
Constructivism Is Difficult

Eric Schechter

In a recent issue of this MONTHLY, Fred Richman [8] discussed existence proofs. Richman's conclusion, as I understood it, was that once a mathematician sees the distinction between constructive and nonconstructive mathematics, he or she will choose the former. That conclusion, if extrapolated further than Professor Richman intended, suggests that any mathematician can learn constructivism easily if he or she so desires.

But in fact constructivism is unusually difficult to learn. Learning most mathematical subjects merely involves adding to one's knowledge, but learning constructivism involves *modifying* all aspects of one's knowledge: theorems, methods of reasoning, technical vocabulary, and even the use of everyday words that do not seem technical, such as "or". I discuss, in the language of mainstream mathematicians, some of those modifications; perhaps newcomers to constructivism will not be so overwhelmed by it if they know what kinds of difficulties to expect.

1. INTRODUCTION. A proof is *nonconstructive* if it asserts the existence of some object without actually constructing or finding that object. Such proofs are used freely in mainstream ("classical") mathematics. *Constructivism* is the practice of avoiding such proofs or at least pointing them out explicitly. Different philosophical views lead to different kinds of constructivism:

- A mathematician who accepts nonconstructive proofs may nevertheless prefer constructive ones when available, because they are more informative. Indeed, most mathematicians are part-time constructivists, at least in this respect: When we are teaching, we try to illustrate abstract concepts with concrete examples.
- On the other hand, some mathematicians consider all nonconstructive proofs unacceptable, for philosophical reasons. At issue are the meaning of "existence" of an object and the meaning of "proof" of a statement.

There are actually several different schools of constructivism, and I do not attempt to distinguish them here; some of them are surveyed in [2]. Constructibility differs slightly from computability; algorithms can be studied in either the classical or constructive setting.

2. DIFFERENT THEOREMS. Constructive mathematics, with its stricter notion of proof, proves fewer theorems than classical mathematics does. A mainstream mathematician who wishes to learn constructivism must go through his or her entire catalog of theorems, reevaluating each one by new criteria.

Perhaps the most blatant example of nonconstructive mathematics is the Axiom of Choice (AC). Here is one of its forms:

Let $\{S_\alpha : \alpha \in A\}$ be a nonempty collection of nonempty sets. Then there exists a function that chooses one member from each of those sets, i.e., a function f satisfying $f(\alpha) \in S_\alpha$ for each $\alpha \in A$.

To accept AC is to agree to the convention that we permit ourselves to use the choice function f in proofs, as though f "exists" in some sense. That agrees with the intuition of the classical mathematician: To "define" $f(\alpha)$, just "pick any member" of S_α .

But that description of f is nonconstructive: It doesn't produce a particular explicit example of f . A particular example can be produced, by means other than the axiom, in some cases but not others. For instance, we can take $f(\alpha)$ to be the smallest member of S_α , if the S_α 's are the nonempty subsets of \mathbb{N} . No analogous selection is available from the nonempty subsets of \mathbb{R} .

Besides choice functions, AC yields many other objects without examples. For instance, every introductory functional analysis textbook uses AC to prove the Hahn-Banach Theorem. That theorem is used in many textbooks to prove $(\ell^\infty)^* \not\cong \ell^1$, but no examples of members of $(\ell^\infty)^* \setminus \ell^1$ are given. It turns out that the examples are omitted because they *must* be omitted. By this we do not mean merely that no one has found an example yet. Rather, it is impossible for an explicit example to be constructed; see [9] for outline of a proof. That fact may be surprising because ℓ^∞ and ℓ^1 seem fairly concrete, but it is less surprising when we note that the definition of $(\ell^\infty)^*$ is more abstract. Some textbooks do not point out the lack of examples, but I think we should mention it to our students; the student who cannot visualize some object will be reassured upon hearing that no one else can visualize it either.

The newcomer to constructivism should not simply discard all nonconstructive theorems. Many of those theorems can be replaced by constructive variants. For instance, constructive variants of the Hahn-Banach Theorem are known for uniformly convex spaces [5] and for separable spaces [1]. Adding a separability hypothesis increases the theorem's complexity but does not greatly reduce its applicability, because a real-world application ultimately must rely on only finitely many measurements. Applied mathematics is sometimes performed in a nonseparable space such as ℓ^∞ or L^∞ , but that may be just for simplicity of notation. The relevant portion may be some separable subspace, which we may be able to identify with sufficient effort.

Newcomers to constructivism must reconsider even the most elementary facts about even the most familiar objects. For instance, the Trichotomy Law of classical analysis states that for any real numbers r and s ,

$$r < s \quad \text{or} \quad r = s \quad \text{or} \quad r > s.$$

But this law is not a part of constructive mathematics, because there is no algorithm with the property that

if we are given constructive descriptions of two real numbers r and s , the algorithm determines whether one number is larger than the other;

see [1] or [2]. However, for most applications the Trichotomy Law can be replaced by the following principle, which *can* be proved constructively:

$$\text{If } r, u, v \in \mathbb{R} \text{ with } u < v, \text{ then } r > u \text{ or } r < v.$$

3. DIFFERENT LOGIC. We begin this section with an old example, but some of our comments on it are new. We wish to prove:

Theorem A. *There exist positive, irrational numbers a and b such that a^b is rational.*

Here is a quick, easy proof:

- If $\sqrt{2}^{\sqrt{2}}$ is rational, use $a = b = \sqrt{2}$.
- If $\sqrt{2}^{\sqrt{2}}$ is irrational, use $a = \sqrt{2}^{\sqrt{2}}$ and $b = \sqrt{2}$; then $a^b = 2$.

This argument would convince most mathematicians that suitable numbers a and b “exist”, but we haven’t found a , so this proof is nonconstructive. (We have found b , but some constructivists would say that we still haven’t *constructed* b , because we don’t know which of the two cases holds. The words “find” and “construct” are often, but not always, used interchangeably.)

On the other hand, the same computations yield this constructive result:

Theorem B. *If we can constructively establish either of the statements “ $\sqrt{2}^{\sqrt{2}}$ is rational” or “ $\sqrt{2}^{\sqrt{2}}$ is irrational”, then we can construct positive irrational numbers a and b such that a^b is rational.*

Some theorems are inherently nonconstructive; for instance, $(\ell^\infty)^* \not\cong \ell^1$ cannot be given a constructive proof. For other theorems, however, a nonconstructive proof can be replaced by a constructive proof. For instance, here are two constructive proofs of our Theorem A:

- It is easy to show that $a = \sqrt{2}$ and $b = \log_2 9$ are both irrational, and $a^b = 3$.
- A long and complicated proof yields the Gel’fond-Schneider Theorem [10]: if q and r are algebraic numbers, $q \notin \{0, 1\}$, and r is irrational, then q^r is transcendental. In particular, $\sqrt{2}^{\sqrt{2}}$ is irrational.

Our first, nonconstructive proof of Theorem A was based on the *Law of the Excluded Middle* (LEM), which states “ P or not- P ” for every proposition P . That principle is a tautology of classical logic, but not of the logic of constructivism. But how can there be different logics? Isn’t logic just common sense? A formalist answer is that we’ve simply decided to use a different set of axioms for our logic. A more intuitive answer is that in classical logic, variables such as P represent *statements* that are true or false in some formal sense. In the logic of constructivism, the variables represent *procedures* (e.g., constructions or proofs) that we can carry out. The computational rules for combining procedures are similar to, but not identical to, the computational rules for combining statements.

Symbolic logic is of interest to logicians, but it is seldom of interest to other, “ordinary” mathematicians, e.g., analysts or topologists. Their prose reasoning is adequate and even preferable for their needs. This is just as true in constructive mathematics; for instance, formal logic is barely mentioned in Bishop’s book on constructive analysis [1]. However, a mainstream mathematician who wishes to learn constructivism may find it helpful to study the underlying formal logic, in order to become more consciously aware of the occurrences of LEM and other nonconstructive forms of reasoning.

The symbolic logic of constructivism was codified by Heyting and Kolmogorov, who named it *intuitionistic logic* because of its origins in Brouwer’s intuitionist philosophy. (My explanation would be simpler if the logic of constructivism were known as “constructive logic”, but that name has been given to a somewhat different, and more complicated, kind of logic; see [6].) An introduction to intuitionistic logic can be found in [3] and [4]. Its theorems are a proper subset of the theorems of classical logic. For instance, the formulas

$$P \vee \neg P, \quad (\neg\neg P) \rightarrow P, \quad (\neg\forall x Q(x)) \leftrightarrow (\exists x \neg Q(x))$$

are theorems of classical logic but not of intuitionistic logic.

4. DIFFERENT VOCABULARY. The newcomer to constructivism must learn new words, because the constructivist sees distinctions overlooked by the classicist. Analogously, eskimos and skiers have distinct words for several different kinds of snow; to the rest of us, snow is just snow. For instance, for two real numbers a and b , some constructivists distinguish between

- *inequality* ($a \neq b$)—we would obtain a contradiction by assuming $a = b$; and
- *apartness* ($a \# b$)—we can construct a positive integer k such that $|a - b| > 1/k$.

Old words must be relearned, because the constructivist viewpoint adds new requirements to many notions. For instance, suppose S is a set of real numbers, and r is a real number. To show constructively that $r = \inf(S)$, we must prove that $r \leq s$ for every $s \in S$, and we must also *construct* numbers $s_1, s_2, s_3, \dots \in S$ satisfying $r > s_k - 1/k$. It is not enough merely to show the classical “existence” of some s_k ’s with that property.

Perhaps the most difficult adjustment for the new constructivist is in the use of the word “or”. Its usage changes because the context is changed. What a constructivist means by “ P or Q ” usually is very close to what a classicist would mean by the longer statement

I can prove at least one of the two statements P , Q , and I know which one.

But if you already know which one, why do you bother to mention the other one at all? Well, in many applications the phrase “ P or Q ” is a hypothesis in some longer statement, so we don’t really “know which one” yet. The following Theorem B’ illustrates this situation.

In a book or paper written by constructivists for constructivists, certain abbreviations become appropriate. In that context, the phrases “we can establish” and “we are able to construct” are always understood and need not be mentioned. Hence Theorem B, stated earlier in this paper, can be stated more briefly as:

Theorem B’. *If $\sqrt{2}^{\sqrt{2}}$ is rational or $\sqrt{2}^{\sqrt{2}}$ is irrational, then there are positive irrational numbers a and b such that a^b is rational.*

Theorem B’ may bewilder mathematicians who are unfamiliar with constructivism, because they see its hypothesis as *trivially true* (by LEM), and therefore superfluous. See [7] for further use of “or” in this fashion.

5. CONCLUSION. Learning constructivism involves not only learning new concepts, but also unlearning old habits and preconceptions. Thus, constructivism may be absorbed more easily by younger mathematicians. We should introduce our students to it.

ACKNOWLEDGMENT. I am grateful to Professor Richman for comments on earlier versions of this paper.

REFERENCES

1. E. Bishop, *Foundations of Constructive Analysis*, McGraw-Hill, New York, 1967. Later revised as *Constructive Analysis* by E. Bishop and D. Bridges, Grundlehren der mathematischen Wissenschaften **279**, Springer-Verlag, Berlin, 1985.

2. D. Bridges and F. Richman, *Varieties of Constructive Mathematics*, London Math. Soc. Lecture Note Ser. **97**, Cambridge University Press, Cambridge, 1987.
3. D. van Dalen, *Logic and Structure*, 3rd ed., Universitext series, Springer-Verlag, Berlin, 1994.
4. A. G. Dragalin, *Mathematical Intuitionism: Introduction to Proof Theory*, Translations of Mathematical Monographs **67**, American Mathematical Society, Providence, 1988.
5. H. Ishihara, On the constructive Hahn-Banach theorem, *Bull. London Math. Soc.* **21** (1989) 79–81.
6. H. Rasiowa, *An Algebraic Approach to Non-Classical Logics*, Studies in Logic and the Foundations of Mathematics **78**, North-Holland, Amsterdam, 1974.
7. F. Richman, Meaning and information in constructive mathematics, *Amer. Math. Monthly* **89** (1982) 386–388.
8. F. Richman, Existence proofs, *Amer. Math. Monthly* **106** (1999) 303–308.
9. E. Schechter, *Handbook of Analysis and its Foundations*, Academic Press, San Diego, 1997.
10. R. Tijdeman, Hilbert's seventh problem: on the Gel'fond-Baker method and its applications, pp. 241–268 in *Mathematical Developments Arising from Hilbert Problems*, F. E. Browder, ed., *Proc. Sympos. Pure Math.* **28**, American Mathematical Society, Providence, 1976.

ERIC SCHECHTER received his Ph.D. from the University of Chicago in 1978. Most of his papers are on differential equations in Banach spaces. His interest in constructivism grew largely from a desire for a better understanding of the Hahn-Banach Theorem; his analysis book includes (among other things) 26 equivalents of that theorem. He is now working on a logic book. His most popular work (judging by the email response) is his web page about common errors in college mathematics; you can reach it from his home page at math.vanderbilt.edu/~schectex.

Vanderbilt University, Nashville, TN 37240
schectex@math.vanderbilt.edu