Constructive Mathematics Without Countable Choice

1 Constructive mathematics

By "constructive mathematics" I mean doing mathematics in the context of intuitionistic logic, that is, without the law of excluded middle. Why should anyone be interested in constructive mathematics, let alone constructive mathematics without choice? The traditional reason is that the law of excluded middle is the source of the nonconstructive aspects of mathematics, so rejecting it allows us to think about mathematics algorithmically without thinking about algorithms.

From a foundational point of view, constructive mathematics, with its modest but substantial body of results, dealing with subjects ranging from measure theory, complex function theory and operators on Hilbert spaces to torsion-free abelian groups, provides a living alternative to the prevailing metaphysics in mathematics. For the practicing mathematician, constructive mathematics is a way of looking at mathematics from a fresh perspective, focusing attention on interesting questions that slip through the cracks of a classical treatment.

For teachers and students, constructive mathematics, with its built-in emphasis on direct arguments, encourages better exposition and understanding. Proof by contradiction is an elaborate dance where everything turns out to be an illusion. Absurdities, instead of indicating faulty reasoning, are actively sought, and any mistake you make helps you toward your goal. Constructive mathematics also highlights the difference between construction and verification, a useful distinction in teaching how to prove. It constantly suggests the question, "What am I given, and what do I have to construct?"

2 Why not countable choice?

The full axiom of choice implies the law of excluded middle [6], [11], so is incompatible with constructive mathematics. The countable axiom of choice (indeed, the axiom of dependent choices), on the other hand, is accepted by all serious schools of constructive mathematics, from Brouwer's intuitionism to Markov's algorithmic mathematics. Why, then, would anyone reject countable choice?

- **The arguments in favor of countable choice are weak.**

Troelstra and van Dalen [23] argue that countable choice follows from the BHK interpretation (Brouwer, Heyting, Kolmogorov) of the logical constants. Under that interpretation, a proof of "$P(x)$ for all x in D" is a construction which transforms each proof that d is in D into a proof of $P(d)$. Troelstra and van Dalen claim that "a method to find, for each m in $\mathbf{N}$, $\ldots$ is nothing else but a description of a function." Of course that's not true for domains other than $\mathbf{N}$, like the real numbers. They explain that by saying that for an arbitrary domain D the proof of $P(d)$ may depend on the proof that d is in D, not just on d itself, so equality might not be respected. But for $\mathbf{N}$ they say that each natural number "represents its own proof of belonging to $\mathbf{N}$" and "a natural number is given to us as such, we do not need a separate proof of this fact."

These last two comments are crucial for the argument, yet it is not clear what to make of them. There is the problem of how to recognize when an element of a set is "given to us as such." Is a sequence of natural numbers, like $a_n = n$, given to us as such? Is the natural number $7^{10^{10}}$ given to us as such?

A constructivist who accepts countable choice must distinguish sharply between full choice, which is rejected, and countable choice, which is accepted. But try explaining to a group of algebraists why countable-rank free modules are projective, while arbitrary free modules are not. I am inclined to go along with the sentiment expressed by Lebesgue in a letter to Borel [10, page 316], "I agree completely with Hadamard when he states that to speak of an infinity of choices without giving a rule presents a difficulty that is just as great whether or not the infinity is denumerable."

Arguments for countable choice are often indistinguishable from arguments for choice. Bishop [1] said that "A choice function exists in constructive mathematics, because a choice is *implied by the very meaning of existence*." He did not argue for countable choice explicitly; rather he employed such subterfuges as writing "for each n there exists an element a_n in A" and then proceeding to treat a_n as a sequence.

Martin-Löf, in his discussion of the axiom of choice in [8], which includes a proof of that axiom in his system, never uses the word "countable." He proves the full axiom of choice. How can that be? The usual explanation is that he is talking about intensional choice functions—the same defense used for Bishop's remark—and that for the natural numbers, there is no distinction between intensional and extensional functions (see Troelstra and van Dalen's argument above). But Martin-Löf does not make that case, and from his comments it is clear that he is talking about extensional choice functions; otherwise his theorem would have no consequences in constructive practice, as he obviously thinks it does. Of course we know that neither Bishop nor Martin-Löf can be asserting what they so clearly seem to be saying, but these fuzzy arguments indicate the difficulty of drawing a line between countable choice and full choice.

Followers of Bishop argue that neither they nor Bishop use countable choice in their arguments. In a note to Chapter 2 of [1], Bishop wrote, "The 'choice' involved in the proof of (b) of Proposition 8 is performed according to a definite rule." The point, however, is that if you routinely pass from "for all m there exists n such that $P(m, n)$" to "there exists f such that $P(m, f(m))$ for all m," then you have accepted the countable axiom of choice, regardless of what rationale you offer for your actions.

Martin-Löf in [8] claims that, "The need for the axiom of choice is clear when developing intuitionistic mathematics at depth, for instance, in finding the limit of a sequence of reals ...". That's only true if a real number is defined to be a Cauchy sequence of rational numbers (a Cauchy real). However, this is the wrong definition: the essence of a real number is that it be approximable by rational numbers, not that it admit a *sequence* of rational approximations. If we adopt the more general notion and define a real number to be, say, a located Dedekind cut, then completeness can be proved without choice—a completeness that is stronger than sequential completeness. The idea is simply that if you have a coherent way to approximate something by real numbers, then you have a coherent way to approximate it by rational numbers (see [17]).

- **It's natural to reject countable choice**

Our computational experience suggests that we should operate within intuitionistic logic. It also suggests that we reject countable choice. Computation is a transformation of data, and the given data must be taken as it comes. We access the mechanisms that produce the data only through the interfaces to them.

Think of a procedure which returns rational approximations to some real number r. Upon being given a positive integer m, it returns a rational number q that is within $1/m$ of r. This procedure need not return the same rational q each time it is presented with the integer m: unknown to us, it could take some other variable into consideration. All that is required of it is that if it ever returns q when presented with m, and q' when presented with m', then $|q - q'| \leq 1/m + 1/m'$. This implementation of a real number does not allow us to construct a Cauchy sequence converging to it (but countable choice would).

This natural computational situation emphasizes a characteristic feature of constructive mathematics: the interplay between the mathematician and the mathematical universe. We are given certain data, some of it in the form of interfaces to (seemingly nondeterministic) procedures with certain properties, and we carry out computations from that point. In this setting there is a clear distinction between a procedure that accepts a natural number m and returns a rational number q with certain properties, and a function that returns a rational number q_m for each natural number m, exactly the distinction that the countable axiom of choice obliterates.

- **The mathematics is better without countable choice**

That is, you formulate and prove better theorems—it's not just that the theorems hold in greater generality because they are proved with fewer assumptions. The prevalence of separability hypotheses and sequential arguments in constructive mathematics is due in large part to the use of the axiom of countable choice. Without this axiom, there is very little that you can do with sequences that you cannot do more generally. Rejecting countable choice forces you to formulate things better, and it makes separability hypotheses pointless for lack of consequences.

In the preface to [1], Bishop says that he is guided by three basic principles, the third of which is "to avoid pseudogenerality," so "separability hypotheses are freely employed." I would suggest that pointless hypotheses are worse than pseudogenerality. Bishop *defines* a Hilbert space to be separable. However, the separability hypothesis doesn't buy much without countable choice. For example, Bishop proves

HSB: Every Hilbert space has a basis.

The proof requires separability and countable choice. This is a fairly useless theorem, like the theorem that every pointwise continuous function on a closed interval is uniformly continuous which Bishop does quite well without (presumably pointwise continuity is also pseudogenerality). One doesn't need HSB to construct bases for the standard Hilbert spaces, like L^2, or even for some nonseparable Hilbert spaces (see below). To be sure there are applications of HSB, such as the closure of the range of an operator has a basis, but one could argue that these applications lead nowhere, and that HSB itself is an example of pseudogenerality. If you want to prove something about Hilbert spaces with bases (countable or otherwise), then include the existence of a basis as a hypothesis.

A side benefit of considering arbitrary bases in Hilbert spaces is that there is no temptation to allow basis elements to be zero, as is done in the sequential treatment in [1]. To avoid that in the sequential treatment, you have to index the basis elements by a detachable subset of $\mathbf{N}$; in the general treatment, a basis is already just a set.

In [4, Theorem 3] the Riesz representation theorem for linear functionals with norms is proved without constructing a sequence, in an arbitrary Hilbert space. (The proof in [2, 2.3 page 419] relies on a familiar kind of sequential argument that requires countable choice.) In the same paper, conditions for a bounded operator to have an adjoint are established in a Hilbert space with a (not necessarily countable) basis. Examples of Hilbert spaces with bases that need not be countable are obtained by taking an arbitrary subset S of the integers and considering $\ell^2(S)$, the space of square-summable complex functions on S. To do this requires a notion of the sum of an arbitrary family of nonnegative real numbers, but that is straightforward.

Of course we don't want to assume the existence of a basis if we don't have to. A basis gives a nicely organized system of finite-dimensional subspaces, but in the absence of a basis, we can still refer to finite-dimensional subspaces. This idea was exploited in [5] to define trace-class operators on a Hilbert space H (actually, for operators from H to K). Instead of assuming that H has a countable basis (e_n), postulating an absolute value $|T|$ for the operator (or an adjoint), and demanding that $\sum \langle |T|\, e_n, e_n \rangle$ converge, a single supremum

$$\|T\|_1 = \sup_{\substack{e,f \in F_k(H) \\ k \in \omega}} |\langle Te_1, f_1 \rangle| + \cdots + |\langle Te_k, f_k \rangle|$$

is considered, where $F_k(H)$ is the set of families of k orthonormal vectors in H. If that supremum exists, then T is of trace class. It follows (with some work) that T has an adjoint, and hence an absolute value.

Along the same lines, we can say that an operator $T : H \to K$ has singular values if the supremum

$$\|T\|_1^{(n)} = \sup_{\substack{e \in F_k(H) \\ f \in F_k(K) \\ k \leq n}} |\langle Te_1, f_1 \rangle| + \cdots + |\langle Te_k, f_k \rangle|$$

exists for each n. The singular values of T are then defined by $\sigma_n = \|T\|_1^{(n)} - \|T\|_1^{(n-1)}$. Compact operators have singular values, and one can prove an approximate singular value decomposition theorem. An exact decomposition is not possible, even in the finite-dimensional case, for the same reason that the exact principle axes theorem does not admit a constructive proof [21].

In algebra, countable choice is used to prove that countable rank free modules are projective. This comes up in proving the (abstract) Hilbert syzygy theorem on the connection between projective dimension of modules over a ring R and modules over the ring of polynomials $R[X]$. The polynomial ring $R[X]$ is a countable rank free R-module and its projectivity over R is central to the proof. Avoiding countable choice leads to a better treatment. The real theorem is about *flat* dimension, a much more tractable concept than projective dimension from a constructive point of view. There is no problem in proving that free modules are flat. The abstract Hilbert syzygy theorem becomes

> $\operatorname{fd} A \leq n$ *for all R-modules A, if and only if* $\operatorname{fd} A \leq n+1$ *for all $R[X]$-modules A*

(see [14]). One need not restrict to finitely presented A, as seems to be the case for projective dimension. Moreover, if R is coherent, then the projective dimension of a finitely presented module is the same as its flat dimension, so the concrete Hilbert syzygy theorem follows from this version of the abstract theorem.

3 Is it feasible?

The fundamental theorem of algebra provided a good test for the feasibility and desirability of constructive mathematics without choice. Choice (or the law of excluded middle) is required even to find the square root of an arbitrary complex number a. The problem is deciding, in advance, which branch to follow if you discover that a is nonzero. From a classical point of view, the difficulty is that the function z^2 does not admit a continuous section in a neighborhood of 0. Thus there are sheaf models of the complex numbers in which the fundamental theorem of algebra is false.

The solution here is to rethink the meaning of the fundamental theorem of algebra [17]. Given a monic polynomial f, we can construct, for each n, a set of Gaussian numbers that approximate the roots of f within $1/n$ in the following sense.

- If S is a $1/m$ approximation, T is a $1/n$ approximation, and $\varepsilon > 0$, then there is a one-to-one correspondence between S and T so that corresponding elements differ by at most $1/m + 1/n + \varepsilon$.
- The polynomial $\prod_{s \in S}(x - s) \to f(x)$ as $n \to \infty$.

In any practical sense, this is a completely satisfactory method for constructing the roots of f. The spectrum of f is a limit of finite sets of complex numbers, but need not itself be a set of complex numbers.

This development of the fundamental theorem of algebra indicates that choiceless constructive mathematics may be feasible and satisfying. The problems raised by choice avoidance call attention to a real distinction between an approximate complete factorization, and an approximation to a complete factorization. Also, one is essentially forced to show that the set of roots of a polynomial over the algebraic numbers (so you *can* construct its roots) is a continuous function of the coefficients, an interesting theorem that is bypassed by traditional constructive treatments (but see Weyl's intuitionistic proof [24]).

It should be noted that Ruitenburg gave a choiceless proof of the fundamental theorem of algebra using Cauchy reals [20]. The difficulty with that approach is that we construct the reals because the rationals are not complete, but then neither are the Cauchy reals! Troelstra and van Dalen [23] are a little misleading on this point. By *defining* a Cauchy sequence of real numbers to be a Cauchy sequence of Cauchy sequences of rational numbers, they can prove the reals are complete without using countable choice. But this definition makes no sense in an arbitrary metric space, and they do remark shortly thereafter that with the natural definition of a Cauchy sequence of real numbers, countable choice is required to prove completeness.

4 Proposed research

4.1 Measure theory and the spectral theorem

These are major challenges for a choiceless development, and may well constitute a large part of the proposed research. Constructive measure theory is pervaded with sequences and appeals to choice. However, an indication that a choiceless approach could succeed is the fact that L^1 of a compact interval can be constructed by completing (not sequentially!) the space of uniformly continuous functions with the L^1 norm. The measurable sets will at worst be represented by the idempotents in the algebra $\bigcup_n L_n^1$ of bounded elements of L^1, where L_n^1 is the completion of the n-bounded uniformly continuous functions under the L^1 norm.

It remains to be seen whether this optimistic view stands up when the details of constructive measure theory are confronted. Countable choice is required in classical measure theory to prove countable additivity, as there are models in which the reals are a countable union of countable sets. But a similar problem caused no difficulty in constructive measure theory *with* countable choice: the recursive reals are a model in which the unit interval can be covered by a sequence of intervals whose total length is bounded by ε (but the partial sums do not converge).

Related to this are the spectral theorem and functional calculus for commuting Hermitian operators on a Hilbert space [1, Thm. 8, Ch. 9]. Here again, the current theory is pervaded by sequences and choice. There is also the question of the role of separability. We can expect a choiceless development of this theory to be accompanied by some surprising insights and a gain in clarity.

4.2 Noetherian rings

Countable choice comes up in showing that an arbitrary image B/A of a Noetherian module B is Noetherian. The usual constructive formulation of the Noetherian condition is in terms of ascending sequences of finitely generated submodules [12], [22], [9]. The problem is in lifting an ascending sequence of finitely generated submodules from B/A to B. If A is finitely generated, which is usually the situation, you can just take complete inverse images; otherwise you need choice.

It's not clear what the proper treatment here is. Perhaps this theorem for an arbitrary submodule A is simply pseudogenerality. On the other hand, if we are trying to break the sequence habit, perhaps we have the wrong definition of Noetherian. An alternative definition of Noetherian, in a constructive context, is based on the idea of Noetherian induction. This definition was developed in [7]. It gets around choice (although that was not its purpose, as far as I know) but constitutes quite a conceptual break with the more usual definition as a chain condition. Further study of the uses of the Noetherian condition in practice will be required to sort this out.

I also intend to pursue the constructive development of the general theory of commutative rings, particularly Noetherian ones. In [15] I showed how to calculate maximal regular sequences, a basic construction in commutative ring theory, for finitely presented algebras over discrete fields, and for coherent, Noetherian, strongly discrete rings that contain an infinite field. This is a large class of rings, but one would really like to have a more general condition not involving a subfield.

Prime ideals play a central role in commutative ring theory. This is awkward for a constructive development because prime ideals are often difficult or impossible to construct, even in as simple a situation as a polynomial ring in one variable over a discrete field. The Krull dimension d may be formulated in terms of the Hilbert characteristic function, allowing it to be computed for a coherent local ring R with a finitely generated maximal ideal M without having to construct prime ideals. An interesting question that arises immediately is whether you can construct a system of parameters for R—a set of d elements of M that generates an ideal containing some power of M.

4.3 Dimension theory

The constructive treatment of covering dimension in [18] relies heavily on countable choice via Theorem 8 of Chapter 4 in [1], which reads

($*$) *If f is a continuous real function on a (pre)compact set, then for all but countably many $r > \inf f$, the set $f^{-1}((-\infty, r])$ is (pre)compact.*

Because ($*$) was available, and because precompact subsets seemed "more constructive" than arbitrary subsets, attention in [18] was restricted to precompact subsets both for covering and being covered. This necessitated continual appeals to ($*$) to enlarge sets slightly to make them precompact.

Rejecting choice requires abandoning $(*)$, but here that may result in a more elegant and straightforward development, with a gain in generality and perspicuity. In conjunction with the formalism developed in [16], which allows us to work with distances to nonlocated sets (infima of bounded sets of real numbers) in much the same way we work with real numbers, we should be able to deal with arbitrary coverings in a natural and constructive way.

4.4 Compactly generated Banach spaces are finite dimensional

Countable choice plays a dominant role in the proof of this curious theorem [19], so it would be an interesting testing ground for choice avoidance. The simplest case is showing that if a generates a closed vector subspace $\mathbf{R}a$ of the real numbers $\mathbf{R}$, then $a = 0$ or $a \neq 0$. The proof of this is improved by avoiding the choice argument of [13]. You start by noting that $\sqrt{|a|}$ is in the closure of $\mathbf{R}a$ (square roots of real numbers are unproblematic—you can find one in the first quadrant), and make the natural moves from that point on. Essentially the same argument proves the theorem for any cyclic Banach space, and for finitely generated Banach spaces it can be modified to show that one of the generators is zero or nonzero, the theorem following by induction. The general case is less clear, and conceivably requires choice.

References

[1] BISHOP, ERRETT, *Foundations of constructive analysis*, McGraw-Hill, 1967.

[2] BISHOP, ERRETT AND DOUGLAS BRIDGES, *Constructive Analysis,* Grundlehren der math. Wissenschaften **279**, Springer-Verlag, 1985.

[3] BRIDGES, DOUGLAS, FRED RICHMAN AND PETER SCHUSTER, Linear independence without choice, (preprint, `www.math.fau.edu/Richman/html/docs.htm`, submitted to *Annals of Pure and Applied Logic*).

[4] BRIDGES, DOUGLAS, FRED RICHMAN AND PETER SCHUSTER, Adjoints, absolute values and polar decompositions (preprint, `www.math.fau.edu/Richman/html/docs.htm`).

[5] BRIDGES, DOUGLAS, FRED RICHMAN AND PETER SCHUSTER, Trace-class operators (preprint, `www.math.fau.edu/Richman/html/docs.htm`).

[6] DIACONESCU, R., Axiom of choice and complementation, *Proc. Amer. Math. Soc.*, **51**(1975) 176-178.

[7] JACOBSSON, C., AND C. LÖFWALL, Standard bases for general coefficient rings and a new constructive proof of Hilbert's basis theorem, *J. Symbolic Comput.* **12**(1991), 337–371.

[8] MARTIN-LÖF, PER, *Intuitionistic type theory*, Bibliopolis, Naples 1984.

[9] MINES, RAY, FRED RICHMAN AND WIM RUITENBURG, *A course in constructive algebra*, Springer-Verlag 1988.

[10] MOORE, GREGORY H., *Zermelo's axiom of choice, its origins development, and influence*, Springer-Verlag, 1982.

[11] MYHILL, JOHN AND NELSON GOODMAN, Choice implies excluded middle, *Zeit. Math. Log.* **23**(1978) 461.

[12] RICHMAN, FRED, Constructive aspects of Noetherian rings, *Proc. Amer. Math. Soc.*, **44**(1974), 436–441.

[13] ————————, Meaning and information in constructive mathematics, *Amer. Math. Monthly*, **89**(1982), 385-388.

[14] ————————, Flat dimension and the Hilbert syzygy theorem, *New Zealand J. Math.*, **26**(1997), 263–273.

[15] ————————, The regular element property, *Trans. Amer. Math. Soc.*, **126**(1998), 2123–2129.

[16] ————————, Generalized real numbers in constructive mathematics, *Indag. Math.*, to appear.

[17] ————————, The fundamental theorem of algebra, a constructive development without choice, (preprint, `www.math.fau.edu/Richman/html/docs.htm`, submitted to *Trans. Amer. Math. Soc.*).

[18] Richman, Fred, Gordon Berg, Henry Cheng and Ray Mines, Constructive dimension theory, *Compositio Math.*, **33**(1976), 1–17.

[19] Richman, Fred, Douglas Bridges, Allan Calder, William Julian and Ray Mines, Compactly generated Banach spaces, *Archiv der Mat.*, **36**(1981), 239–243.

[20] Ruitenburg, Wim, Constructing roots of polynomials over the complex numbers, *Computational aspects of Lie group representations and related topics, CWI Tract* **84**(1991), 107–128.

[21] Scedrov, Andre, Diagonalization of continuous matrices as a representation of intuitionistic reals, *Ann. Pure Appl. Logic* **30**(1986), 201–206.

[22] Seidenberg, Abraham, What is Noetherian? *Rend. Sem. Math. e Fis. Milano*, **44**(1974), 55–61.

[23] Troelstra, Anne S. and Dirk van Dalen, *Constructivism in mathematics, an introduction*, Volume 1, North-Holland 1988.

[24] Weyl, Hermann, Randbemerkungen zu Hauptproblemen der Mathematik, *Math. Z.* **20**(1924), 131–150.