Existence Proofs

Fred Richman

The American Mathematical Monthly, Vol. 106, No. 4. (Apr., 1999), pp. 303-308.

Stable URL:
http://links.jstor.org/sici?sici=0002-9890%28199904%29106%3A4%3C303%3AEP%3E2.0.CO%3B2-E

The American Mathematical Monthly is currently published by Mathematical Association of America.

Existence Proofs

Fred Richman

> [If] a proof convinces you that there is a root of an equation (without giving you any idea *where*)—how do you know that you understand the proposition that there is a root? Ludwig Wittgenstein

The proposition that mathematics is grounded on computation would seem to be quite uncontroversial; the only people I have heard argue against it are mathematicians. By "mathematics" I mean pure mathematics—theorems and proofs. Many pure mathematicians think that they engage in a high art form that is incompatible with strong links to computation, the nerdy province of bookkeepers, statisticians, and calculators. This attitude goes hand in glove with the practically unquestioned acceptance of nonconstructive existence proofs in modern mathematics.

Constructive mathematicians are unsatisfied by nonconstructive existence proofs: proofs that attempt to convince you of the existence of a number, or of some more complicated mathematical object, without giving any method for computing it. The difference between constructive mathematics and classical mathematics is that when a constructive mathematician says there is a number that satisfies a given equation, or has some other property, he has an algorithm in his pocket for computing that number. The pocket of a classical mathematician, who makes the same statement, might contain only a derivation of a contradiction from the assumption that every number fails to have the property.

The quotation from Wittgenstein [**12**, p. 282] at the head of this article suggests that the peculiar nature of such a proof should cause us to reconsider the meaning of the proposition it proves. Did we really understand what was meant by "there is a root" if we can be convinced of its truth by an argument that does not provide a method for finding a root? What are these numbers that exist without our being able to construct them? I am reminded of the Nobel laureate Eugene Wigner's response to the question of whether there are any inherently unknowable laws of physics: he said, "I don't know of any."

I will illustrate the idea of a nonconstructive proof by several examples. The first is of a theorem that admits a celebrated constructive proof (the Euclidean algorithm). Many textbooks give both proofs, but the constructive one is usually presented as a method of computation rather than a proof.

Theorem 1. *There exist integers s and t such that $437s + 323t$ is positive and divides 437 and 323.*

Of course the numbers 437 and 323 are not special; I choose them to make sure we see immediately which are the constants and which are the variables. It follows immediately from the theorem that the number $437s + 323t$ is the greatest common divisor of 437 and 323. Here is a nonconstructive proof of the theorem.

Proof: Consider the set of positive integers

$$S = \{d : d > 0 \text{ and } d = 437s + 323t \text{ for some } s \text{ and } t\}.$$

The set S is nonempty because 437 and 323 are clearly in it. Therefore, by what is often called the "well-ordering principle," there is a least element d_0 of S. As $d_0 \in S$ we can write $d_0 = 437s_0 + 323t_0$. The division algorithm enables us to write $437 = qd_0 + r$, where $0 \leq r < d_0$. If $r > 0$, then $r = 437(1 - qs_0) + 323(-qt_0)$ is in S. But d_0 is the smallest element of S, so $r = 0$. Thus d_0 divides 437; similarly d_0 divides 323. ∎

That was a proof in the spirit of the Wittgenstein quotation. What are these integers s and t that we have shown to exist? We have not been given a clue as to how to find them. They were constructed by choosing the smallest element of the set S. But how do we find that smallest element? What is the basis for the well-ordering principle?

A proof of the well-ordering principle, as applied to a set S of positive integers that contains 323, might go as follows. If $1 \in S$, then 1 is the smallest element of S. If $1 \notin S$, but $2 \in S$, then 2 is the smallest element of S. If $1 \notin S$ and $2 \notin S$, but $3 \in S$, then 3 is the smallest element of S. We can write this as a computer program as follows: for $i = 1$ to 323 do if $i \in S$ return i.

What does this proof of the well-ordering principle prove? It proves that a nonempty detachable subset of the positive integers has a least element. A subset S of a set X is called *detachable* if you can tell (there is an algorithm for telling) whether or not any given element of X is in S. The problem with the set S in the proof of Theorem 1 is that we have not established that it is detachable; for example, how do we decide whether or not $1 \in S$? In fact S is detachable, but to prove that we usually invoke Theorem 1!

A constructive proof of Theorem 1 might go like this. Consider the following table.

437	1	0
323	0	1
114	1	−1
95	−2	3
19	3	−4

Each row represents values of d, s, and t such that $d = 437s + 323t$. This is clear for the first two rows. Each subsequent row R_{n+1} is computed from the previous two rows R_{n-1} and R_n by setting $R_{n+1} = R_{n-1} - m_n R_n$. We choose the m_n so as to make the first entry of R_{n+1} positive, and as small as possible. The first entries of the rows must decrease, unless the first entry of R_n divides the first entry of R_{n-1}, in which case we stop. The equation $d = 437s + 323t$ is inherited by R_{n+1} from R_n and R_{n-1}. Moreover, as $R_{n-1} = m_n R_n + R_{n+1}$, any number that divides the first entries of two consecutive rows, divides the first entry of the row before them. So 19 divides all the numbers above it; in particular, 437 and 323.

Errett Bishop, the mathematician most responsible for the recent renaissance of constructive mathematics, formulated four principles of constructive mathematics in [**2**]:

(A) *Mathematics is common sense.*
(B) *Do not ask whether a statement is true until you know what it means.*
(C) *A proof is any completely convincing argument.*
(D) *Meaningful distinctions deserve to be maintained.*

Constructivists think that a proof of the existence of a mathematical object should tell you how to construct that object; this follows from their belief that other interpretations of the phrase "there exists," in a mathematical context, are either incomprehensible, or can be formulated in more descriptive ways (principles B and D).

These are controversial, even revolutionary, ideas. Constructivists want to make fundamental changes in the way we view mathematics; they want to change the rules by which the game of mathematics is played. But most people aren't interested in changing the rules. For one reason, most people like the rules as they are. In fact, successful mathematicians have a vested interest in keeping the rules as they are. Why should a champion chess player be interested in changing the rules of chess? Another reason is that, after all, aren't the present rules of mathematics correct? How can there be any serious alternatives?

Constructive mathematics is not just an idea but a substantial body of results. In [**1**] Bishop developed much of analysis along constructive lines. This classic book has been revised and extended [**3**]. A corresponding constructive development of algebra was carried out in [**9**]. For expository articles on constructive mathematics see [**4**], [**5**], [**8**], [**10**], and [**11**].

Let's look at another example: consider the decimal expansion

$$\pi = 3.14159265358979323846264338327 95\ldots.$$

This decimal expansion suits our purposes because it is a familiar example of a computable infinite sequence, and very little is known about it. The theorem I want to consider says that

Theorem 2. *There is a digit that appears infinitely often in the decimal expansion of* π.

How can we establish such a theorem in light of the fact that so little is known about the decimal expansion of π? As you may have already realized, the proof that I have in mind gives you no idea as to which digit appears infinitely often. It is a genuine indirect proof, as opposed to the usually cited examples of proof by contradiction—proofs that $\sqrt{2}$ is not rational, or that the number of primes is not finite—where the very meaning of the theorem requires that a contradiction be derived.

Let's look at a proof of this theorem.

Proof: Suppose each digit occurs only a finite number of times in the decimal expansion of π: so 0 occurs n_0 times, 1 occurs n_1, times, etc. Then compute

$$n_0 + n_1 + \cdots + n_9 + 1$$

places in the decimal expansion of π. But there are only $n_0 + n_1 + \cdots + n_9$ digits available to fill up these places, which is absurd, so our original assumption that each digit occurred only finitely many times must be false. ■

What has been proved? Let P_i denote the proposition that the digit i appears (only) a finite number of times in the decimal expansion of π. Then we have shown that the proposition

$$A = P_0 \text{ and } P_1 \text{ and } \cdots \text{ and } P_9$$

leads to a contradiction. That is, we have proved the negation $\neg A$ of the proposition A. What we wanted to show, on the other hand, was that $\neg P_i$ holds for some digit i, that is we wanted to prove the proposition

$$B = \neg P_0 \text{ or } \neg P_1 \text{ or } \cdots \text{ or } \neg P_9.$$

According to the usual laws of logic (DeMorgan's law), B is the same as $\neg A$, which we *have* proved. The constructivist grants that $\neg A$ has been proved, but denies that B has been proved. He wants to draw a distinction between B, which asserts the existence of a digit with the property $\neg P$, and $\neg A$, which merely says that it is impossible for all digits to have the property P (*meaningful distinctions deserve to be maintained*). So we see what the rules are that constructivists want changed: no less than the rules of logic.

The controversy over nonconstructive techniques in mathematics goes back at least to the beginning of this century. David Hilbert and L.E.J. Brouwer were the principal participants in this controversy. Brouwer founded the philosophy of mathematics called *intuitionism* [**6**]. Most constructive mathematicians, of whatever school, consider Brouwer to be a spiritual ancestor. Hilbert was the foremost mathematician in Germany, and possibly in the world, in the early twentieth century. Many American mathematicians trace their mathematical lineage to Hilbert because of the German mathematicians who settled in the United States prior to World War II, and the Americans who went to Germany in the early part of the century to study with Hilbert and his students.

Hilbert used nonconstructive techniques to solve a well-known problem concerning the construction of a finite set of polynomials with certain properties. The problem had been solved by P. Gordan, for the case of polynomials in two variables, by explicitly exhibiting the finite set of polynomials. Hilbert solved the general case, but his proof gave no clue how to construct the required polynomials. This appears to be the first use of a nonconstructive proof to establish the existence of mathematical objects that were expected to be constructed explicitly. The reaction of Gordan, perhaps in reference to proofs of the existence of God, was "That's not mathematics, that's theology."

Since then, Hilbert's approach has so dominated mathematical thinking that alternatives are not considered seriously. But Hilbert and Brouwer had one thing in common: they both thought that nonconstructive techniques needed justification. Brouwer thought that the answer was to use only constructive techniques. Hilbert did not want to abandon his nonconstructive techniques; instead, he proposed to show that you couldn't get into trouble using them.

Hilbert's program was to show that if a theorem proved by nonconstructive means predicted the result of a computation, then it would predict the correct result. Thus even if the smallest digit that occurred infinitely often in the decimal expansion of π were just a fiction, we could treat it as reality and still never say anything that was verifiably false. An analogy is the introduction of a square root of -1, which in some sense is simply a fiction, but it helps us to prove things about, and to discover properties of, the numbers that we actually believe in.

Of course the proof that you don't get into trouble using nonconstructive techniques must use only constructive techniques in order to be convincing. Hilbert's program was utterly demolished by Kurt Gödel, who showed in the thirties that not only couldn't you prove such a thing constructively, but you couldn't even prove it using only the nonconstructive techniques that you were attempting to justify. Remarkably, this had no effect on the acceptability of nonconstructive techniques!

My final example of a nonconstructive existence proof comes from the theory of computable functions. That theory was developed to clarify the idea of what it means for a function to be computable. This is not an idea that would occur to a constructivist, for whom computability is part of the intuitive idea of a function: if there exists y such that $f(x) = y$, then we must be able to construct that y if we

are given x. A constructivist can certainly entertain the idea of looking at a restricted class of functions that are computed in a special way, but would be unlikely to call that class "the computable functions."

In the theory of computable functions, a function from the natural numbers $0, 1, 2, \ldots$ to the natural numbers is called *computable* if there exists a computer program to compute it. It is well known that, for all but the most anemic programming languages, this notion is independent of the particular language used. In the standard theory, nonconstructive proofs of the existence of a program are allowed.

It will be convenient to abbreviate the following statement by P_n.

P_n : There are (at least) n consecutive 4's in the decimal expansion of π.

Consider the function f defined by setting $f(n) = 1$ if P_n and $f(n) = 0$ otherwise. Clearly $f(0) = f(1) = 1$. I asked my computer to give me 100 places of π, and I see that there are 4's in places 59 and 60 (if I haven't miscounted), so $f(2) = 1$. But what, for example, is $f(12)$? No one knows; no one even knows of a computation that would resolve this question. If you compute a billion places in the expansion of π you might discover that $f(12) = 1$, but no such computation would tell you that $f(12) = 0$. A constructivist would say that we have not defined f for all n because we have not shown how to compute $f(n)$ in general.

The orthodox view, however, is that not only have we defined f for all n, but that f is computable! To verify the first claim, consider the set

$$G = \{(n, i) : i = 1 \text{ and } P_n, \text{ or } i = 0 \text{ and not } P_n\}.$$

This is the graph of f. To show that f is defined at n means to show that there exists i in the set $\{0, 1\}$ so that (n, i) is in G. It is clear that if $(n, 1)$ is not in G, then $(n, 0)$ is in G, because either statement is equivalent to denying P_n. But this constitutes a nonconstructive proof that there exists i for which (n, i) is in G, as it is impossible that neither $(n, 1)$ nor $(n, 0)$ is in G.

So suppose that f is defined for all n—the constructivist will have to imagine that he can consult an oracle to determine the value of $f(n)$. Now we want to show that there is a program that computes f. This requires a second nonconstructive argument. The gimmick is that we do not have to produce the program; we merely have to show that f cannot be different from each computable function.

Consider the following collection of functions, one for each natural number m:

$$g_m(n) = \begin{cases} 1, & \text{if } n < m \\ 0, & \text{if } n \geq m, \end{cases}$$

together with the function g_∞ which is identically equal to 1. Certainly g_∞ is computable, and g_m is computable for each finite m: the program simply compares n with the fixed number m and returns 1 or 0 as appropriate. Suppose $f(n) = 0$ for some n. Let m be the first place where $f(m) = 0$. Then $f = g_m$ because $f(x) \geq f(y)$ whenever $x \leq y$. Thus if $f \neq g_m$ for each finite m, then $f(n)$ cannot be 0 for any n, so $f = g_\infty$. Therefore f cannot be different from each computable function. That's as far as a constructivist can go, even with an oracle that tells him what $f(n)$ is.

The final (nonconstructive) step in the argument is to apply an infinite version of the DeMorgan's law used in the second example. We have shown that the proposition

$$f \neq g_\infty \text{ and } f \neq g_0 \text{ and } f \neq g_1 \text{ and } \cdots$$

is false. We conclude that the proposition

$$f = g_\infty \text{ or } f = g_0 \text{ or } f = g_1 \text{ or } \cdots$$

is true—if a collection of propositions cannot all be false, then one of them is true. So there exists m such that $f = g_m$, whence f is computable.

I will let Wittgenstein have the last word. Georg Cantor laid the foundation for the theory of transfinite sets upon which twentieth century mathematics is based. Hilbert [7], referring to objections of the intuitionists, said "No one shall drive us out of the paradise which Cantor created for us." Wittgenstein [**13**, p. 103] later wrote in response to this,

> I would say, "I wouldn't dream of trying to drive anyone out of this paradise." I would try to do something quite different: I would try to show you that it is not a paradise—so that you'll leave of your own accord. I would say, "You're welcome to this; just look about you."

REFERENCES

1. Bishop, Errett, *Foundations of constructive analysis*, McGraw-Hill, 1967.
2. _____, Schizophrenia in contemporary mathematics, American Mathematical Society, 1973, in *Errett Bishop: Reflections on him and his research*, AMS Contemporary mathematics series, volume 39, 1985.
3. _____, and Douglas Bridges, *Constructive analysis*, Springer-Verlag, 1985.
4. Bridges, Douglas, and Ray Mines, What is constructive mathematics?, *The Mathematical Intelligencer* **6** (1984) 32–38.
5. Calder, Allan, Constructive mathematics, *Scientific American*, Oct. 1979, 146–171.
6. Heyting, A., *Intuitionism, an introduction*, North-Holland, 1971.
7. Hilbert, David, On the infinite (1925), in *Philosophy of mathematics: selected readings*, Paul Benacerraf and Hilary Putnam (eds.), Cambridge University Press, 1983.
8. Mandelkern, Mark, Brouwerian counterexamples, *Math. Mag.* **62** (1989) 3–27.
9. Mines, Ray, Fred Richman, and Wim Ruitenburg, *A course in constructive algebra*, Springer-Verlag, 1988.
10. Richman, Fred, Meaning and information in constructive mathematics, *Amer. Math. Monthly*, **89** (1982) 385–388.
11. _____, Interview with a constructive mathematician, *Modern Logic* **6** (1996) 247–271.
12. Wittgenstein, Ludwig, *Remarks on the foundations of mathematics*, (1942–1944), edited by G.H. von Wright et al., MIT Press, 1978.
13. _____, *Wittgenstein's lectures on the foundations of mathematics, Cambridge* 1939, edited by Cora Diamond, Cornell University Press, 1976.

FRED RICHMAN has an AB from Princeton University and a PhD from the University of Chicago. After teaching at New Mexico State University for many years, dividing his research time between infinite abelian group theory and constructive mathematics, he worked off and on at TCI Software Research as a developer of Scientific Word and Scientific WorkPlace. He has coauthored texts on modern algebra, mathematics for liberal arts students, constructive algebra, and varieties of constructive mathematics.

Florida Atlantic University, Boca Raton FL 33431
richman@fau.edu