

PLURALISM IN MATHEMATICS

E.B. Davies

Department of Mathematics

King's College

Strand

London WC2R 2LS

EBDRS1.pdf, 29 July 2004

Abstract

We defend pluralism in mathematics, and in particular Errett Bishop's constructive approach to mathematics, on pragmatic grounds, avoiding the philosophical issues which have dissuaded many mathematicians from taking it seriously. We also explain the computational value of interval arithmetic.

keywords: pluralism, constructive mathematics, interval arithmetic

1 Introduction

Errett Bishop's book 'Foundations of Constructive Analysis' appeared in 1967 and started a new era in the development of constructive mathematics. His account of the subject was entirely different from, and far more systematic than, Brouwer's programme of intuitionistic mathematics. The latter attracted a few adherents in the 1920s and 1930s, but was widely rejected because of its conflicts with the dominant classical view of the subject.

Unfortunately Bishop's book was ignored by most mathematicians, who assumed that the issues involved had all been settled, and that he could not have anything interesting to say. My task in this meeting is to try to persuade you that his programme provides valuable insights into matters which should be of concern to anyone who has even indirect involvement in computation.

In this paper I will *not* discuss the philosophical issues relating to Bishop's work, which are treated at some length in Billinge [2003] and Davies [2004], beyond saying that one can admire his mathematical contributions without adopting his philosophical position. Briefly, I defend what I call pluralism in mathematics – the view that classical mathematics, constructive mathematics, computer assisted

mathematics and various forms of finitistic mathematics can co-exist. I revive Carnap's dictum that one must decide the framework of discourse *before* questions about existence and truth make sense; see Carnap [1950]. In different frameworks the answer to a question may be different, but this in no way implies that one or the other is 'right'. This position is strongly anti-Platonistic.

From Chapter 2 onwards Bishop [1967] is completely recognizable as rigorous pure mathematics. Many well-known theorems appear, sometimes in forms which are not the usual ones, although trivially equivalent to them from a classical point of view. A few theorems are simply absent. The value of Bishop's efforts may not be immediately clear to everyone, in spite of what he writes in his first chapter. I will show that the differences between classical and constructive mathematics always correspond to situations in which real difficulties arise in numerical computation for suitable examples. Classical mathematics may provide a more flexible context for proving the existence of entities, but constructive mathematics provides a systematic approach to understanding why computational solutions of problems are sometimes not easy to obtain.

This is not to say that constructive mathematics is simply a part of numerical analysis. Numerical analysts have more problems than do constructivists. Bishop [1967, p. 140] gives a straightforward proof of the fundamental theorem of algebra, with one small but interesting proviso, even though it is known that the zeros of a polynomial of moderate degree (say at least 50) are highly unstable with respect to very small changes in the coefficients. Nevertheless, one can gain many insights into the differences between classical and constructive mathematics by considering numerical examples.

2 What is Constructive Mathematics?

It has often been said that Bishop rejected the law of the excluded middle, but a more useful description of the situation is that he gave the symbol \exists a different meaning from the usual one. In classical mathematics \exists refers to Platonic existence, but Bishop used it to refer to the possibility of constructing the relevant quantity. In classical mathematics \exists may be defined in terms of \forall : the expression $\exists xA$ is logically equivalent to $\neg(\forall a\neg A)$. In constructive mathematics \exists is a new quantifier with stricter conditions for its application. All of the differences between classical and constructive mathematics follow from this fact. We wish to emphasize that every theorem in Bishop's constructive mathematics is also a theorem in classical mathematics. Constructive mathematicians simply have to work harder to prove theorems, because their criteria for existence are stricter.

Starting from the natural numbers, Bishop constructed the real number system and established many of its familiar properties. However, in his approach, one cannot assert for every real x that either $x = 0$ or $x \neq 0$. Nor is it the case that

every bounded set of real numbers has a least upper bound. The reason for this is best illustrated with an example. For each positive integer n we define the number a_n to be 1 if the n th digit in the decimal expansion of π is the start of a sequence of a thousand consecutive sevens; otherwise we put $a_n = 0$. For each n the value of a_n can be determined within a finite length of time. However, the least upper bound A of the sequence is currently unknown. Platonically speaking either $A = 0$ or $A = 1$, even if we do not know which is the case, but Bishop would say that such a statement has no content: it simply reformulates the question in a trivial manner.

If we put

$$s = \sum_{n=1}^{\infty} a_n (-1/3)^n$$

then s exists constructively, because it can be calculated with any assigned accuracy in a finite length of time. However, we do not know whether or not $s = 0$. Even if it were eventually proved that every sequence of digits occurs somewhere in the decimal expansion of π , so that $s \neq 0$, whether s is positive or negative seems to be far beyond reach. The development of constructive mathematics has to reflect the existence of an infinite number of questions of the same type.

It is very surprising that Bishop could develop so much of Analysis without using the least upper bound principle or the law of the excluded middle. His book contains fully worked out accounts of the theory of Banach spaces, the spectral theorem, integration theory and Haar measure, among other topics. Set theory is substantially different from the classical version, particularly when considering the complement of a subset. Point set topology is not easy to develop (see Bridges and Luminița [2003] for one approach to this), but Bishop contains a rich constructive version of the theory of metric spaces. Compactness is defined in Bishop [1967, p. 88] using the notion of total boundedness, which is the way compactness is often proved in applications. This is not a coincidence.

3 What Significance do the Differences Have?

My goal in this section is to explain why a classical mathematician should be interested in Bishop's programme. My thesis is that by adopting the constructive framework one can handle certain numerical problems on a systematic basis, whereas classical mathematicians have to deal with them piecemeal, and try to remember whether or when they are using results which are not computationally feasible.

Producing examples to demonstrate the differences between classical and constructive mathematics often exploits the difference between recursive and recursively enumerable subsets of \mathbf{N} . Another method is to define a sequence whose behaviour

as $n \rightarrow \infty$ depends upon whether some famous conjecture is true or false.¹ We adopt a third strategy, showing that the impossibility of proving something in constructive mathematics is regularly *associated* with the extreme difficulty of showing it numerically for quite ordinary functions. We emphasize that the functions considered below do not, strictly speaking, provide examples of the constructive phenomenon, but we feel that in spite of this they explain why the constructive phenomenon exists.

Let us start with the intermediate value theorem for continuous functions of a single real variable. Bishop [1967, p. 5] explains why this theorem cannot be proved in a constructive framework. In the context of constructive mathematics one cannot find a value of the variable x for which $f(x) = c$ by the method of bisection, because being able to evaluate $f(x) - c$ with arbitrary accuracy does not imply that one can determine whether it is positive or negative.

Slight modifications of the intermediate value theorem are, however, valid constructively. If $f : [a, b] \rightarrow \mathbf{R}$ is continuous and $f(a) < c < f(b)$ then, given $\varepsilon > 0$, one can constructively find $x \in (a, b)$ such that $|f(x) - c| < \varepsilon$. In addition the intermediate value theorem itself may be proved constructively under a mild extra condition on f (being locally non-constant), which is almost always satisfied. See Bridges [1998] and Bishop-Bridges [1985].

We explain the differences between the classical and constructive versions of the intermediate value theorem by means of two examples, one ersatz and one genuine. Let

$$f(x) = \log(1 + x^{45}) \tag{1}$$

on $(-1, 1)$. *Given the formula* it is evident that the only solution of $f(x) = 0$ is $x = 0$, but one would need to calculate to very high accuracy to determine from its numerical values that the function is not identically zero throughout the interval $(-1/4, 1/4)$. However many digits one uses in the numerical computation, a similar difficulty arises if one replaces 45 by a larger number. In applied situations a closed formula is frequently not available, and the above problem may be a pressing one.

A genuine example presenting exactly the same difficulties is obtained as follows. Let

$$g(x) = \begin{cases} x + 1 & \text{if } x \leq -1 \\ 0 & \text{if } -1 < x < 1 \\ x - 1 & \text{if } 1 \leq x. \end{cases}$$

Then one cannot constructively solve $g(x) = c$ if c is an extremely small number for which one does not know whether $c = 0$, $c < 0$ or $c > 0$.

In constructive mathematics every non-negative continuous function on a closed bounded interval has a non-negative infimum; Bishop [1967, p. 35] provides a

¹If one uses the Goldbach conjecture, for example, then one puts $a_n = 0$ if $2n$ may be written as the sum of two primes, and $a_n = 1$ if it cannot.

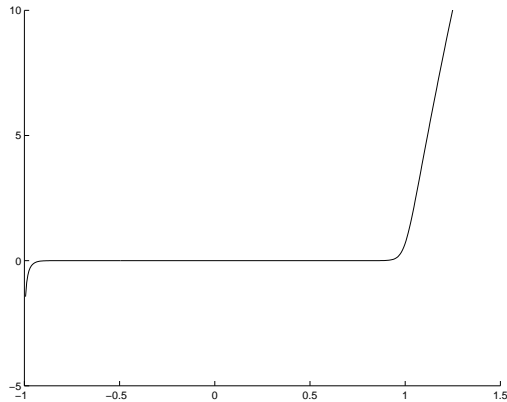


Figure 1: Graph of the function $f(x)$ defined by (1)

procedure for computing this with arbitrary accuracy. This does not imply that one can always determine whether the minimum is zero or positive, nor does it imply that one can find a point at which the infimum is achieved. Both can be a serious problem in numerical analysis as well. If $\varepsilon > 0$ is sufficiently small it is difficult to show *purely numerically* that the polynomial

$$p(x) = x^4 - 2\pi^2 x^2 + \pi^4 + \varepsilon(x - 2)^2 \quad (2)$$

never vanishes, and also difficult to determine whether its minimum value occurs near $x = \pi$ or $x = -\pi$. For functions arising in applied mathematics that are not given by explicit formulae this can again be a serious problem.

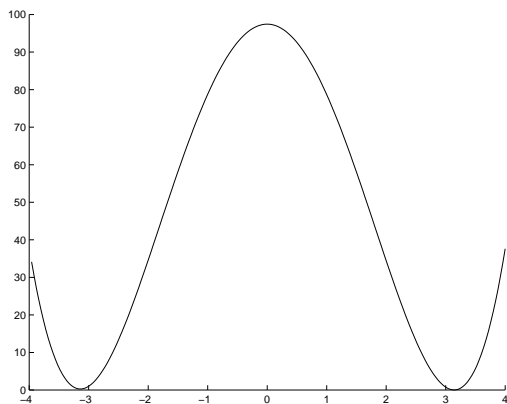


Figure 2: Graph of the function $p(x)$ defined by (2) with $\varepsilon = 0.01$

The suggestion that classical mathematics takes priority over constructive mathematics because of the Putnam-Quine argument about the indispensability of the former in the physical sciences are not convincing, for reasons spelled out in Davies [2003a, 2003b]. The differences between the two systems are unimportant from the

point of view of physicists; this is why Newton, Laplace, Maxwell and other scientists were able to develop very successful mathematically based theories long before the real number system was formalized at the end of the nineteenth century.

Hellman has tried to put some flesh on the Putnam-Quine argument in several recent papers, in which he claims that there are no constructive versions of some key results in mathematical physics. We start with Hellman [1993a], which deals with Gleason's theorem, considered by some (but not the author of this paper) to be of considerable importance in the foundations of quantum theory. This concerns the (non-distributive) lattice \mathcal{L} of closed subspaces of a Hilbert space \mathcal{H} of dimension greater than 2. In this lattice the analogues of set-theoretic complements are orthogonal complements. Gleason's theorem states that if μ is a normalized, countably additive measure on \mathcal{L} in a suitable sense, then there exists a non-negative, self-adjoint operator S on \mathcal{H} with trace 1 such that

$$\mu(L) = \text{trace}(SP_L)$$

for all $L \in \mathcal{L}$, where P_L is the orthogonal projection with range L . Hellman showed that a different version of Gleason's theorem cannot be proved in constructive mathematics. Nevertheless, Gleason's original version of the theorem, stated above, is constructively valid; see Richman and Bridges [1999] and Richman [2000]. The difference between the two versions relates to the validity of the principal axes theorem, discussed below.

In Hellman [1993b] the author showed that one version of the spectral theorem for unbounded self-adjoint operators is not constructively valid. However, in his original book Bishop [1967, p. 275] had already proved a different version, for a commuting sequence of bounded self-adjoint operators, which is completely acceptable even to classical mathematicians. After Hellman's paper appeared, Ye [1998] published a constructive version of the spectral theorem for an unbounded self-adjoint operator. Hellman's focus on the issue of domains and unboundedness is misguided because an unbounded operator becomes bounded as soon as one makes its domain into a Banach space by endowing it with the graph norm

$$\| \|f\| \| = \sqrt{\|f\|^2 + \|Af\|^2} .$$

The difficulty of determining whether a vector lies in the domain of an unbounded self-adjoint operator is not just a problem for constructivists. The classical theory progressed much more rapidly after it was realized that it was sufficient to specify an explicit domain of essential self-adjointness (a so-called 'core' of the operator) or even a core for the associated quadratic form; see Davies [1980]. A considerable amount is now known about the spectral theory of whole classes of differential operators whose domains cannot be identified as standard function spaces; see, for example, Davies [1989].

The principal axes theorem is the classical result that every self-adjoint matrix has a complete orthonormal set of eigenvectors. The theorem is not constructively valid,

for good reasons: the eigenvectors may change extremely rapidly as a parameter passes smoothly through a critical value. We invite the reader to check this for the elementary example

$$A_s = \begin{bmatrix} \varepsilon & s \\ s & -\varepsilon \end{bmatrix}$$

where $\varepsilon = 10^{-100}$ and s passes through the value 0. The classical version of the spectral theorem provides no insight into the existence of these computational problems, and does not suggest how they might be overcome. One *can* understand such problems in classical terms, but the constructive approach provides a systematic framework for doing so.

Finally in Hellman [1998] the author showed that the Hawking-Penrose singularity theorem is not constructively valid. This is an interesting observation, since the theorem has been extremely influential in the subject. It is embarrassing for a radical constructivist but not for a pluralist. It remains extremely hard to say much about the nature of the singularities: one certainly cannot identify them as black holes without the benefit of a ‘cosmic censorship hypothesis’. It is very likely that any detailed classification of the singularities will also be constructively valid.

The constructive version of the Hahn-Banach theorem in Bishop [1967, p. 263] applies to normable linear functionals, and Bishop needs to distinguish these from the more general bounded linear functionals. The following example explains why. Take the sequence a defined above using the digits of π and put $b_n = 1$ if $a_n = 1$ and also $a_m = 0$ for all $m < n$; otherwise put $b_n = 0$. The sequence b identifies the first occurrence of a thousand consecutive sevens in the decimal expansion of π , if such a sequence exists. Classically $b \in l^2(\mathbf{N})$ but constructively we cannot assert this (in 2004), because we are not able to evaluate $\|b\|_2$ with arbitrary accuracy. Even if we were assured that a sequence of a thousand consecutive sevens existed, we would still not be able to place b in $l^2(\mathbf{N})$ constructively unless we were given some information about how large n had to be for $b_n = 1$.² Nevertheless the formula

$$\phi(c) = \sum_{n=1}^{\infty} c_n b_n$$

is constructively well defined for all $c \in l^2(\mathbf{N})$ and defines a bounded linear functional ϕ on $l^2(\mathbf{N})$.

A linear operator A on a Banach space \mathcal{B} is said to be bounded if there exists a constant c such that $\|Ax\| \leq c\|x\|$ for all $x \in \mathcal{B}$. Its norm is the smallest such constant if that exists. The constructive distinction between bounded and normable operators is related to the fact that there is no effective classical algorithm for determining the norms of bounded operators on most Banach spaces. This is

²Taken literally this sentence is nonsensical. Classically there is no problem in asserting that $b \in l^2(\mathbf{N})$. Constructively one could not be given an assurance about the existence of n for which $b_n = 1$ without also being given the relevant information about how large n had to be. This is a familiar problem when one tries to compare two incommensurate frameworks.

why finding the best constants for various Sobolev embeddings in L^p spaces and other operators of importance in Fourier analysis has occupied mathematicians for decades. The same problem occurs for large finite matrices – standard software packages only provide routines for computing the norm of a sufficiently large $n \times n$ matrix with respect to the l^p norm on \mathbf{C}^n for $p = 1, 2, \infty$.

4 Computer Assisted Mathematics

Over the last thirty years a number of famous problems in pure mathematics have been solved by methods that conform entirely to the standards of classical mathematics, except for the fact that they involve a large amount of computation. The first example that we know about was the disproof of a conjecture of Euler that

$$a^5 + b^5 + c^5 + d^5 = e^5$$

has no solutions for which a, b, c, d, e are all positive integers. Eventually Lander and Parkin [1967] discovered the identity

$$27^5 + 84^5 + 110^5 + 133^5 = 144^5 \tag{3}$$

by a computer search. As far as we know there is still no ‘theoretical’ proof of the solubility of Euler’s equation: simply checking the truth of the above identity is the only method.

Because it may not be well-known to this audience, I will concentrate on the increasing use of controlled numerical methods to *prove* the existence of, and describe, the solutions of various problems in analysis. Interval arithmetic provides a new version of finite arithmetic. It has an ancient history, but is slowly becoming more important in connection with computer assisted proofs of theorems in analysis. See Markov and Okumura [1999], Moore [1979], Plum [2001], Plum and Wieners [2002], Breuer et al. [2003] and the Interval Computations Web-Site [2004] for some of the many applications. Its basic entities may be written in the form

$$x = 1.24706_{296}^{304} e3$$

where one imposes some upper bound on the number of significant digits allowed. The interpretation of this expression is as the interval $x = [\underline{x}, \bar{x}]$ where $\underline{x} = 1247.06296$ and $\bar{x} = 1247.06304$, but the definitions of the basic operations of arithmetic on the entities do not depend logically upon this intuition, nor upon any commitment to the existence of the real number system. To add two entities one adds the lower bounds and then rounds *down* to the prescribed number of digits, and also adds together the two upper bounds and rounds *up* to the prescribed number of digits. If one ignores the rounding procedure then $u + v = w$ where

$$\begin{aligned} \underline{w} &= \underline{u} + \underline{v}, \\ \bar{w} &= \bar{u} + \bar{v}. \end{aligned}$$

The definition of multiplication is similar, but more complicated. One puts $uv = w$ where

$$\begin{aligned}\underline{w} &= \min\{\underline{uv}, \underline{u}\bar{v}, \bar{u}\underline{v}, \bar{u}\bar{v}\}, \\ \bar{w} &= \max\{\underline{uv}, \underline{u}\bar{v}, \bar{u}\underline{v}, \bar{u}\bar{v}\}.\end{aligned}$$

One identifies an integer n with the interval $[n, n]$ and writes $u \sim v$ if the two intervals overlap, i.e. if

$$\max\{\underline{u}, \underline{v}\} \leq \min\{\bar{u}, \bar{v}\}.$$

One puts $x > 0$ if $\underline{x} > 0$, and $x < 0$ if $\bar{x} < 0$; if neither of these holds then $x \sim 0$. One might define π in the system by $\pi \sim 3.14159265\frac{4}{3}$, without commitment to the existence of an ‘exact’ value. In interval arithmetic

$$(x - 1)^2 \sim x^2 - 2x + 1$$

but the two are not equal. One needs to choose the ‘right’ way of evaluating an expression to minimize the width of the interval produced.

In interval arithmetic the standard functions such as sin, log etc. take intervals to intervals. The programming languages have to be fairly subtle to achieve this. For example when obtaining

$$\sin([1.57, 1.58]) \subseteq [0.999957, 1]$$

the programming language must take into account the fact that $1.57 < \pi/2 < 1.58$. There is no requirement for the initial interval to be small. Thus

$$\cos([0, 2]) \subseteq [-0.416147, 1].$$

A systematic description of interval arithmetic has been completed, and programming languages using it are readily available. It allows a rigorous approach to global optimization and makes a small but vital contribution to the solution of certain well-known, non-linear, elliptic, partial differential equations. See Plum [2001], Plum and Wieners [2002], Breuer et al. [2003] and references there. In these papers the authors start by searching for approximate solutions on an experimental basis. They then prove various analytical results which establish that one can use a contraction mapping principle to prove the existence of true solutions close to the approximate solutions, *provided* certain inequalities hold. Finally they verify the inequalities rigorously using interval arithmetic.

It seems clear that this approach to non-linear PDEs will expand rapidly over the coming decades. Those of us who feel uneasy about computer assisted proofs will either have to come to terms with them, or assign an ever-increasing fraction of their subject to a new category. Computers are fallible in different ways from mathematicians, but both are subject to independent verification. Absolute certainty is a chimera in both contexts.

References

- [2003] Billinge H: Did Bishop have a philosophy of mathematics? *Phil. Math.* 11 (2003) 176-194.
- [1967] Bishop E: 'Foundations of constructive analysis', McGraw-Hill, 1967.
- [1985] Bishop E, Bridges D: 'Constructive Analysis', *Grundlehren der math. Wiss.* vol. 279, Springer-Verlag, Heidelberg, 1985.
- [2003] Breuer B, McKenna P J, Plum M: Multiple solutions for a semilinear boundary value problem: a computational multiplicity proof. *J. Diff. Eqns.* 195 (2003) 243-269.
- [1998] Bridges D: Constructive truth in practice, pp. 53-69 in 'Truth in Mathematics', eds. H G Dales and G Olivieri, Clarendon Press, Oxford, 1998.
- [2003] Bridges D, Luminița V: Apartness spaces as a framework for constructive topology. *Ann. Pure Appl. Logic* 119 (2003) 61-83.
- [1950] Carnap R: Empiricism, Semantics, and Ontology. *Revue Internationale de Philosophie* 4 (1950) 20-40. See also Supplement to 'Meaning and Necessity: A Study in Semantics and Modal Logic', enlarged edition, Univ. of Chicago Press, Chicago, 1956.
- [1980] Davies E B: One-Parameter Semigroups. *LMS Monographs* vol. 15. Academic Press, London, 1980.
- [1989] Davies E B: Heat Kernels and Spectral Theory. *Cambridge Tracts in Math.* vol. 92. Cambridge Univ. Press, Cambridge, 1989.
- [2003a] Davies E B: Empiricism in Arithmetic and Analysis. *Philosophia Mathematica* (3) 11 (2003) 53-66.
- [2003b] Davies E B: Quantum mechanics does not require the continuity of space. *Stud. Hist. Phil. Mod. Phys.* 34 (2003) 319-328.
- [2004] Davies E B: A defence of pluralism in mathematics. Preprint 2004.
- [1993a] Hellman G: Gleason's theorem is not constructively provable. *J. Phil. Logic* 22 (1993) 193-203.
- [1993b] Hellman G: Constructive mathematics and quantum mechanics: unbounded operators and the spectral theorem. *J. Phil. Logic* 22(1993) 221-248.
- [1998] Hellman G: Mathematical constructivism in space-time. *Brit. J. Phil. Sci.* 49 (1998) 425-450.
- [2004] Interval Computations Web-Site.
<http://www.cs.utep.edu/interval-comp/main.html>

- [1967] Lander L J, Parkin T R: A counterexample to Euler's sum of powers conjecture. *Math. Comput.* 21 (1967) 101-103.
- [1999] Markov S, Okumura K: The Contribution of T. Sunaga to Interval Analysis and Reliable Computing. pp. 167-188 in: T. Csendes (ed.), 'Developments in Reliable Computing', Kluwer, Dordrecht, 1999.
- [1979] Moore R E: 'Methods and Applications of Interval Analysis'. SIAM, Philadelphia, 1979.
- [2001] Plum M: Computer-assisted enclosure methods for elliptic differential equations. *Lin. Alg. Appl.* 324 (2001) 147-187.
- [2002] Plum M, Wieners C: New solutions of the Gelfand problem. *J. Math. Anal. Appl.* 269 (2002) 588-606.
- [2000] Richman F: Gleason's theorem has a constructive proof. *J. Phil. Logic* 29 (2000), 425-431.
- [1999] Richman F, Bridges D: A constructive proof of Gleason's theorem. *J. Funct. Anal.* 162 (1999) 287-312.
- [1998] Ye, Feng: On Errett Bishop's Constructivism – Some Expositions, Extensions and Critiques. Ph. D. thesis, Princeton University, 1998.