excluded middle, reformulate the theorem to a classically equivalent but constructively valid variant, weaken the conclusion, or strengthen the assumptions.

In the present case it is not possible to eliminate excluded middle. Consider a function which *hovers* around zero:



An arbitrarily small perturbation of the function may cause large movements of its zeroes. To a constructive mathematician such a discontinuous phenomenon signals that something is amiss, which indeed it is, but we are not going to prove this. The intermediate value theorem fails constructively not because there are too few zeroes, but because they may be too unstable.

Before reformulating the theorem, we verify that the closed interval is still connected in constructive mathematics.

Proposition 5.2. *If A and B are inhabited open subsets of $[0,1]$ such that $A \cup B = [0,1]$, then $A \cap B$ is inhabited.*

Proof. Because A and B are inhabited, there are $a_0 \in A$ and $b_0 \in B$. Define sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$, starting from a_0 and b_0, letting $c_n = \frac{1}{2}(a_n + b_n)$, and setting

$$(a_{n+1}, b_{n+1}) = \begin{cases} (c_n, b_n) & \text{if } c_n \in A, \\ (a_n, c_n) & \text{if } c_n \in B. \end{cases} \tag{6}$$

The two cases in (6) cover all possibilities because $A \cup B = [0,1]$, but they need not be disjoint. At each step we have to *choose* one, and since the next choice depends on the previous ones, we here rely on a generalization of countable choice known as the *dependent choice*. By construction $a_n \in A$ and $b_n \in B$ for all $n \in \mathbb{N}$. Both sequences converge to a common point x, which is in A or in B. If $x \in A$, then, since A is open and $(b_n)_{n\in\mathbb{N}}$ converges to x, there is $b_m \in A$, and we are done. If $x \in B$, we argue symmetrically. □

Notice how we stated connectedness in a *positive* way that allows us to deduce existence of a point. The common definition of connectedness "there is no partition of $[0,1]$ into two open inhabited sets" is classically equivalent to our definition but has a constructively weaker conclusion, namely that something does not exist. Constructive mathematics encourages a positive attitude: instead of saying that A is not empty, say that A has an element; instead of saying that x is not zero, say that x is negative or positive; instead of saying that a finite set is one that is not infinite, say that it is one that is in bijection with $\{0, \ldots, n\}$ for some $n \in \mathbb{N}$; etc.

Here is a reformulation of the intermediate value theorem which is classically equivalent to Theorem 5.1 but has a constructive proof.

Theorem 5.3. *If $f : [0,1] \to \mathbb{R}$ is continuous and, for every $x \in [0,1]$, either $f(x) < 0$ or $f(x) > 0$, then f is either everywhere positive or everywhere negative.*

Proof. Actually, let us prove a more general statement: if A and B are disjoint open sets such that $A \cup B = [0,1]$, then either $A = [0,1]$ or $B = [0,1]$, which is yet another form of connectedness of $[0,1]$. One of the sets contains 0, say $0 \in A$. Then B is not inhabited, or else $A \cap B$ would be inhabited as well by Proposition 5.2. A set which is not inhabited is empty, therefore $A = [0,1]$. If $0 \in B$, we exchange the roles of A and B.

The theorem is obtained when we take the sets $A = \{y \in [0,1] \mid f(y) < 0\}$ and $B = \{z \in [0,1] \mid f(z) > 0\}$. □

We may weaken the intermediate value theorem to an approximate version.

Theorem 5.4 (Approximate intermediate value theorem)**.** *If $f : [0,1] \to \mathbb{R}$ is continuous, $f(0) \leq 0 \leq f(1)$, and $\epsilon > 0$, then $|f(x)| < \epsilon$ for some $x \in [0,1]$.*

Proof. Apply Proposition 5.2 to $\{x \in [0,1] \mid f(x) < \epsilon\}$ and $\{x \in [0,1] \mid -\epsilon < f(x)\}$, and take into account that every real number is smaller than ϵ or greater than $-\epsilon$. □

The approximate intermediate value theorem faithfully reflects computational practice: numerical bisection algorithms stop as soon as the desired accuracy ϵ is achieved, lest numerical errors send them on a wild goose chase. This is no accident, for where numerical analysts run into numerical instability and topologists detect discontinuity, constructive mathematicians find logical impossibility.

Lastly, we can strengthen the assumptions of the theorem and keep the conclusion unchanged. A simple requirement is that $f : [0,1] \to \mathbb{R}$ be strictly monotone (if $x < y$, then $f(x) < f(y)$).

Theorem 5.5 (Monotone intermediate value theorem)**.** *If $f : [0,1] \to \mathbb{R}$ is continuous, strictly monotone, and $f(0) < 0 < f(1)$, then $f(x) = 0$ for a unique $x \in [0,1]$.*

Proof. This time it is easiest to construct x as a Dedekind cut. We outline a proof here and refer the interested readers to [27] for details and an exhaustive study of the constructive intermediate value theorem. The sets

$$L = \{y \in [0,1] \mid f(y) < 0\} \qquad \text{and} \qquad U = \{z \in [0,1] \mid 0 < f(z)\}$$

form a Dedekind cut in the constructive sense, because they are

(1) *bounded*: $\exists y \in \mathbb{R}\,.\, y \in L$ and $\exists z \in \mathbb{R}\,.\, z \in U$;
(2) *rounded*: $y \in L \Leftrightarrow \exists y < z\,.\, z \in L$ and $z \in U \Leftrightarrow \exists y < z\,.\, y \in U$;
(3) *disjoint*: $L \cap U = \emptyset$; and
(4) *located*: $y < z \Rightarrow y \in L \vee z \in U$.

The reals are Dedekind complete; therefore, there is a unique $x \in [0,1]$ such that $y < x < z$ for all $y \in L$ and $z \in U$. Because f is continuous and strictly monotone, it follows that $f(x) = 0$. □

The monotone intermediate value theorem suffices in many cases, for instance as soon as we are able to isolate a zero in an interval where f is strictly monotone.

It should be clear from our short excursion into constructive analysis that lack of excluded middle requires certain changes in formulations of theorems, but for the most part it allows us to keep familiar concepts and techniques. In contrast, giving up the axiom of choice requires a rethinking of fundamental notions.

5.2. **Mathematics without choice.** Tychonoff's theorem [29] states that the product of any family of compact topological spaces is compact. It has many applications throughout mathematics and is equivalent to the axiom of choice. To make it constructive, we need to rethink the notion of space itself.

Which is more important in a topological space, its points or its opens? This is a silly question, but you would probably answer that the open sets are the essence of

topology. This is true in a precise mathematical sense: the points of a sufficiently nice space can be reconstructed from the open sets, but not vice versa. To see how this works, recall that in a topological space (X, τ_X) the open neighborhoods of a point $x \in X$ form a *completely prime filter*, i.e., the set $\mathcal{F}_x = \{U \in \tau_X \mid x \in U\}$ is:

(1) *upward closed*: if $U \subseteq V \in \tau_X$ and $U \in \mathcal{F}_x$, then $V \in \mathcal{F}_x$;
(2) *downward directed*: any finite collection of sets in $\mathcal{F}_x$ has a lower bound in $\mathcal{F}_x$;
(3) *completely prime*: for any set I, if $\bigcup_{i\in I} U_i \in \mathcal{F}_x$, then $U_j \in \mathcal{F}_x$ for some $j \in I$.

If X is Hausdorff (and more generally sober), then every completely prime filter arises as the neighborhood filter of a unique point. And a continuous map $f : X \to Y$ takes a completely prime filter $\mathcal{F}$ in τ_X to the completely prime filter $\{V \in \tau_Y \mid f^*(V) \in \mathcal{F}\}$ in τ_Y, which means that we can recover f from its inverse image map $f^* : \tau_Y \to \tau_X$. The upshot is that the points are not essential, and that spaces and continuous maps can be defined purely in terms of the order-theoretic structure of opens.

Definition 5.6. A *locale* is a partially ordered set $(\tau, \leq_\tau)$ with finite infima and all suprema satisfying the distributivity law $u \wedge \bigvee_{i\in I} v_i = \bigvee_{i\in I}(u \wedge v_i)$. The elements of τ are called the *opens* and the completely prime filters the *points*. A *locale map* from $(\tau, \leq_\tau)$ to $(\sigma, \leq_\sigma)$ is given by a map $f^* : \sigma \to \tau$ in the opposite direction that preserves finite infima and all suprema.

The definition captures the essential properties of open sets, namely closure under finite intersections and arbitrary unions. We need not posit the existence of the least and the largest opens because they are the empty infimum and the empty supremum, respectively. (One is tempted to write "Five stages of accepting the empty set".) We think of $f^* : \sigma \to \tau$ as the inverse image map, although the opens of a locale need not be actual subsets of anything and f^* need not be induced by a map on points. Quite the opposite—f^* induces a map on the points by taking a completely prime filter $\mathcal{F} \subseteq \tau$ to the completely prime filter $\{u \in \sigma \mid f^*(u) \in \mathcal{F}\} \subseteq \sigma$.

Every topological space (X, τ_X) can be construed as a locale $(\tau_X, \subseteq)$ of its open sets, and every continuous map $f : X \to Y$ as a locale map $f^* : \tau_Y \to \tau_X$. A locale which arises as a topology of a space is called *spatial*. There are nonspatial locales, for instance the locale $\mathrm{Reg}(\mathbb{R})$ of *regular* open subsets of $(\mathbb{R}, \mathcal{O}(\mathbb{R}))$ (with the Euclidean topology $\mathcal{O}(\mathbb{R})$), i.e., those open sets which are the interiors of their closures. In the locale $(\mathrm{Reg}(\mathbb{R}), \subseteq)$ the infima and suprema are computed as

$$U \wedge V = U \cap V \qquad \text{and} \qquad \textstyle\bigvee_{i\in I} U_i = \mathrm{reg}(\bigcup_{i\in I} U_i),$$

where $\mathrm{reg}(U)$ is the interior of the closure of U. The locale $\mathrm{Reg}(\mathbb{R})$ is a sublocale of $\mathbb{R}$ because the map $\mathrm{reg} : \mathcal{O}(\mathbb{R}) \to \mathrm{Reg}(\mathbb{R})$ is onto, and is thus a locale embedding (remember that maps between locales are backwards). It is quite rich in its structure, but it has no points! Briefly, the regular open sets form a complete Boolean algebra, with topological exterior acting as negation. A completely prime filter $\mathcal{F}$ in $\mathrm{Reg}(\mathbb{R})$ would have to have a least element, namely the exterior of the largest open that is not in $\mathcal{F}$. Such a least element must be the empty set because in $\mathbb{R}$ every inhabited regular open can be decomposed into two disjoint inhabited regular opens. However, no filter contains the empty set.

Doing topology without ever mentioning the points is possible but takes a bit of exercise. For instance, a locale is *compact* when it has the finite subcover property: if $\bigvee_{i\in I} u_i = 1$, then $u_{i_1} \vee \cdots \vee u_{i_n} = 1$ for some $i_1, \ldots, i_n \in I$. Some locale constructions are best described algebraically: the *product* $(\tau, \leq_\tau) \times (\sigma, \leq_\sigma)$ is the locale generated by pairs (u, v) where $u \in \tau$ and $v \in \sigma$, subject to equations

$$(u_1, v_1) \wedge (u_2, v_2) = (u_1 \wedge u_2, v_1 \wedge v_2)$$

and

$$(\bigvee_{i\in I} u_i, \bigvee_{j\in J} v_j) = \bigvee_{i\in I} \bigvee_{j\in J} (u_i, v_j).$$

An analogous construction works for products of any family of locales.

Why are we bothering with all this? Because we can trade the points for Tychonoff's theorem.

Theorem 5.7 (Localic Tychonoff)**.** *The product of any family of compact locales is compact, constructively.*

Proof. See [32]. A proof which uses classical logic but avoids the axiom of choice can be found in [15]. □

Were the constructive Tychonoff's theorem the only advantage of locales, there would be little reason to accept them as a new notion of spaces. That they play a significant role in mathematics was explained expertly by Peter Johnstone [14]. There is little we can add, except for mentioning a recent development which shows that locales ought to be taken seriously in classical mathematics, too.

Measure theorists have come to accept that not all sets of reals are Lebesgue measurable [33]. While it is possible to trade the axiom of choice for measurability of all sets [25], the accepted norm is to keep choice and deal with the phenomenon of nonmeasurability. A surprising alternative was proposed by Alex Simpson [23]: if we use locales instead of spaces, we can have both, the axiom of choice and an isometry-invariant measure on all sublocales of $\mathbb{R}^n$, which agrees with the Lebesgue measure on the measurable sets. This is possible because there are more sublocales than subspaces—earlier we saw that $\mathbb{R}$ has a nontrivial sublocale without points—and so the usual attempt at decomposition of a space into nonmeasurable parts does not work anymore. Such parts have no common points but still overlap on the "topological" glue that constitutes the space, so additivity of measure does not apply.

The passage from spaces to locales has at least two additional benefits [23]: the Banach–Tarski "paradoxical" decomposition of the unit ball [2] does not work with locales even though we keep the axiom of choice, and the naive definition of random reals does. Let us briefly explain the latter one. An appealing intuitive definition of a *random* real $x \in [0, 1]$ is that it has no special properties, i.e., that it avoids every null set. This will not do, of course, since a singleton is a null set, and so no particular real can be random. Thus, if we attempt to define the space of random reals as the intersection of all full subsets of $[0, 1]$, we are left with an empty space. However, the intersection of all *sublocales* of $[0, 1]$ with full measure is a nontrivial sublocale of $[0, 1]$ whose measure is still full. This is the locale of random sequences. It is nonspatial, has no points, it disrupts the traditional notion of space, and it captures our intuitive ideas about randomness in a way that usual spaces cannot.

If you are still reading these lines, you have perhaps advanced by one or two stages, and my work is done. In the unlikely event that you have reached the

terminal stage, I invite you to leave the overcrowded Cantor's paradise and explore the rest of mathematics.

Acknowledgment

The author thanks the editor, Mark Goresky, for kind and sustained encouragement that kept the author going until the paper took an acceptable form.

About the author

Andrej Bauer is professor of computational mathematics at the University of Ljubljana, Slovenia. He works in the areas of constructive and computable mathematics, mathematical principles of programming languages, type theory, and homotopy type theory.

References

[1] Steve Awodey and Michael A. Warren, *Homotopy theoretic models of identity types*, Math. Proc. Cambridge Philos. Soc. **146** (2009), no. 1, 45–55, DOI 10.1017/S0305004108001783. MR2461866

[2] Stefan Banach and Alfred Tarski, *Sur la décomposition des ensembles de points en parties respectivement congruentes*, Fundamenta Mathematicae **6** (1924), 244–277.

[3] Andrej Bauer, *An injection from the Baire space to natural numbers*, Math. Structures Comput. Sci. **25** (2015), no. 7, 1484–1489, DOI 10.1017/S0960129513000406. MR3391059

[4] Michael J. Beeson, *Foundations of constructive mathematics. Metamathematical studies*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 6, Springer-Verlag, Berlin, 1985. MR786465

[5] Errett Bishop, *Foundations of constructive analysis*, McGraw-Hill Book Co., New York-Toronto-London, 1967. MR0221878

[6] Paul Cohen, *The independence of the continuum hypothesis*, Proc. Nat. Acad. Sci. U.S.A. **50** (1963), 1143–1148. MR0157890

[7] Paul J. Cohen, *The independence of the continuum hypothesis. II*, Proc. Nat. Acad. Sci. U.S.A. **51** (1964), 105–110. MR0159745

[8] Radu Diaconescu, *Axiom of choice and complementation*, Proc. Amer. Math. Soc. **51** (1975), 176–178. MR0373893

[9] Eduardo J. Dubuc, *C^∞-schemes*, Amer. J. Math. **103** (1981), no. 4, 683–690, DOI 10.2307/2374046. MR623133

[10] Ryszard Engelking, *General topology*, 2nd ed., Sigma Series in Pure Mathematics, vol. 6, Heldermann Verlag, Berlin, 1989. Translated from the Polish by the author. MR1039321

[11] Paul R. Halmos, *Naive set theory*, The University Series in Undergraduate Mathematics, D. Van Nostrand Co., Princeton, N.J.-Toronto-London-New York, 1960. MR0114756

[12] J. M. E. Hyland, *The effective topos*, The L. E. J. Brouwer Centenary Symposium (Noordwijkerhout, 1981), Stud. Logic Foundations Math., vol. 110, North-Holland, Amsterdam-New York, 1982, pp. 165–216. MR717245

[13] Thomas Jech, *Set theory*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2003. The third millennium edition, revised and expanded. MR1940513

[14] Peter T. Johnstone, *The point of pointless topology*, Bull. Amer. Math. Soc. (N.S.) **8** (1983), no. 1, 41–53, DOI 10.1090/S0273-0979-1983-15080-2. MR682820

[15] Peter T. Johnstone, *Stone spaces*, Cambridge Studies in Advanced Mathematics, vol. 3, Cambridge University Press, Cambridge, 1986. Reprint of the 1982 edition. MR861951

[16] André Joyal, *Une théorie combinatoire des séries formelles* (French, with English summary), Adv. in Math. **42** (1981), no. 1, 1–82, DOI 10.1016/0001-8708(81)90052-9. MR633783

[17] Elisabeth Kübler-Ross, *On death and dying*, New York, Routledge, 1969.

[18] Saunders Mac Lane and Ieke Moerdijk, *Sheaves in geometry and logic: A first introduction to topos theory*, Springer-Verlag, 1992.

[19] F. William Lawvere, *Qualitative distinctions between some toposes of generalized graphs*, Categories in computer science and logic (Boulder, CO, 1987), Contemp. Math., vol. 92, Amer. Math. Soc., Providence, RI, 1989, pp. 261–299, DOI 10.1090/conm/092/1003203. MR1003203

[20] Andrew M. Pitts, *Nominal sets. Names and symmetry in computer science*, Cambridge Tracts in Theoretical Computer Science, vol. 57, Cambridge University Press, Cambridge, 2013. MR3113350

[21] Fred Richman, *Intuitionism as generalization*, Philos. Math. (2) **5** (1990), no. 1-2, 124–128. MR1061335

[22] Giuseppe Rosolini, *Un modello per la teoria intuizionista degli insiemi*, Atti degli Incontri di Logica Matematica (Siena) (C. Bernardi and P. Pagli, eds.), 1982, pp. 227–230.

[23] Alex Simpson, *Measure, randomness and sublocales*, Ann. Pure Appl. Logic **163** (2012), no. 11, 1642–1659, DOI 10.1016/j.apal.2011.12.014. MR2959665

[24] Raymon M. Smullyan, *This book needs no title: A budget of living paradoxes*, Prentice-Hall, 1980.

[25] Robert M. Solovay, *A model of set-theory in which every set of reals is Lebesgue measurable*, Ann. of Math. (2) **92** (1970), 1–56. MR0265151

[26] Ernst Specker, *Nicht konstruktiv beweisbare Sätze der Analysis* (German), J. Symbolic Logic **14** (1949), 145–158. MR0031447

[27] Paul Taylor, *A lambda calculus for real analysis*, J. Log. Anal. **2** (2010), Paper 5, 115, DOI 10.4115/jla.2010.2.5. MR2678927

[28] Anne S. Troelstra and Dirk van Dalen, *Constructivism in mathematics, an introduction*, Studies in Logic and the Foundations of Mathematics., no. 121 and 123, North-Holland, 1988.

[29] A. Tychonoff, *Über die topologische Erweiterung von Räumen* (German), Math. Ann. **102** (1930), no. 1, 544–561, DOI 10.1007/BF01782364. MR1512595

[30] The Univalent Foundations Program, *Homotopy type theory: Univalent foundations of mathematics*, Institute for Advanced Study, 2013.

[31] Jean van Heijenoort, editor, *From Frege to Gödel. A source book in mathematical logic, 1879–1931*, Harvard University Press, Cambridge, MA, 2002. Reprint of the third printing of the 1967 original [MR0209111 (35 #15)]. MR1890980

[32] Steven Vickers, *Some constructive roads to Tychonoff*, *From sets and types to topology and analysis*, Oxford Logic Guides, vol. 48, Oxford Univ. Press, Oxford, 2005, pp. 223–238, DOI 10.1093/acprof:oso/9780198566519.003.0014. MR2188646

[33] Giuseppe Vitali, *Sul problema della misura dei gruppi di punti di una retta*, Tech. report, Bologna, 1905.

[34] Vladimir Voevodsky, *The equivalence axiom and univalent models of type theory* (talk at CMU on February 4, 2010), ArXiv e-print 1402.5556 (2010).

[35] Vladimir Voevodsky, *Univalent foundations* (talk notes), `https://www.math.ias.edu/vladimir/sites/math.ias.edu.vladimir/files/2011_UPenn.pdf`, September 2011.

[36] Philip Wadler, *Propositions as types*, Communications of the ACM **58** (2015), no. 12, 75–84.

Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia

E-mail address: `Andrej.Bauer@andrej.com`