CUBO *a Mathematical Journal*
Vol. 6, № 1, (209-258). March 2004.

Constructivity in Mathematics

Douglas S. Bridges
Department of Mathematics & Statistics
University of Canterbury
Private Bag 4800, Christchurch
New Zealand
douglas.bridges@canterbury.ac.nz

ABSTRACT

This is an expanded version of the lectures that I delivered to the Department of Pure Mathematics at the University of Calcutta in the period 27 January–6 February 1998. I am most grateful to that University for inviting me to be its Rani and Asutosh Ganguli Visiting Professor for 1998; to Professor Mihir Chakroborty, Head of its Department of Pure Mathematics, for generating so much interest in my visit and for his many kindnesses to me during my memorable stay in his city; and to the graduate students in his Department, for making me so welcome and for showing such enthusiasm for my subject matter.

1 A fixed–point theorem

The main theme of these lectures is a revolution in mathematics, one that was started by L.E.J. Brouwer (1881–1966) in 1907 and looked doomed to failure until a dramatic intervention by Errett Bishop sixty years later. I first set the scene by describing Brouwer's most famous result, ironically one in classical topology and definitely *not*, at least as it stands, a part of Brouwer's revolution.

Consider[1] a closed disc B (that is, one that contains its bounding circle) in the

[1] In the first part of these lectures I shall try to be as informal and mathematically undemanding as possible, since not all members of the audience are trained mathematicians. The mathematical content and pace of the lectures will increase when we start to discuss varieties of constructive mathematics in more technical detail.

plane. To each point x of B let there be assigned a unique corresponding point y of B. We may picture this assignment by the line segment $\overrightarrow{xy}$ drawn from x to y, and think of x as being moved to the position y. Notice that, although a given x has a unique corresponding y, the same y may correspond to different choices of x. Also, there is nothing to forbid the situation where $y = x$; when that occurs, we call x a **fixed point** of the assignment.

We make one additional requirement: namely, that points which start close together end up close together. Mathematicians describe this requirement by saying that the assignment—or, to give it its proper mathematical name, the **mapping**—is **continuous**. According to **Brouwer's fixed–point theorem**, such a mapping as we have described always has at least one fixed point in B [28].

It may surprise you to learn that this theorem has many applications within mathematics. For example, it is used, in one form or another, by mathematical economists to prove that, under reasonable conditions, an economy has an equilibrium—a state in which supply and demand are balanced and hence producers and consumers are satisfied. Indeed, Brouwer's fixed–point theorem is equivalent, mathematically, to the existence of such an equilibrium. Thus it would seem beneficial—maybe, in the light of world-wide economic difficulties, advisable—to seek some method of computing the position of a fixed-point of our mapping on B.

Although there are elementary proofs of this theorem, none of them is easy to understand without a substantial mathematical background, and the shortest proofs require some quite heavy mathematical artillery. One feature of all these proofs is that they do not provide the means of computing the fixed points.[2] What they actually do (although this limitation is heavily disguised) is to prove that it is impossible that there not be a fixed point; in other words, they prove the statement

not not(there exists a fixed point).

Here, at last, we reach the distinction in meaning that formed the basis of Brouwer's revolution: the distinction between

- **idealistic existence**, where we prove the impossibility of the non–existence of the object in question (in the foregoing case, a fixed point) and conclude that the object does exist after all, and
- **constructive existence**, where in order to prove that our object exists, we must provide a method for finding it.

With the publication of his doctoral thesis [27] in Amsterdam, in 1907, Brouwer began a mathematical career largely devoted to his philosophy of **intuitionism**, in which mathematics is regarded as a free creation of the human mind, and the objects—mental constructs—of mathematics come into existence precisely when they

[2] The work of Scarf [61] and others would seem to contradict this statement; but what Scarf has done is to show how to compute an approximate fixed point, a point x that is close to the corresponding y. There is no guarantee that an approximate fixed point, however small the difference between x and y, is close enough to an exact fixed point to allow us to use the approximation with impunity.

are constructed. Actually, on the advice of his doctoral supervisor Korteweg, Brouwer took a few years' leave from the exposition of intuitionism, to establish a formidable reputation in traditional—or as we now call it, **classical**—mathematics, thereby ensuring that his intuitionistic views would gain some respect, even if, as history shows, they were not widely accepted. It was during that leave that Brouwer proved, among other important results, his fixed-point theorem.[3]

2 Intuitionistic Logic

For Brouwer, mathematics took precedence over logic. In order to describe the logic used by the intuitionist mathematician—a logic different from the **classical logic** normally used in mathematics—it was necessary first to analyse the mathematical processes of the mind, from which analysis the logic could be extracted. In 1930, Brouwer's most famous pupil, Arend Heyting (1898–1980), published a set of formal axioms which so clearly characterise the logic used by the intuitionist that they have become universally known as the axioms for **intuitionistic logic** [39]. These axioms capture the intuitionistic, or **constructive**, interpretations of the various connectives

$$\vee \text{(or)}, \wedge \text{(and)}, \Rightarrow \text{(implies)}, \neg \text{(not)}$$

and quantifiers

$$\exists\text{(there exists)}, \quad \forall\text{(for all/each)}$$

which we now outline:

- $P \vee Q$: either we have a proof of P or else we have a proof of Q.

- $P \wedge Q$: we have both a proof of P and a proof of Q.

- $P \Rightarrow Q$: by means of an algorithm—that is, a finite, computational procedure—we can convert any proof of P into a proof of Q.[4]

- $\neg P$: assuming P, we can derive a contradiction (such as $0 = 1$); equivalently, we can prove $(P \Rightarrow (0 = 1))$.

- $\exists x\, P(x)$: we have an algorithm which computes an object x and demonstrates that $P(x)$ holds.

- $\forall x \in A\, P(x)$: we have an algorithm which, applied to an object x and a proof that $x \in A$, demonstrates that $P(x)$ holds.

---

[3] For fascinating accounts of Brouwer's life and work see [33, 65].

[4] This interpretation of implication, while more natural than the classical one of **material implication** in which $(P \Rightarrow Q)$ is equivalent to $(\neg P \vee Q)$, has not completely satisfied all researchers using constructive logic. Shortly before he died, Bishop communicated to me his dissatisfaction with the standard constructive interpretation of implication. Unfortunately, he left nothing more than very rudimentary sketches of his ideas for its improvement.

In intuitionistic logic, even for a decidable property $P(n)$ of natural numbers n the property

$$\forall n\, P(n) \vee \neg\forall n\, P(n)$$

need not hold; so, in turn, the **law of excluded middle (LEM)**

$$P \vee \neg P$$

fails. As a result, many classical results cannot be proved constructively, since they would imply LEM or some other manifestly nonconstructive principle.

To illustrate this point, consider the following simple statement, the **limited principle of omniscience (LPO)**:

$$\forall \mathbf{a} \in \{0,1\}^{\mathbf{N}}\, (\mathbf{a} = \mathbf{0} \vee \mathbf{a} \neq \mathbf{0}),$$

where $\mathbf{a} = (a_0, a_1, a_2, \ldots)$, $\mathbf{N} = \{0, 1, 2, \ldots\}$ is the set of natural numbers, $\{0,1\}^{\mathbf{N}}$ is the set of all binary sequences, and

$$\begin{aligned} \mathbf{a} = \mathbf{0} \quad &\Leftrightarrow \quad \forall n\, (a_n = 0), \\ \mathbf{a} \neq \mathbf{0} \quad &\Leftrightarrow \quad \exists n\, (a_n = 1). \end{aligned}$$

In words, LPO states that for each binary sequence (a_n), either $a_n = 0$ for all n or else there exists n such that $a_n = 1$. Of course, this is a triviality from the viewpoint of classical logic; but its intuitionistic interpretation is not so simple. That interpretation says that there is an algorithm which, applied to any binary sequence $\mathbf{a}$, either verifies that all the terms of the sequence are 0 or else computes the index of a term equal to 1. Anyone familiar with computers ought to be highly sceptical about such an algorithm, since in the case $\mathbf{a} = \mathbf{0}$ it would normally need to test all the infinitely many terms a_n in order to come up with the correct decision.

In classical recursion theory—the classical version of computability theory, in which all computations are performed by Turing machines—we can prove that the recursive interpretation of LPO is false, since it would entail the decidability of the halting problem; see [19], pages 52–53. (There is a point worth noting here: the classical invalidity of the recursive interpretation of LPO is not a matter of logic, since it can be demonstrated even with classical logic.)

For these reasons alone, we may feel justified in not accepting LPO, or any classical proposition that intuitionistically implies LPO, as a valid principle of intuitionistic mathematics. But we have another reason for not doing so: it can be shown that there are models of **Heyting arithmetic**—Peano arithmetic with intuitionistic logic—in which LPO is false; so LPO cannot be derived in Heyting arithmetic; see [19, 34]. Since LPO is a special case of the law of excluded middle, we are forced, in turn, to reject the latter from intuitionistic mathematics. A similar informal analysis leads us to reject both the classical rule

$$\neg\neg P \Rightarrow P$$

that forms the basis of proof-by-contradiction, and the **lesser limited principle of omniscience (LLPO)**,

> *For each binary sequence* **a** *with at most one term equal to* 1, *either* $a_{2n} = 0$ *for all* n *or else* $a_{2n+1} = 0$ *for all* n,

which is easily seen to be a consequence of LPO.

The exclusion of such principles from intuitionistic mathematics has serious consequences for mathematical practice. For example, we cannot hope to prove intuitionistically the simple statement

$$\forall x \in \mathbb{R}\ (x = 0 \vee x \neq 0),$$

where $\mathbb{R}$ denotes the set of real numbers, and $x \neq 0$ means that we can compute a rational number strictly between 0 and x (which is not the same, constructively, as proving that $\neg(x = 0)$). To see this, consider any binary sequence **a**, and use it to define the binary expansion of a real number

$$x = \sum_{n=0}^{\infty} a_n 2^{-n}.$$

If $x = 0$, then $\mathbf{a} = \mathbf{0}$. If $x \neq 0$, we can compute a positive integer N such that $x > 2^{-N}$; by testing the terms $a_1, \ldots, a_N$, we can decide whether or not there exists n such that $a_n = 1$. Thus the above statement about real numbers implies LPO and is therefore essentially nonconstructive.[5] A similar argument, using the real number

$$\sum_{n=0}^{\infty} (-1)^n a_n 2^{-n},$$

shows that the statement

$$\forall x \in \mathbb{R}\ (x \geqslant 0 \vee x \leqslant 0)$$

implies LLPO and is therefore essentially nonconstructive.

The following elementary classical statements also turn out to be nonconstructive.

- Each real number x is either rational or irrational (in the sense that $x \neq r$ for each rational number r). To see this, consider

$$x = \sum_{n=0}^{\infty} (1 - a_n)/n!$$

  where **a** is any increasing binary sequence.

- Each real number x has a binary expansion. Note that the standard interval-halving argument for 'constructing' binary expansions does not work, since we cannot necessarily decide, for a given number x between 0 and 1, whether $x \geqslant \frac{1}{2}$ or $x \leqslant \frac{1}{2}$. In fact, the existence of binary expansions is equivalent to LLPO.

- $\forall x, y \in \mathbb{R}\ (xy = 0 \Rightarrow (x = 0 \vee y = 0))$. This clearly has implications for the theory of integral domains!

[5] This is not so surprising when you consider the problem of *underflow*, which can cause a computer to register a small, nonzero number as 0.

3 Fundamental principles of intuitionism

I would now like briefly to describe, in turn, the three main modern varieties of constructive mathematics—that is, mathematics in which only constructive existence, and not idealistic existence, is used.

I have already introduced the first of these, **intuitionistic mathematics** (**INT**), which, you will recall, is based on Brouwer's intuitionistic philosophy of mathematics as a free creation of the human mind. Introspection led Brouwer not only to his informal explanation of the logical processes used in intuitionistic mathematics, but also to adopt certain nonclassical principles. The first of these, the **principle of continuous choice**, has two parts:

- Every function from $\mathbf{N^N}$ (the set[6] of sequences of natural numbers) to $\mathbf{N}$ is continuous relative to the metric ρ defined on $\mathbf{N^N}$ by
$$\rho(\mathbf{a}, \mathbf{b}) = \inf\left\{2^{-n} : \forall k \leqslant n\,(a_k = b_k)\right\}.$$
Thus if f is a function from $\mathbf{N^N}$ to $\mathbf{N}$, then
$$\forall \mathbf{a} \in \mathbf{N^N}\ \exists n \in \mathbf{N}\ \forall \mathbf{b} \in \mathbf{N^N}\ (\forall k \leqslant n\ \ (a_k = b_k) \Rightarrow (f(\mathbf{a}) = f(\mathbf{b}))).$$
- If $P \subset \mathbf{N^N} \times \mathbf{N}$, and for each $\mathbf{a} \in \mathbf{N^N}$ there exists $n \in \mathbf{N}$ such that $(\mathbf{a}, n) \in P$, then there exists a **continuous choice function** $f : \mathbf{N^N} \to \mathbf{N}$ such that $(\mathbf{a}, f(\mathbf{a})) \in P$ for all $\mathbf{a} \in \mathbf{N^N}$.

It follows from this principle that every function from a nonempty[7] complete separable metric space into a metric space is continuous, and hence that any linear mapping of a separable Banach space into a normed space is bounded ([19], pages 109–110). We can also prove that LLPO, and therefore LPO, is false. To this end, suppose that LLPO holds, and define $P \subset \mathbf{N^N}$ as follows. Given $\mathbf{a} \in \mathbf{N^N}$, define $\mathbf{b} \in \mathbf{N^N}$ by setting

$$b_n = \begin{cases} 1 & \text{if } a_n \neq 0 \text{ and } \forall k < n\,(a_k = 0) \\ \\ 0 & \text{otherwise.} \end{cases}$$

LLPO ensures that either $b_{2n} = 0$ for all n or $b_{2n+1} = 0$ for all n. Let $(\mathbf{a}, 0) \in P$ in the former case, and $(\mathbf{a}, 1) \in P$ in the latter. For each $i \in \mathbf{N}$ define $\mathbf{a}^i \in \mathbf{N^N}$ by

$$\mathbf{a}^i_n = \begin{cases} 1 & \text{if } i = n \\ \\ 0 & \text{if } i \neq n. \end{cases}$$

By the principle of continuous choice, there exists a continuous function $f : \mathbf{N^N} \to \mathbf{N}$ such that $(\mathbf{a}, f(\mathbf{a})) \in P$ for all $\mathbf{a} \in \mathbf{N^N}$. Since $(\mathbf{a}^{2i}, 0) \notin P$ and $(\mathbf{a}^{2i+1}, 1) \notin P$, we see

[6] It makes for easier explanation if I couch Brouwer's work in terms of set theory, rather than his own notion of a 'spread'.

[7] When I say that a set S is **nonempty**, I mean that we can construct a point of S; the intuitionists refer to such a set S as **inhabited**. Being inhabited is a stronger property than $\neg(S = \emptyset)$.

that $f(\mathbf{a}^{2i}) = 1$ and $f(\mathbf{a}^{2i+1}) = 0$. But the sequence $(\mathbf{a}^i)_{i=0}^{\infty}$ converges to $\mathbf{0}$ in the metric on $\mathbf{N}^{\mathbf{N}}$, so f is not continuous at $\mathbf{0}$. This contradiction shows that LLPO is false.

We need some more definitions in order to formulate Brouwer's second important principle. For any set X let X^* denote the set of all finite (possibly empty) sequences in X. We say that a subset S of $\mathbf{N}^*$ is

- **detachable** if for each $x \in \mathbf{N}^*$ either $x \in S$ or $x \notin S$;
- a **fan** if it is closed under restriction—that is, $(x_0, \ldots, x_k) \in S$ whenever $(x_0, \ldots, x_n) \in S$ and $-1 \leqslant k < n$. (The case $k = -1$ corresponds to the empty restriction of x.)

The set 2^* is a fan, called the **complete binary fan**, which can be represented in the obvious way as a tree, from which the name 'fan' was derived. A **path** in a fan σ is a sequence s, finite or infinite, such that each restriction of s is in σ. We say that a path s is **blocked** by a subset B of σ if some restriction of s belongs to B; if no restriction of s is in B we say that s **misses** B. A subset B of a fan σ is called a **bar** for σ if each infinite path of σ is blocked by B; a bar B is a **uniform bar** if there exists n such that each path of length n is blocked by B.

The following principle, Brouwer's **fan theorem**,[8] is of major importance in INT:

Every detachable bar of a fan is a uniform bar.

Under the hypotheses of the principle of continuous choice, the fan theorem is equivalent to the **intuitionistic uniform continuity theorem**,

Every real-valued function on a compact interval in $\mathbf{R}$ *is uniformly continuous,*

a result so patently at odds with classical analysis as to suggest to some mathematicians that intuitionism is false. The apparent absurdity of the intuitionistic uniform continuity theorem is, however, illusory, as is suggested by the following more careful re-statement of it:

Every intuitionistically defined function from an intuitionistic compact interval to the intuitionistic real line is, intuitionistically, uniformly continuous.

In fact, there is a strong case for saying that, except at certain levels of formalism, INT and classical mathematics (CLASS) are incomparable, and that it is not possible to capture fully the spirit and meaning of intuitionistic statements, such as our reformulated intuitionistic uniform continuity theorem, within a classical framework.

[8]The classical contrapositive of the fan theorem is **König's Lemma:** If, for each n, there exists a path of length n in σ that misses B, then there exists an infinite path that misses B. It is an exercise to show that the fan theorem entails LPO.

It is fair to say that the attempts to justify Brouwer's fan theorem from an intuitionistic standpoint have not been altogether successful. Brouwer himself made such an attempt, based on another principle—'bar induction'—whose justification seems equally elusive.

For more details about INT, see [19, 32, 34, 67]. An extremely readable, if somewhat outdated, introduction to intuitionism is given by Heyting in [40].

4 Recursive constructive mathematics

The second of the main varieties of modern constructivism is the **recursive constructive mathematics** (RUSS) initiated by Markov in the late 1940's and subsequently developed by him and his followers, primarily in the former Soviet Union, and in particular in Leningrad (now St Petersburg) and Moscow[9] [46]. In this variety the objects are defined by means of Gödel–numberings, and the procedures are all recursive; the main distinction between RUSS and the classical recursive analysis developed after, in 1936, the work of Turing, Church, and others clarified the nature of computable processes, is that the logic used in RUSS is intuitionistic. Thus RUSS may be described as recursive mathematics with intuitionistic logic.

One obstacle faced by the mathematician attempting to come to grips with RUSS is that, expressed in the language of recursion theory, it is not easily readable; indeed, on opening a page of Kushner's excellent lectures [45], one might be forgiven for wondering whether this is analysis or logic. Fortunately, we can get to the heart of RUSS by an axiomatic approach, due to Fred Richman [52], which I now outline.

Recall that a **partial function** $f : X \to Y$ is a function from a subset of X called the domain of f, and written $\mathrm{dom}(f)$, into Y; that $f(x)$ is **defined** if $x \in \mathrm{dom}(f)$, and **undefined** if $x \notin \mathrm{dom}(f)$; and that f is **total** if $\mathrm{dom}(f) = X$. Experience with recursion theory leads us to the following axiom about computable partial functions:

> **CPF:** *There is an enumeration of the set of all partial functions from* N *to* N *that have countable domains.*

Note that, for us, a set S is **countable** if there exists a function from a detachable subset of N onto S. The empty set is countable; and a nonempty set is countable if and only if it is the range of a function with domain N.

In the remainder of this section, and wherever we discuss RUSS in these lectures, we shall assume that

$$\varphi_0, \varphi_1, \varphi_2, \dots$$

is a fixed enumeration of the set of computable partial functions from N to N, and that

$$D_0, D_1, D_2, \dots$$

[9] When, in the mid-1980's, Fred Richman and I first used the term "RUSS" to signify the mathematics of the Markov school, it was suggested that "SOV", might be a more appropriate designation. We believe that subsequent political developments have validated the wisdom of our choice.

is a fixed enumeration of the corresponding domains. It can easily be shown that for each m there exists a sequence $(D_m(n))_{n=0}^{\infty}$ of finite subsets of $\mathbf{N}$ such that $D_m(0) \subset D_m(1) \subset \cdots$ and

$$\operatorname{dom}(\varphi_m) = \bigcup_{n=0}^{\infty} D_m(n).$$

Indeed, if we think of φ_m as the partial function computed by the mth Turing machine in some effective enumeration of the set of Turing machines (see, for example, [13]), then we can take $D_m(n)$ to be the set of those $k \in \mathbf{N}$ such that the Turing machine computes $\varphi_m(k)$ in $n+1$ steps. For convenience we take $D_m(n) = \emptyset$ if $n < 0$.

The following is perhaps the fundamental result in RUSS.

Proposition 1 For each total function $g : \mathbf{N} \to \{0,1\}$ there exists $m \in \mathbf{N}$ such that $g(m) = 0$ if and only if $\varphi_m(m)$ is undefined.

An immediate consequence of this is that there is no total function $g : \mathbf{N} \to \{0,1\}$ such that for each m, $g(m) = 1$ if and only if $\varphi_m(m)$ is defined; this is the expression within RUSS of the undecidability of the halting problem (cf. [13], Chapter 4).

Proposition 1 is very easy to prove using CPF. Given g, we see that

$$g^{-1}(0) = \bigcup_{n=0}^{\infty} \left(g^{-1}(0) \cap \{0,1,\ldots,n\}\right)$$

is countable. Now choose m such that φ_m is a partial function with domain $g^{-1}(0)$; then $\varphi_m(m)$ is defined if and only if $g(m) = 0$. ∎

We can now prove that both LPO and LLPO are false within RUSS. For example, assuming LPO, let $\lambda : \{0,1\}^{\mathbf{N}} \to \{0,1\}$ be such that $\lambda(\mathbf{a}) = 1$ if $a_n = 1$ for some n, and $\lambda(\mathbf{a}) = 0$ if $a_n = 0$ for all n. Define a total function $\mu : \mathbf{N} \to \{0,1\}^{\mathbf{N}}$ such that $\mu(m)_k = 1$ if and only if $m \in D_m(k)$, and let $g = \mu \circ \lambda$. Then $g(m) = 1$ if and only if $\varphi_m(m)$ is defined. This contradicts the fundamental result, Proposition 1. The proof that LLPO is false in RUSS can be found on pages 53–54 of [19].

Another consequence of CPF is that Brouwer's principle of continuous choice is false. For, by CPF, for each $\mathbf{a} \in \mathbf{N}^{\mathbf{N}}$ there exists n such that $\mathbf{a} = \varphi_n$. Let

$$P = \left\{(\mathbf{a}, n) \in \mathbf{N}^{\mathbf{N}} \times \mathbf{N} : \mathbf{a} = \varphi_n\right\}.$$

If there is a continuous function $f : \mathbf{N}^{\mathbf{N}} \to \mathbf{N}$ such that $(\mathbf{a}, f(\mathbf{a})) \in P$ for each $\mathbf{a} \in \mathbf{N}^{\mathbf{N}}$, then each nonzero $\mathbf{a}$ sufficiently close to $\mathbf{0}$ satisfies $f(\mathbf{a}) = f(\mathbf{0})$ and therefore $\mathbf{a} = \varphi_{f(\mathbf{0})} = \mathbf{0}$. This is clearly false.

What can we say about continuity in RUSS? Before answering this question, I must say something about **Markov's Principle (MP):**

$$\forall \mathbf{a} \in \{0,1\}^{\mathbf{N}} \ (\neg(\mathbf{a} = \mathbf{0}) \Rightarrow \mathbf{a} \neq \mathbf{0}).$$

In words: for each binary sequence (a_n), if it is impossible that $a_n = 0$ for all n, then there exists n such that $a_n = 1$. This principle embodies the notion of an unbounded search—given that $\mathbf{a} \neq 0$, to find n such that $a_n = 1$, we simply test the terms $a_0, a_1, \ldots$ in turn, being guaranteed by Markov's Principle that we will eventually, although we do not know in advance when, find a term equal to 1. For this reason it is viewed with caution by most practitioners of RUSS and is rejected outright by most other constructive mathematicians.[10]

With the (essential) help of MP, we can prove **Čeitin's theorem** ([19], Chapter 3):

Every function from $\mathbb{R}$ *to* $\mathbb{R}$ *is continuous.*

As with the intuitionistic uniform continuity theorem, we have to be careful about our interpretation of this statement; otherwise, it will appear to contradict classical mathematics. The full interpretation says that every recursively defined function from the recursive real line to the recursive real line is recursively continuous; from a classical viewpoint this theorem is acceptable, as it does not say anything about functions defined also on the nonrecursive real numbers.

The classical uniform continuity theorem, however, does not hold in RUSS, in which there exists a continuous function $f : [0,1] \to \mathbb{R}$ that is not uniformly continuous. The construction of such a function, which I shall not describe, depends on a famous result, **Specker's theorem** [64], showing that the monotone sequence principle of classical analysis does not hold in RUSS:

Theorem 2 *There exists an increasing sequence* (r_n) *of rational numbers in the Cantor set*

$$C = \left\{ \sum_{n=1}^{\infty} c_n 3^{-n} : \forall n \, (c_n \in \{0,2\}) \right\}$$

such that

$$\forall x \in \mathbb{R} \; \exists \delta > 0 \; \exists N \in \mathbb{N} \; \forall n \geqslant N \, (|x - r_n| \geqslant \delta).$$

In other words, although the sequence (r_n) *is increasing and bounded above, it is eventually bounded away from any given (recursive) real number; whence its classical limit is a nonrecursive real number.*

I shall return to prove Specker's theorem in Section 9, once I have discussed our next variety of constructive mathematics.

5 Bishop's constructive mathematics

The final variety that I want to talk about, and the one that will occupy most of the rest of these lectures, is **Bishop's constructive mathematics (BISH)**, which first appeared in Errett Bishop's ground-breaking monograph *Foundations of Constructive Analysis* [4]. In that book, which was born after a remarkably short period of

[10] MP is inconsistent with Brouwer's theory of the creating subject, itself a controversial extension of Brouwer's basic intuitionism; see [34].

gestation, Bishop disclosed, by thoroughgoing constructive means but without resorting to either Brouwer's principles or the formalism of recursive function theory, a vast panorama of mathematics, covering elementary real and complex analysis, metric and normed spaces, abstract measure and integration, the spectral theory of selfadjoint operators on a Hilbert space, Haar measure and duality on locally compact groups, and Banach algebras. At a stroke, Bishop's work gave the lie to Hilbert's dismissal of constructive mathematics in the words

> 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. [41].

Bishop's approach to mathematics has one great advantage over INT and RUSS: his results and proofs are valid, *mutatis mutandis*, in classical mathematics, INT, RUSS, and all reasonable models of computable mathematics—such as, for example, Weihrauch's TTE [71]. Every proof of a theorem T in BISH is also (or, in the recursive setting, can routinely be translated into) a proof of T in each of INT, RUSS, and CLASS. Thus BISH can be regarded as the common constructive core of these other varieties of mathematics, and each of INT, RUSS, and CLASS as a model of BISH.

This has important consequences in practice. For example, since

> *Every function from* $[0,1]$ *to* $\mathbb{R}$ *is uniformly continuous*

holds in the model INT, we cannot disprove it in BISH. On the other hand,

> *There exists a continuous mapping of* $[0,1]$ *into* $\mathbb{R}$ *that is not uniformly continuous*

holds in the model RUSS, so we cannot prove the uniform continuity theorem in BISH. (Since uniform, rather than pointwise, continuity seems to be what is needed for most computations with functions on a compact interval, Bishop freely uses uniform continuity hypotheses that hold automatically in INT and CLASS.) For another example, the statements

> *Every compact subset of* $\mathbb{R}$ *is Lebesgue measurable*

and

> *There exists a compact subset of* $\mathbb{R}$ *that does not have Lebesgue outer measure*

hold in CLASS and RUSS, respectively (for the latter see page 64 of [19]); so neither of them can be proved or disproved within BISH.

The foundations of BISH are close to those of INT.[11] Indeed, Bishop says from the outset that

[11] Bishop's position may have been closer to that of an intuitionist than would appear from most of his writings: see pages 357–360 of [4].

> The primary concern of mathematics is number, and this means the positive integers. We feel about number the way Kant felt about space. The positive integers and their arithmetic are presupposed by the very nature of our intelligence and, we are tempted to believe, by the very nature of intelligence in general. The development of the positive integers from the primitive concept of the unit, the concept of adjoining a unit, and the process of mathematical induction carries complete conviction. In the words of Kronecker, the positive integers were created by God. ([4], page 2)

Building on the set of positive integers, using intuitionistic logic and a primitive, unspecified concept of 'algorithm', Bishop systematically introduces mathematics at higher and higher levels of abstraction. To do so, he also needs notions of 'set' and 'function'. For him, a set S is given by two pieces of data:

- ▷ a property that enables members of S to be constructed using objects that have already been constructed (note this last phrase, which rules out the possibility of impredicative definitions and therefore of Russell–type paradoxes), and
- ▷ an equivalence relation $=_S$ of **equality**[12] between members of S.

The use of equivalence relations rather than intensional equality—identity of description—is common, but often goes unnoticed, in classical mathematics. For example, we call the rational numbers $1/2$ and $3/6$ equal, even though, strictly speaking, they are equivalent and not intensionally identical.

In more advanced work—even, as we shall see, on the real line—we frequently need a set S to be equipped with an **inequality** relation $\neq$ describing what it means for two elements of S to be unequal, or distinct. Such a relation must satisfy the following two properties:

$$x \neq y \;\Rightarrow\; \neg(x = y),$$
$$x \neq y \;\Rightarrow\; y \neq x.$$

One such inequality relation is defined by setting $x \neq y$ if and only if $\neg(x = y)$; but, as Markov's Principle suggests, this inequality is normally too weak for practical purposes.

Note that as we shall see later, the equality and inequality relations on the real line $\mathbb{R}$ are not decidable; but when restricted to the set $\mathbb{Q}$ of rational numbers, they are decidable.

Naturally, Bishop requires functions to be given by algorithms and to respect equality. Thus a **function** f from a set A to a set B is an algorithm that, applied to any element a of A, produces an element $f(a)$ of B, such that f is **extensional:** if $a = a'$ in A, then $f(a) = f(a')$ in B. If A and B have inequality relations, then we may require f to be **strongly extensional**, in the sense that if $f(a) \neq f(a')$ in B, then $a \neq a'$ in A.

[12] When the meaning is clear from the context, I shall write $=$, rather than $=_S$, to denote the equality on a set S.

With these definitions at hand, we can now prove the result of Goodman and Myhill[13] [37]:

Theorem 3 *The axiom of choice implies the law of excluded middle.*

Proof. Let P be any constructively meaningful statement, and define the set A to consist of the two elements 0 and 1, together with the equality relation such that

$$0 =_A 1 \text{ if and only if } P \text{ holds.}$$

(We could have defined A in more classical terms as a set of equivalence classes under the equivalence relation

$$0 \sim 1 \text{ if and only if } P \text{ holds,}$$

but it is more in keeping with Bishop's approach to proceed as we have done.) Let B be the set $\{0, 1\}$ with the standard equality, and let

$$S = \{(0,0), (1,1)\} \subset A \times B,$$

where the equality on S is derived in the usual way from those on A and B :

$$(x, y) =_S (x', y') \text{ if and only if } x =_A x' \text{ and } y =_B y'.$$

Suppose that there exists a function $f : A \to B$ such that $(x, f(x)) \in S$ for all $x \in A$. If $f(0) = 1$ or $f(1) = 0$, then $0 =_A 1$, and hence P holds; if $f(0) = 0$ and $f(1) = 1$, then $\neg(0 =_A 1)$, and hence P is false. Thus we have derived $P \vee \neg P$. ■

On page 9 of [4], Bishop remarks that

> the axiom of choice...is not a real source of nonconstructivity in classical mathematics. A choice function exists in constructive mathematics, because a choice *is implied by the very meaning of existence.*

How do we square this with the foregoing Goodman–Myhill theorem? It is true that if to each element x of a set A there corresponds an element y of a set B such that the property $P(x, y)$ holds, then it is implied by the meaning of existence in constructive mathematics that there is a finite routine for computing an appropriate $y \in B$ from a given $x \in A$; however, this computation may depend not only on the value a but also on the information that shows that a belongs to the set A. The computation of the value at a of a function $f : A \to B$ would depend only on a, and not on the proof that a belongs to A (functions are extensional). So Bishop's remark is correct if he admits functions whose value depends on both a and a proof that $a \in A$, but is not correct if, as most constructive practitioners do, he only admits extensional functions.

[13] Since the solution to Problem 2 on page 58 of [4] uses an argument very similar to that of the Goodman–Myhill proof, it is reasonable to suggest that Bishop may have been aware of their result when he published his book.

The axiom of choice will hold constructively if the set A is one for which no computation is necessary to demonstrate that an element belongs to it; Bishop calls such sets **basic sets**. Most constructivists would regard the set $\mathbf{N}^+$ of positive integers as a basic set, a belief that is reflected in the acceptance of the **principle of countable choice**:

> If to each positive integer n there corresponds an element y of a set A such that $P(n, y)$, then there is a function $f : \mathbf{N}^+ \to A$ such that $P(n, f(n))$ for each $n \in \mathbf{N}^+$.

In fact, many constructive proofs use the stronger **principle of dependent choice**:

> If $a \in A$, and to each $x \in A$ there corresponds $y \in A$ such that $P(x, y)$, then there exists a function $f : \mathbf{N}^+ \to A$ such that $f(1) = a$ and $P(f(n), f(n+1))$ for each $n \in \mathbf{N}^+$.

There is growing evidence that many applications of even these weaker choice principles can be avoided in BISH (see, for example, [24, 25, 26, 56]).

6 A modern view of BISH

In the last few years Fred Richman has advocated the view,[14] based on his experience as a practitioner of constructive mathematics for more than a quarter of a century, that Bishop's mathematics is simply mathematics with intuitionistic logic [53, 55]. On the one hand, the original constructivists' desire for algorithmic interpretability forces us to use intuitionistic logic; on the other, the exclusive use of intuitionistic logic seems to result, inevitably, in arguments that are entirely algorithmic in character. Is this Bishop's 'secret still on the point of being blabbed' ([4], epigraph):

> *algorithmic mathematics is equivalent to mathematics that uses only intuitionistic logic?*

If this is the case—and all the evidence of our constructive mathematical practice suggests that it is—then we can carry out our mathematics using intuitionistic logic on any reasonably defined mathematical objects, not just some class of 'constructive' objects. Constructive mathematics becomes a matter of epistemology, rather than ontology. However, this does not preclude the ontological possibility that, as Brouwer maintained, mathematical objects are mental constructs.

> *I adopt Richman's viewpoint for the remainder of this paper, which is written entirely within the framework of BISH except where it is clearly stated otherwise.*

[14] This view of BISH is prefigured in [9].

From that viewpoint, CLASS, RUSS, and INT can each be regarded as BISH plus some additional principles. In CLASS, the principle is the law of excluded middle, added to the intuitionistic logic of BISH. In RUSS, it is the axiom CPF characterising computable partial functions from $\mathbb{N}$ to $\mathbb{N}$. To obtain INT, we add to BISH the principle of continuous choice and the fan theorem.

We obtain alternative information about the BISH–INT and BISH–CLASS borderlines from the following result of Julian and Richman ([42]; see also pages 127–128 of [19]).[15]

Theorem 4 *Let C be the fan consisting of all finite sequences in $\{-1, 1\}$. If B is a detachable bar of C, then there exists a nonnegative uniformly continuous function f on $[0,1]$ such that*

- *$f(x) > 0$ for all x if and only if B is a bar for C;*
- *$\inf f > 0$ if and only if B is a uniform bar for C.*

Conversely, if f is a nonnegative uniformly continuous mapping on $[0,1]$, then there exists a detachable subset B of C satisfying these two conditions.

An immediate corollary of this theorem is a clarification of the status of the fan theorem.

Corollary 5 *The following statements are equivalent within BISH.*

— *Every detachable bar of C is a uniform bar.*

— *Every uniformly continuous map of $[0,1]$ into the positive real line has positive infimum.*[16]

On the other hand, as it can be shown that, under CPF, there exists a detachable bar for C that is not uniform (page 112 of [19]), we have

Corollary 6 *If CPF holds, then there exists a uniformly continuous mapping f of $[0,1]$ into the positive real line such that* $\inf f = 0$.

We conclude from these two corollaries that, within BISH, we can prove neither the proposition

> *Every uniformly continuous, positive-valued function on $[0,1]$ has positive infimum*

[15] Such results can be regarded as contributions to the *reverse mathematics* of Simpson and others [63].

[16] The sup and inf of a uniformly continuous real-valued function on a compact metric space always exist, although they may not be attained, in BISH.

nor its negation. If we add that (classically true) proposition to BISH as an axiom, then we are half-way to INT, it only being necessary to add the principle of continuous choice to go the whole way.

This completes my introduction to the foundations of constructive mathematics. Since I am more concerned, in these lectures and professionally, with the practice, rather than the formalism, of constructive mathematics, I shall say nothing about formal systems for BISH other than to point you towards Myhill's intuitionistic ZF [49] and constructive set theory [50]; Martin-Löf's theory of types [47], which has had such an influence on the 'proofs as programs' activity of several computer science research groups ([31], [38]); and other systems of Friedman [36], Feferman [35], and Bridges [12]. The standard reference for a formal development of INT is [44]. Excellent general references for the formal foundations of BISH, RUSS, and INT are [2] and [67].

7 The real line

Up to this stage, I have referred to the real numbers and their constructive properties without actually giving a precise definition of what it means, constructively, to be a real number. The time has come to set matters right.

Bishop introduced the real numbers as Cauchy sequences of rationals with highly specific convergence rates. Others have used Dedekind's approach to constructing reals [54], and there is even an axiomatic development [14]. I prefer here to introduce an interval-arithmetic development of the real numbers, based on one found in a recent book by Aberth [1].

By a **real number** we mean a nonempty subset $\mathbf{x}$ of $\mathbf{Q} \times \mathbf{Q}$ such that for all elements (q, q') and (r, r') of $\mathbf{x}$,

- ▷ $q < q'$,
- ▷ the closed intervals $[q, q']$ and $[r, r']$ (intervals in $\mathbf{Q}$) intersect, and
- ▷ for each positive rational ε there exists (q, q') in $\mathbf{x}$ such that $q' - q < \varepsilon$.

The underlying intuition here is that the elements of $\mathbf{x}$ are the rational endpoints of closed intervals, with one point, namely $\mathbf{x}$, common to all those intervals. Any rational number q gives rise to a canonical real number

$$\mathrm{q} = \{(q, q)\}$$

with which the original rational q is identified.

Two real numbers $\mathbf{x}$ and $\mathbf{y}$ are

- ▶ **equal**, written $\mathbf{x} = \mathbf{y}$, if for all $(q, q') \in \mathbf{x}$ and all $(r, r') \in \mathbf{y}$, the intervals $[q, q']$ and $[r, r']$ intersect;
- ▶ **unequal** (or **distinct**), written $\mathbf{x} \neq \mathbf{y}$, if there exist $(q, q') \in \mathbf{x}$ and $(r, r') \in \mathbf{y}$ such that the intervals $[q, q']$ and $[r, r']$ do not intersect.

It is almost immediate that $\neq$ satisfies the defining properties of an inequality relation. Let's check that equality is an equivalence relation. It is trivial that it is reflexive and symmetric, so only transitivity has to be handled. Let $\mathbf{x} = \mathbf{y}$ and $\mathbf{y} = \mathbf{z}$, and suppose that for some $(q, q') \in \mathbf{x}$ and $(r, r') \in \mathbf{z}$ we have $[q, q'] \cap [r, r'] = \emptyset$. We may assume without loss of generality that $q' < r$. Let ε be the positive rational number $r - q'$, and choose $(s, s') \in \mathbf{y}$ such that $s' - s < \varepsilon$. Then the interval $[s, s']$ cannot intersect *both* $[q, q']$ and $[r, r']$. This is absurd since $\mathbf{y} = \mathbf{x}$ and $\mathbf{y} = \mathbf{z}$. Thus

$$\neg([q, q'] \cap [s, s'] = \emptyset).$$

Since we are working with intervals in $\mathbb{Q}$, we can turn this round to construct a point of $[q, q'] \cap [s, s']$, as follows.

Lemma 7 *Let I, J be closed bounded intervals in $\mathbb{Q}$ such that $\neg(I \cap J = \emptyset)$. Then there exists $r \in \mathbb{Q}$ such that $r \in I \cap J$.*

Proof. Let $I = [a, b]$ and $J = [c, d]$. If $b < c$, then $I \cap J = \emptyset$, a contradiction. Hence $c \leqslant b$. Likewise, $a \leqslant d$. If $c < a$, then $I \subset J$. If $a \leqslant c$, then either $b < d$ and therefore $b \in I \cap J$, or else $d \leqslant b$ and therefore $d \in I \cap J$. ∎

Taken with the equality and inequality we have defined above, the collection of real numbers forms a set—the **real line** $\mathbb{R}$.

Let $\mathbf{x}, \mathbf{y}$ be real numbers. We say that $\mathbf{x}$ is **greater than** $\mathbf{y}$, and that $\mathbf{y}$ is **less than** $\mathbf{x}$, if there exist $(q, q') \in \mathbf{x}$ and $(r, r') \in \mathbf{y}$ such that $q > r'$; we then write $\mathbf{x} > \mathbf{y}$ or, equivalently, $\mathbf{y} < \mathbf{x}$. On the other hand, we say that $\mathbf{x}$ is **greater than or equal to** $\mathbf{y}$, and that $\mathbf{y}$ is **less than or equal to** $\mathbf{x}$, if for all $(q, q') \in \mathbf{x}$ and all $(r, r') \in \mathbf{y}$ we have $q' \geqslant r$; we then write $\mathbf{x} \geqslant \mathbf{y}$ or, equivalently, $\mathbf{y} \leqslant \mathbf{x}$. Clearly, $\mathbf{x} \geqslant \mathbf{x}$ and $\mathbf{x} \not> \mathbf{x}$ (that is, $\neg(\mathbf{x} > \mathbf{x})$).

The following properties of the inequality relations on $\mathbb{R}$ are relatively easy to establish:

$$\begin{aligned}
\mathbf{x} = \mathbf{y} &\Leftrightarrow \mathbf{x} \geqslant \mathbf{y} \wedge \mathbf{y} \geqslant \mathbf{x},\\
\mathbf{x} \neq \mathbf{y} &\Leftrightarrow \mathbf{x} > \mathbf{y} \vee \mathbf{x} < \mathbf{y},\\
\mathbf{x} > \mathbf{y} &\Rightarrow \mathbf{y} \not> \mathbf{x},\\
\mathbf{x} \not> \mathbf{y} &\Rightarrow \mathbf{y} \geqslant \mathbf{x},\\
\mathbf{x} > \mathbf{y} &\Rightarrow \mathbf{x} \geqslant \mathbf{y}.
\end{aligned}$$

In connection with the second last of these, note that the statement

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}\ (\neg(\mathbf{x} \geqslant \mathbf{y}) \Rightarrow \mathbf{y} > \mathbf{x})$$

implies Markov's Principle. To see this, let (a_n) be an increasing binary sequence (a_n) such that $\neg\forall n\,(a_n = 0)$, and define a real number by

$$\mathbf{x} = \left\{\left(0, \frac{1}{n}\right) : a_n = 0\right\} \cup \left\{\left(\frac{1}{n}, \frac{1}{n}\right) : a_n = 1 - a_{n-1}\right\}. \tag{1}$$

Then $\neg(0 \geqslant \mathbf{x})$: for if $0 \geqslant q$ for all $(q, q') \in \mathbf{x}$, then $a_n = 0$ for all n, a contradiction. However, if $\mathbf{x} > 0$, then there exists $(q, q') \in \mathbf{x}$ such that $q > 0$; whence $(q, q') = \left(\frac{1}{n}, \frac{1}{n}\right)$ for an (unique) n such that $a_n = 1 - a_{n-1}$.

Here is the transitivity of inequality, stated without proof:

$$(\mathbf{x} > \mathbf{y} \geqslant \mathbf{z} \Rightarrow \mathbf{x} > \mathbf{z}) \wedge (\mathbf{x} \geqslant \mathbf{y} > \mathbf{z} \Rightarrow \mathbf{x} > \mathbf{z}). \tag{2}$$

The classical law of trichotomy,

$$\forall \mathbf{x} \in \mathbb{R}\ (\mathbf{x} \geqslant \mathbf{0} \Rightarrow \mathbf{x} > \mathbf{0} \vee \mathbf{x} = \mathbf{0}),$$

implies LPO. For, given an increasing binary sequence (a_n), and defining the real number $\mathbf{x}$ as at (1), we routinely check that $\mathbf{x} \geqslant \mathbf{0}$; that if $\mathbf{x} = \mathbf{0}$, then $a_n = 0$ for all n; and that if $\mathbf{x} > \mathbf{0}$, then there exists n such that $a_n = 1$. Even the weaker form of trichotomy,

$$\forall \mathbf{x} \in \mathbb{R}\ (\mathbf{x} \geqslant \mathbf{0} \vee \mathbf{x} \leqslant \mathbf{0}),$$

is nonconstructive, as it implies LLPO. To prove this, given a binary sequence (a_n) at most one term equal to 1, define a real number by

$$\mathbf{x} = \left\{ \left(-\frac{1}{n}, \frac{1}{n} \right) : \forall k \leqslant n\, (a_k = 0) \right\} \cup \left\{ \left((-1)^n \frac{1}{n}, (-1)^n \frac{1}{n} \right) : a_n = 1 \right\}.$$

If $\mathbf{x} \geqslant \mathbf{0}$, then it is impossible that $a_n = 1$ for an odd n, so $a_n = 0$ for all odd n. Likewise, if $\mathbf{x} \leqslant \mathbf{0}$, then $a_n = 0$ for all even n.

Before you despair entirely, let me give you a surprisingly powerful constructive substitute for these two inadmissible versions of trichotomy. This result shows that, while exact splits of the real line cannot be carried out constructively, overlapping splits can.

Proposition 8 *If* $\mathbf{a} > \mathbf{b}$, *then for all* $\mathbf{x}$ *either* $\mathbf{a} > \mathbf{x}$ *or* $\mathbf{x} > \mathbf{b}$.

Proof. There exist $(q, q') \in \mathbf{a}$ and $(r, r') \in \mathbf{b}$ such that $r' < q$. Given a real number $\mathbf{x}$, we can find $(s, s') \in \mathbf{x}$ such that $s' - s < q - r'$. If $s' < q$, then $\mathbf{x} < \mathbf{a}$; if $s' \geqslant q$, then $r' < s$ and so $\mathbf{b} < \mathbf{x}$. ∎

For convenience, let's dispose here of a couple of simple results.

Lemma 9 *If* $(q, q') \in \mathbf{x}$, *then* $q \leqslant \mathbf{x} \leqslant q'$.

Proof. Rational arithmetic shows that for each $(r, r') \in \mathbf{x}$, since $[r, r']$ meets $[q, q']$, we have $r' \geqslant q$ and $q' \geqslant r$. ∎

Lemma 10 *For each real number* $\mathbf{x}$ *there exist rational numbers* q, q' *such that* $q < \mathbf{x} < q'$.

Proof. Let (r, r') be any element of $\mathbf{x}$. By Lemma 9, $r \leqslant \mathbf{x} \leqslant r'$. Choosing q, q' in $\mathbb{Q}$ with $q < r \leqslant r' < q'$, we see from (2) that $q < \mathbf{x} < q'$. ∎

Let S be a set of real numbers. We say that a real number $\mathbf{b}$ is an **upper bound** of/for S if $\mathbf{s} \leqslant \mathbf{b}$ for all $\mathbf{s} \in S$; and that $\mathbf{b}$ is the (perforce unique) **least upper**

bound of S if it is an upper bound for S and for each $\mathbf{x} < \mathbf{b}$ there exists $\mathbf{s} \in S$ such that $\mathbf{x} < \mathbf{s}$. In the latter event we also call $\mathbf{b}$ the **supremum** of S and we denote it by $\sup S$.

To see that the classical least–upper–bound principle implies the law of excluded middle, let P be any constructively meaningful proposition and define

$$S = \{\mathbf{x} \in \mathbb{R} : \mathbf{x} = \mathbf{0} \vee (\mathbf{x} = \mathbf{1} \wedge P)\}.$$

This set is nonempty (it contains $\mathbf{0}$) and bounded above by $\mathbf{1}$. Suppose it has a least upper bound $\mathbf{b}$. By Proposition 8, either $\mathbf{b} > \mathbf{0}$ or $\mathbf{b} < \mathbf{1}$. In the first case, by the definition of 'supremum', there exists $\mathbf{s} \in S$ with $\mathbf{s} > 0$; whence $\mathbf{s}$ must equal $\mathbf{1}$, so $\mathbf{1} \in S$ and therefore P holds. On the other hand, if $\mathbf{b} < \mathbf{1}$, then $\mathbf{1} \notin S$ and so $\neg P$ holds.

We define a set S of real numbers to be **order located** if for all rational numbers a, b with $a < b$, either $\mathbf{x} \leqslant b$ for all $\mathbf{x}$ in S or else there exists $\mathbf{x} \in S$ with $a < \mathbf{x}$. This definition enables us to state the **constructive least–upper–bound principle**, which is also known as the **Dedekind (order) completeness of $\mathbb{R}$**.

Theorem 11 *Let S be a nonempty set of real numbers that is both bounded above and order located. Then the least upper bound of S exists.*

Proof. Let B be the set of upper bounds for S, and define

$$\xi = \{(q, q') \in \mathbb{Q} \times \mathbb{Q} : \exists \mathbf{s} \in S \; \exists \mathbf{b} \in B \;\; (q \leqslant \mathbf{s} \leqslant \mathbf{b} \leqslant q')\}.$$

Taken with Lemma 9 and (2), the hypotheses ensure that ξ is nonempty. If (q, q') and (r, r') belong to ξ, then there exist $\mathbf{s}_1, \mathbf{s}_2 \in S$ and $\mathbf{b}_1, \mathbf{b}_2 \in B$ such that $q \leqslant \mathbf{s}_1 \leqslant \mathbf{b}_1 \leqslant q'$ and $r \leqslant \mathbf{s}_2 \leqslant \mathbf{b}_2 \leqslant r'$. Then $\mathbf{s}_1 \leqslant \mathbf{b}_2$, so $q \leqslant \mathbf{b}_2 \leqslant r'$; similarly, $r \leqslant q'$. Rational arithmetic shows that there exists a rational number in $[q, q'] \cap [r, r']$. To complete the proof that ξ is a real number, we show that for each rational $\varepsilon > 0$ there exists $(q, q') \in \xi$ with $q' - q < \varepsilon$. To this end, fix (a, a') in ξ. If $a = a'$, then there is nothing to prove; so we may assume that $a < a'$. Construct rational numbers $a_0 = a < a_1 < a_2 < \cdots < a_n = a'$ such that $a_i - a_{i-1} < \varepsilon/2$ for $1 \leqslant i \leqslant n$. Since S is order located, either $a_2 \in B$ or else $a_1 < \mathbf{s}$ for some element $\mathbf{s}$ of S. In the first case, $(a_0, a_2) \in \xi$ and $a_2 - a_0 < \varepsilon$. In the second, either $a_3 \in B$ and therefore $(a_1, a_3) \in \xi$ and $a_3 - a_1 < \varepsilon$; or else $a_2 < \mathbf{s}$ for some element $\mathbf{s}$ of X. Carrying on in this way, since $a_n \in B$ we can be sure of finding $k < n - 1$ such that $(a_k, a_{k+2}) \in \xi$ and $a_{k+2} - a_k < \varepsilon$. Thus ξ is indeed a real number.

To show that ξ is an upper bound for S, consider any (q, q') in ξ and any $\mathbf{s}$ in S. There exists $\mathbf{b} \in B$ such that $q \leqslant \mathbf{b} \leqslant q'$. For any (r, r') in $\mathbf{s}$ we have $r \leqslant \mathbf{s}$, by Lemma 9, and therefore $r \leqslant \mathbf{b}$; whence $r \leqslant q'$. It follows that $\mathbf{s} \leqslant \xi$.

Finally, if $\mathbf{x} < \xi$, then we can find $(q, q') \in \xi$ and $(r, r') \in \mathbf{x}$ such that $q > r'$. It follows from Lemma 9 that $\mathbf{x} \leqslant r' < q$. By definition of ξ, there exists $\mathbf{s} \in S$ with $q \leqslant \mathbf{s}$; whence $\mathbf{x} < \mathbf{s}$, by (2). This completes the proof that ξ is the least upper bound for S. ■

Let S be a set of real numbers. We say that a real number $\mathbf{b}$ is a **lower bound** of/for S if $\mathbf{b} \leqslant \mathbf{s}$ for all $\mathbf{s} \in S$; and that $\mathbf{b}$ is the (perforce unique) **greatest lower bound** of S if it is a lower bound for S and for each $\mathbf{x} > \mathbf{b}$ there exists $\mathbf{s} \in S$ such that $\mathbf{x} > \mathbf{s}$. In the latter event we also call $\mathbf{b}$ the **infimum** of S and we denote it by $\inf S$.

In order to establish a greatest lower bound principle analogous to Theorem 11, we define the **negative** of a real number $\mathbf{x}$ to be

$$^{-}\mathbf{x} = \{(q, q') \in \mathbb{Q} \times \mathbb{Q} : (^{-}q', ^{-}q) \in \mathbf{x}\},$$

where, for a rational number q, the expression ^{-}q denotes the negative defined in the usual way.

Corollary 12 *Let S be a nonempty set of real numbers that is bounded below and satisfies the property: for all rational numbers a, b with $a < b$, either a is a lower bound for S or else there exists $\mathbf{s} \in S$ with $\mathbf{s} < b$. Then the greatest lower bound of S exists.*

Proof. Apply Theorem 11 to the set

$$T = \{^{-}\mathbf{s} : \mathbf{s} \in S\},$$

(which is nonempty and bounded above) to construct its supremum $\mathbf{b}$. Then $^{-}\mathbf{b}$ is the infimum of S. ∎

The **absolute value** of the real number $\mathbf{x}$ is the set of rational pairs of the form

$$(\min\{|q|, |q'|\}, \max\{|q|, |q'|\}) \tag{3}$$

with (q, q') in X, where in (3) the absolute values are taken in $\mathbb{Q}$.

Lemma 13 *For each real number $\mathbf{x}$ there exists a positive integer n such that $|\mathbf{x}| < n$.*

Proof. Pick (q, q') in $\mathbf{x}$, choose a positive integer n such that $\max\{|q|, |q'|\} < n$, and apply Lemma 9 and (2). ∎

Lemma 14 *For each real number $\mathbf{x}$ there exist a positive integer N such that* $\max\{|q|, |q'|\} < N$ *for all $(q, q') \in \mathbf{x}$ with $q' - q < 1$.*

Proof. Pick n as in Lemma 13, and set $N = n + 1$. If $(q, q') \in \mathbf{x}$ and $q' - q < 1$, then

$$\min\{|q|, |q'|\} \leqslant |\mathbf{x}| \leqslant \max|q|, |q'|$$

by Lemma 9, and

$$0 \leqslant \max\{|q|, |q'|\} - \min\{|q|, |q'|\} < 1.$$

Hence

$$\max\{|q|, |q'|\} < |\mathbf{x}| + 1 < N,$$

as required. ∎

We now introduce the arithmetic operations on real numbers. Given real numbers **x** and **y**, we need to find the appropriate ways of combining the rational pairs (representing intervals) that constitute the real number **x** with those that constitute **y**, in order to create the rational pairs that represent $\mathbf{x} \circ \mathbf{y}$, where $\circ$ stands for any of the operations $+, -, \times, \div$. We begin with the easy definitions of $+$ and $-$, for the moment leaving aside the more complicated ones for $\times$ and $\div$.

We define the **sum** $\mathbf{x}+\mathbf{y}$ and **difference** $\mathbf{x}-\mathbf{y}$ of the real numbers $\mathbf{x}, \mathbf{y}$ to be, respectively,

$$\begin{aligned} \mathbf{x}+\mathbf{y} &= \{(s,s') : \exists (q,q') \in \mathbf{x}\, \exists (r,r') \in \mathbf{y}\ (s = q+r \wedge s' = q'+r')\}, \\ \mathbf{x}-\mathbf{y} &= \{(s,s') : \exists (q,q') \in \mathbf{x}\, \exists (r,r') \in \mathbf{y}\ (s = q-r \wedge s' = q'-r')\}. \end{aligned}$$

Let's verify, for example, that $\mathbf{x}+\mathbf{y}$ is a real number. Let $(q_1,q_1') \in \mathbf{x}$, and $(r_1,r_1') \in \mathbf{y}$. Certainly, $q_i+r_i \leqslant q_i'+r_i'$. Moreover, given a positive rational number ε, we can arrange that $q_1'-q_1 < \varepsilon/2$ and $r_1'-r_1 < \varepsilon/2$, so $(q_1+r_1') - (q_1+r_1) < \varepsilon$. It remains, then, to show that if also $(q_2,q_2') \in \mathbf{x}$ and $(r_2,r_2') \in \mathbf{y}$, then the intervals $[q_i+r_i, q_i'+r_i']$ $(i=1,2)$ in $\mathbb{Q}$ intersect. This is easy: there exist rational numbers ξ, η such that $q_i \leqslant \xi < q_i'$ and $r_i \leqslant \eta \leqslant r_i'$ $(i=1,2)$; whence

$$q_i + r_i \leqslant \xi + \eta \leqslant q_i' + r_i'\ (i=1,2).$$

Thus $\mathbf{x}+\mathbf{y}$ is a real number.

Since, as is routinely verified, $\mathbf{x}-\mathbf{y} = \mathbf{x}+(^{-}\mathbf{y})$, we adopt the normal convention of writing $-\mathbf{x}$ instead of $^{-}\mathbf{x}$.

We next define the **product** $\mathbf{x} \times \mathbf{y}$, usually denoted by $\mathbf{xy}$, to be the set of all rational pairs (s,s') such that there exist $(q,q') \in \mathbf{x}$ and $(r,r') \in \mathbf{y}$ with

$$s = \min\{qr, qr', q'r, q'r'\}, \quad s' = \max\{qr, qr', q'r, q'r'\}.$$

Certainly, $s \leqslant s'$. Compute a positive integer N such that if $(q,q') \in \mathbf{x}$ and $q'-q < 1$, then $\max\{|q|, |q'|\} < N$, and such that if $(r,r') \in \mathbf{y}$ and $r'-r < 1$, then $\max\{|r|, |r'|\} < N$. Given a rational ε with $0 < \varepsilon < 1$, we can arrange that $q'-q < \varepsilon/2N$ and $r'-r < \varepsilon/2N$; then

$$s' - s < (q'-q)\max\{|r|, |r'|\} + (r'-r)\max\{|q|, |q'|\} < \varepsilon.$$

For $i = 1,2$ let $(q_i,q_i') \in \mathbf{x}$, $(r_i,r_i') \in \mathbf{y}$, and

$$s_i = \min\{q_ir_i, q_ir_i', q_i'r_i, q_i'r_i'\}, \quad s_i' = \max\{q_ir_i, q_ir_i', q_i'r_i, q_i'r_i'\}.$$

Pick ξ in $[q_1,q_1'] \cap [q_2,q_2']$ and η in $[r_1,r_1'] \cap [r_2,r_2']$; it is easy to verify that $\xi\eta \in [s_1,s_1'] \cap [s_2,s_2']$.

When dealing with division, we consider two real numbers $\mathbf{x}, \mathbf{y}$ with $\mathbf{y} \neq \mathbf{0}$. Let's illustrate the definition with the case $\mathbf{y} > \mathbf{0}$. To construct a rational pair (s,s') in the **quotient** $\mathbf{x}/\mathbf{y}$, we first pick $(q,q') \in \mathbf{x}$ and $(r,r') \in \mathbf{y}$ such that $r > 0$. If $q \geqslant 0$, we set $s = q/r'$, $s' = q'/r$; if $q < 0 < q'$, we set $s = q/r$, $s' = q'/r$; and if $q' \leqslant 0$, we set $s = q'/r$, $s' = q/r'$. Pairs (s, s') constructed according to these rules, and only such pairs, belong to $\mathbf{x}/\mathbf{y}$. We omit the detailed argument showing that $\mathbf{x}/\mathbf{y}$ is indeed a real number.

As classically, we say that a sequence $(\mathbf{x}_n)_{n=1}^{\infty}$ of real numbers **converges** to a real number $\mathbf{x}_\infty$, called the **limit** of the sequence, if

$$\forall \varepsilon > 0\ \exists N\ \forall n \geqslant N\ \left(|\mathbf{x}_\infty - \mathbf{x}_n| < \varepsilon\right).$$

We then write

$$\mathbf{x}_n \to \mathbf{x}_\infty \text{ as } n \to \infty$$

or

$$\mathbf{x}_\infty = \lim_{n\to\infty} \mathbf{x}_n.$$

The uniqueness of the limit, and the basic algebraic properties of sequences of real numbers, will be assumed without proof, since the proofs are virtually the same as their classical counterparts.

By a **Cauchy sequence** of real numbers we mean a sequence $(\mathbf{x}_n)$ such that

$$\forall \varepsilon > 0\ \exists N\ \forall m, n \geqslant N\ \left(|\mathbf{x}_m - \mathbf{x}_n| < \varepsilon\right).$$

Every convergent sequence of real numbers is a Cauchy sequence; the converse statement is the so-called **(Cauchy) completeness of $\mathbb{R}$**.

Theorem 15 *Every Cauchy sequence of real numbers converges to a real number.*

Proof. Let $(\mathbf{x}_n)_{n=1}^{\infty}$ be a Cauchy sequence of real numbers, and, using the principle of countable choice, compute a function $k \rightsquigarrow n_k$ from $\mathbf{N}^+$ to $\mathbf{N}^+$ such that

$$\forall k\ \forall m, n \geqslant n_k\ \left(|\mathbf{x}_m - \mathbf{x}_n| < 2^{-k}\right).$$

Again using countable choice, construct a sequence $((q_k, q_k'))_{k=1}^{\infty}$ such that

$$\forall k\ \left((q_k, q_k') \in \mathbf{x}_{n_k} \wedge q_k' - q_k < 2^{-k}\right).$$

Set

$$r_k = q_k - 2^{-k},\ r_k' = q_k' + 2^{-k}.$$

Then for all $n \geqslant n_k$,

$$r_k \leqslant \mathbf{x}_{n_k} - 2^{-k} < \mathbf{x}_n \leqslant \mathbf{x}_{n_k} + 2^{-k} \leqslant r_k'. \tag{4}$$

It follows that for all $j \geqslant k$,

$$\mathbf{x}_{n_j} \in \left[r_j, r_j'\right] \cap \left[r_k \cap r_k'\right].$$

Since $r'_k - r_k < 2^{-k+1} \to 0$ as $k \to \infty$, we conclude that

$$\mathbf{x}_\infty = \{(r_k, r'_k) : k \geqslant 1\}$$

is a real number. From (4) we have $\mathbf{x}_n \in [r_k, r'_k]$ for all $n \geqslant n_k$. It follows that

$$\forall k \, \forall n \geqslant n_k \, \left(|\mathbf{x}_n - \mathbf{x}_\infty| \leqslant r'_k - r_k < 2^{-k+1}\right).$$

Hence $\mathbf{x}_n \to \mathbf{x}_\infty$ as $n \to \infty$. ■

For the remainder of these lectures we shall drop the use of boldface type to denote real numbers; it has served its purpose to signal a distinction between a rational number and a real number—a set of special pairs of rational numbers. We shall also assume, without further comment, basic properties of real numbers—for example, $x^2 \geqslant 0$ for all real x —that can easily be deduced from the foregoing results. Note, however, that although

$$\forall x \in \mathbb{R} \, \left(x^2 = 0 \Rightarrow x = 0\right)$$

and

$$\forall x, y \in \mathbb{R} \, \left((x \neq 0 \wedge xy = 0) \Rightarrow y = 0\right),$$

the statement

$$\forall x, y \in \mathbb{R} \, \left(xy = 0 \Rightarrow x = 0 \vee y = 0\right)$$

implies LLPO.

The **complex plane** $\mathbb{C}$ consists of all **complex numbers**—ordered pairs (x, y) of real numbers—with addition and multiplication defined by

$$\begin{aligned}(x, y) + (x', y') &= (x + x', y + y'), \\ (x, y) \times (x', y') &= (xx' - yy', xy' + x'y),\end{aligned}$$

and equality and inequality defined by

$$\begin{aligned}(x, y) = (x', y') &\Leftrightarrow x = x' \wedge y = y', \\ (x, y) \neq (x', y') &\Leftrightarrow x \neq x' \vee y \neq y'.\end{aligned}$$

We embed $\mathbb{R}$ as a subset of $\mathbb{C}$ in the usual way, by identifying the real number x with the complex number $(x, 0)$. The pair $\mathrm{i} = (0, 1)$ then has the special property that $\mathrm{i}^2 = -1$; and every complex number $z = (x, y)$ can be written in the form $x + \mathrm{i}y$, with **real part** $\operatorname{Re} z = x$ and **imaginary part** $\operatorname{Im} z = y$. We shall assume basic properties of $\mathbb{C}$ as they are needed. The same goes for the Euclidean spaces $\mathbb{R}^n$ and $\mathbb{C}^n$, which are now defined in the standard ways.

Constructive proofs of typical applications of the completeness of $\mathbb{R}$ usually pose no problems. For example, if (a_n) and (b_n) are sequences of real numbers such that $0 < a_n \leqslant b_n$ for each n, and if $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ converges (*comparison test*). For, given $\varepsilon > 0$, since the partial sums of a convergent series form a Cauchy sequence, we can find N such that

$$0 < \sum_{n=j}^{k} b_n < \varepsilon$$

whenever $k > j \geqslant N$. For all such j, k we then have $0 < \sum_{n=j}^{k} a_j < \varepsilon$. So the partial sums of $\sum_{n=1}^{\infty} a_n$ form a Cauchy sequence, which converges by the completeness of $\mathbb{R}$. A particular case of this occurs when $b_n = r^n$ for some fixed r with $|r| < 1$: in that case, $\sum_{n=1}^{\infty} b_n$ is a geometric series, and $\sum_{n=1}^{\infty} a_n$ converges to a sum at most $r/(1-r)$.

We shall see in the next section how completeness is used in constructive mathematics to prove propositions that are immediate classical consequences of omniscience principles.

8 Applications of the completeness of $\mathbb{R}$

For a first application of the completeness of $\mathbb{R}$ consider the classical intermediate value theorem:

If $f : [0,1] \to \mathbb{R}$ is continuous and $f(0)f(1) < 0$, then there exists ξ strictly between 0 and 1 such that $f(\xi) = 0$.

To see that this theorem is nonconstructive as it stands, let $(a_n)_{n=0}^{\infty}$ be a binary sequence with at most one term equal to 1, let

$$a = \sum_{n=0}^{\infty} (-1)^n 2^{-n} a_n,$$

and define a continuous (actually, uniformly continuous) $f : [0,1] \to \mathbb{R}$ such that

$$f(0) = -1,\ f(1) = 1, f\left(\tfrac{1}{3}\right) = a = f\left(\tfrac{2}{3}\right),$$

and f is linear on each of the intervals $\left[0, \frac{1}{3}\right], \left[\frac{1}{3}, \frac{2}{3}\right], \left[\frac{2}{3}, 1\right]$. Suppose that $f(\xi) = 0$. Then either $\xi > \frac{1}{3}$ or $\xi < \frac{2}{3}$. In the first case we have $\neg (a > 0)$ and therefore $a \leqslant 0$; in the second, $\neg (a < 0)$ and so $a \geqslant 0$. Thus the classical intermediate value theorem implies LLPO.

There are two standard elementary ways of proceeding from the hypotheses to the conclusion of the classical intermediate value theorem. In the first of these we define

$$\xi = \sup \{x \in [0,1] : f(x) < 0\}$$

and use the continuity of f at ξ to show that $f(\xi) \leqslant 0$ and $f(\xi) \geqslant 0$; this argument fails constructively at the definition of ξ as a supremum. The second way is an interval-halving argument which at first sight looks constructive—indeed, it is the basis for a numerical method of root-finding; but this, too, is nonconstructive, since it depends on the law of trichotomy

$$\forall x \in \mathbb{R}\ (x > 0 \vee x = 0 \vee x < 0),$$

which is equivalent to LPO. Fortunately, there is a constructive intermediate value theorem; in fact, there are several which are classically equivalent to the classical intermediate value theorem but are constructively distinct. Here are two of them

Theorem 16 *Let $f : [0,1] \to \mathbb{R}$ be continuous and such that $f(0)f(1) < 0$. Then for each $\varepsilon > 0$ there exists $x \in (0,1)$ such that $|f(x)| < \varepsilon$.*

Theorem 17 *Let $f : [0,1] \to \mathbb{R}$ be continuous and* **locally nonzero**, *in the sense that*

$$\forall x \in (0,1)\ \forall \varepsilon > 0\ \exists x'\ \left(|x - x'| < \varepsilon \wedge f(x') \neq 0\right).$$

If also $f(0)f(1) < 1$, then there exists $\xi \in (0,1)$ such that $f(\xi) = 0$.

The proofs of these theorems use an approximate interval–halving argument, which I illustrate by proving Theorem 17. Setting $a_1 = 0$ and $b_1 = 1$, and assuming without loss of generality that $f(0) < 0$ and $f(1) > 0$, we choose $x' \in \left(\frac{1}{3}, \frac{2}{3}\right)$ such that $f(x') \neq 0$. If $f(x') < 0$, we put $a_2 = x'$ and $b_2 = 1$; if $f(x') > 0$, we put $a_2 = a_1$ and $b_2 = x'$. Carrying on in this way, we construct a sequence $([a_n, b_n])_{n=1}^{\infty}$ of compact subintervals of $[0,1]$ such that

- $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$,
- $b_{n+1} - a_{n+1} \leqslant \frac{1}{3}(b_n - a_n)$,
- $f(a_n) < 0$ and $f(b_n) > 0$.

Then (a_n) is a Cauchy sequence, whose limit ξ exists by Theorem 15. It is easy to show that $f(\xi) = 0$. ■

The intermediate value theorem illustrates a common phenomenon in constructive mathematics, in which one classical theorem may split into several constructive ones, each of which is classically equivalent to the original.

Another application of the completeness of $\mathbb{R}$ is rather more amusing, and demonstrates that a classically trivial result may require some ingenuity to establish within BISH. Consider, for any real number a, the linear subset

$$\mathbb{R}a = \{ax : x \in \mathbb{R}\}$$

of $\mathbb{R}$. If $a = 0$, then $\mathbb{R}a = \{0\}$; whereas if $a \neq 0$, then $\mathbb{R}a = \mathbb{R}$. In either case, $\mathbb{R}a$ is both finite–dimensional and closed in $\mathbb{R}$. What happens if we do not know whether $a = 0$ or $a \neq 0$?

Proposition 18 *The following are equivalent statements.*

- *$\mathbb{R}a$ is finite–dimensional.*
- *$a = 0$ or $a \neq 0$.*
- *$\mathbb{R}a$ is closed in $\mathbb{R}$.*

Proof. We have already observed that the second statement implies both the first and the third. If $\mathbb{R}a$ is finite–dimensional, then either its dimension is 0, in which case $a = 0$, or else its dimension is 1; in the latter case there exists ξ such that $a\xi = 1$, and therefore $a \neq 0$. Hence the first statement implies the second. It remains to prove that the third implies the second. To this end, assume that $\mathbb{R}a$ is closed in $\mathbb{R}$, and, using dependent choice, construct a decreasing binary sequence $(\lambda_n)_{n=1}^{\infty}$ such that

$$\begin{aligned} \lambda_n = 1 &\Rightarrow |a| < 1/n^2, \\ \lambda_n = 0 &\Rightarrow |a| > 1/(n+1)^2. \end{aligned}$$

This construction is possible since either $1/n^2 > |a|$ or $|a| > 1/(n+1)^2$. The completeness of $\mathbb{R}$ ensures that the series $\sum_{n=1}^{\infty} \lambda_n a$ converges by comparison with the known convergent series $\sum_{n=1}^{\infty} 1/n^2$. Since $\mathbb{R}a$ is a closed subset of $\mathbb{R}$, the sum of the series $\sum_{n=1}^{\infty} \lambda_n a$ has the form ξa for some $\xi \in \mathbb{R}$. Choose $N > |\xi|$, and consider λ_N. If $\lambda_N = 0$, then $a \neq 0$ and we are finished; so we may assume that $\lambda_N = 1$. Suppose there exists $m \geqslant N$ such that $\lambda_{m+1} = 1 - \lambda_m$. Then $\lambda_1 = \lambda_2 = \cdots = \lambda_m = 1$, and $\lambda_k = 0$ for all $k > m$; so

$$\xi a = \sum_{n=1}^{\infty} \lambda_n a = \sum_{n=1}^{m} \lambda_n a = ma$$

and therefore $\xi = m$. (Note that as $\lambda_{m+1} = 0$, $|a| > 1/(m+2)^2$ and we can divide by a.) This is absurd, as $m \geqslant N > |\xi|$. We conclude that if $\lambda_N = 1$ for our special choice of N, then $\lambda_m = 1$ for all $m \geqslant N$ and hence for all m; it follows that $|a| < 1/m^2$ for all m, and therefore that $a = 0$. Thus, by the careful construction of a series whose convergence depends on the completeness of $\mathbb{R}$, and an equally careful estimation using its sum, we have been able to show that if $\mathbb{R}a$ is closed, then either $a = 0$ or $a \neq 0$. ∎

Since spaces of the type $\mathbb{R}a$ do not occur very often in advanced analysis, Proposition 18 may seem rather insignificant. But our result about $\mathbb{R}a$ suggested, and is a special case of, the more general theorem,

A Banach space with a compact generating set is finite-dimensional,

whose original constructive proof required several applications of completeness similar to, and in one case generalising, the one we have just used (see [58], or pages 41–44 of [19]; for a newer, alternative proof see [22]). In turn, this theorem enables us to prove that if the range of a compact linear mapping between normed spaces is complete, then that range is finite–dimensional, a result whose standard classical proof depends on a nonconstructive version of the Open Mapping Theorem from functional analysis (see Theorem 4.18 of [60], and Chapter 2 of [19]).

There are many situations where an application of completeness, and the construction of some auxiliary binary sequence (λ_n), similar to the one used in the proof of Proposition 18 has enabled the constructive mathematician to circumvent omniscience principles. The completeness usually has to be added to the hypotheses of what would otherwise be a trivial classical theorem. In our discussion of $\mathbb{R}a$, although the completeness of $\mathbb{R}$ is used to establish the convergence of the series $\sum_{n=1}^{\infty} \lambda_n a$, we really need $\mathbb{R}a$ to be complete in order to ensure that the series converges to a sum that belongs to $\mathbb{R}a$; the required completeness is implicitly contained in the hypothesis that $\mathbb{R}a$ is a closed subset of $\mathbb{R}$.

Among the interesting constructive theorems whose proofs use such applications of completeness are the following:

- *If f is a nonnegative Lebesgue integrable function that is positive throughout a set of positive measure, then* $\int f > 0$ ([8] Chapter 6, (4.13); see also Section 10 below).

- *A linear mapping T of a normed space X onto a Banach space is* **well-behaved**, *in the sense that if $x \in X$ and $x \neq y$ for each element of the kernel of T, then* $Tx \neq 0$ [18].

- *Let F be a finite-dimensional subspace of a real normed space X, and let $a \in X$ have* **at most one best approximation** *in F, in the following sense: if x, x' are distinct elements of F, then*

$$\max\{\|a - x\|, \|a - x'\|\} > \rho(a, F) \equiv \inf\{\|a - y\| : y \in F\}.$$

Then there exists a unique element $b \in F$ such that $\|a - b\| = \rho(a, F)$ [11].

As a footnote to Proposition 18, let me add Fred Richman's alternative, choice-free proof of the final implication. Assume that $\mathbb{R}a$ is closed in $\mathbb{R}$. For each $\varepsilon > 0$ we have either $\sqrt{|a|} < \varepsilon$ or else $|a| > 0$; in the latter case,

$$\sqrt{|a|} = \pm \frac{a}{\sqrt{|a|}} \in \mathbb{R}a.$$

Since $\varepsilon > 0$ is arbitrary, we now see that $\sqrt{|a|} \in \overline{\mathbb{R}a} = \mathbb{R}a$ and hence that there exists r such that $\sqrt{|a|} = ra$. Choosing a positive integer $N > r$, we have either $|a| > 0$ or $|a| < 1/N^2$. In the latter case, if $a \neq 0$, then

$$|r|\,|a| = |ra| = \sqrt{|a|} = \frac{|a|}{\sqrt{|a|}} > N\,|a|,$$

so $|r| > N$, a contradiction; whence $a = 0$.

9 Specker's Theorem

In this part of the lectures I shall present Richman's proof of Specker's theorem (page 218). This requires some definitions and a very useful lemma, whose proof, originating with Bishop ([4], page 177, Lemma 7), seems to be the first instance of the use of an auxiliary binary sequence (λ_n) such as that in the proof of Proposition 18 above.

A set S is called **finitely enumerable** if there exist $n \in \mathbf{N}$ and a mapping f of $\{1, 2, \ldots, n\}$ onto S; if the mapping f is one–one, then S is said to be **finite**. The empty set is finite, since it corresponds to the case $n = 0$ of the definition. If every finitely enumerable set is finite, then LPO holds.

Now let (X, ρ) be a metric space, and S a subset of X. We say that S is

- **located** (in X) if the **distance**

$$\rho(x, S) = \inf \{\rho(x, s) : s \in S\}$$

  exists for each $x \in X$;

- **totally bounded** if for each $\varepsilon > 0$ there exists a finite subset F of S, called a **finitely enumerable** ε**–approximation** to S, such that

$$\forall s \in S \ \exists x \in F \ \ (\rho(s, x) < \varepsilon);$$

- **compact** if it is both totally bounded and complete.

The statement

Every nonempty (inhabited) subset of $\mathbb{R}$ *is located*

entails LPO. For if $a \in \mathbb{R}$ and $\mathbb{R}a$ is located, then either $\rho(1, \mathbb{R}a) > 0$ or else $\rho(1, \mathbb{R}a) < 1$. In the first case we have $\neg (a \neq 0)$ and so $a = 0$; in the second, choosing ξ such that $|1 - a\xi| < 1$, we see that $a\xi \neq 0$ and hence that $a \neq 0$.

We have already observed that not every nonempty bounded subset of $\mathbb{R}$ has a supremum. However, things improve dramatically if we replace 'bounded' by 'totally bounded'.

Proposition 19 *Every nonempty totally bounded subset of* $\mathbb{R}$ *has a supremum and an infimum.*

Proof. Let S be a nonempty totally bounded subset of $\mathbb{R}$. We first consider the case where $S = \{x_1, \ldots, x_n\}$ is finitely enumerable. Given real numbers α, β with $\alpha < \beta$, we apply Proposition 8 n times to prove that either $x_k < \beta$ for each k or else there exists j such that $x_j > \alpha$. It follows from Theorem 11 that $\sup S$ exists.

Now consider the general case. Again let α, β be real numbers with $\alpha < \beta$, but this time write $\varepsilon = \frac{1}{2}(\beta - \alpha)$ and construct a finitely enumerable ε–approximation $\{x_1, \ldots, x_n\}$ to S. By the first part of the proof

$$\sigma = \sup \{x_1, \ldots, x_n\}$$

exists. By Proposition 8, either $\sigma > \alpha$ or $\sigma < \alpha + \varepsilon$. In the first case there exists j such that $x_j > \alpha$. In the second, consider any $x \in S$. Choosing j such that $|x - x_j| < \varepsilon$, we have

$$x \leqslant x_j + |x - x_j| < \sigma + \varepsilon < \alpha + 2\varepsilon = \beta.$$

So in this case, β is an upper bound for S. It follows from Theorem 11 that $\sup S$ exists. Similar arguments show that $\inf S$ exists. ∎

It is relatively straightforward to prove that a uniformly continuous function between metric space maps totally bounded sets onto totally bounded sets. Thus if $f : X \to \mathbb{R}$ is uniformly continuous and X is totally bounded, then, by Proposition 19,

$$\inf f = \inf_{x \in X} f(x)$$

and

$$\sup f = \sup_{x \in X} f(x)$$

both exist.

Proposition 20 *A totally bounded subset of a metric space is located.*

Proof. Let S be a totally bounded subset of a metric space X, and let $a \in X$. Then the mapping $x \rightsquigarrow \rho(a, x)$ is uniformly continuous on S, so, by Proposition 19, $\rho(a, S)$ exists. ∎

Why do we define compactness as we do, rather than in the traditional classical ways? It is easy to see that the sequential compactness of the closed interval $[0, 1]$ implies LPO. On the other hand, whereas the Heine–Borel–Lebesgue theorem holds in INT ([67], page 305, 3.5), it is false in RUSS ([19], page 60), so it cannot be proved in BISH.

We are now able to deal with Bishop's lemma, referred to above.

Lemma 21 *Let S be a nonempty complete located subset of a metric space X. For each $a \in X$ there exists $b \in S$ such that if $\rho(a, b) > 0$, then $\rho(a, S) > 0$.*

Proof. Construct an increasing binary sequence $(\lambda_n)_{n=1}^{\infty}$ such that

$$\begin{aligned} \lambda_n = 0 \quad &\Rightarrow \quad \rho(a, S) < 1/n, \\ \lambda_n = 1 \quad &\Rightarrow \quad \rho(a, S) > 1/(n+1). \end{aligned}$$

If $\lambda_1 = 1$, choose $s \in S$ and set $s_n = s$ for all n. If $\lambda_n = 0$, choose $s_n \in S$ such that $\rho(a, s_n) < 1/n$. If $\lambda_n = 1 - \lambda_{n-1}$, put $s_k = s_{n-1}$ for all $k \geqslant n$. Then (s_n) is a Cauchy sequence in S; in fact,

$$\rho(s_m, s_n) \leqslant \tfrac{2}{n} \quad (m \geqslant n).$$

Since S is complete, (s_n) converges to a limit b in S; letting $m \to \infty$ in the last display, we see that

$$\rho(b, s_n) \leqslant \tfrac{2}{n} \quad (n \geqslant 1).$$

Now suppose that $\rho(a, b) > 0$, and choose a positive integer N such that $\rho(a, b) > 3/N$. If $\lambda_N = 0$, then $\rho(a, s_N) < 1/N$ and so

$$\rho(a, b) \leqslant \rho(a, s_N) + \rho(s_N, b) < \tfrac{1}{N} + \tfrac{2}{N} = \tfrac{3}{N},$$

a contradiction. Hence $\lambda_N = 1$ and therefore $\rho(a, S) > 0$. ∎

Note that this lemma is a classical triviality: for, classically, we may take $b = a$ if $a \in S$, and b as any element of S if $a \notin S$.

Before proving Specker's Theorem, let us remind ourselves of some basic facts about the Cantor set

$$C = \left\{ \sum_{n=1}^{\infty} c_n 3^{-n} : \forall n \, (c_n \in \{0, 2\}) \right\}.$$

Lemma 22 *If a and b are two numbers in C that differ in their mth ternary digits, then $|a - b| \geqslant 3^{-m}$.*

Proof. Write

$$a = \sum_{n=1}^{\infty} a_n/3^n, \quad b = \sum_{n=1}^{\infty} b_n/3^n,$$

where $a_n, b_n \in \{0, 2\}$ for each n, and $a_m \neq b_m$. Let k be the smallest value of j such that $a_j \neq b_j$; then $k \leqslant m$, $a_j = b_j$ $(1 \leqslant j \leqslant k - 1)$, and $|a_k - b_k| = 2$. Hence

$$\begin{aligned} |a - b| &= \left| (a_k - b_k) \, 3^{-k} + \sum_{n=k+1}^{\infty} (a_n - b_n) \, 3^{-n} \right| \\ &\geqslant |a_k - b_k| \, 3^{-k} - \sum_{n=k+1}^{\infty} |a_n - b_n| \, 3^{-n} \\ &\geqslant 2.3^{-k} - \sum_{n=k+1}^{\infty} 2.3^{-n} \\ &= 2.3^{-k} - 2.3^{-k-1} \sum_{n=0}^{\infty} 3^{-n} \\ &= 2.3^{-k} - 2.3^{-k-1} \frac{1}{1 - \frac{1}{3}} = 2.3^{-k} - 3^{-k} = 3^{-k} \geqslant 3^{-m}. \end{aligned}$$ ∎

Using this lemma, we can easily show that C is closed in $\mathbb{R}$ and therefore complete. On the other hand, for each positive integer N the set

$$\left\{ \sum_{n=1}^{N} c_n 3^{-n} : c_n \in \{0, 2\} \text{ for } 1 \leqslant n \leqslant N \right\}$$

is a finite 3^{-N}-approximation to C; so C is totally bounded and hence compact.

We now give the

Proof of Specker's Theorem. Assuming CPF, and using the notation of Section 4, define rational numbers $r_n \in (0,1)$ as follows:

$$r_n = \sum_{i=1}^{n} s_i(m)3^{-m},$$

where

$$s_n(m) = \begin{cases} 2 & \text{if } m \in D_m(n) \text{ and } \varphi_m(m) = 0 \\ 0 & \text{otherwise.} \end{cases}$$

First consider any real number x in the Cantor set. There exists m such that φ_m is a total partial mapping of $\mathbb{N}$ into $\{0,2\}$ and

$$x = \sum_{n=1}^{\infty} \varphi_m(n)3^{-n}.$$

Since $\varphi_m(m)$ is defined, there exists $N \geqslant m$ such that if $m \in D_m(N)$, then

$$\varphi_m(m) = 2 - s_n(m) \quad (n \geqslant N).$$

Thus, for all $n \geqslant N$, the elements x and r_n of C differ in the mth ternary place. It follows from Lemma 22 that

$$|x - r_n| \geqslant 3^{-m} \quad (n \geqslant N).$$

Now consider any real number x. By Bishop's lemma, there exists $b \in C$ such that if $x \neq b$, then $\rho(x, C) > 0$. By the first part of this proof, there exist $\delta' > 0$ and N such that

$$|b - r_n| \geqslant \delta' \quad (n \geqslant N).$$

Either $|x - b| < \delta'/2$ or $x \neq b$. In the first case, for all $n \geqslant N$ we have

$$|x - r_n| \geqslant |b - r_n| - |x - b| > \tfrac{\delta'}{2}.$$

In the second case,

$$|x - r_n| \geqslant \rho(x, C) > 0$$

for all n. So in either case, the sequence (r_n) is eventually bounded away from x. ∎

It is a simple matter to extend Specker's Theorem to make (r_n) a strictly increasing sequence of rational numbers in $(0,1)$. With such a sequence we can produce *recursive counterexamples* to certain classical propositions. Here are some.

For each positive integer n, let I_n be a closed subinterval of $(0,1)$ centred at r_n, such that $|I_n| < 1/n$, and such that I_m and I_n are disjoint when $m \neq n$. Define a continuous function $f_n : [0,1] \to \mathbb{R}^{0+}$ such that $f_n(x) = 0$ for x outside I_n, $f_n(r_n) = 1$, and f_n is linear on each half of I_n. Then $f = \sum_{n=1}^{\infty} f_n$ is a continuous mapping on $[0,1]$ that is not uniformly continuous.

For a second example, define a continuous function $g_n : [0,1] \to \mathbb{R}$ such that $g_n(x) = 1$ if x is outside I_n, $g_n(r_n) = 1 - n^{-1}$, and g_n is linear on each half of I_n. Then $g = \sum_{n=1}^{\infty} g_n$ is a continuous, positive-valued function on $[0,1]$ whose infimum is 0. (As we noted in Section 4, under CPF there exists a *uniformly* continuous, positive-valued mapping whose infimum on $[0,1]$ is 0.)

10 Some recent developments

I would now like to sketch for you some aspects of constructive mathematics that have been developed since the publication of Bishop's book [4].

10.0.1 Ring theory

It is high time that I addressed constructive problems in algebra. Whereas, in classical algebra, the splitting field associated with a given polynomial is unique up to isomorphism, in the constructive approach the uniqueness of splitting fields for polynomials over countable discrete fields is equivalent to LLPO. Nevertheless, such a splitting field does exist ([48], page 152).

Classically, a ring is (left) Noetherian if each of its left ideals is finitely generated; constructively, even the field $\mathbf{Z}_2$ fails to satisfy this definition! Indeed, given a syntactically correct statement P, consider the ideal I of $\mathbf{Z}_2$ generated by the set

$$\{n : n = 0 \vee (n = 1 \wedge P)\}.$$

Note that for this set to be constructively finite we must be able to tell how many distinct elements it has. Suppose I is generated by a finite set F. If F has a single element, then $I = (0)$ and $\neg P$; whereas if F contains two elements, then $1 \in F$ and P.

It might be suspected that, since we cannot prove constructively that $\mathbf{Z}_2$ is Noetherian in the usual sense, there will be no constructive version of the Hilbert Basis Theorem. This suspicion is doubtless reinforced by recollection of the furore that arose after Hilbert's original, highly nonconstructive proof of that theorem, about which the invariant-theorist Gordan commented,

> Das ist nicht Mathematik; das ist Theologie.

But the real constructive problem lies with the definition of Noetherian. Mines et al. [48] define a ring R to be **Noetherian** if for each ascending chain

$$J_1 \subset J_2 \subset J_3 \subset \cdots$$

of finitely generated left ideals in R there exists n such that $J_n = J_{n+1}$. This definition of Noetherian is classically equivalent to the standard classical one, is satisfied by the ring $\mathbf{Z}_2$, and leads to the following constructive version of the **Hilbert Basis Theorem**:

If R is a coherent Noetherian ring, then so is $R[x]$ ([19], pages 91–97),

where, as usual, $R[x]$ is the ring of polynomials over R. (Regarding coherence, suffice it to say that it is a property that holds automatically for a Noetherian ring in classical mathematics.)

The Hilbert Basis Theorem illustrates the point that many—if not most—classical results can be re-cast, sometimes with additional hypothesis that are trivially satisfied in classical mathematics, into forms that are constructively provable; moreover, those forms are often, and ideally, equivalent to the original classical theorem with classical logic.

For additional results on ring theory, see [15, 48].

10.0.2 The existence of adjoints

In constructive mathematics, as the following Brouwerian example shows, we have no guarantee that a (bounded linear) operator on a Hilbert space H will have an adjoint. Let $(e_n)_{n=1}^{\infty}$ be the standard orthonormal basis of the separable complex Hilbert space l^2. Given a binary sequence $(a_n)_{n=1}^{\infty}$ with at most one term equal to 1, define a mapping $A : l^2 \to l^2$ by setting

$$Ax = \left(\sum_{n=1}^{\infty} a_n \langle x, e_n \rangle \right) e_1. \tag{5}$$

To see that this is a good definition, let $\varepsilon > 0$ and choose N such that $|\langle x, e_n \rangle| < \varepsilon$ for all $n \geqslant N$; then, since $a_n = 1$ for at most one n, we have

$$\sum_{n=N+1}^{M} |a_n \langle x, e_n \rangle| \leqslant \varepsilon \quad (M > N).$$

Hence the partial sums of the series $\sum_{n=1}^{\infty} |a_n \langle x, e_n \rangle|$ form a Cauchy sequence in $\mathbb{R}$, so the series $\sum_{n=1}^{\infty} a_n \langle x, e_n \rangle$ converges absolutely in $\mathbb{C}$, and A is well defined. It is easy to show that A is a bounded linear mapping of l^2 into itself. Suppose that its adjoint A^* exists. Either $\|A^* e_1\| > 0$ or else $\|A^* e_1\| < 1$. In the first case, choosing N such that $\langle e_1, Ae_N \rangle = \langle A^* e_1, e_N \rangle \neq 0$, we have $a_N = 1$; in the second case we have $a_n = 0$ for all n. Thus if every operator on l^2 has an adjoint, then LPO holds.

At this point you are probably asking, "What is the problem, constructively, with the classical method of obtaining $A^* y$ for any $A \in \mathcal{B}(H)$ and $y \in H$: namely, apply the Riesz Representation Theorem to the bounded linear functional $x \rightsquigarrow \langle Ax, y \rangle$?". In order to apply the Riesz Representation Theorem constructively to a linear functional f on H, we need to know that f is not just bounded but **has a norm**, in the sense that

$$\sup \{|f(x)| : x \in H, \|x\| \leqslant 1\}$$

exists ([8], page 419, (2.3)); since the classical least upper bound principle does not hold constructively, we may not be able to find the supremum in question. However, the classical proof of the existence of A^*y will work for us if we know that

$$\sup\{|\langle Ax, y\rangle| : y \in H,\ \|y\| \leqslant 1\}$$

exists for each $y \in H$. This observation leads to the following result.

Proposition 23 *A bounded linear operator A on a Hilbert space has an adjoint if and only if PA has an adjoint for each 1-dimensional projection P on H.*

Proof. Since 'only if' is trivial, we consider only 'if'. Suppose that PA has an adjoint for each 1-dimensional projection P. Let $c > 0$ be a bound for A, let y be any element of H, and let a, b be real numbers with $0 \leqslant a < b$. Either $\|y\| < b/c$, in which case we have

$$|\langle Ax, y\rangle| \leqslant \|Ax\|\,\|y\| \leqslant b$$

whenever $\|x\| \leqslant 1$, or else $y \neq 0$. In the latter case, with P the projection of H on the 1-dimensional subspace $\mathbb{C}y$, we see that

$$\sup_{\|x\|\leqslant 1} |\langle Ax, y\rangle| = \sup_{\|x\|\leqslant 1} |\langle x, (PA)^* y\rangle| = \|(PA)^* y\|$$

exists. It follows from the constructive least-upper-bound principle that $\sup_{\|x\|\leqslant 1} |\langle Ax, y\rangle|$ exists, and therefore that the linear functional $x \rightsquigarrow \langle Ax, y\rangle$ has a norm, for all $y \in H$. Thus the Riesz Representation Theorem can be used to produce A^*y as in the standard classical proof. ∎

There are more interesting conditions that ensure the existence of A^*. We state two of them without proof.

Proposition 24 *A bounded operator A on a Hilbert space H has an adjoint if and only if the image, under A, of the unit ball in H is located* [57].

In order to state the second, we introduce the **weak-operator topology** on the set $\mathcal{B}(H)$ of bounded linear operators on the Hilbert space H : namely, the weakest topology with respect to which the mappings $T \rightsquigarrow \langle Tx, y\rangle$ are continuous at 0 for all $x, y \in H$. Classically, the unit ball $\mathcal{B}_1(H)$, consisting of those elements of $\mathcal{B}(H)$ that have bound 1, is weak-operator compact ([43], page 306); moreover, for each $A \in \mathcal{B}(H)$ the mapping $T \rightsquigarrow AT$ of $(\mathcal{B}(H), \tau_w)$ to $(\mathcal{B}(H), \tau_w)$ is weak-operator continuous, and therefore weak-operator uniformly continuous[17] on $\mathcal{B}_1(H)$ (to prove the weak-operator continuity of this mapping, simply use the identity

$$\langle ATx, y\rangle = \langle Tx, A^*y\rangle,$$

whose validity depends on the existence of A^*). Constructively, $\mathcal{B}_1(H)$ is totally bounded [16], but may not be complete [10]. Also—this is left as an exercise—defining A as at (5), we can show that the τ_w-continuity of the mapping $T \rightsquigarrow AT$ at 0 implies LPO. In view of all this, the following criterion, established in [17], for the existence of an adjoint should produce little surprise.

[17] The weak-operator topology, being locally convex, has a natural associated uniform structure.

Proposition 25 *Let H be a Hilbert space, let $A \in \mathcal{B}(H)$, and let f_A be the linear mapping $T \rightsquigarrow AT$ of $(\mathcal{B}_1(H), \tau_w)$ into $(\mathcal{B}(H), \tau_w)$. Then the following are equivalent conditions.*

- *f_A is continuous at* 0.
- *f_A is uniformly continuous on $(\mathcal{B}_1(H), \tau_w)$.*
- *f_A maps totally bounded subsets of $(\mathcal{B}_1(H), \tau_w)$ to totally bounded subsets of $(\mathcal{B}(H), \tau_w)$.*
- *A has an adjoint.*

10.0.3 Integration theory

The original development of constructive integration theory in [4] was somewhat clumsy, and was superseded by a very elegant one, due to Bishop and his student Henry Cheng, published in [7] and, in a more refined form, in Chapter 6 of [8]. I would like to sketch the final form of that theory here, referring you to [8] for the justification of results stated here without proof.

We begin with a nonempty set X carrying an inequality relation $\neq$ that satisfies two extra properties:

▷ **cotransitivity:** $x \neq y \Rightarrow \forall z \in X \ (x \neq z \vee z \neq y)$;

▷ **tightness:** $\neg(x \neq y) \Rightarrow x = y$.

We also require

- a set L of strongly extensional functions—the **integrable functions**—from X into $\mathbb{R}$ that contains the mapping $x \rightsquigarrow 1$ (identified with the constant 1) and is closed under the pointwise operations of addition and multiplication-by-scalars (so L is a vector space under these operations);
- a linear mapping $I : L \to \mathbb{R}$, called the **integral,** satisfying the following axioms.

I1 If $(f_n)_{n=0}^{\infty}$ is a sequence of nonnegative elements of L such that $\sum_{n=1}^{\infty} I(f_n)$ converges to a sum less than $I(f_0)$, then there exists $x \in X$ such that $\sum_{n=1}^{\infty} f_n(x)$ converges and is less than $f_0(x)$.

I2 There exists $\varphi \in L$ such that $I(\varphi) = 1$.

I3 $I(\min\{f, n\}) \to I(f)$ and $I\left(\min\left\{|f|, n^{-1}\right\}\right) \to 0$ as $n \to \infty$.

The triple (X, L, I)—or, loosely, X itself—is then called an **integration space.**

The most important example of an integration space occurs when X is a **locally compact** metric space (that is, one in which every bounded set is contained in a compact set);[18] L is the set $C(X)$ of all uniformly continuous mappings of X into $\mathbb{R}$ that vanish on the **metric complement**

$$-K = \{x \in X : \rho(x, K) > 0\}$$

of some compact set K; and I is a **positive measure** on X—that is, a mapping of X into $\mathbb{R}$ such that if $f \in C(X)$ and $f \geqslant 0$, then $I(f) \geqslant 0$. The proof that these data do define an integration space is not an easy one (see [8], pages 220–221).

The first problem of integration theory is to extend the integral to the largest possible class of functions on X. To this end, we say that a partial function $f : X \to \mathbb{R}$ is **integrable** if there exists a sequence $(f_n)_{n=1}^{\infty}$ in L, called a **representation** of f, such that

- $\sum_{n=1}^{\infty} I(|f_n|)$ converges in $\mathbb{R}$, and
- $f(x) = \sum_{n=1}^{\infty} f_n(x)$ whenever $\sum_{n=1}^{\infty} |f_n(x)|$ converges.

It can then be shown that the **integral** of f, defined by

$$I(f) = \sum_{n=1}^{\infty} I(f_n),$$

is independent of the representation (f_n) of f in L; and that, denoting the set of all integrable functions on X by $L_1(X)$ or simply L_1, we obtain an integration space (X, L_1, I). This process of extending the integral cannot be continued to enlarge the class L_1 : if $(f_n)_{n=1}^{\infty}$ is a sequence in L_1 such that $\sum_{n=1}^{\infty} I(|f_n|)$ converges, then $\sum_{n=1}^{\infty} f_n$ is also in L_1.

By a **full set** we mean a subset of X that contains the domain of some function in L_1. It turns out that a subset F is full if and only if it contains the intersection of the domains of countably many integrable functions, and that the intersection of a sequence of full sets is therefore full.

The equality relation on L_1 is defined by setting $f = g$ if and only if $I(|f - g|) = 0$. It then turns out that $f = g$ in L_1 if and only if $f(x) = g(x)$ on a full set.

The Lebesgue integral occurs as the special case where $X = \mathbb{R}$ and the original integral, from which the extended one is constructed, is the Riemann integral on $C(\mathbb{R})$.

Classically, the important sets in integration theory are the complements of sets of measure 0. To accommodate successfully the rather negative notion of complement, Bishop introduced the idea of a **complemented set**, which is a pair $\mathbf{A} = (A^1, A^0)$ of subsets of X such that $x \neq y$ for all $x \in A^1$ and $y \in A^0$. The **characteristic function** of this (complemented) set is the mapping $\chi_{\mathbf{A}} : A^1 \cup A^0 \to \mathbb{R}$ defined by

$$\chi_{\mathbf{A}}(x) = \begin{cases} 1 & \text{if } x \in A^1 \\ 0 & \text{if } x \in A^0. \end{cases}$$

[18] Note that this definition of locally compact is more restrictive than the classical one.

Operations on complemented sets are defined using the corresponding operations on the characteristic functions:

$$\begin{aligned}\chi_{\mathbf{A}\vee\mathbf{B}} &= \max\{\chi_{\mathbf{A}},\chi_{\mathbf{B}}\},\\ \chi_{\mathbf{A}\wedge\mathbf{B}} &= \min\{\chi_{\mathbf{A}},\chi_{\mathbf{B}}\},\\ \chi_{\mathbf{A}-\mathbf{B}} &= \chi_{\mathbf{A}}(1-\chi_{\mathbf{B}}),\end{aligned}$$

with corresponding definitions for $\bigvee_{n=1}^{\infty}\mathbf{A}_n$ and $\bigwedge_{n=1}^{\infty}\mathbf{A}_n$ when $(\mathbf{A}_n)$ is a sequence of complemented sets. A complemented set $\mathbf{A}$ is said to be **integrable** if $\chi_{\mathbf{A}}$ is an integrable function; in which case $A^1 \cup A^0$ is a full set, and the **measure** of $\mathbf{A}$ is defined to be

$$\mu(\mathbf{A}) = I(\chi_{\mathbf{A}}).$$

Among the useful properties of integrable sets are the following.

Proposition 26 *Let* $\mathbf{A}$ *be an integrable set. If* $\mu(\mathbf{A}) > 0$, *then* A^1 *is nonempty. If* $\mu(\mathbf{A}) = 0$, *then* A^0 *is a full set.*

Proof. Suppose first that $\mu(\mathbf{A}) > 0$. Using axiom I2, choose φ such that $I(\varphi) = 1$, and for each positive integer n let $f_n = 0\varphi$. Then

$$\sum_{n=1}^{\infty} I(f_n) = \sum_{n=1}^{\infty} 0I(\varphi) = 0 < I(\chi_{\mathbf{A}}).$$

By axiom I1, there exists x such that $\chi_{\mathbf{A}}(x) > 0$. Then $\chi_{\mathbf{A}}(x) = 1$ and $x \in A^1$.

Now suppose that $\mu(\mathbf{A}) = 0$. Then

$$\sum_{n=1}^{\infty} I(n\chi_{\mathbf{A}}) = \sum_{n=1}^{\infty} n\mu(\mathbf{A}) = 0,$$

so $\sum_{n=1}^{\infty} n\chi_{\mathbf{A}}$ is in L_1. The domain D of this function is therefore a full subset of $A^1 \cup A^0$. But D cannot contain any point of A^1, since the series $\sum_{n=1}^{\infty} n$ diverges; hence $D \subset A^0$ and therefore A^0 is a full set. ∎

Proposition 27 *Let* $(\mathbf{A}_n)_{n=1}^{\infty}$ *be a sequence of integrable sets such that* $A_1^1 \subset A_2^1 \subset A_3^1 \subset \cdots$, *where* $\mathbf{A}_n = (A_n^1, A_n^0)$. *Then* $\bigvee_{n=1}^{\infty}\mathbf{A}_n$ *is integrable if and only if* $l = \lim_{n\to\infty}\mu(\mathbf{A}_n)$ *exists, in which case* $\mu(\bigvee_{n=1}^{\infty}\mathbf{A}_n) = l$ ([8], page 234, (3.9)).

The construction of a rich supply of integrable functions is a very difficult one and is solved by a special theory of profiles, introduced in [7]; the technical details of this construction are found in [8] (Chapter 6, Section 4). The main result is the following fundamental theorem.

Theorem 28 *Let* (X, L_1, I) *be an integration space, and* f *an integrable function. Then for all but countably many* $r > 0$, *the complemented sets*

$$[f > r] = (\{x : f(x) > r\}, \{x : f(x) \leqslant r\})$$

and

$$[\![f \geqslant r]\!] = (\{x : f(x) \geqslant r\}, \{x : f(x) < r\})$$

are integrable and have the same measure. Call such values of r **admissible.** *Then for each admissible r and each* $\varepsilon > 0$*, there exists* $\delta > 0$ *such that*

$$|\mu([\![f > r]\!]) - \mu([\![f > r']\!])| < \varepsilon$$

whenever $r' > 0$ *is admissible and* $|r - r'| < \delta$.

As a final illustration of the completeness technique that we used in earlier sections, let me now prove the following result.

Proposition 29 *If f is a nonnegative integrable function on X that is positive throughout an integrable set of positive measure, then* $I(f) > 0$.

Proof. Let $\mathbf{A}$ be an integrable set of positive measure on which f is everywhere positive. Replacing f by $f\chi_{\mathbf{A}}$, we may assume that $\mathbf{A} = \{x \in \mathbb{R} : f(x) > 0\}$. Using Theorem 28, choose a sequence $(r_n)_{n=1}^{\infty}$ of admissible values for f that decreases strictly to 0, and for each n set

$$\mathbf{A}_n = [\![f \geqslant r]\!].$$

Since $I(f) \geqslant r_n\mu(\mathbf{A}_n)$, it will suffice to find n such that $\mu(\mathbf{A}_n) > 0$. To this end, construct an increasing binary sequence (λ_n) such that

$$\lambda_n = 0 \;\Rightarrow\; \mu(\mathbf{A}_n) < n^{-1},$$
$$\lambda_n = 1 \;\Rightarrow\; \mu(\mathbf{A}_n) > (n+1)^{-1}.$$

We may assume that $\lambda_1 = 0$. If $\lambda_n = 0$, set $\mathbf{B}_n = \mathbf{A}_n$; if $\lambda_n = 1 - \lambda_{n-1}$, set $\mathbf{B}_k = \mathbf{A}_{n-1}$ for all $k \geqslant n$. Then

$$B_1^1 \subset B_2^1 \subset B_3^1 \subset \cdots \tag{6}$$

and

$$0 \leqslant \mu(\mathbf{B}_m) - \mu(\mathbf{B}_n) \leqslant \tfrac{2}{n} \quad (m \geqslant n).$$

So $(\mu(\mathbf{B}_n))_{n=1}^{\infty}$ is a Cauchy, and therefore convergent, sequence of real numbers. It follows from (6) and Proposition 27 that

$$\mathbf{B} = \bigvee_{n=1}^{\infty} \mathbf{B}_n$$

is an integrable set with measure equal to $\lim_{n\to\infty} \mu(\mathbf{B}_n)$. Either $\mu(\mathbf{B}) > 0$ or $\mu(\mathbf{B}) < \mu(\mathbf{A})$. In the first case we can find $n > 1$ such that $\mu(\mathbf{B}_n) > n^{-1}$. If $\lambda_n = 0$, then $\mu(\mathbf{B}_n) = \mu(\mathbf{A}_n) < n^{-1}$, a contradiction; so $\lambda_n = 1$. In the case $\mu(\mathbf{B}) < \mu(\mathbf{A})$ we have $\mu(\mathbf{A} - \mathbf{B}) > 0$, so there exists a point $\xi \in (\mathbf{A} - \mathbf{B})^1$; then $f(\xi) > 0$ and $\xi \in B^0$. Choose a positive integer N such that $f(\xi) > 1/N$; then $\xi \in A_N^1$. If $\lambda_N = 0$, then $A_N^1 = B_N^1 \subset B^1$, so $\xi \in B^1$, a contradiction; it follows that $\lambda_N = 1$. ∎

Note that the last proposition is a trivial consequence of Markov's Principle. For if $I(f) = 0$, then, since $f \geqslant 0$, we have $f(x) = 0$ on a full set, which contradicts the hypotheses of Proposition 29; hence $\neg(I(f) = 0)$ and therefore, by Markov's Principle, $I(f) > 0$.

I think we have seen enough of integration theory to give a taste of its constructive development. Suffice it, then, to say that there are constructive analogues of standard theorems of classical integration theory, such as the dominated convergence theorem, Egorov's theorem, and Fubini's theorem. The full development is found in Chapter 6 of [8]; the Radon-Nikodým theorem and the theory of the L_p spaces is found in Chapter 8 of that book.

10.0.4 The Riemann mapping theorem

My next example of constructive analysis is the improved, definitive version of the **Riemann mapping theorem** due to Bishop and Cheng.[19]

Recall the classical Riemann mapping theorem:

> *If U is a proper, open, and simply connected subset of the complex plane $\mathbb{C}$, then there exists an analytic equivalence of U with the open unit disc D: that is, a one-one analytic mapping f of D onto U with analytic inverse* ([59], Theorem 14.8).

In order to provide motivation for the hypotheses of the constructive Riemann mapping theorem, I shall give the details of a Brouwerian counterexample to the classical theorem. Given a binary sequence (a_n), define $U = \bigcup_{n=1}^{\infty} S_n$, where

- $S_n = D$ if $a_n = 0$, and
- S_n is the open disc with centre 0 and radius 2 if $a_n = 1$.

Then U is open and simply connected, and is clearly a proper subset of $\mathbb{C}$. Suppose that f is an analytic equivalence of U with D; we may assume that $f(0) = 0$. Either $|f'(0)| > 1$ or $|f'(0)| < 2$. In the first case, choosing $r \in (0, 1)$ such that $1/r < |f'(0)|$ and then using standard estimates, we obtain

$$\sup\{|f(z)| : |z| = r\} > 1.$$

Hence there exists z such that $|z| = r$ and $|f(z)| > 1$. In turn, there exists n such that $f(z) \in S_n$, so $a_n = 1$.

In the case $|f'(0)| < 2$, consider any positive integer k. If $a_k = 1$, then U is the open disc with centre 0 and radius 2. It then follows from the maximal derivative property of the Riemann mapping ([59], pages 273–275)—which holds constructively provided that the mapping f exists—that $|f'(0)| = 2$. This contradiction implies that $a_k = 0$. Hence, in this case, $a_n = 0$ for all n.

[19] The paper [30] in which this work on the Riemann mapping theorem first appeared contains many errors; a corrected treatment of the theorem is found in the final section of Chapter 5 of [8].

Thus the classical form of the Riemann mapping theorem entails LPO. How, if at all, can we recover from this situation?

To answer this, consider the following pathological features of our Brouwerian counterexample:

- We cannot pin down the boundary of the domain U.
- For each point $z \in U$ we cannot tell the minimum distance we need to travel from z in order to reach the outside of U; equivalently, we cannot compute the radius of the largest ball centred on z and lying inside U.

Perhaps if we were to add hypotheses that ensure that neither of these pathologies can occur, we would be able to recover a constructively valid form of the Riemann mapping theorem. To this end, Bishop and Cheng introduced the following notion of approximate border for a proper bounded[20] open subset U of $\mathbb{C}$.

Let z_0 be any point—which we refer to as a *distinguished point*—of U, and let $\varepsilon > 0$. An ε-**border of** U **relative to** z_0 is a finitely enumerable subset B of the **complement** of U,

$$\sim U = \{z \in \mathbb{C} : \forall u \in U \, (z \neq u)\},$$

such that if γ is a path in $\mathbb{C}$ with left endpoint z_0, and γ keeps at least ε away from B, then γ lies in U. For example, if the positive integer N is sufficiently large, then the points

$$x_k = \left(1 + \frac{\varepsilon}{2}\right) \exp\left(\frac{k\pi \mathrm{i}}{N}\right) \quad (k = 1, 2, ..., 2N)$$

form an ε-border of the disc D relative to 0, since the union of the open discs with centres x_k and radius ε contains the annulus $\left\{z : 1 - \frac{\varepsilon}{4} < |z| < 1\right\}$.

We say that U is **mappable** if

▷ it is simply connected and

▷ there exists $z_0 \in U$ such that for each $\varepsilon > 0$ there is an ε-border of U relative to z_0.

(It can be shown that any point of U will then serve as the distinguished point.) Thus a mappable set is one whose border is approximated arbitrarily closely by finitely enumerable subsets of the complement.

Turning to the second pathological feature of our Brouwerian counterexample, we say that U has the **maximal extent property** if there is a function ρ from U into the positive real line such that for each $z \in U$,

▶ the disc with centre z and radius $\rho(z)$ lies in U and

▶ any disc with centre z and radius greater than $\rho(z)$ intersects $\sim U$.

We are now able to state the constructive **Riemann mapping theorem.**

[20] The general case, in which U may or may not be bounded, is handled in Chapter 5 of [8].

Theorem 30 *The following are equivalent conditions on a proper, open, and simply connected subset U of $\mathbf{C}$:*

- *U is mappable.*
- *U has the maximal extent property.*
- *U is analytically equivalent to D.*

For the long, difficult proof and other material on complex analysis I refer you to Chapter 5 of [8]. Note also the fascinating elementary constructive proof of the Jordan curve theorem in [3].

10.0.5 Apartness spaces

Let me give, as a final example of modern constructive mathematics, an introduction to a very recent development: the theory of apartness spaces.

At the start of Chapter 3 of [4], Errett Bishop stated that

> Very little is left of general topology after that vehicle of classical mathematics has been taken apart and reassembled constructively. With some regret, plus a large measure of relief, we see this flamboyant engine collapse to constructive size.

To be fair, I should point out that Bishop later amended his views on constructive general topology: in a letter to me (dated 14 April 1975) he wrote

> The problem of finding a suitable constructive framework for general topology is important and elusive.

He even produced an unpublished manuscript containing an unusual, but ultimately not satisfactory, development of topology based on a hierarchy of increasingly complicated types of space [6] (see also [5], pages 28–29). However, it is reasonable to suggest that his original views on topology led people to believe that constructive topology, other than intuitionistic topology [66, 70], was a non–starter.

At the beginning of 2000, my then doctoral student, Luminiţa Viţă, and I resurrected an idea that I had first explored, unsuccessfully, some twenty–five years earlier: an axiomatic treatment of spaces with a relation of apartness between subsets, which correspond, under classical logic, to the proximity spaces discussed in [29, 51]. This time, two heads definitely proved better than one, for we managed to find a set of axioms from which the theory flowed very naturally. For simplicity of exposition, I shall describe a restricted version of the theory, dealing with apartness between points and subsets of a space.

Let X be a set with a binary relation $\neq$ of inequality that is **nontrivial** in the sense that there exist x, y in X with $x \neq y$. A subset S of X has two natural complementary subsets:

▷ **the logical complement**

$$\neg S = \{x \in X : \forall y \in S \ \neg(x = y)\};$$

▷ **the complement**

$$\sim S = \{x \in X : \forall y \in S \ (x \neq y)\}.$$

We are interested in a set X that carries a nontrivial inequality $\neq$ and a relation $\mathbf{apart}(x, S)$ between points x and subsets S of X. If $\mathbf{apart}(x, S)$, then we say that the point x is **apart from** the set S. For convenience we introduce the **apartness complement**

$$-S = \{x \in X : \mathbf{apart}(x, S)\}$$

of S; and, when A is also a subset of X, we write

$$A - S = A \cap -S.$$

We assume that the following axioms are satisfied.

A1 $x \neq y \Rightarrow \mathbf{apart}(x, \{y\})$

A2 $\mathbf{apart}(x, A) \Rightarrow x \notin A$

A3 $\mathbf{apart}\,(x, A \cup B) \Leftrightarrow \mathbf{apart}(x, A) \wedge \mathbf{apart}(x, B)$

A4 $x \in -A \subset \sim B \Rightarrow \mathbf{apart}(x, B)$

A5 $\mathbf{apart}\,(x, A) \Rightarrow \forall y \in X \ (x \neq y \vee \mathbf{apart}(y, A))$

We then call X an **apartness space**, and the data defining the relations $\neq$ and **apart** the **apartness structure** on X. It readily follows from the axioms that for each $A \subset X$,

$$-A \subset \sim A \subset \neg A. \tag{7}$$

We say that the point $x \in X$ is **near** the set $A \subset X$, and we write $\mathbf{near}(x, A)$, if

$$\forall S \ (x \in -S \Rightarrow \exists y \in A - S).$$

If X is an apartness space, and Y is a subset of X upon which the induced inequality is nontrivial, then there is a natural apartness structure induced on Y by that on X. Taken with that structure, Y is called an apartness **subspace** of X.

In the corresponding classical development [29], nearness is taken as the primitive notion and apartness is defined as the negation of nearness. It is easy to see, using the classical axioms for nearness, that our definition of nearness is classically equivalent to the negation of apartness; but, as we prove in a moment, this equivalence does not hold constructively.

Our canonical example of an apartness space is a metric space (X, ρ), in which the inequality and apartness are defined by

$$x \neq y \Leftrightarrow \rho(x, y) > 0$$

and

$$\mathbf{apart}(x, A) \Leftrightarrow \exists r > 0 \forall y \in A \; (\rho(x, y) \geqslant r).$$

It is routine to verify axioms A1–A5 in this case. We call this apartness structure the **metric apartness structure** corresponding to the metric ρ, and we refer to X as a **metric apartness space.** The apartness complement

$$-S = \{x \in X : \exists r > 0 \forall y \in S \; (\rho(x, y) \geqslant r)\}$$

is then also called the **metric complement** of S in X.

Among the elementary deductions we can make from the axioms are the following.

$\neg(\mathbf{near}(x, A) \wedge \mathbf{apart}(x, A))$

$\mathbf{apart}(x, \emptyset)$

$\mathbf{near}(x, A) \Rightarrow \exists y \in A$

$x \in A \Rightarrow \mathbf{near}(x, A)$

$-A \subset {\sim}A$

$(\mathbf{apart}(x, A) \wedge B \subset A) \Rightarrow \mathbf{apart}(x, B)$

$(\mathbf{near}(x, A) \wedge A \subset B) \Rightarrow \mathbf{near}(x, B)$

$(\mathbf{near}(x, A) \wedge \mathbf{apart}(x, B)) \Rightarrow \mathbf{near}(x, A - B)$.

Another example of an apartness structure occurs when (X, τ) is a T_1 topological space.[21] If $x \in X$ and $A \subset X$, we define

$$\mathbf{apart}(x, A) \Leftrightarrow \exists U \in \tau \; (x \in U \subset {\sim}A)$$

and then introduce **near** (x, A) as above. The relation **apart** satisfies A1, since we are assuming that X is a T_1 space, and it clearly satisfies axioms A2–A4; but to make X into an apartness space we also need to postulate axiom A5. We then call this apartness structure on X the **topological apartness structure** corresponding to τ.

A subset S of an apartness space X is said to be **nearly open** if there exists a family $(A_i)_{i \in I}$ such that $S = \bigcup_{i \in I} -A_i$. The nearly open sets form a topology—the **apartness topology**—on X for which the apartness complements form a basis. Every nearly open set in a topological apartness space is open.

We say that a topological apartness space X is **topologically consistent** if every open subset X is nearly open. An example communicated to us privately by Jeremy Clark shows that we cannot prove that every topological apartness space is topologically consistent. However, there is a natural condition, one that always holds

[21] That is, for all $x, y \in X$ with $x \neq y$ there exist $U, V \in \tau$ such that $x \in U$ and $y \notin U$, and $y \in V$ and $x \notin V$.

classically, that ensures topological consistency: an apartness space X is said to be **locally decomposable** if

$$\forall x \in X\ \forall S \subset X\ \ (x \in -S \Rightarrow \exists T\ \ (x \in -T \wedge \forall y \in X\ \ (y \in -S \vee y \in T))). \quad (8)$$

It is easy to show that metric spaces are locally decomposable and hence topologically consistent.

We say that a subset S of an apartness space X is **nearly closed** if

$$\forall x \in X\ \ (\mathbf{near}(x, S) \Rightarrow x \in S).$$

Both X and $\emptyset$ are nearly closed. The intersection of any family of nearly closed sets is nearly closed, but even in $\mathbb{R}$ we cannot expect a constructive proof that the union of two nearly closed sets is nearly closed ([69], Section 6). If S is a nearly open subset of an apartness space X, then its logical complement equals its complement and is nearly closed.

There are at least three natural types of continuity for functions between apartness spaces. We say that a mapping $f : X \rightarrow Y$ between apartness spaces is

- **nearly continuous** if

$$\forall x \in X\ \ \forall A \subset X\ \ \ (\mathbf{near}(x, A) \Rightarrow \mathbf{near}(f(x), f(A)));$$

- (apartness) **continuous** if

$$\forall x \in X\ \ \forall A \subset X\ \ (\mathbf{apart}(f(x), f(A)) \Rightarrow \mathbf{apart}(x, A));$$

- **topologically continuous** if $f^{-1}(S)$ is nearly open in X for each nearly open $S \subset Y$.

The following conditions are equivalent on a mapping $f : X \rightarrow Y$ between apartness spaces.

(i) f is nearly continuous.

(ii) For each nearly closed subset T of Y, $f^{-1}(T)$ is nearly closed.

(iii) For each subset S of X, $f(\overline{S}) \subset \overline{f(S)}$, where the closures are relative to the apartness topologies on X, Y respectively.

For mappings between metric spaces, apartness continuity is equivalent to the usual notion of ε–δ continuity.

Apartness continuity is a natural extension of the notion of strong extensionality. In fact, strong extensionality holds even for nearly continuous functions. To see this, let $f : X \rightarrow Y$ be nearly continuous and let $x, y \in X$ be such that $f(x) \neq f(y)$. Define

$$A = \{z \in X : z = x \vee (z = y \wedge x \neq y)\}.$$

Note that $x \in A$. To show that $\mathbf{near}(y, A)$, consider any $U \subset X$ such that $y \in -U$; by axiom A5, either $x \neq y$ and therefore (by the definition of A) $y \in A - U$; or else $x \in -U$ and so $x \in A - U$. Using the near continuity of f, we obtain $\mathbf{near}(f(y), f(A))$. Since $f(x) \neq f(y)$ and therefore (by axiom A1) $f(y) \in -\{f(x)\}$, it follows that there exists $z \in A$ such that $f(z) \in -\{f(x)\}$; whence $f(z) \neq f(x)$, by (7). Then $\neg(z = x)$, so we must have $z = y$ and $x \neq y$, as required.

A topologically continuous mapping between apartness spaces is both continuous and nearly continuous; so every topologically continuous mapping between apartness spaces is strongly extensional. Every continuous mapping from an apartness space into a locally decomposable apartness space is topologically continuous and hence nearly continuous. However, nearness continuity, although classically equivalent to apartness continuity, is a constructively weaker notion.

At this point I shall stop, leaving the interested reader to pursue the notion of apartness, including the more interesting set–set version, through the succession of papers [69, 62, 21, 20, 68].

Acknowledgements: In addition to the acknowledgements made in my introduction, I wish to thank my Canterbury colleagues Bill Taylor, for his many helpful suggestions about the paper, and Luminiţa Vîţă, for proofreading my draft. My indebtedness to my co-workers, including my graduate students, over the past twenty-three years is too great to be measured; to one and all I say a heartfelt "Thank you".

References

[1] O. Aberth, *Computable Calculus*, Academic Press, San Diego, CA, 2001.

[2] Michael J. Beeson, *Foundations of Constructive Mathematics*, Springer–Verlag, Heidelberg, 1985.

[3] G. Berg, W. Julian, R. Mines, and F. Richman, 'The constructive Jordan curve theorem', Rocky Mountain J. Math. **5**(2), 225–236.

[4] Errett Bishop, *Foundations of Constructive Analysis*, McGraw–Hill, New York, 1967.

[5] Errett Bishop, *Schizophrenia in Contemporary Mathematics*, Amer. Math. Soc. Colloquium Lectures, Univ. of Montana, Missoula, 1973; reprinted in *Errett Bishop: Reflections on Him and His Research* (M. Rosenblatt, ed.), Contemporary Math. **39**, 1–32, Amer. Math. Soc., 1985.

[6] Errett Bishop, 'The neat category of stratified spaces', unpublished manuscript, University of California, San Diego, 1970 (?).

[7] Errett Bishop and Henry Cheng, *Constructive Measure Theory*, Mem. Amer. Math. Soc. **116**, 1972.

[8] Errett Bishop and Douglas Bridges, *Constructive Analysis*, Grundlehren der math. Wissenschaften **279**, Springer–Verlag, 1985.

[9] Douglas Bridges, *Constructive Mathematics—Its Set Theory and Practice*, D. Phil thesis, Oxford University, 1975.

[10] Douglas Bridges, 'On weak operator compactness of the unit ball of $L(H)$', Zeit. math. Logik Grundlagen Math. **24**, 493–494, 1978.

[11] Douglas Bridges, 'A constructive proximinality property of finite–dimensional linear spaces', Rocky Mountain J. Math. **11**(4), 491–497, 1981.

[12] 'A constructive Morse theory of sets', in *Mathematical Logic and its Applications* (D. Skordev, ed.), Plenum Publishing Corp., New York, 61–79, 1987.

[13] Douglas Bridges, *Computability—A Mathematical Sketchbook*, Graduate Texts in Mathematics **146**, Springer–Verlag, Heidelberg–Berlin–New York, 1994.

[14] Douglas Bridges, 'Constructive Mathematics: a Foundation for Computable Analysis', Theoretical Computer Science **219** (1–2), 95-109, 1999.

[15] Douglas Bridges, 'Prime and maximal ideals in constructive ring theory', Comm. Algebra. **29**(7), 2787–2803, 2001.

[16] Douglas Bridges and Luminiţa Dediu (Vîţă), 'Weak–operator continuity and the existence of adjoints', Math. Logic Quarterly **45**(2), 203–206, 1999.

[17] Douglas Bridges and Luminiţa Dediu (Vîţă), 'The weak–operator sequential continuity of left multiplication', Proc. Koninklijke Nederlandse Akad. Wetenschappen (Indag. Math.) **11**(1), 39–42, 2000.

[18] Douglas Bridges and Hajime Ishihara, 'Linear mappings are fairly well–behaved', Arch. Math. **54**, 558–562, 1990.

[19] Douglas Bridges and Fred Richman, *Varieties of Constructive Mathematics*, London Math. Soc. Lecture Notes 97, Cambridge University Press, 1987.

[20] Douglas Bridges and Luminiţa Vîţă, 'Cauchy nets in the constructive theory of apartness spaces', Scientiae Mathematicae Japonicae **6**, 123–132, 2002.

[21] Douglas Bridges and Luminiţa Vîţă, 'Apartness spaces as a foundation for constructive topology', Ann. Pure Appl. Logic. **119**(1–3), 61–83, 2002.

[22] Douglas Bridges, Hajime Ishihara, and Luminiţa Vîţă, 'A new constructive version of Baire's theorem', preprint, University of Canterbury, Christchurch, New Zealand, 2002.

[23] Douglas Bridges, William Julian, and Ray Mines, 'A constructive treatment of open and unopen mapping theorems', Zeit. math. Logik Grundlagen Math. **35**, 29–43, 1989.

[24] Douglas Bridges, Fred Richman, and Peter Schuster, 'A weak countable choice principle', Proc. Amer. Math. Soc. **128**(9), 2749–2752, 2000.

[25] Douglas Bridges, Fred Richman, and Peter Schuster, 'Adjoints, absolute values, and polar decompositions', J. Operator Theory **44**, 243–254, 2000.

[26] Douglas Bridges, Fred Richman, and Peter Schuster, 'Linear independence and choice', Annals of Pure and Applied Logic **101**(1), 95–102, 1999.

[27] L.E.J. Brouwer, *Over de Grondslagen der Wiskunde*, Doctoral Thesis, University of Amsterdam, 1907. Reprinted with additional material (D. van Dalen, ed.) by Matematisch Centrum, Amsterdam, 1981.

[28] L.E.J. Brouwer, 'Über Abbildungen von Mannigfaltigkeiten', Math. Ann. **17**, 97–115, 1911–12.

[29] P. Cameron, J.G. Hocking, and S.A. Naimpally, *Nearness—a better approach to topological continuity and limits* (Parts 1 and 2), Mathematics Report #18–73, Lakehead University, Canada, 1973.

[30] Henry Cheng, 'A constructive Riemann mapping theorem', Pacific J. Math. **44**, 435–454, 1973.

[31] R.L. Constable et al., *Implementing Mathematics with the Nuprl Proof Development System*, Prentice-Hall, Englewood Cliffs, New Jersey, 1986.

[32] D. van Dalen, *Brouwer's Cambridge Lectures on Intuitionism*, Cambridge University Press, Cambridge, 1981.

[33] D. van Dalen, *Mystic, Geometer, and Intuitionist* (Volume 1), Oxford, The Clarendon Press, 1999.

[34] M.A.E. Dummett, *Elements of Intuitionism*, Clarendon Press, Oxford, 1977.

[35] S. Feferman, 'Constructive theories of functions and classes' in: *Logic Colloquium '78* (M. Boffa, D. van Dalen, K. McAloon, eds) 159–174, North-Holland, Amsterdam, 1979.

[36] Harvey Friedman, 'Set theoretic foundations for constructive analysis', Ann. of Math. (2) **105**, 1–28, 1977.

[37] N. D. Goodman and J. Myhill, 'Choice Implies Excluded Middle', Zeit. Logik und Grundlagen der Math. **24**, 461.

[38] S. Hayashi and H. Nakano, *PX: A Computational Logic*, MIT Press, Cambridge MA, 1988.

[39] A. Heyting, 'Die formalen Regeln der intuitionistischen Logik', Sitzungsber. preuss. Akad. Wiss. Berlin, 42–56, 1930.

[40] A. Heyting, *Intuitionism—An Introduction* (Third Edition), North Holland, 1971.

[41] D. Hilbert, 'Über das Unendliche', Mathematische Annalen **95** , 161–190; translated in *Philosophy of Mathematics* (P. Benacerraf and H. Putnam, eds), 183–201, Cambridge University Press, Cambridge, 1964.

[42] William H. Julian and Fred Richman, 'A uniformly continuous function on [0, 1] that is everywhere different from its infimum', Pacific J. Math. **111**, 333–340, 1984.

[43] Richard V. Kadison and John R. Ringrose, *Fundamentals of the Theory of Operator Algebras* (Vol. I), Academic Press, New York, 1983.

[44] S.C. Kleene and R.E. Vesley, *The Foundations of Intuitionistic Mathematics*, North–Holland, Amsterdam, 1965.

[45] B. Kushner, *Lectures on Constructive Mathematical Analysis*, Amer. Math. Soc., Providence RI, 1985.

[46] A.A. Markov, *Theory of Algorithms* (Russian), Trudy Mat. Istituta imeni V.A. Steklova **42** (Izdatel'stvo Akademii Nauk SSSR, Moskva), 1954.

[47] Per Martin–Löf, 'An Intuitionistic Theory of Types: Predicative Part', in *Logic Colloquium 1973* (H.E. Rose and J.C. Shepherdson, eds), 73–118, North–Holland, Amsterdam, 1975.

[48] Ray Mines, Fred Richman and Wim Ruitenburg, *A Course in Constructive Algebra*, Universitext, Springer–Verlag, Heidelberg, 1988.

[49] John Myhill, 'Some properties of intuitionistic Zermelo–Fraenkel set theory', in: *Cambridge Summer School in Mathematical Logic* (A. Mathias and H. Rogers, eds), 206–231, Lecture Notes in Mathematics **337**, Springer–Verlag, Heidelberg, 1973.

[50] John Myhill, 'Constructive set theory', J. Symbolic Logic **40**, 347–382, 1975.

[51] S.A. Naimpally and B.D. Warrack, *Proximity Spaces*, Cambridge Tracts in Math. and Math. Phys. **59**, Cambridge at the University Press, 1970.

[52] Fred Richman, 'Church's thesis without tears', J. Symbolic Logic **48**, 797–803, 1983.

[53] Fred Richman, 'Intuitionism as a Generalization', Philosophia Math. **5**, 124–128, 1990.

[54] Fred Richman, 'Generalized real numbers in constructive mathematics', Proc. Koninklijke Nederlandse Akad. Wetenschappen (Indag. Math.) **9**, 595-606, 1998.

[55] Fred Richman, 'Interview with a Constructive Mathematician', Modern Logic **6**, 247–271, 1996.

[56] Fred Richman, 'The fundamental theorem of algebra: a constructive development without choice', Pacific J. Math. **196**, 213-230, 2000.

[57] Fred Richman, 'Adjoints and the image of the ball', Proc. Amer. Math. Soc., **129**, 1189-1193, 2001.

[58] Fred Richman, Douglas Bridges, Allan Calder, William Julian, and Ray Mines, 'Compactly generated Banach spaces', Arch. Math. **36**, 239–243, 1981.

[59] Walter Rudin, *Real and Complex Analysis*, McGraw–Hill, New York, 1970.

[60] Walter Rudin, *Functional Analysis*, McGraw–Hill, New York, 1973.

[61] H.E. Scarf, *The Computation of Economic Equilibria*, Cowles Foundation, Yale University Press, New Haven, Conn., 1973.

[62] Peter Schuster, Luminiţa Vîţă, and Douglas Bridges, 'Apartness as a Relation Between Subsets', in: *Combinatorics, Computability and Logic* (Proceedings of DMTCS'01, Constanţa, Romania, 2–6 July 2001; C.S. Calude, M.J. Dinneen, S. Sburlan (eds.)), 203–214, DMTCS Series **17**, Springer–Verlag, London, 2001.

[63] S. Simpson, *Subsystems of Second Order Arithmetic*, Springer–Verlag, Heidelberg, 1998.

[64] Ernst Specker, 'Nicht konstruktiv beweisbare Sätze der Analysis', J. Symbolic Logic **14**, 145–158, 1949.

[65] W.P. van Stigt, *Brouwer's Intuitionism*, North–Holland, Amsterdam, 1990.

[66] Anne S. Troelstra, *Intuitionistic General Topology*, Ph.D. Thesis, University of Amsterdam, 1966.

[67] A.S. Troelstra and D. van Dalen, *Constructivity in Mathematics: An Introduction* (two volumes), North Holland, Amsterdam, 1988.

[68] Luminiţa Vîţă, 'Proximal convergence implies uniform convergence', to appear in Math. Logic Quarterly **49**, 2003.

[69] Luminiţa Vîţă and Douglas Bridges, 'A constructive theory of point–set nearness', to appear in *Topology in Computer Science: Constructivity; Asymmetry and Partiality; Digitization* (Proc. Dagstuhl Seminar 00231, 4-9 June 2000); R. Kopperman, M. Smyth, D. Spreen, eds.), special issue of Theoretical Computer Science, 2003.

[70] Frank A. Waaldijk, *Constructive Topology*, Ph.D. thesis, University of Nijmegen, Netherlands, 1996.

[71] Klaus Weihrauch, 'A foundation for computable analysis', in ***Combinatorics, Complexity, & Logic*** (Proceedings of Conference in Auckland, 9–13 December 1996; D.S. Bridges, C.S. Calude, J. Gibbons, S. Reeves, I.H. Witten, eds), Springer–Verlag, Singapore, 1996.

[72] Klaus Weihrauch, ***Computable Analysis***, Springer–Verlag, Heidelberg, 2000.