University of Leeds
Department of Pure Mathematics

Constructing fixed points and economic equilibria

A thesis submitted in accordance with the requirements of the Degree of Doctorate in Philosophy at the University of Leeds by

Matthew Ralph John Hendtlass

Supervisors: Prof. Michael Rathjen and Dr. Peter Schuster

April 2013

Dedicated to my fiancée

Mayra Montalvo Ballesteros

...who made even Leeds a sunny place.

Abstract

Constructive mathematics is mathematics with intuitionistic logic (together with some appropriate, predicative, foundation)—it is often crudely characterised as mathematics without the law of excluded middle. The intuitionistic interpretation of the connectives and quantifiers ensure that constructive proofs contain an inherent algorithm which realises the computational content of the result it proves, and, in contrast to results from computable mathematics, these inherent algorithms come with fixed rates of convergence.

The value of a constructive proof lies in the vast array of models for constructive mathematics. Realizability models and the interpretation of constructive **ZF** set theory into Martin Löf type theory allows one to view constructive mathematics as a high level programing language, and programs have been extracted and implemented from constructive proofs. Other models, including topological forcing models, of constructive set theory can be used to prove metamathematical results, for example, guaranteeing the (local) continuity of functions or algorithms. In this thesis we have highlighted any use of choice principles, and those results which do not require any choice, in particular, are valid in any topos.

This thesis looks at what can and cannot be done in the study of the fundamental fixed point theorems from analysis, and gives some applications to mathematical economics where value is given to computability.

Acknowledgements

I would like to acknowledge the help and encouragement of my teachers, Douglas Bridges and Peter Schuster in particular, and give them my hearty thanks. My equally ardent thanks belongs to the people and departments that hosted me, and made me feel welcome, during my studies: Helmut Schwichtenberg at Ludwig Maximillias Universität, Hajime Ishihara at the Japan advanced institute of science and technology, Vasco Brattka at the University of Cape Town, and Douglas et al at the University of Canterbury. Finally I must thank fellow (former) students Ray-Ming Chen, Mayra Montalvo Ballesteros, Andrew W. Swan, and Pedro Valencia Vizcaino for frequent interesting and useful conversations (and in Andrew's case, at least, for being inspirational and awesome), and Albert Ziegler and Jacob Cook for the less frequent, but equally interesting, conversations allowed for by different offices and different ages.

I thank the university of Leeds for supporting me with a fully funded international student scholarship, and my father for being born in Middlesbrough (allowing me to claim a British passport). I apologise to my lead supervisor Michael Rathjen that we did not find many of the mutual interests I know we must share, and that no research came from those we did find, and I thank him for, nonetheless, treating me with as much kindness and attention as his other, more diligent, students.

Preface

As mathematicians we habitually rewrite the past, to write a paper reflecting a true development of the results and proofs therein would typically lead to a large garbled mass of ideas. I restrain myself from rewriting the construction of this thesis, and apologise for any mess. I decided to begin 'writing up' upon returning to Leeds at the start of 2012, much earlier than I had previously intended so I could (unsuccessfully) apply for my dream post doc. Thus committed to finishing my PhD studies, I gathered together the various work I had done over my time at Leeds and took the largest subset (an admirably large one) to which I could apply a uniform subject—and so we have 'Constructing fixed points and economic equilibria'.

The work of Chapter 3, on fixed point theorems, was done almost entirely during my first year. The section on the intermediate value theorem without choice was the first piece of work I did when I arrived at Leeds and was hurriedly written up as [55]; in the intervening time I have given it periodic thought, and Section 3.1 gives a much better picture than that paper. At Peter's encouragement (which was prompted by a question of Andrea Cantini) I then worked on the fixed point theorems of Schauder and Kakutani, which is contained in [56, 57], and though I stopped thinking about these problems some time ago, I am still not entirely happy with the contents.

The work on constructing equilibria was started after reading [105], and feeling the paper had little interest even to the constructive mathematician. In response I quickly wrote and submitted [58] to the same journal, and Section 4.2 is a carbon copy of this. I was encouraged to work on McKenzie's theorem on the existence of equilibria by Douglas Bridges, and Naz Miheisi's contribution to [59] was forcing me to sit down and work through the details of the general proof outline I had formed. The work on demand functions (Section 2.2), the most recent of the main content, had the original purpose of making the thesis more self contained by developing the material borrowed from [27], but ended up generalising Bridges' results.

The first chapter represents the bulk of the work specifically devoted to writing this thesis. Its dual purpose of justifying the study of mathematics from a constructive perspective and presenting a broad overview of the mathematics and metamathematics which might interest and benefit the practising constructivist[1] (this one at least), received much help from my

[1] Anyone interested in constructive techniques, with or without a philosophical commitment to constructivism.

mathematical brothers and sisters (and here I include my mathematical half-sister Mayra). In particular, the extension of the result of [63] to infinitary systems was a goal set, and pursued, with Pedro Valencia to help us understand the details of Ishihara's work. The work of Section 1.3 on models of **IZF** was done independently after I learnt of Cohen's weak forcing and its relationship to Kripke models; after this I found that forcing models for **IZF** are discussed in [73] (there is, however, in this paper a small error in the soundness proof of set induction), and this was extended, in particular to **CZF**, in Ray-Ming Chen's thesis [38]. I am not aware of forcing semantics for topological models being presented in the literature, but it is easily derived from the more general presentation of Grayson [54] and others.

Most of the thesis is quite elementary and can be read alone. The first chapter provides a relatively broad introduction to 'aspects of constructive mathematics and metamathematics', particularly its purpose and formalisation. The second chapter provides a few results used in the later chapters; these two later chapters represent the body of the thesis. Chapter three first looks at the question 'when can we construct (classically existing) fixed points (without choice)?', and then gives approximate constructive versions of Brouwer's fixed point theorem and its two main generalisations: first to infinite dimensional spaces (Schauder's fixed point theorem), and then to set valued mappings (Kakutani's fixed point theorem). Schauder's fixed point theorem is applied to give an approximate version of Peano's theorem on the existence of solutions to differential equations. The final chapter applies the approximate fixed point theorems of chapter three to equilibrium problems from mathematical economics. We make comments throughout on the constructive reverse mathematics of our results.

I hope that constructivists, both experienced and budding, will find the first chapter interesting, although it could undoubtedly be greatly improved. The latter chapters are perhaps of less general interest, but the work on questions from mathematical economics should interest some in that field (and has few prerequisites), and the work on the intermediate value theorem without choice, though simple, might appeal more generally because it holds more generally. Our treatment of the intermediate value theorem also highlights some of the difficulties, and benefits, of working in an inuitionistic setting without choice. After writing the first chapter, and after gaining an appreciation for choice-free mathematics from Peter Schuster, I made an effort to indicate the use of choice throughout the thesis—I hope that I have at least overestimated the choice needed at each stage—and I have tried to indicate the prevalent use of choice in Bishop's constructive mathematics, and some of the drawbacks for its acceptance or rejection.

Leeds, September 2012 Matthew Hendtlass

Contents

Chapter 1

Constructive mathematics

This introduction is not really designed for this thesis. The contents of this thesis are largely elementary, and consequently the only introduction it requires is a brief note on the practice—and possibly philosophy—of constructive mathematics. In contrast this introduction, as much for myself as any reader, will give a brief nontechnical introduction to aspects of logic which may interest a practising (philosophical or pragmatic) constructivist. In short, it is an attempt (i) to provide a simple introduction to what I should have known or studied when I arrived at Leeds—assuming I had decided to remain a student of constructive mathematics—and (ii) to turn the writing of this thesis into more than mere farce (a job it did admirably). Some results—none of which are by any means exciting—are original; no effort is made to point these out or, more generally, to include any historical commentary, but general references will be given.

This chapter consists of five sections. The first introduces constructive mathematics, in particular intuitionistic logic, informally and then discusses formal intuitionistic logic and briefly mentions some related models of computation. Section 1.2 gives background on the three main set theories for constructive mathematics; and models of the strongest of these theories are introduced in the next section. The fourth section takes up the argument, begun in Section 1.1, that some aspects of constructive mathematics should be of interest to a large class of classical mathematicians. The final section introduces some of the fundamental definitions of constructive mathematics, and outlines the programme of constructive reverse mathematics.

1.1 What is constructive mathematics?

Constructive mathematics reinterprets the quantifier 'there exists' as 'we can construct'. As a result of this stronger interpretation of existence, a constructive proof is algorithmic. If upon seeing a statement "there exists x such that $\varphi(x)$", you think "so what does x look like?," then you should have some small sympathy with constructive mathematics; constructive mathematicians are not (necessarily) fostering a different notion of truth (they need not reject Plato's paradise), they just concern themselves with different questions ("can we construct?", "is there an algorithmic process such that...?", etc.) in the appropriate framework for such questions—intuitionistic logic. This distinction in logic, from that of the classical—that is traditional—mathematician, is not important, it is merely a natural consequence of a distinction in intent.

Mathematics based on intuitionistic logic was first put forward by Brouwer as (part of) a philosophy of mathematics in opposition to Hilbert's formalism. Brouwer's philosophy of intuitionism was plagued by mysticism and received little help from the very public and demeaning controversy with Hilbert (see [44] for an historical perspective on Brouwer and his intuitionism). From the mathematical perspective, there are two major criticisms of Brouwer's work on intuitionism:

(i) Brouwer's system was inconsistent with classical mathematics, in particular in its assertion that all functions are continuous;

(ii) Brouwer focused heavily on what he considered false assertions of classical mathematics, showing that many fundamental theorems, for instance his own fixed point theorem, implied principles which are provably false in intuitionism.

The second criticism led to the belief, still commonly held, that in the words of Hilbert

> "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."

Through the first halve of the 20$^{\text{th}}$ century these faults seemed terminal, and constructive mathematics languished.[2] It was not until 1967, with Erret Bishop's publication of 'Foundations of constructive analysis' [16], were Hilbert's accusations dismissed. Foundations of constructive analysis gave an account of the core of modern mathematics on the basis of intuitionistic logic, together with a minimal informal set theory and dependent choice. Gone was the inconsistency with classical logic, and gone too were any serious doubts of the strength of the constructive approach to mathematics. Shortly after this Per Martin-Löf independently provided a firm philosophical foundation, with the introduction of his intuitionistic type theory [81] in 1972, for constructive mathematics sufficient for the formalisation of the work of Bishop. Mathematically strong, but proof theoretically weak, constructive set theories emerged soon after.

Intuitionistic logic will be introduced formally in the next section; but in the practice of constructive mathematics, like classical mathematics, we are guided by an understanding of what constitutes a proof, and not by any formal system.[3] In the constructive setting this understanding is expressed by the Brouwer-Heyting-Kolmogorov (**BHK**) interpretation of the logical connectives and quantifiers:

- $P \vee Q$: either we have a proof of P or we have a proof of Q;
- $P \wedge Q$: we have both a proof of P and a proof of Q;
- $P \to Q$: we can convert, in a systematic way, any proof of P into a proof of Q;
- $\neg P$: we can derive a contradiction from P;
- $\exists_{x \in A} P(x)$: we have an algorithm constructing an object $x \in A$ together with a proof of $P(x)$;

[2] Although Brouwer's student Arend Heyting's work legitimised the study of intuitionistic systems within mathematical logic.

[3] Brouwer and Bishop each made clear their belief that mathematics is primary to logic, but formal systems are needed for the metamathematics of intuitionistic systems if nothing else.

- $\forall_{x \in A} P(x)$: we can convert an object x and a proof that $x \in A$, to a proof that $P(x)$ holds.

If we have a symbol $\perp$ for absurdity, then the condition for $\neg P$ is just the special case of $P \to Q$ where $Q = \perp$. We could also, for instance, define $\perp$ as $0 = 1$.

It is often suggested that, since intuitionistic logic gives non-classical meanings to the connectives and quantifiers, we should use different symbols.[4] But this full array of logical symbols really belongs to the constructive setting: the logical symbols of classical mathematics are $\neg, \wedge, \forall$, the others are merely convenient (and intuitive) notation (we, classically, write $\exists_x \varphi x$ for $\neg\forall_x \neg\varphi x$, and so forth). Intuitionistic logic gives the same meaning to these fundamental symbols, but gives a stronger, independent meaning to $\vee, \to, \exists$. This simple observation is the essence of the Gödel-Gentzen double-negation translation (introduced below) of classical logic into intuitionistic logic. Constructive mathematics gives a positive meaning to each of the connectives and quantifiers, as opposed to classical mathematics which, as witnessed by the double negation interpretation, gives negative definitions to $\vee, \to, \exists$.

In practice the intent of constructive mathematics manifests itself most prominently in the rejection of the law of excluded middle,[5] or equivalently proof by contradiction. Constructive mathematics is often, falsely, presented as mathematics without the **LEM**, but if we are to conform to the **BHK** interpretation of the connectives and quantifiers, then we must exclude weak instances, or fragments, of the **LEM**. We list some of the commonly encountered fragments of the law of excluded middle, which we must be careful to avoid; Bishop called these the *principles of omniscience.*

[4]This is done in proof decorating, which is a method for extracting computational information from classical proofs [99].

[5]More accurately, constructive mathematics excludes **LEM**: since constructive mathematics is a subsystem of classical mathematics, constructive mathematics makes no assertion on the truth of **LEM**.

- **The limited principle of omniscience (LPO)**: For any binary sequence $(a_n)_{n\geqslant 1}$, either $a_n = 0$ for all n or there exists n such that $a_n = 1$.
- **The weak limited principle of omniscience (WLPO)**: For any binary sequence $(a_n)_{n\geqslant 1}$, either $a_n = 0$ for all n or not $a_n = 0$ for all n.
- **The lesser limited principle of omniscience (LLPO)**: For any binary sequence $(a_n)_{n\geqslant 1}$ with at most one non-zero term, either $a_n = 0$ for all even n or $a_n = 0$ for all odd n.
- **Markov's Principle (MP)**: For any binary sequence $(a_n)_{n\geqslant 1}$, if it is impossible for $a_n = 0$ for all n, then there exists n such that $a_n = 1$.

LPO is the restriction of **LEM** to Π^0_1 sentences; **WLPO** the restriction of the *weak law of excluded middle*—for any proposition A, either A is false or $\neg A$ is false—to Π^0_1 sentences; **LLPO** is the restriction of *De Morgan's law*—$\neg(\neg A \wedge B) \leftrightarrow \neg(A \vee \neg B)$; and **MP** is an instance of proof by contradiction. It is easily seen that the first three principles are ordered by (strictly) decreasing strength and that **WLPO** + **MP** is equivalent to **LPO**. We shall meet other, more subtle, omniscience principles later. Although these principles are rejected as inherently nonconstructive, they do sometimes play a part in fully constructive proofs via Ishihara's tricks [32, Section 3.2]; this is related to the recent work of Martín Escardó [50] showing that some initially surprising instances of excluded middle are constructively valid.

Examples of propositions which are Π^0_1 include Fermat's last theorem, the Riemann hypothesis, and the Goldbach conjecture. For instance, define a binary sequence $(a_n)_{n\geqslant 1}$ such that

$$\begin{aligned} a_n = 0 \quad &\Rightarrow \quad 2n \text{ is the sum of two primes;} \\ a_n = 1 \quad &\Rightarrow \quad \text{the Goldbach conjecture fails for } 2n. \end{aligned}$$

Note that this binary sequence is constructively well defined since we can decide for each n whether $a_n = 0$ or $a_n = 1$ (at least in theory—any calculations are left as an exercise for the reader). Applying **WLPO** to the sequence $(a_n)_{n\geqslant 1}$ tells us whether Goldbach's conjecture is true or false; **LPO** goes one better by either asserting that Goldbach's conjecture is true, or by providing an explicit counterexample. Of course classically each of these assertions is trivial, but with a constructive reading of the connective $\vee$, proving 'the Goldbach conjecture is true or the Goldbach conjecture is false' is to decide the conjecture. In constructive mathematics we seek (computational) proof rather than truth.

A very popular area of current research in the constructive community is the classification of classical theorems by the fragment of the law of excluded middle required, in addition to the techniques available to the constructivist, in order to prove them. This research programme is dubbed constructive reverse mathematics after the related branch of mathematical logic, reverse mathematics, initiated by Harvey Friedman and Stephen Simpson in the 70's. Any theorem which is consistent with constructive mathematics is within the realm of constructive reverse mathematics, in particular classically false results from computable analysis and Brouwer's intuitionism. In section 1.5 we discuss constructive reverse mathematics in a little more detail.

Reverse mathematics is so called because, in addition to proving a theorem φ in the appropriate system T (forward mathematics), one must also show that this system is the weakest sufficient to prove φ; that is, we must show that φ implies the axioms of T (over an appropriate base theory)—this latter task is the reversal. In the (informal) constructive setting, reversals date back to Brouwer's work, so called *Brouwerian (counter)examples*, showing classical theorems to be constructively unprovable. We give a simple, but fundamental, example: the statement

(*) 'Every nonempty set has an element'

implies, and hence is equivalent to, the law of excluded middle. Fix a sentence φ, and define

$$S = \{x : x = 0 \wedge \varphi\} \cup \{x : x = 1 \wedge \varphi\}.$$

If S were empty, then we would have $\varphi \wedge \neg\varphi$, which is absurd and hence S is nonempty. But if $x \in S$, then either $x > 0$ or $x < 1$, allowing us to decide whether φ or $\neg\varphi$ holds. Hence (*) implies the law of excluded middle.

This Brouwerian example leads us to isolate the more positive notion of being nonempty expressed by (*): a set S is *inhabited* if there exists x such that $x \in S$. A slightly slicker, but less intuitive, proof of this Brouwerian example uses the set

$$S = \{x : x = 0 \wedge (\varphi \vee \varphi)\}.$$

There are also a number of so called 'semi-constructive' principles which have been isolated in the literature, and which play a part in the constructive reverse mathematics of computable analysis and Brouwer's intuitionism. Key among these are

(i) those introduced by Brouwer in order to prove that all real valued functions on $[0, 1]$ are uniformly continuous, and

(ii) those which characterise the Russian school of computable analysis.

The former principles are

- Brouwer's continuity principles

  BCP : (1) Any function from $\mathbf{N}^{\mathbf{N}}$ to $\mathbf{N}$ is continuous;
  (2) If $P \subset \mathbf{N}^{\mathbf{N}} \times \mathbf{N},$ and for each $\mathbf{a} \in \mathbf{N}^{\mathbf{N}}$ there exists $n \in \mathbf{N}$ with $(\mathbf{a}, n) \in P,$ then there is a function $f : \mathbf{N}^{\mathbf{N}} \to \mathbf{N}$ such that $(\mathbf{a}, f(\mathbf{a})) \in P$ for all $a \in \mathbf{N}^{\mathbf{N}}$;

- the contrapositive of (weak) König's lemma, Brouwer's fan theorem (**FT**): if every branch of a binary tree T is finite, then the tree is finite.

Different variants of Brouwer's fan theorem are formed by restricting the complexity of (non)-membership in the tree. The appropriate notation and definitions are introduced in Section 2.1.

The principles which characterise computable analysis are Markov's principle together with

> **CPF**: there is an enumeration of the set of partial functions from **N** to **N** with countable domains.

CPF (which stands for countable partial functions) is a simple consequence of the *Church-Turing thesis*—a function is computable if and only if it is computable by a Turing machine—together with a focus on computable functions.

Brouwer isolated the notion of the 'creating subject' which was central to his philosophy of mathematics. In particular, the creating subject allowed Brouwer to argue that Markov's principle implies the law of excluded middle, his motivation being to prove Markov's principle to be false. Myhill introduced a principle, which he called Kripke's schema, to formalise the creating subject arguments of Brouwer. *Kripke's schema* says

> **KS**: For every proposition P, there exists a binary sequence $(a_n)_{n\geqslant 1}$ such that if P is true, then there exists n such that $a_n = 1$, and if P is false, then $a_n = 0$ for all n;

$$\exists_{\alpha\in 2^{\mathbf{N}}}(P \leftrightarrow \exists_{n\in\mathbf{N}}\alpha(n) = 1).$$

Although Kripke's schema is not of great mathematical interest, Peter Schuster has shown it to have a few mathematically interesting equivalents [79, 98].

We will often refer to: constructive mathematics with dependent choice as **BISH** in honour of Erret Bishop; constructive mathematics augmented by Brouwer's continuity principles and fan theorem as **INT**; and constructive mathematics augmented by **CPF** and Markov's principle as **RUSS**. Together with classical mathematics **CLASS** (**BISH** plus **LEM** and the full axiom of choice), **INT** and **RUSS** are the historically important extensions of constructive mathematics. See [23], from which we have borrowed this notation, for an introduction to these systems and there interrelations.

Constructive mathematics without the axiom of dependent choice is also of interest, primarily because it has a wide array of models.[6] As a result, we will be interested in a number of choice axioms of varying strength, including the following.

The axiom of choice

AC: If a is an inhabited set and ψ is a (class) function with domain a such that for each $x \in a$ there exists y such that $\psi(x, y)$, then there exists a *choice function* $f : a \to b$ with $\psi(x, f(x))$ for each $x \in a$;

$$\forall_{x \in a} \exists_y \psi(x, y) \to \exists_b \exists_{f \in b^a} \forall_{x \in a} \psi(x, f(y)).$$

The axiom of dependent choice

DC: If a is a set, $x_0 \in a$, S is a subset of $a \times a$, and for each $x \in a$ there exists $y \in a$ such that $(x, y) \in S$, then there exists a sequence $(x_n)_{n \in \mathbf{N}}$ in X such that $x_0 = a$ and $(x_n, x_{n+1}) \in S$ for each n;

$$\forall_a \forall_{R \subset a \times a} (\forall_{x \in a} \exists_{y \in a} xRy \wedge x_0 \in a \to \exists_{f \in a^{\mathbf{N}}} (f(0) = x_0 \wedge \forall_{n \in \mathbf{N}} f(n) R f(n+1))).$$

The axiom of relativised dependent choice

RDC: Dependent choice for a class relation ψ on a class φ:

$$\forall_x ((\varphi(x) \to \exists_y (\varphi(y) \wedge \psi(x, y))) \to (\varphi(x_0) \to \exists_b \exists_{f \in b^{\mathbf{N}}} (f(o) = x_o \wedge \forall_{n \in \mathbf{N}} (\varphi(n) \wedge \psi(f(n), f(n+1)))))).$$

The axiom of countable choice

CC: If b is an inhabited set and S is a subset of $\mathbf{N} \times b$ such that for

[6] Also Fred Richman has voiced concerns over the constructive validity of choice axioms [93].

each n there exists $y \in b$ with $(n, y) \in S$, then there exists a function $f : \mathbf{N} \to b$ such that $(n, f(n)) \in S$ for each n;

$$\forall_{n\in\mathbf{N}}\exists_{y\in b}(n, y) \in S \to \exists_{f\in b^{\mathbf{N}}}\forall_{n\in\mathbf{N}}(n, f(n)) \in S.$$

The axiom of unique choice

AC!: The axiom of choice where the (class) relation ψ is restricted to a (class) function;

$$\forall_{x\in a}\exists!_y\psi(x, y) \to \exists_b\exists_{f\in b^a}\forall_{x\in a}\psi(x, y).$$

The axiom of weak countable choice

WCC: If $(a_n)_{n\in\mathbf{N}}$ is a sequence of inhabited sets at most one of which is not a singleton—if $n \neq n'$, then one of $A_n, A_{n'}$ is a singleton—then there is a choice function $f : \mathbf{N} \to \bigcup a_n$ with $f(n) \in a_n$ for each n.

Restrictions of countable choice such as

AC$_{\omega,2}$: The axiom of choice with the additional restriction that

$$|\{y \in Y : (n, y) \in S\}| \leqslant 2$$

for each n.

We can also form new choice principles by restricting the complexity of the classes from which elements are chosen. This is done using the arithmetic hierarchy from second order arithmetic. A formula φ is a Δ_0 formula if it contains no quantifiers. In the classical setting we can then define the arithmetically hierarchy for formulas inductively by

- $\Sigma_0 \equiv \Pi_0 \equiv \Delta_0$;
- a formula is Σ_{n+1} if it is of the form $\exists_x\varphi$ for a Π_n formula φ;
- a formula is Π_{n+1} if it is of the form $\forall_x\varphi$ for a Σ_n formula φ.

Then any formula with n quantifiers is (classically) in $\Sigma_n \cup \Pi_n$. A *set is* Π_n (resp. Σ_n) if it is defined by a Π_n (resp. Σ_n) formula. In the constructive context, where not every formula is equivalent to a formula in which all quantifiers occur at the front, we must be more careful; we leave the formulation of an appropriate definition of the arithmetic hierarchy as an exercise (the solution can be found in [64]). We shall only make use of Π_1 sets, which we define directly: a subset S of a set X is a *Π_1-set* if it is the intersection of countably many decidable subsets of X. We are particularly interested in the restriction of $\mathbf{AC}_{\omega,2}$ to Π_1 subsets of $\mathbf{N} \times Y$, denoted Π^0_1-$\mathbf{AC}_{\omega,2}$.

The axiom of unique choice, and the axiom of finite choice—

$$\forall_{m \in n} \exists_y \psi(m, y) \to \exists_b \exists_{f \in b^n} \forall_{m \in n} \psi(m, f(m))$$

for any positive integer n—are valid in **CZF**. The relationships between these choice principles are summarised in Figure 1.

Constructive reverse mathematics is normally, in contrast to classical reverse mathematics, done over a base system including the axiom of dependent choice, but it is also interesting, and more in line with the classical programme, to classify the choice principles required to prove a theorem. An additional motivating factor towards considering choice principles is that many of the common models of constructive set theory do not, in general, satisfy any axiom of choice beyond unique and finite choice.

In Section 1.5 on reverse mathematics we summarise some of the relations between the logical and choice principles introduced so far.

Why pursue constructive mathematics?

Of course constructive mathematics has its origins in philosophy, first in the bizarre and confusing philosophy of Brouwer[7] [36] and then in the pragmatic, and comparably unobjectionable (though aggressively pushed), philosophy of Bishop [17, 15]. Briefly, when asked the question "When does an object exist?": Brouwer answered when it can be 'mentally constructed'; while

[7]See [83] for an interesting, if somewhat acerbic, exposition of the philosophical misconceptions of the founders of constructive, particularly intuitionistic, mathematics.

Bishop suggested that the correct question is “when does an object have a concrete existence?” Many philosophical motivations have been put forward for studying constructive mathematics in the intervening years, and if this is your cup of tea, then please have a sip. If you a not a tea drinker, then rest assured there are non-philosophical motivations for the study of constructive mathematics.

Developments in the 20^{th} century, both mathematical and technological, have provided many reasons to study mathematics with intuitionistic logic. The development and proliferation of computers, and the application of mathematics throughout the sciences, has made the question “when can we, and when can we not, construct?” of fundamental importance, and ‘constructive mathematics’ provides a natural and elegant domain in which to address this question. Perhaps more importantly (at least for those of us with an inclination towards more pure mathematics), intuitionistic systems have come to prominence in mathematics in a natural way through the category theoretic generalization of set theory: the logic of topos theory[8] is intuitionistic.

More precisely, just as the category of sets forms a topos, general toposes can be seen as the universe of sets of some *local set theory*. Toposes can then be regarded as models of local set theories in such a way that we have both soundness—theorems of a local set theory T are true in every model of T—and completeness—any proposition validated by all models of a local set theory T is a theorem of T. Those theorems which hold in all local set theories coincide precisely with (some higher order) intuitionistic logic [10]. Moreover, the standard constructive set theories, introduced in the next section, can be interpreted into toposes (with possibly a little extra structure).

[8] Very briefly, a topos is the natural category theoretic generalization of the category of sets, the standard universe of **ZFC**, which is sufficient for the development of mathematics; each topos defines a mathematical framework. In particular, in every topos one can identify objects which behave like the empty set, the naturals, the reals, and powersets—toposes are impredicative. See, for example, [67] for an introduction to topos theory.

Many further motivations for constructive mathematics have been put forth and we briefly mention three; the latter two are merely consequences of the fact that constructive mathematics admits many models.

- ▶ Martin Löf type theory provides a firm and informative foundation for constructive mathematics with a complete, and unobjectionable, philosophical justification—although one might disagree with his interpretations of the connectives and quantifiers. In contrast the justifications for **ZFC**, the standard foundation for classical mathematics, is generally far more personal and intuitive.

- ▶ Fred Richman has emphasised the idea that constructive mathematics is a generalisation of classical mathematics: any classical theorem φ corresponds to the constructive theorem $\mathbf{LEM} \to \varphi$ (and possibly a stronger theorem $\mathbf{OP} \to \varphi$ for some fragment $\mathbf{OP}$ of $\mathbf{LEM}$), so constructive mathematics reveals a finer structure than classical mathematics. Of course this is just the observation that constructive mathematics (appropriately formalised) is a subsystem of classical mathematics (**ZFC**, say). One must still argue, as Richman has, that classical mathematics obscures meaningful distinctions which constructive mathematics makes apparent. In addition, the wealth of interesting non-classical toposes, and other models of constructive mathematics, and their use, for instance, in the foundations of quantum theory [60], make such an argument.

  Related to the arguments of Richman are the ideas of mathematical precision and pragmatism. Constructive mathematics can also provide a weak foundation for reverse mathematics with the coding power to express much of core mathematics; indeed, constructive mathematics is put forward by Simpson [100] as roughly akin to the base system, RCA_0, of classical reverse mathematics. It is also the case that a constructive proof is more informative; not only do we, by using fewer assumptions (or rather axioms), prove a stronger result, it is also possible to extract more information, both computational and noncomputational, from a constructive proof.

- Constructive mathematics has a close relationship with classical computability theory. Any positive result from constructive mathematics can be translated directly into the computability context, allowing one to avoid the tedious formal details associated with the need for coding and so forth in computability theory. Many of the negative results, Brouwerian examples, also transfer over to computable mathematics: if, for example, from an algorithmic proof of φ you could give an algorithm for **LLPO**, then, since it can be shown no such algorithm exists, φ is noncomputable.

  The relationship between constructive mathematics and computable mathematics is in fact tighter than these simple observations suggest:

  > "Computable mathematics is the realizability[9] interpretation of constructive mathematics."
  >
  > Andrej Bauer, [7]

  Although Bauer's comment is an embellishment, it has been suggested, by Douglas Bridges among others, that many of the positive results of computable mathematics can be obtained in this way.

These points are discussed further in section 1.4.

The development of constructive mathematics (by Bishop [16] for example) does not require any formal system of logic; it is done in the typical informal style of the working (classical) analyst, algebraist, topologist... and this is the style adopted in this thesis. We give some basic examples of the informal practice of constructive mathematics in section 1.5, this section also furnishes us with many of the common positive definitions that are necessary for the profitable application of constructive methods.

[9] Realizability is a semantics for constructive mathematics based on trying to formalise the notions of proof and algorithm in the **BHK** interpretation of the logical symbols. We briefly introduce realizability in section 1.3

For a development of constructive analysis see [16, 18, 23, 32], constructive algebra has also seen great progress [75, 87] in particular in the pursuit of a revised Hilbert's programme [41, 42].

For those who still disparage the study of mathematics from a constructive point of view (in addition to the classical approach), we finish with this remarkable (classical) theorem:

Theorem *[**ZF**] Constructive mathematics is universal, in that it is the mathematics for every mathematician.

Proof. Given a mathematician M, we have to cases.

Case 1: M is a pessimist. Then he is skeptical and doubts the consistency of **ZFC**; whence he falls back upon the firm foundation of Martin-Löf's type theory, and is a constructivist.

Case 2: M is an optimist. As an optimist, M prefers the positive constructive approach to mathematics to the negative approach of classical mathematics. ■

1.1.1 Intuitionistic logic

It has come time for us to be a little more precise, but we will endeavor to keep the unsightly and cumbersome details of formal logic in the background.

We introduce two (equivalent) formulations of first order intuitionistic logic: the intuitionistic quantifier calculus **IQC**. The first, due to Gentzen,[10] is formulated in a way which reflects arguments of mathematical practice, and is aptly named *natural deduction*. Natural deduction can be seen as an attempt to make the BHK-interpretation precise. We first recall the basic definitions of formal logic that we shall need.

The language $\mathcal{L}$ of propositional logic contains:

[10] A few years after introducing natural deduction, Gentzen gave another formulation of logic: the sequent calculus. Gentzen's sequent calculus is less natural than natural deduction, but is generally superior for proof theoretic investigations.

- the logical symbols $\wedge, \vee, \to, \perp, \forall, \exists$;
- countably many variables $v_0, v_1, \ldots$;
- countably many n-ary relation symbols $R^n_0, R^n_1, \ldots$ for each positive natural number n;
- countably many n-ary function symbols $f^n_0, f^n_1, \ldots$ for each n; in particular, 0-ary function symbols are (called) constants.

Falsity $\perp$ is identified as the unique 0-ary relation. *Terms* of $\mathcal{L}$ are defined inductively: variables and constants of $\mathcal{L}$ are terms; and if $t_1, t_2, \ldots, t_n$ are terms and f is an n-ary function symbol of $\mathcal{L}$, then $f(t_1, \ldots, t_n)$ is a term. *Atomic formulas* are those expressed by relations: if $t_1, t_2, \ldots, t_n$ are terms and R is an n-ary function symbol of $\mathcal{L}$, then $R(t_1, \ldots, t_n)$ is an atomic formula; in particular, $\perp$ is an atomic formula. The *formulas* of $\mathcal{L}$ are constructed from atomic formulas using the logical symbols:

(i) atomic formulas are formulas;

(ii) if φ, ψ are formulas, then

$$\varphi \wedge \psi, \varphi \vee \psi, \text{ and } \varphi \to \psi$$

are formulas;

(iii) if φ is a formula and v is a variable, then $\exists_v \varphi$ and $\forall_v \varphi$ are formulas.

In first order logic there is an important distinction between the variables in a formula introduced within atomic formulas and those introduced by (iii); the former are called *free variables*, while the later are *bound variables* (being bounded by the quantifier). Formally we define the free variables of φ, denoted $\mathrm{FV}(\varphi)$, by induction on the construction of φ:

- if φ is atomic, then $\mathrm{FV}(\varphi)$ is the set of variables occurring in φ;
- $\mathrm{FV}(\varphi \wedge \psi), \mathrm{FV}(\varphi \vee \psi), \mathrm{FV}(\varphi \to \psi) = \mathrm{FV}(\varphi) \cup \mathrm{FV}(\psi)$;
- $\mathrm{FV}(\exists_v \varphi), \mathrm{FV}(\forall_v \varphi) = \mathrm{FV}(v) \setminus \{\varphi\}$.

A *sentence* is a formula with no free variables. We define $\neg A$ as A implies absurdity $A \to \bot$, and we write $A \leftrightarrow B$ for $A \to B \wedge B \to A$. By $A[x/y]$ we denote the formula given by replacing every occurrence of x in A by y; if x only occurs bound in A, then we make no distinction between A and $A[x/y]$.

We can now introduce Gentzen's natural deduction. Natural deduction is built around the application of valid rules, and an argument is given by a valid finite sequence of these rules; we denote arguments by $\mathcal{D}_1, \mathcal{D}_2, \ldots$. Our most basic rule, which must be used at the beginning of any argument, is the introduction of assumptions, or suppositions:

$$\text{Ass.} \ \frac{\mathcal{D}_1}{A}.$$

The other basic rules are those corresponding to the logical connectives and quantifiers. For each logical symbol $*$ we have an introduction rule $*I$ and an elimination rule $*E$.

Introduction rules	Elimination rules
$\wedge\mathrm{I} \ \dfrac{\begin{matrix}\mathcal{D}_1 & & \mathcal{D}_2\\ A & & B\end{matrix}}{A \wedge B}$	$\wedge\mathrm{E}_r \ \dfrac{\begin{matrix}\mathcal{D}_1\\ A \wedge B\end{matrix}}{A} \qquad \wedge\,\mathrm{E}_l \ \dfrac{\begin{matrix}\mathcal{D}_1\\ A \wedge B\end{matrix}}{B}$
$\to \mathrm{I} \ \dfrac{\begin{matrix}\mathcal{D}_1\\ B\end{matrix}}{A \to B}$	$\to \mathrm{E} \ \dfrac{\begin{matrix}\mathcal{D}_1 & & \mathcal{D}_\in\\ A \to B & & A\end{matrix}}{B}$
$\vee\mathrm{I}_l \ \dfrac{\begin{matrix}\mathcal{D}_1\\ A\end{matrix}}{A \vee B} \qquad \vee\,\mathrm{I}_r \ \dfrac{\begin{matrix}\mathcal{D}_1\\ B\end{matrix}}{A \vee B}$	$\vee\mathrm{E} \ \dfrac{\begin{matrix}\mathcal{D}_1 & \mathcal{D}_2 & \mathcal{D}_3\\ A \vee B & A \to C & B \to C\end{matrix}}{C}$

$$\forall\mathrm{I} \quad \frac{\begin{array}{c}\mathcal{D}_1\\ A\end{array}}{\forall_y A[x/y]} \qquad\qquad \forall\mathrm{E} \quad \frac{\begin{array}{c}\mathcal{D}_1\\ \forall_x A\end{array}}{A[x/t]}$$

where x bound in $\mathcal{D}_1$ — where t is free for x in a

$$\exists\mathrm{I} \quad \frac{\begin{array}{c}\mathcal{D}_1\\ A[x/t]\end{array}}{\exists_x A} \qquad\qquad \exists\mathrm{E} \quad \frac{\begin{array}{c}\mathcal{D}_1\\ \exists_y A[x/y]\end{array} \qquad \begin{array}{c}\mathcal{D}_2\\ C\end{array}}{C}$$

$(\mathcal{D}_1, A)$ satisfy condition 1 — $(\mathcal{D}_2, C, A)$ satisfy condition 2

Conditions 1 and 2 are the natural variable restrictions, but before we can state them precisely, we need a definition. The assumptions $\mathrm{Ass}(\mathcal{D})$ of a deduction $\mathcal{D}$ are defined inductively by the following equations.

(i) The assumptions for an argument ending in an application of the assumption rule are $\mathrm{Ass}(\mathcal{D}_1) \cup \{A\}$.

(ii) The assumptions for an argument ending with $\forall$I are $\mathrm{Ass}(\mathcal{D}_1) \setminus \{A\}$.

(iii) The assumptions for an argument ending with $\exists$E are

$$\mathrm{Ass}(\mathcal{D}_1) \cup \mathrm{Ass}(\mathcal{D}_2) \setminus \{A\}.$$

(iv) For a argument ending in one of the other rules the assumptions are $\mathrm{Ass}(\mathcal{D}_1)$, $\mathrm{Ass}(\mathcal{D}_1) \cup \mathrm{Ass}(\mathcal{D}_2)$, or $\mathrm{Ass}(\mathcal{D}_1) \cup \mathrm{Ass}(\mathcal{D}_2) \cup \mathrm{Ass}(\mathcal{D}_3)$ depending on which deductions appear in the rules.

Condition 1, in $\forall$I, says that x is not free in any assumption of $\mathcal{D}_1$ and either x and y are identical or y is not free in A. Condition 2, in $\exists$E, says that A is the only assumption of $\mathcal{D}_2$ in which x may be free, x is not free in C, and either x and y are identical or y is not free in A.

A *deduction* is a triple $(\Gamma, \varphi, \mathcal{D})$ where Γ is a finite set of formula, φ is a single formula, and $\mathcal{D}$ is an argument ending with φ the assumptions of which form a subset of Γ. The argument $\mathcal{D}$ is said to be a proof of 'φ is derivable from Γ', in symbols $\Gamma \vdash_m \varphi$.

What we have really just described is natural deduction for *minimal logic*—this is what the m subscript denotes. To get *intuitionistic logic* (**IQC**) we must interpret $\perp$ by the *intuitionistic absurdity rule*:

$$\perp_i \;\; \cfrac{\begin{matrix}\mathcal{D}_1\\ \perp\end{matrix}}{A}.$$

The intuitionistic absurdity rule does not effect the assumptions of an argument. A deduction $(\Gamma, \varphi, \mathcal{D})$ using the assumption rule, logical rules, and the intuitionistic absurdity rule is said to be an intuitionistic deduction of $\Gamma \vdash_i \varphi$, φ is intuitionistically provable from Γ. *Classical deduction*, denoted $\Gamma \vdash_c \varphi$, is defined similarly by adding the law of excluded middle in the guise of proof by contradiction. The *absurdity rule* is the same as in the intuitionistic case,

$$\perp_c \;\; \cfrac{\begin{matrix}\mathcal{D}_\infty\\ \perp\end{matrix}}{A},$$

but we are allowed to eliminate $\neg A$ from the assumptions of $\mathcal{D}_1$: the assumptions of an argument ending in an application of $\perp_c$ are $\mathrm{Ass}(\mathcal{D}_1) \setminus \{\neg A\}$. Classically, we can define some of the connectives and quantifiers in terms of $\perp$ and the other connectives and quantifiers using the following classical equivalences.[11]

$$\begin{array}{rclcrcl} A \vee B & \leftrightarrow & \neg(\neg A \wedge \neg B) & \quad & A \wedge B & \leftrightarrow & \neg(\neg A \vee \neg B) \\ A \to B & \leftrightarrow & \neg A \vee B & & & & \\ \forall_x Ax & \leftrightarrow & \neg\exists_x \neg Ax & & \exists_x Ax & \leftrightarrow & \neg\forall_x \neg Ax. \end{array}$$

We will sometimes use $\vdash$ to represent any of $\vdash_m, \vdash_i, \vdash_c$ when the intended

[11]In fact we need only the NAND connective given by $A\textsc{nand}B \;\leftrightarrow \neg(A \wedge B)$ in order to define classical propositional logic.

meaning is clear from context. Under standard terminology, which we now adopt, the argument $\mathcal{D}$ is itself said to be a *deduction* of $\Gamma \vdash \varphi$, and Γ is generally taken to be the set of assumptions open in $\mathcal{D}$. We write deductions in the obvious linear form and label stages in the deduction by the rules they apply, the previous lines to which the rules are applied, and the assumptions, in square brackets, of the current subdeduction; this is demonstrated by the following examples. We first note that, by the assumption introduction and $\to$I rules, there is a derivation of $\Gamma \cup \{A\} \vdash B$ if and only if there is a derivation of $\Gamma \vdash A \to B$.

We begin with an almost trivial example, $\vdash_m A \to \neg\neg A$:

1.	A	Ass.	$[1]$
2.	$\neg A$	Ass.	$[1,2]$
3.	$\bot$	$1,2 \to$ E	$[1,2]$

Here we have proved $\{A, \neg A\} \vdash_m \bot$, which is equivalent, by our previous remark, to $\vdash_m A \to (\neg A \to \bot)$, which by definition is $\vdash_m A \to \neg\neg A$. We adopt the convention that on the left of the turnstile we write a formula φ for the singleton $\{\varphi\}$ and we interpret comma as union, so for example Γ, φ represents $\Gamma \cup \{\varphi\}$.

We also make use of the standard mathematical practice of appealing to a previously proved result within a separate proof: $\vdash_m \neg\neg\neg A \to \neg A$

1.	$\neg\neg\neg A$	Ass.	$[1]$
2.	A	Ass.	$[1,2]$
3.	$\neg\neg A$	2 Theorem	$[1,2]$
4.	$\bot$	$1,3 \to$ E	$[1,2]$

Now for some more substantial examples. We present a proof, in minimal logic, that $\vdash_m \neg\neg A \wedge \neg\neg B \to \neg\neg(A \wedge B)$.

1.	$\neg\neg A \wedge \neg\neg B$	Ass.	$[1]$
2.	$\neg(A \wedge B)$	Ass.	$[1, 2]$
3.	A	Ass.	$[1, 2, 3]$
4.	B	Ass.	$[1, 2, 3, 4]$
5.	$\neg\neg A$	$1 \ \wedge$E	$[1, 2, 3, 4]$
6.	$\neg\neg B$	$1 \ \wedge$E	$[1, 2, 3, 4]$
7.	$A \wedge B$	$3, 4 \ \wedge$I	$[1, 2, 3, 4]$
8.	$\perp$	$2, 7 \ \to$E	$[1, 2, 3, 4]$
9.	$\neg B$	$4, 8 \ \to$I	$[1, 2, 3]$
10.	$\perp$	$6, 9 \ \to$E	$[1, 2, 3]$
11.	$\neg A$	$9, 10 \ \to$I	$[1, 2]$
12.	$\perp$	$5, 11 \ \to$E	$[1, 2]$

A more complicated example:

$$\vdash_m (\neg(A \vee B) \to C) \to (((\neg A \to C) \vee (\neg B \to C)) \to C \to C).$$

In the following proof, we tacitly repeat the assumptions on lines 3,4 by not removing them from our running assumptions on line 6.

1.	$\neg(A \vee B) \to C$	Ass.	$[1]$
2.	$((\neg A \to C) \vee (\neg B \to C)) \to C$	Ass.	$[1, 2]$
3.	$\neg A$	Ass.	$[1, 2, 3]$
4.	$\neg B$	Ass.	$[1, 2, 3, 4]$
5.	$A \vee B$	Ass.	$[1, 2, 3, 4, 5]$
6.	$\perp$	$3, 4, 5 \ \vee$E	$[1, 2, 3, 4, 5]$
7.	$\neg(A \vee B)$	$5, 6 \ \to$I	$[1, 2, 3, 4]$
8.	C	$1, 7 \ \to$E	$[1, 2, 3, 4]$
9.	$\neg B \to C$	$4, 8 \ \to$I	$[1, 2, 3]$
10.	$(\neg A \to C) \vee (\neg B \to C)$	$10 \ \vee\mathrm{I}_l$	$[1, 2, 3]$
11.	C	$2, 10 \ \to$E	$[1, 2, 3]$
12.	$\neg A \to C$	$3, 11 \ \to$I	$[1, 2]$

13.	$(\neg A \to C) \vee (\neg B \to C)$	12 $\vee \mathrm{I}_r$	$[1, 2]$
14.	C	$2, 13 \to \mathrm{E}$	$[1, 2]$

We now extend the language $\mathcal{L}$ for propositional logic to $\mathcal{L}'$ by

(i) no longer restricting ourselves to only countably many variables, relations, and function symbols, and

(ii) adding symbols $\bigvee, \bigwedge$ for infinitary disjunction and conjunction; this allows us to make a step toward second order logic, without leaving the comfort of a first order system.

Formulas of $\mathcal{L}'$ are defined as before except we allow additional constructions for our new logical symbols: if $(\varphi_i)_{i \in I}$ is a set of formulas, then $\bigvee_{i \in I} \varphi_i$ and $\bigwedge_{i \in I} \varphi_i$ are also formulas; to keep things tidy we will suppress mention of the index set I. We then extend our minimal, intuitionistic, and classical natural deduction systems by replacing the rules for disjunction and conjunction by:

$$\bigwedge \mathrm{I} \ \frac{\begin{matrix}\mathcal{D}_i \\ A_i\end{matrix}}{\bigwedge A_i} \qquad \bigwedge \mathrm{E}_j \ \frac{\begin{matrix}\mathcal{D}_1 \\ \bigwedge A_i\end{matrix}}{A_j}$$

$$\bigvee \mathrm{I}_j \ \frac{\begin{matrix}\mathcal{D}_1 \\ A_j\end{matrix}}{\bigvee A_i} \qquad \bigvee \mathrm{E} \ \frac{\begin{matrix}\mathcal{D}_1 \\ \bigvee A_i\end{matrix} \qquad \begin{matrix}\mathcal{D}_i \\ A_i \to C\end{matrix}}{C}$$

where deductions ending with applications of rules $\bigwedge \mathrm{E}_j$, $\bigvee \mathrm{I}_j$ have assumptions $\mathrm{Ass}(\mathcal{D}_1)$; deductions ending with applications of $\bigwedge \mathrm{I}$ have assumptions $\bigcup\{\mathrm{Ass}(\mathcal{D}_i) : i \in I\}$; and deductions ending with applications of $\bigvee \mathrm{E}$ have assumptions $\mathrm{Ass}(\mathcal{D}_1) \cup \bigcup\{\mathrm{Ass}(\mathcal{D}_i) : i \in I\}$. Note that proofs can now have infinite width; in our linear notation this is represented by lines indexed by elements of the possibly infinite index set I. We will also use $\vdash_m, \vdash_i, \vdash_c$ for infinitary derivations in minimal, intuitionistic, and classical natural deduction. Again we leave the details of how deductions are written to a few examples. We also now allow the assumption set Γ to be infinite.

The extension of **IQC** to infinitary logic allows us to make a move toward second order arithmetic, and thus to express concepts which cannot be defined in pure first order logic, for example torsion groups.

As an example of an infinitary deduction we prove

$$\vdash_m \left(\neg \bigwedge A_i \to B\right) \to \bigwedge (\neg A_i \to B)\,.$$

1.	$\neg \bigwedge A_i \to B$	Ass.	$[1]$
2.	$\bigwedge A_i$	Ass.	$[1,2]$
$3i$.	$\neg A_i$	Ass.	$[1,2,3i]$
$4i$.	A_i	$2\ \bigwedge$E	$[1,2,3i]$
$5i$.	$\bot$	$3i,4i\ \to$E	$[1,2,3i]$
$6i$.	$\neg \bigwedge A_i$	$2,5i\ \to$I	$[1,3i]$
$7i$.	B	$1,6i\ \to$E	$[1,3i]$
$8i$.	$\neg A_i \to B$	$3i,7i\ \to$I	$[1]$
9.	$\bigwedge (\neg A_i \to B)$	$8i\ \bigwedge$I	$[1]$

And an example with disjunction: $\vdash_m \neg\neg \bigvee \neg\neg A_i \to \neg\neg \bigvee A_i$.

1.	$\neg\neg \bigvee \neg\neg A_i$	Ass.	$[1]$
2.	$\neg \bigvee A_i$	Ass.	$[1,2]$
$3i$.	A_i	Ass.	$[1,2,3i]$
$4i$.	$\bigvee A_i$	$2\ \bigwedge$E	$[1,2,3i]$
$5i$.	$\bot$	$2,4i\ \to$E	$[1,2,3i]$
$6i$.	$\neg A_i$	$3i,5i\ \to$I	$[1,2]$
$7i$.	$\bigvee \neg\neg A_i$	Ass.	$[1,2,7]$
$8i$.	$\neg\neg A_i$	Ass.	$[1,7,8i]$
$9i$.	$\bot$	$6i,8i\ \to$E	$[1,2,7,8i]$
$10i$.	$\neg\neg A_i \to \bot$	$6i,8i\ \to$E	$[1,2,7]$
$11i$.	$\neg \bigvee \neg\neg A_i$	$7,10i\ \to$I	$[1,2]$
12.	$\bot$	$1,11i\ \to$E	$[1,2]$

It is trivially true that if $\Gamma \vdash_m \varphi$, then $\Gamma \vdash_i \varphi$, and if $\Gamma \vdash_i \varphi$, then $\Gamma \vdash_c \varphi$.

Using the Gödel-Gentzen negative translation we have a partial converse: we can embed classical infinitary first order logic into minimal infinitary first order logic. The Gödel-Gentzen negative translation A^G of a formula A is defined inductively by

$$\begin{array}{ll} \bot^G = \bot & P^G = \neg\neg P \text{ for } P \text{ prime, } P \neq \bot \\ (\bigwedge A_i)^G = \bigwedge A_i^G & (A \to B)^G = A^G \to B^G \\ (\bigvee A_i)^G = \neg \bigwedge \neg A_i^G & \\ (\forall_x A)^G = \forall_x A^G & (\exists_x A)^G = \neg\forall_x \neg A^G. \end{array}$$

Since $\bot$ has no special meaning in minimal logic, we can replace it by an arbitrary proposition: the $\Box$-translation $A^\Box$ of a formula A is $A^G[\bot/\Box]$. We adopt the shorthand that $\neg_\Box A$ stands for $A \to \Box$, and for a set Γ of formulas, we define $\Gamma^\Box = \{\varphi^\Box : \varphi \in \Gamma\}$. We say that a set of formulas Γ is (intuitionistically) closed under $\Box$ if for each $A \in \Gamma$, $\Gamma \vdash_i A^\Box[\Box/B]$ for any formula B such that no free variable of B is a bound variable of a formula in Γ.

Proposition 1 *For any formula A, $\vdash_m \neg_\Box\neg_\Box A^\Box \leftrightarrow A^\Box$, and if $\Gamma \vdash_c A$, then $\Gamma^\Box \vdash_m A^\Box$.*

Proof. By induction on the definition of A and the length of the derivation of $\Gamma \vdash_c A$ respectively. We argue informally.

Since $\bot$ has no special meaning in minimal logic, the prime case follows from $\vdash_m \neg A \leftrightarrow \neg\neg\neg A$. We have the following inductive cases; the right to left directions are trivial.

1. $A = \bigwedge A_i$: If $\vdash_m \neg_\Box\neg_\Box A^\Box$, then $\vdash_m \neg_\Box\neg_\Box A_i^\Box$ for each i, so for each i $\vdash_m A_i^\Box$ by the induction hypothesis. Hence $\vdash_m \bigwedge A_i^\Box$; that is, $\vdash_m A^\Box$.

2. $A = \bigvee A_i$ and $\mathbf{A} = \exists_\mathbf{x}\mathbf{A'}$ follow from $\vdash_m \neg_\Box\neg_\Box\neg_\Box A \leftrightarrow \neg_\Box A$.

3. $A = A_1 \to A_2$: By the induction hypothesis it suffices to show

$$\neg_\Box\neg_\Box(A_1^\Box \to A_2^\Box), A_1^\Box \vdash_m \neg_\Box\neg_\Box A_2^\Box.$$

We have

1.	$\neg_\Box \neg_\Box (A_1^\Box \to A_2^\Box)$	Ass.	$[1]$
2.	$A_1^\Box$	Ass.	$[1,2]$
3.	$\neg_\Box A_2^\Box$	Ass.	$[1,2,3]$
4.	$A_1^\Box \to A_2^\Box$	Ass.	$[1,2,3,4]$
5.	$A_2^\Box$	$2,4 \to \mathrm{E}$	$[1,2,3,4]$
6.	$\Box$	$3,5 \to \mathrm{E}$	$[1,2,3,4]$
7.	$\neg_\Box (A_1^\Box \to A_2^\Box)$	$4,6 \to \mathrm{I}$	$[1,2,3]$
8.	$\Box$	$1,7 \to \mathrm{E}$	$[1,2,3]$

4. $A = \forall_x A'$: If $\vdash_m \neg_\Box \neg_\Box \forall_x A'$, then $\vdash_m \neg_\Box \neg_\Box A'[x/t]$ for suitable t; by the induction hypothesis $\vdash_m A'^\Box$, and therefore $\vdash_m \forall_x A'^\Box$.

Suppose that we have a classical proof of A from assumptions Γ; we prove the second part of the proposition by induction on the length of this derivation. The translations of $\vee, \exists$ are essentially just their classical definitions; since the deduction rules for $\vee, \exists$ can classically be seen as derived rules, we may assume, without loss of generality, that the last rule applied is one of $\wedge, \forall, \to$ introduction or elimination, or an application of the classical absurdity rule. The rules for $\wedge, \forall, \to$ translate directly, so it only left to verify the case where the deduction ends in an application of the classical absurdity rule:

$n.$	$\neg A$	Ass.	$S \cup \{n\}$
	$\ldots$		
$m.$	$\perp$		$S' \cup \{n\}$
$m+1.$	A	$n, m \perp_c$	S'

By the induction hypothesis we have a corresponding argument from $\neg A^\Box$

to $\perp$, so we get the following deduction.

$n.$	$\neg A^{\Box}$	Ass.	$S \cup \{n\}$
	$\dots$		
$m.$	$\perp$		$S' \cup \{n\}$
$m+1.$	$\neg\neg A^{\Box}$	$n, m \to$ I	S'
$m+2.$	$\neg\neg A^{\Box} \to A^{\Box}$	$m+1$ Theorem	S'
$m+3.$	$A^{\Box}$	$m=1, m+2 \to$ E	S'

■

As an example of deductions in the intuitionistic and classical natural deduction systems and the relation between them, we extend Ishihara's result from [63] to the infinitary systems; essentially we take Ishihara's proof and replace instances of conjunctives and disjunctives by their infinitary versions where we can.

We define classes $\mathcal{Q}, \mathcal{R}, \mathcal{J}, \mathcal{K}$ by a simultaneous recursion as follows. Let P range over atomic formulas, Q_i range over $\mathcal{Q}$, R_i range over $\mathcal{R}$, J_i range over $\mathcal{J}$, and K_i range over $\mathcal{K}$. Then $\mathcal{Q}, \mathcal{R}, \mathcal{J}, \mathcal{K}$ are generated by the clauses

$$\begin{aligned} P, \bigwedge Q_i, \bigvee Q_i, \forall_x Q, \exists_x Q, J \to Q &\quad \in \mathcal{Q}; \\ P, \bigwedge R_i, R_1 \vee R_2, \forall_x R, \qquad , J \to R &\quad \in \mathcal{R}; \\ P, J_1 \wedge J_2, \bigvee J_i, \qquad , \exists_x J, R \to J &\quad \in \mathcal{J}; \\ J, \bigwedge K_i, \forall_x K, Q \to K &\quad \in \mathcal{K}. \end{aligned}$$

The infinitary version of Ishihara's theorem is

Theorem 2 *If Γ is closed under $\Box$, $A \in \mathcal{K}$, and $\Gamma \vdash_c A$, then $\Gamma_i \vdash A$.*

The proof of Theorem 2 gives an effective method to convert proofs in infinitary first order classical logic to proofs in infinitary first order intuitionistic logic.

Lemma 3 *In the following 3,6 are provable in infinitary minimal logic, 8 in intuitionistic logic, and the remainder in minimal logic.*

(1) $A \to \neg_\Box\neg_\Box A$

(2) $(A \to B) \to (\neg_\Box\neg_\Box A \to \neg_\Box\neg_\Box B)$

(3) $\neg_\Box\neg(\bigwedge A_i) \to \bigwedge \neg_\Box\neg A_i$

(4) $\neg_\Box\neg_\Box A \wedge \neg_\Box\neg_\Box B \to \neg_\Box\neg_\Box(A \wedge B)$

(5) $\neg_\Box\neg(A \vee B) \to \neg_\Box\neg_\Box(\neg_\Box\neg A \vee \neg_\Box\neg B)$

(6) $\neg_\Box\neg_\Box \bigvee \neg_\Box\neg_\Box A_i \to \neg_\Box\neg_\Box \bigvee a_i$

(7) $\neg_\Box\neg(A \to B) \to (\neg_\Box\neg_\Box A \to \neg_\Box\neg B)$

(8) $(\neg_\Box\neg A \to \neg_\Box\neg_\Box B) \to \neg_\Box\neg_\Box(A \to B)$

(9) $\neg_\Box\neg\forall_x A \to \forall_x\neg_\Box\neg A$

(10) $\neg_\Box\neg_\Box\exists_x\neg_\Box\neg_\Box A \to \neg_\Box\neg_\Box\exists_x A$

Proof. The example deductions prove 3-6 (recall that $\perp$ has no special meaning in minimal logic). We only prove 8 since it requires the intuitionistic absurdity rule.

1.	$\neg_\Box\neg A \to \neg_\Box\neg_\Box B$	Ass.	$[1]$
2.	$\neg_\Box(A \to B)$	Ass.	$[1, 2]$
3.	$\neg A$	Ass.	$[1, 2, 3]$
4.	A	Ass.	$[1, 2, 3, 4]$
5.	$\perp$	$3, 4 \to$ E	$[1, 2, 3, 4]$
6.	B	$5\perp_i$	$[1, 2, 3, 4]$
7.	$A \to B$	$4, 6 \to$ I	$[1, 2, 3]$
8.	$\Box$	$2, 7 \to$ E	$[1, 2, 3]$
9.	$\neg_\Box\neg A$	$3, 8 \to$ I	$[1, 2]$
10.	$\neg_\Box\neg_\Box B$	$1, 9 \to$ E	$[1, 2]$
11.	B	Ass.	$[1, 2, 11]$
12.	A	Ass.	$[1, 2, 11, 12]$
13.	$A \to B$	$11, 12 \to$ I	$[1, 2, 11]$
14.	$\Box$	$2, 13 \to$ E	$[1, 2, 11]$

15.	$\neg_{\Box} B$	$11, 14 \to I$	$[1,2]$
16.	$\Box$	$10, 15 \to \mathrm{E}$	$[1,2]$

■

Proposition 4 *If $A \in \mathcal{Q}$, then $\vdash_i A \to A^{\Box}$; in particular, $\mathcal{Q}$ is closed under $\Box$. If $A \in \mathcal{R}$, then $\vdash_i \neg_{\Box}\neg A \to A^{\Box}$. If $A \in \mathcal{J}$, then $\vdash_i A^{\Box} \to \neg_{\Box}\neg_{\Box} A$.*

Proof. By simultaneous induction; we argue informally in intuitionistic logic. The induction for formulas in $\mathcal{Q}$ are straightforward, but we, nonetheless, prove three cases. If $Q = \bigvee Q_i$, we have the following derivation

1.	$\bigvee Q_i$	Ass.	$[1]$
$2i$.	Q_i	Ass.	$[1, 2i]$
$3i$.	$Q_i \to Q_i^{\Box}$	Induction hypothesis	$[1, 2i]$
$4i$.	$Q_i^{\Box}$	$2i, 3i \to \mathrm{E}$	$[1, 2i]$
$5i$.	$\bigvee Q_i^{\Box}$	$4i \vee \mathrm{I}$	$[1, 2i]$
$6i$.	$Q_i \to \bigvee Q_i^{\Box}$	$2i, 5i$	$[1]$
7.	$\bigvee Q_i^{\Box}$	$1, 3i \vee \mathrm{E}$	

Formally step $3i$ represents a deduction of $Q_i \to Q_i^{\Box}$ which exists by the induction hypothesis step. If $Q = \exists_x Q'$, then $Q[x/y]$ holds for some y. By the induction hypothesis $(Q[x/y])^{\Box} = Q^{\Box}[x/y]$ holds, so $\exists_x Q^{\Box}$, and hence $\neg_{\Box}\neg_{\Box}\exists_x Q^{\Box}$, holds. Now suppose $Q = J \to Q_1$, where $J \in \mathcal{J}$ and $Q_1 \in \mathcal{Q}$. Then $Q \to Q^{\Box}$, by the induction hypothesis in $\mathcal{Q}$, so $J \to Q^{\Box}$. Thus

$$J^{\Box} \to \neg_{\Box}\neg_{\Box} J \to \neg_{\Box}\neg_{\Box} Q^{\Box} \to Q^{\Box},$$

where the first implication is by the induction hypothesis in $\mathcal{J}$, the second follows from $J \to Q^{\Box}$, and the last is an application of Proposition 1.

The inductions on the formation of formulas in $\mathcal{R} \cup \mathcal{J}$ all follow a similar pattern, and we give a representative sample. The base case $R = \bot$ is the only place (other than Lemma 3 (8)) that we need the intuitionistic absurdity rule; and the case $J = \bot$ is an application of $\vdash_m A \to (B \to A)$.

$\bigwedge R_i \in \mathcal{R}$:

$$\neg\Box\neg \bigwedge R_i \to \bigwedge \neg\Box\neg R_i \to \bigwedge R_i^{\Box}.$$

by Lemma 3, and the induction hypothesis.

$R_1 \vee R_2 \in \mathcal{R}$:

$$\neg\Box\neg(R_1 \vee R_2) \to \neg\Box\neg\Box(\neg\Box\neg R_1 \vee \neg\Box\neg R_2) \to \neg\Box\neg\Box(R_1^{\Box} \vee R_2^{\Box}).$$

$J \to R \in \mathcal{R}$:

$$\begin{array}{rcll} \neg\Box\neg(J \to R) & \to & (\neg\Box\neg\Box J \to \neg\Box\neg R) & \text{by Lemma 3} \\ & \to & (J^{\Box} \to \neg\Box\neg R) & \text{induction on } J \\ & \to & (J^{\Box} \to R^{\Box}) & \text{induction on } R. \end{array}$$

$\exists_x J \in \mathcal{J}$:

$$\neg\Box\neg\Box\exists_x J^{\Box} \to \neg\Box\neg\Box\exists_x\neg\Box\neg\Box J \to \neg\Box\neg\Box\exists J$$

by first the induction hypothesis in $\mathcal{J}$ and then by Lemma 3.

$R \to J \in \mathcal{J}$:

$$\begin{array}{rcll} (R \to J)^{\Box} & = & R^{\Box} \to J^{\Box} & \\ & \to & (\neg\Box\neg R \to J^{\Box}) & \text{induction on } R \\ & \to & (\neg\Box\neg R \to \neg\Box\neg\Box J) & \text{induction on } J \\ & \to & \neg\Box\neg\Box(R \to J) & \text{by Lemma 3.} \end{array}$$

■

We now give **the proof of Theorem 2**:

Proof. The proof is by induction on the length of a formula in $\mathcal{K}$. Relabeling variables, if necessary, we may assume that all variable conditions are satisfied. With Γ a set of formulas closed under $\Box$, we have four cases to consider.

1. $J \in \mathcal{J}$: If $\Gamma \vdash_c A$, then by Lemma 1, $\Gamma^{\Box} \vdash_i A^{\Box}$. By Proposition 4, $\Gamma^{\Box} \vdash_i \neg_{\Box}\neg_{\Box}A$, so[12] $\Gamma^{\Box}[\Box/A] \vdash_i (\neg_{\Box}\neg_{\Box}A)[\Box/A]$. Noting that $(\neg_{\Box}\neg_{\Box}A)[\Box/A]$—that is, $(A \to A) \to A$—is equivalent over minimal logic to A, it follows from Γ being closed under $\Box$ that $\Gamma \vdash_i A$.

2. $\bigwedge K_i$: If $\Gamma \vdash_c \bigwedge K_i$, then $\Gamma \vdash_c K_i$ for each i. By the induction hypothesis, for all i, $\Gamma \vdash_i K_i$; whence $\Gamma \vdash_i \bigwedge K_i$.

3. $\forall_x K$: If $\Gamma \vdash_c \forall_x K$, then $\Gamma \vdash_c K$, so by the induction hypothesis $\Gamma \vdash_i K$, and finally $\Gamma \vdash_i \forall_x K$.

4. $Q \to K$: Suppose that $\Gamma \vdash_c Q \to K$ for some $Q \in \mathcal{Q}$ and $K \in \mathcal{K}$. By Proposition 4, $\Gamma \cup \{Q\}$ is closed under $\Box$, so we can apply the induction hypothesis to a classical proof of K from $\Gamma \cup \{Q\}$ to show that $\Gamma \cup \{Q\} \vdash_i K$. Hence $\Gamma \vdash_i Q \to K$.

■

A Hilbert style system

We also give a Hilbert style system—a system with axioms and rules, but no introduction and elimination of assumptions—for intuitionistic logic. There are six groups of axiom schemata:

$$
\begin{array}{ll}
\wedge - \text{Ax.} & A \wedge B \to A, \quad A \wedge B \to B, \\
 & A \to (B \to (A \wedge B)); \\
\vee - \text{Ax.} & A \to A \vee B, \quad B \to (A \vee B), \\
 & (A \to C) \to ((B \to C) \to (A \vee B \to C)); \\
\to -\text{Ax.} & A \to (B \to A), \\
 & (A \to (B \to C)) \to ((A \to B) \to (A \to C)); \\
\perp - \text{Ax.} & \perp \to A; \\
\exists - \text{Ax.} & A[x/t] \to \exists_x A \quad t \text{ free for } x \text{ in } A, \\
 & \exists_x(A \to B) \to (\exists_y A[x/y] \to B) \\
 & \qquad x \notin \mathrm{FV}(B), y \in \mathrm{FV}(A) \Rightarrow x = y;
\end{array}
$$

[12] The obvious substitution lemma applied here is proved by another straightforward, but unsightly, induction on the length of a deduction.

$$\begin{aligned}\forall - \text{Ax.}\quad & \forall_x A \to A[x/t] \quad t \text{ free for } x \text{ in } A,\\ & \forall_x (B \to A) \to (B \to \forall_y A[x/y])\\ & \qquad x \notin \mathrm{FV}(B), y \in \mathrm{FV}(A) \Rightarrow x = y;\end{aligned}$$

together with modus ponens and the rule of generalisation: if $\vdash A$, then $\vdash \forall_x A$. The relationship between these axioms and the rules of intuitionistic natural deduction should be clear, and indeed the theorems provable from intuitionistic natural deduction and this intuitionistic Hilbert style system coincide.

These axioms allow us to construct models of intuitionistic logic by interpreting proof and algorithm by computable functions. For example, the $\wedge$ axioms correspond to pairing and projection, and the $\to$ axioms correspond to the function sending (x, y) to x and the function sending (x, y, z) to the value given by the kth Turing machine applied to j, where k is the output of the xth Turing machine applied to z and j is the output of the yth Turing machine applied to z; in the lambda notation (introduced below)

$$\mathbf{k} = \lambda xy.x \quad \text{and} \quad \mathbf{s} = \lambda xyz\{\{x\}(z)\}(\{y\}(z)).$$

Equality axioms

In contrast to the classical case, in constructive mathematics equality is normally given as a defined notion, and it is natural to think of equality as mathematical rather than logical. For first order intuitionistic logic with equality we add a symbol $=$ together with the *axioms of equality*:

E0 Replacement, $\forall_{x,y}(\varphi(x) \wedge x = y \to \varphi(y))$;

E1 Identity, $\forall_x(x = x)$;

E2 Symmetry, $\forall_{x,y}(x = y \to y = x)$;

E3 Transitivity, $\forall_{x,y,z}(x = y \wedge y = z \to x = z)$.

If in addition we have sets and the notion of membership, then we add

E4 $\forall_{x,y,z}(x = y \wedge x \in z \to y \in z)$;

E5 $\forall_{x,y,z}(x = y \wedge z \in x \to z \in y)$.

Models of computation

The λ-*calculus* is a theory of functions as rules, in contrast to the standard set theoretic treatment of functions as graphs—regarding functions as rules stresses their computational content. Using his λ-calculus, Church proposed a notion of 'effectively computable' and formulated the Church-Turing thesis: "a function is (algorithmically) computable if and only if it is computable by a Turing machine, or equivalently a term of the λ-calculus." Terms of the λ-calculus are defined over an alphabet with countably many variables $v_1, v_2, \ldots$, the *abstractor* λ, and parenthesis. The set Λ of λ terms is given inductively by the following rules.

(i) For each i, $v_i \in \Lambda$.

(ii) If $M \in \Lambda$, then $\lambda x.M \in \Lambda$ for any variable x.

(iii) If M, N are λ terms, then the *application* MN of M to N is a λ term.

A few remarks are in order. The term $\lambda x.M$ produced by rule (ii)—instances of which are known as *abstraction*—intuitively corresponds to the function $x \mapsto M$, and MN is the result of applying the function M to input N—the λ-calculus makes no distinction between functions and input, this corresponds to the fact that in computer science programs and data are the same, each being represented by a binary string. In particular, functions can be applied to themselves, which is not possible with the graph conception of function (by cardinality considerations). *Bound variables* are those which follow a λ. Terms M, N which can be converted into one another by renaming bound variables are said to be α-*equivalent*; it is customary to treat α-equivalent terms as identical. In the λ-calculus we can restrict our consideration to single valued functions by the process of *currying* where we associate the functions

$$\lambda(x, y).M \quad \text{and} \quad \lambda x.(\lambda y.M).$$

Anticipating the Curry-Howard correspondence (see section 1.4), currying corresponds to the equivalence between $A \wedge B \to C$ and $A \to B \to C$.

The λ-*calculus* consists of formulas $M = N$ for λ terms M, N with axioms saying that $=$ is an equivalence relation and that abstraction and application respect $=$, together with the axiom for β-*conversion*

$$(\lambda x.M)N = M[x := N],$$

where *substitution* $M[x := N]$ is defined in the natural way and with the requisite care. Replacing $(\lambda x.M)N$ by $M[x := N]$ is called β-reduction; β-reduction captures the idea of function application. Terms to which we cannot apply β-reduction are called *normal*.

In general λ-terms cannot be reduced to normal terms. This defect can be repaired by assigning to each term a type restricting the terms to which it can be applied. In this context application is a partial relation, so we must distinguish cases when application is undefined:

- we write $fa\downarrow$ to indicate that f is defined at a;
- we write $f(a) \simeq g(b)$ for $f(a)\downarrow$ if and only if $g(a)\downarrow$ and $f(a) = g(b)$ if either, and hence both, are defined.

Both the typed and untyped λ-calculus can be seen as a simple programing language, and several common programing languages have features inspired by the λ calculus. In section 1.4 we briefly introduce a foundational system for constructive mathematics, Martin-Löf type theory, which extends the typed λ-calculus and which can be seen as a high level programing language. See [6] for a (very) comprehensive introduction to the λ-calculus.

We introduce a second model of computation, a generalisation of Curry's combinatory algebras [43], which is of great importance for the metamathematics of constructive theories. A *partial combinatory algebra* (pca) is a set $\mathcal{A}$ together with a binary operation $\cdot$, called *application*,[13] and distinguished elements $\mathbf{s}, \mathbf{k}$ such that for all $a, b, c \in \mathcal{A}$

[13] We write ab for $a \cdot b$.

(pca1) $\mathbf{k}ab \simeq a$;
(pca2) $\mathbf{s}ab\downarrow$;
(pca3) $\mathbf{s}abc \simeq ac(bc)$.

A pca is *nontrivial* if $\mathbf{s} \neq \mathbf{k}$, or equivalently if it has two distinct elements. *Terms* are defined inductively from a countable set of variables and the elements of $\mathcal{A}$ by application

if a, b are terms, then (ab) is a term.

The terms of a pca are called *combinators*.

The definition of pca's is motivated by the following result, which asserts the λ-*completeness* of the combinators $\mathbf{s}, \mathbf{k}$.

Theorem 5 *Every λ term is extensionally equivalent to some combination of* $\mathbf{s}, \mathbf{k}$.

The proof is by induction on the construction of λ terms and is a good exercise in understanding the combinators. For example, the identity $i = \lambda x.x$ is given by $\mathbf{skk}$:

$$(\mathbf{skk})(x) \;=\; \mathbf{k}x(\mathbf{k}x) \;=\; x.$$

A more interesting example is the construction of a fixed point combinator; $\mathbf{ssk}(\mathbf{s}(\mathbf{k}(\mathbf{ss}(\mathbf{s}(\mathbf{ssk}))))\mathbf{k})$ can be verified as a fixed point combinator by an unsightly computation. The pairing and projection functions

$$\mathbf{p} = \lambda x, y(\lambda z.zxy)$$
$$\pi_0 = \lambda x.x\mathbf{k} \;\text{ and }\; \pi_1 = \lambda x.x(\mathbf{k}i)$$

are of particular importance.

The λ-calculus can be viewed as a pca with

$$\begin{aligned}\mathbf{k} &= \lambda xy.x,\\ \mathbf{s} &= \lambda xyz.xz(yz).\end{aligned}$$

Equivalently, the set of Turing machines can be made into a pca, but to define application we first need some notation. We fix an enumeration $(\varphi_n)_{n\in\mathbf{N}}$ of all Turing machines, and we denote by $\{e\}n$ the result of running the e^{th} Turing machine on input n. Then the set of Turing machines can be represented as a pca with elements natural numbers and application given by $mn = \{m\}n$.

1.2 Constructive set theory

In his seminal monograph [16], Bishop laid down a naive constructive set theory and proceeded to develop (constructive) mathematics in the informal rigorous style of the Bourbaki school. Like the classical mathematician, he was sure that his work could be done in a formal set theory, but was not concerned with it. This attitude is still dominant in Bishop's followers, but logicians, both mathematical and philosophical, have stepped forward to provide rigorous foundations for constructive mathematics, and the metamathematics rigorous foundations allow.

We first give a brief description of Bishop's naive set theory, before we discuss the formal set theoretical foundations for constructive mathematics.

1.2.1 Naive constructive set theory

The mathematics in this thesis is primarily in the informal style of Bishop's constructive mathematics. In 'Foundations of constructive analysis', Erret Bishop set down a naive set theory and proceeded to develop a large body of core mathematics intuitionistically in this informal setting. We briefly sketch Bishop's notion of set.

A set is a well behaved collection of objects. In order for a collection A to form a set we must give

(i) a description of how to construct the elements of A, this construction relying only on sets we have already—at least in theory—constructed prior to A;

(ii) a description of what it means for two elements of A to be equal.

A few remarks. Firstly, Bishop's universe of sets is *predicative*, the construction of a set depends only on previously constructed sets. Secondly, a set is not given only by its members: for Bishop an equivalence relation, normally called equality, satisfying the equality axioms is a necessary part of the description of a set. Of course we must have at least one set to begin the set construction process; reflecting Kronecker—"God made the integers, all else is the work of man"—Bishop takes as his starting point the set of natural numbers.

Let us consider the collection of subsets of a (previously constructed) set X, the *powerclass* of X. How does one construct a subset S of X? The simplest way is to take the elements of X in turn and to indicate whether or not they are in S—this gives us the collection of all decidable subsets of X, or equivalently the collection of all functions from X to $2 = \{0, 1\}$. Defining equality extensionally ($f = g$ if $f(x) = g(x)$ for every $x \in X$), we conclude that the collection of functions from the set X to the set $\{0, 1\} = \{n \in \mathbf{N} : n = 0 \vee n = 1\}$ (where equality on $\{0, 1\}$ is the restriction of equality on $\mathbf{N}$) is a set. This argument justifies the exponentiation axiom: for all sets a, b the collection b^a of functions from a to b is a set. In contrast, since not all sets are decidable—for example $\{n \in \mathbf{N} : n = 0 \wedge P\}$ for some undecided proposition P—it is inconceivable that we can fully describe the general construction of a subset of X. So we cannot constructively justify the powerset axiom: the powerclass of a set is a set. Indeed, the subsets of $\{1\}$ of the form

$$\{n \in \mathbf{N} : n = 0 \wedge P\}$$

for some appropriate P—for example if P contains a quantification over all sets—will depend on sets yet to be constructed; whence accepting the powerset axiom leads to impredicativity.

For Bishop, the notion of function and set were distinct, both sets and functions were primitive; functions were not just particular types of sets. A function f from X to Y is an algorithm which applied to an element x of X outputs an element $f(x)$ of Y and which respects equality; that is, f is *extensional*: if $x =_X x'$, then $f(x) =_Y f(x')$. Since our construction of sets is 'algorithmic', it is reasonable to formalise functions in the standard set theoretic way, although **MLTT**, introduced in Section 1.4, is perhaps more in tune with Bishop's views in this respect.

1.2.2 Set theoretical foundations for constructive mathematics

Foundational systems for constructive mathematics have taken two (distinct, but overlapping) paths:

1. the type theoretic path, its philosophical motivations culminating in Martin Löf type theory; and

2. the set theoretic path motivated by the success of set theory as a foundation for classical mathematics.

While **MLTT** seems to be the most philosophically acceptable foundational framework for constructive mathematics,[14] we will concern ourselves with constructive set theories, which offer the advantage that tools for mathematics and metamathematics are already well developed for set theory; we give a brief discussion of the motivations and structure of **MLTT** in section 1.4. Constructive set theories also make no explicit mention of the notions of constructive object and construction, and may thus be more appealing to those of us with no philosophical commitment to constructivism. Having settled on having a set theoretical foundation for constructive mathematics, there are two natural ways to approach the design of a constructive set theory:

(i) we could attempt to formalise the informal set theory adopted by the Bishop school—this approach, which seems to be favoured by practising constructivists, was taken by Myhill in his seminal paper [88];

(ii) the second approach is to start with classical **ZF**(**C**) set theory and make it intuitionistic—this approach leads to Intuitionistic **ZF** set theory **IZF** and Aczel's Constructive **ZF** set theory **CZF**.

Again, we favour the second approach, because it has the advantages of 2 above, and in the case of **CZF** can be justified by the solid constructive foundation of **MLTT**.

[14] Constructive mathematics has been developed directly in **MLTT**, primarily as formal topology [95].

We make use of the following standard notation. We use $\forall_{x\in a}\varphi x$ as shorthand for $\forall_x(x \in a \to \varphi x)$, and $\exists_{x\in a}\varphi$ for $\exists_x(x \in a \wedge \varphi x)$. Formulas containing only quantifiers of the form $\forall_{x\in a}$ or $\exists_{x\in a}$ are called *bounded formulas*: the range of the quantifiers is bounded to the set a.

The three constructive set theories that we will adopt for our mathematics and metamathematics are all subsystems of classical **ZFC** set theory, which consists of first order classical logic with two nonlogical binary relation symbols $\in, =$ together with seven axioms

Extensionality: If sets a and b have the same elements, then $a = b$,

$$\forall_{a,b}\left(\forall_x\left(x \in a \leftrightarrow x \in b\right) \to a = b\right).$$

Pairing: For any sets a and b there exists a set $\{a, b\}$ which contains precisely a and b,

$$\forall_{a,b}\exists_c\left(\forall_x\left(x \in c \leftrightarrow x = a \vee x = b\right)\right).$$

Union: For any set X there exists a set $\bigcup X$ such that if $x \in a \in X$, then $x \in \bigcup X$,

$$\forall_a\exists_c\left(\forall_x\left(x \in c \leftrightarrow \exists_y \in a(x \in y)\right)\right).$$

Power set: Far any set X there exists a set $\mathcal{P}(X)$ consisting of all subsets of X,

$$\forall_a\exists_c\left(\forall_x\left(x \in c \leftrightarrow x \subset a\right)\right).$$

Infinity: There exists an infinite set,

$$\exists_a\left(\left(\exists_x x \in a\right) \wedge \left(\forall_{x\in a}\exists_{y\in a} x \in y\right)\right).$$

Foundation: For every non-empty set a there exists $x \in a$ which is minimal (such that $x \cap a = \emptyset$),

$$\forall_a \left(\exists_x (x \in a) \to \exists_{x \in a} \forall_{y \in a} (y \notin x)\right).$$

Choice: For any set a and any function F with domain a, if for all $x \in a$ there exists $y \in F(a)$, then there exists a function $f : a \to \cup\{F(x) : x \in a\}$ such that $f(x) \in F(x)$ for each $x \in a$; f is called a choice function.

and two axiom schema

Separation: For any set a and any formula φ, the set $\{x \in a : \varphi(x)\}$—that contains all $x \in a$ satisfying φ—exists,

$$\forall_a \exists_c \forall_x \left(x \in c \leftrightarrow x \in a \wedge \varphi(x)\right),$$

for all formula $\varphi(x)$ with $c \notin \mathrm{FV}(\varphi)$.

Replacement: If F is a class function,[15] then for any set a the image of a under F is also a set,

$$\forall_{x \in a} \exists !_y \varphi(x, y) \to \exists_c \forall_y \left(y \in c \leftrightarrow \exists_{x \in a} \varphi(x, y)\right),$$

for all formula $\varphi(x, y)$ with $y \notin \mathrm{FV}(\varphi)$.

In pursuit of a constructive foundation, our first step is to replace the classical logic of **ZFC** with intuitionistic logic. However, since the axioms of **ZFC** were formulated in the classical environment, we must check that none of these axioms reintroduce unacceptable fragments of **LEM**. Myhill showed that the foundation axiom and the axiom of choice, the latter was together

[15] A formula Ψ is a *class function* if it is *functional*:

$$\forall_x (\varphi(x) \to \exists !_y \Psi(x, y))$$

for some formula φ.

with Goodman [53] and was based on an argument of Diaconescu from category theory [48], both imply the law of excluded middle, and consequently must be rejected in their full generality.

Theorem 6 *Assume that Separation, extensionality, collection, and pairing hold. Then the axiom of foundation and the axiom of choice each imply* ***LEM****.*

Proof. Foundation: Fix a formula P and consider the set

$$S = \{x \in \mathbf{N} : x = 0 \wedge P\} \cup \{1\}.$$

By foundation there exists an element u of S such that $z \notin u$ for all $z \in S$. Either $u = 0$ and P holds, or $u = 1$ and $\neg P$ holds.

Choice: Fix a formula P, let X be the set $\{0, 1\}$ with equality[16] $=_X$ satisfying

$$0 =_X 1 \Leftrightarrow P,$$

and let $S = \{(0, 0), (1, 1)\}$. Then S is a subset of $X \times \{0, 1\}$, where $\{0, 1\}$ is given the standard equality $=$, and for all $x \in X$ there exists $y \in \{0, 1\}$ such that $(x, y) \in S$. Applying the axiom of choice we 'construct' a pair $(f(0), f(1)) \in \{0, 1\}$ such that $(0, f(0)), (1, f(1)) \in S$ and if $0 =_X 1$, then $f(0) = f(1)$. If $f(0) = 0$ and $f(1) = 1$, then $\neg(f(0) = f(1))$, so $\neg(0 =_X 1)$ and hence $\neg P$. Otherwise either $f(0) = 1$, so $(0, 1) = (0, f(0)) \in S$, or $f(1) = 0$ and $(1, 0) = (1, f(1)) \in S$. In both cases we have that $0 =_X 1$ and therefore P. ■

The classical nature of **AC** rests on the extensionality of the choice function: it is through this extensionality that we are able to decide what S looks like, and hence whether or not P holds. To illustrate this, we show the constructive inadmissibility of $\mathbf{AC}_{\mathbf{R},2}$:[17]

[16] If one insists on the use of set equality, then we can define X to be $\{a = \{x : x = 0 \wedge P\}, 1\}$ with set equality. For then $a = 1$ if and only if P holds.

[17] The proof is a small variation of a Brouwerian example extracted by Martin Escardo from [5].

For any binary predicate φ on $\mathbf{R} \times \{0, 1\}$, if for each real number x there exists $i \in \{0, 1\}$ such that $\varphi(x, i)$, then there exists a function $f : \mathbf{R} \to \{0, 1\}$ such that $\varphi(x, f(x))$ for all $x \in \mathbf{R}$.

Proposition 7 ***AC***$_{\mathbf{R},2}$ *implies* ***WLPO****.*

Proof. For each $x \in \mathbf{R}$ either $x > 0$ or $x < 1$. Using $\mathbf{AC}_{\mathbf{R},2}$ we construct a function $f : \mathbf{R} \to \{0, 1\}$ such that

$$\begin{aligned} f(x) = 0 \quad &\Rightarrow \quad x < 1, \\ f(x) = 1 \quad &\Rightarrow \quad x > 0. \end{aligned}$$

Then $f(0) = 0$ and $f(1) = 1$, so f is nonconstant. Using an interval halving argument, we can construct $x \in [0, 1]$ such that f is discontinuous at x; without loss of generality, $f(x) = 0$. Given an increasing binary sequence $\alpha \in 2^{\mathbf{N}}$ we inductively construct a sequence $(y_n)_{n\in\mathbf{N}}$ as follows. If $\alpha(n) = 0$, then $y_n = x$ and if $\alpha(n-1) = 1$, then $y_n = y_{n-1}$. If $\alpha(n) = 1 - \alpha(n-1)$, then we let $y_n \in (x - 2^{-n}, x + 2^{-2})$ be such that $f(y_n) = 1$. Then $(y_n)_{n\in\mathbf{N}}$ is Cauchy and hence converges to a limit $y \in \mathbf{R}$. If $f(y) = 0$, then $\alpha(n) = 0$ for all n, and if $f(y) = 1$ then it is not the case that $\alpha(n) = 0$ for each n. ∎

Bishop gave the following argument for accepting some choice even when working constructively: if we have a proof of $\forall_{x\in X}\exists_{y\in Y}\varphi(x, y)$, then, by the **BHK** interpretation, we have an algorithm which given $x \in X$ computes $y \in Y$ such that $\varphi(x, y)$; this algorithm is precisely the function the axiom of choice asserts to exist. This apparent contradiction—**AC** implies **LEM**, but choice can be validated under the **BHK** interpretation—is reconciled by the fact that Bishop's algorithm may fail to be extensional; for example the algorithm present in the latter half of the previous theorem. But Bishop's argument will validate choice when intensional and extensional equality coincide on the choice domain, so the constructive mathematician may happily adopt the axiom of countable choice

CC: If a is an inhabited set and S a subset of $\mathbf{N} \times Y$ such that for each positive integer n there exists $x \in a$ such that $(n, x) \in S$, then there is a function $f : \mathbf{N} \to a$ for which $(n, f(n)) \in S$ for each $n \in \mathbf{N}$.

This argument can be made rigorous in Martin-Löf type theory [81] and can be extended to justify the stronger principle of dependent choice

DC: If a is a set, S is a subset of $a \times a$; and for each $x_0 \in a$ there exists $y \in a$ such that $(x, y) \in S$, then there exists a sequence $(x_n)_{n \in \mathbf{N}}$ in X such that $(x_n, x_{n+1}) \in S$ for each positive integer n.

The foundation axiom is replaced by the classically equivalent, but constructively weaker, form in which it is most often used:

Set induction: A property that holds for a set a whenever it holds for each element of a, holds for all sets,

$$\forall_a \left(\forall_{x \in a} \varphi(x) \to \varphi(a)\right) \to \forall_a \varphi(a)$$

for any formula $\varphi(x)$.

In contrast, in place of the replacement scheme we use collection, which is a theorem of **ZF**, but which cannot be proved from the other axioms of our first constructive set theory.

Collection: Every total class relation with domain a set can be refined to a total class relation with set-sized codomain,

$$\forall_{x \in a} \exists_y \varphi(x, y) \to \exists_b \forall_{x \in a} \exists_{y \in b} \varphi(x, y).$$

This leads us to the system of intuitionistic Zermelo-Fraenkel set theory, which consists of *intuitionistic* logic with two binary relation symbols $\in, =$ together with the axioms and axiom schema of

IZF: Extensionality, Pairing, Union, Powerset, Infinity, *Set induction*, Separation, and *Collection*.

The double negation interpretation of classical logic into intuitionistic logic was extended by Friedman to give an interpretation of **ZF** in **IZF**. It follows that **ZF** and **IZF** are equiconsistant and therefore have the same proof theoretic strength. **IZF** holds in any Grothendieck or realizability topos, and any topos which has small limits.

There is another aspect to the constructivist philosophy which we have thus far ignored: a constructive universe should be built from the bottom up, rather than existing as a completed whole like the classical **ZF** universe. A definition of an object A is *impredicative* if it is given by a quantification over a collection of objects of which A is a member; so impredicative definitions make tacit appeal to a completed static mathematical universe. The rejection of impredicativity is clear in the intuitionistic philosophy put forth by Brouwer, where the universe is dynamic, but is generally less acknowledged, and not considered important, by the Bishop school [91], although it is clearly part of the informal set theory put forth by Bishop. The neglect of predicativity issues by Bishop's followers is likely due to the fact that impredicative definitions are generally easy to avoid in the (informal) practice of constructive (or classical) mathematics. Some consideration will convince the reader that the sources of impredicativity in **IZF** are the powerset axiom, and the schema of separation.

The impredicative nature of separation results from its application to formulas which may be defined only in terms of objects not yet constructed. This is circumvented by restricting separation to bounded formula:

> **Bounded separation** For any set a and any bounded formula φ, the set $\{x \in a : \varphi(x)\}$—that contains all $x \in a$ satisfying φ—exists,
>
> $$\forall_a \exists_c \forall_x \left(x \in c \leftrightarrow x \in a \wedge \varphi(x) \right),$$
>
> for all bounded formula $\varphi(x)$ with $c \notin \mathrm{FV}(\varphi)$.

The impredicative nature of the powerset axiom is a little more subtle, but is essentially the same: in the constructive context the power class of 1 contains far more than just $0, 1$; it contains sets of the form $\{x : x = 0 \wedge P\}$ for any formula P, in particular, formulas which may not be definable over the sets which we have already constructed. The fundamental issue at hand is that constructively membership of a set is generally not *decidable*. If, in powerset, we restrict ourselves to subsets which are decidable, then we get the axiom of exponentiation with the domain $Y = \{0, 1\}$.

Before giving the definition of exponentiation we need to code functions as sets. A *relation* is a triple (R, a, b) with $R \subset a \times b$; the *domain* a and *codomain*, or *range*, b of R are normally suppressed. We often write xRy for $(x, y) \in R$. If for all $x \in a$ there exists a $y \in b$ such that xRy, then R is said to be *total*. A *function* f is a relation such that if $(x, y), (x, z) \in f$, then $y = z$; we write $f(x) = y$ for $(x, y) \in f$ and we write $\mathrm{Fun}(f, X, Y)$, or $f : X \to Y$, to express that f is a function with domain X and codomain Y. The *image* of a relation R is the set

$$\mathrm{image}(R) = \{y \in b : \exists_{x \in a}(x, y) \in R\}.$$

If (R, a, b) is a total relation, any subset S of b for which $R \cap (a \times S)$ is total is called a *partial image* of R. The *image of a function* f is its unique partial image, and for any subset S of X we denote by $f(S)$ the image of f restricted to S.

> **Exponentiation** If a and b are sets, then the collection b^a of functions from a to b is a set,
>
> $$\forall_{a,b}\exists_c(f \in c \leftrightarrow \mathrm{Fun}(f, a, b)).$$

Exponentiation is classically equivalent to the powerset axiom, since every set is decidable, so each subset can be associated with its characteristic function.

We can now define *constructive set theory* (**CST**), which is closely related to Myhill's original formalisation of Bishop's constructive mathematics.[18] **CST** consists of the axioms

> **CST**: Extensionality, Pairing, Union, Exponentiation, Infinity, Set induction, *Bounded* separation, Replacement, and *dependent choice*.

Aczel's Constructive **ZF** set theory has more complicated origins. In the early 1970's Per Martin Löf introduced his (intentional) intuitionistic type theory, a constructive alternative to set theory based on the Curry-Howard isomorphism.[19] **MLTT** is generally accepted as the best foundational framework to make the ideas of constructive mathematics precise. In a series of three papers [1, 2, 3], Peter Aczel showed how a particular type of **MLTT** can be interpreted as a universe of sets in such a way that the Curry-Howard isomorphism validates the axioms of constructive set theory. The constructive set theory so interpreted—Aczel's **CZF**—is very roughly akin to intersecting **MLTT** with **ZF**;[20] Aczel chose not to include any choice axiom in **CZF** to ensure a wealth of interesting models, in particular any topos.

As a result of these complicated origins, some aspects of **CZF**, in particular the subset collection axiom, may look a little unnatural. Before we describe **CZF** we need a few more definitions. A set a is said to be *inductive*, written Ind(a), if $\emptyset \in a$ and $x \cup \{x\}$, the *successor* of x, is in a whenever $x \in a$. Classically, the axiom of infinity is required to define infinite sets, in particular the natural numbers. The first new aspect of **CZF** is the axiom of strong infinity which explicitly defines the natural numbers.

[18]Myhill's constructive set theory [88] was three sorted, having sorts for numbers, partial functions, and sets.

[19]The Curry-Howard isomorphism identifies a proposition with the type of its proofs, effectively relating programs with proofs. It is a generalisation of the correspondence between natural deduction and Church's lambda calculus given by the **BHK** interpretation with 'algorithm' interpreted as 'object of the lambda calculus.'

[20]There are principles which hold in both **MLTT** and **ZF**, for example countable choice on the natural numbers.

Strong Infinity: There is minimal inductive set,

$$\exists_a(\mathrm{Ind}(a) \wedge \forall_b(\mathrm{Ind}(b) \rightarrow a \subset b)).$$

We also add a strong form of collection, also a theorem of **ZFC**, which, in addition to collection, asserts that the set b is a subset of the image of (R, a, b); that is, that b is a partial image of R. Roughly this says that b is slim or almost optimal.

Strong Collection:

$$\forall_{x\in a}\exists_y\varphi(x,y) \rightarrow \exists_b(\forall_{x\in a}\exists_{y\in b}\varphi(x,y) \wedge \forall_{y\in b}\exists_{x\in a}\varphi(x,y)).$$

The final, and most artificial, piece of the **CZF** is the subset collection axiom, which is a predicative powerset-like axiom strictly between powerset and exponentiation, over the other axioms of **CZF** [76].

Subset collection For sets a, b,

$$\exists_c\forall_u(\forall_{x\in a}\exists_{y\in b}\psi(x,y,u) \rightarrow \exists_{d\in c}(\forall_{x\in a}\exists_{y\in b}\psi(x,y,u)\wedge\forall_{y\in b}\exists_{x\in a}\psi(x,y,u))).$$

The formula ψ in the statement of subset collection represents a class of total relations, with domain a and codomain b, indexed by some collection of sets (namely, the class $C(a,b,\psi) = \{u : \forall_{x\in a}\exists_{y\in b}\psi(x,y,u)\}$). Subset collection states that for all sets a, b and formula ψ there exists a set c such that for each $u \in C(a,b,\psi)$, there is a partial image of u in c. Since images of functions are unique, fixing sets a, b and taking ψ to be $\mathrm{Fun}(u,a,b)$, we recover the exponentiation scheme for a, b.

We now define Aczel's Constructive **ZF** set theory [4], which is the most popular of the constructive set theories. **CZF** consists of the axioms

CZF: Extensionality, Pairing, Union, *Subset collection*, *Strong infinity*, Set induction, Bounded separation, and *Strong collection.*

CST minus the axiom of dependent choice is a subsystem of **CZF**. Since **CZF** can be interpreted into Martin Löf type theory, it is predicative, and hence, proof theoretically, much weaker than **IZF**. **CZF** also holds in any Π-pretopos, and hence any topos.

Now that we have a constructive set theory which is predicative, possibly foundation and the axiom of choice no longer imply the law of excluded middle? This is the case, but the arguments of Theorem 6 still show that Foundation and choice imply $P \vee \neg P$ for any bounded formula P, and are therefore still unacceptable.

Ordinals

Ordinals are not very well behaved in the constructive setting. In particular they cannot be linearly ordered, and even defining ordinals in **CZF** takes a little care. Since they still provide a ranking of the universe, which allows us define classes by transfinite recursion on the ordinals, and prove properties using transfinite induction on the ordinals, it is worth the effort. But we must be careful to avoid making the nonconstructive case distinctions common in classical set theory.

An *ordinal* is a transitive set of transitive sets; we denote the collection of ordinals by **ORD**. Any element of an ordinal is also an ordinal, and the statement $\mathrm{Ord}(\alpha)$, that α is an ordinal, is Δ_0. While the ordinals are not as courteous in the constructive setting as we have come to expect from them, they are still closed under the successor map, $a \mapsto a \cup \{a\}$, and unions, and most importantly they still allow for definitions by recursion.

Proposition 8 *For any class V and any total class function G with domain $V^n \times \mathbf{ORD} \times V$, there exists a class function $F : V^n \times \mathbf{ORD} \to V$ such that*

$$F(\mathbf{x}, \alpha) = G(\mathbf{x}, \alpha, \{< z, F(\mathbf{x}, z) >\}),$$

for each $\mathbf{x} \in V^n$ and each ordinal α.

Proof. Let Ψ be the predicate given by $\Psi(\mathbf{x}, f, \alpha)$ if and only if f is a function whose domain is an ordinal containing α and $f(\beta) = G(\mathbf{x}, \beta, f \restriction \beta)$ for all $\alpha \in \mathrm{dom}(f)$. We define F by

$$F(\mathbf{x}, \alpha) = y \ \leftrightarrow \ \exists_f (\Psi(\mathbf{x}, f, \alpha) \wedge f(\alpha) = y).$$

In order to complete the proof we must show that F is a total class function with domain $V^n \times \mathbf{ORD}$, and that if $\Psi(\mathbf{x}, \alpha, f)$ and $\beta \in \mathrm{dom}(f)$, then $f(\beta) = G(\mathbf{x}, \beta, f \restriction \beta)$.

We proceed by set induction. Fix $\mathbf{x}, \alpha$ and suppose that $\forall_{\beta \in \alpha} \exists!_f \Psi(\mathbf{x}, f, \beta)$. By replacement there exists a set A such that

$$\forall_{\beta \in \alpha} \exists_{f \in A} \Psi(\mathbf{x}, f, \beta) \quad \text{and} \quad \forall_{f \in A} \exists_{\beta \in \alpha} \Psi(\mathbf{x}, f, \beta).$$

Define a relation f_0 with ordinal domain by

$$f_0 = \bigcup A,$$

and set

$$f = f_0 \cup \{ < \alpha, G(\mathbf{x}, \alpha, \{ < \beta, f_0(\beta) >: \beta \in \alpha \} \}.$$

Then f is a function with domain α: for if $g, g' \in A$ and $\beta \in \mathrm{dom}(g) \cap \mathrm{dom}(g')$, then

$$g(\beta) = G(\mathbf{x}, \beta, g \restriction \beta) = G(\mathbf{x}, \beta, g' \restriction \beta) = g'(\beta),$$

where the middle equality is by a straightforward set induction. This completes the proof that F is a total class function. That $f(\beta) = G(\mathbf{x}, \beta, f \restriction \beta)$ for all $\beta \in \mathrm{dom}(f)$ follows from the construction. ■

Real numbers

Classically, the real numbers are fully characterised (up to isomorphism) as a complete totally ordered field. In particular, the constructions of Dedekind and Cauchy coincide. In the setting of constructive mathematics without choice (either **CZF** or **IZF**) the situation is more complicated (and more interesting). The standard constructions of the natural numbers, integers, and rationals are valid in the constructive setting.

In [16] Bishop presents the reals as Cauchy sequences with a fixed modulus of convergence: a sequence $(x_n)_{n\in\mathbf{N}}$ is a *regular Cauchy sequence* if

$$|x_n - x_m| \leqslant \frac{1}{n} + \frac{1}{m}$$

for all positive integers n, m. The collection of regular Cauchy sequences of rational numbers form a set by exponentiation and bounded separation. For regular Cauchy sequences of rationals $x = (x_n)_{n\in\mathbf{N}}, y = (y_n)_{n\in\mathbf{N}}$ we define

- an equivalence relation, equality, by $x = y$ if for all $n \in \mathbf{N}$

$$|x_n - y_n| \leqslant 2/n;$$

- a partial order $<$, by $x < y$ if there exists n such that $x_n + 1/n < y_n - 1/n$.

The *Cauchy reals* $\mathbf{R}^C$ is the set of equivalence classes of regular Cauchy sequences under equality. However, in the absence of choice, it is more natural to work with regular Cauchy sequences of rationals with the equivalence relation $=$ rather than equivalence classes. In an abuse of notation we also call this latter set the Cauchy reals.

The algebraic operations are essentially defined componentwise, but we must pass to a subsequence of the componentwise sequence in order to maintain regularity:

$$\begin{aligned} x + y &= (x_{2n} + y_{2n})_{n\in\mathbf{N}}; \\ -x &= (-x_n)_{n\in\mathbf{N}}; \\ xy &= (x_{2Mn}y_{2Mn})_{n\in\mathbf{N}}, \text{ where } M = \max\{x_0, y_0\} + 2. \end{aligned}$$

If $x > 0$, then there exists n such that $x_n - 1/n > 0$, let n_0 be the least such n and let $k = x_{n_0} - 1/n_0$. Then

$$x^{-1} \;=\; (1/x_{Dn+n_0})_{n\in\mathbf{N}},$$

where D is the least integer greater than $1/k^2$. Associating each rational q with the constant sequence $\mathbf{q} = (q)_{n\in\mathbf{N}}$ gives an embedding of the field $(\mathbf{Q}, 0, 1, +, \cdot)$ into the field $(\mathbf{R}_C, 0, 1, +, \cdot)$.

The other common construction of the reals is by Dedekind cuts. A *Dedekind cut* is an inhabited set X of rational numbers such that

1. X is bounded above: there exists $r \in \mathbf{Q}$ such that $r \geqslant q$ for all $q \in L$;
2. X is downward closed: if $q \in L$ and $q' < q$, then $q' \in L$;
3. X is located: for any rationals r, r', if $r < r'$, then either $r \in L$ or $r' \notin L$.

The collection of Dedekind cuts, the *Dedekind reals* $\mathbf{R}^D$, forms a set in **CZF** [4], but is not a set in **CZF** with subset collection replaced by exponentiation [78].

Let X, Y be Dedekind reals. Addition and additive inverse on the Dedekind reals are given by

$$X + Y = \{a + b : a \in X, b \in Y\} \quad\text{ and }\quad -X = \{-q : q \in \mathbf{Q} \setminus X\}.$$

For positive Dedekind reals X, Y we have

$$XY = \{ab : a \in X, b \in Y\}^{\leqslant} \quad\text{ and }\quad X^{-1} = \{q^{-1} : q \in \mathbf{Q} \setminus X\}^{\leqslant},$$

where $S^{\leqslant}$ is the *downward closure* $\{q \in \mathbf{Q} : \exists_{s\in S} q < s\}$ of S. Multiplication is extended to all Dedekind reals by writing[21]

$$X = \max\{1, X + 1\} - \max\{1, -X + 1\}$$

[21] The operation $(x, y) \mapsto \max\{x, y\}$ is given by union.

and using the identity $(a-b)(c-d) = ac-(bc+ad)+bd$, and multiplicative inverse for negative reals is given by $X^{-1} = -(-X)^{-1}$.

To any regular Cauchy sequence x of rationals we can associate the Dedekind real

$$X_x = \{q \in \mathbf{Q} : \exists_n q < x_n + 2/n\}.$$

As a consequence in any context in which the Dedekind and Cauchy reals are not isomorphic, like **CZF** and **IZF**, results about Dedekind reals are stronger than the corresponding results about Cauchy reals. With $\mathbf{AC}_{\omega,2}$ we can show that $x \mapsto X_x$ is invertible: we can associate to any Dedekind real X (the equivalence class of) a regular Cauchy sequence x_X. Using $\mathbf{AC}_{\omega,2}$ we construct a function $f : \mathbf{Q} \times \mathbf{Q} \to \{0,1\}$ such that for all $r, q \in \mathbf{Q}$

$$\begin{aligned} f(q,r) = 0 &\quad\Rightarrow\quad q \in X; \\ f(q,r) = 1 &\quad\Rightarrow\quad r \notin X. \end{aligned}$$

For $n \in \mathbf{N}$ let $x_n = k/n$ where k is the smallest integer such that

$$f\left(\frac{k-1}{n}, \frac{k}{n}\right) = 0 \text{ and } f\left(\frac{k}{n}, \frac{k+1}{n}\right) = 0.$$

Then $x_X = (x_n)_{n\in\mathbf{N}}$ is a regular Cauchy sequence. The mapping $x \mapsto L_x$ is an embedding of the Cauchy reals into the Dedekind reals which preserves the field structure; and with $\mathbf{AC}_{\omega,2}$ this map induces an isomorphism between the Cauchy and Dedekind reals. It is shown in [39] that $\mathbf{AC}_{\omega,2}$ is not required to prove that the Dedekind and Cauchy reals are isomorphic. However, for any intuitionistic theory T without some form of countable choice, there exists a model of T in which

> There exists a Dedekind real X such that $X \neq x$ for each Cauchy real x.

It is also not possible, in the absence of choice, to show that the Cauchy real numbers are Cauchy complete—that is, to show that any Cauchy sequence of regular Cauchy sequences of rational numbers converges to a Cauchy real. In the next section we give a model of **IZF** in which (i) there are Dedekind reals with no associated Cauchy real, (ii) the Dedekind reals are not countable (in a constructively strong sense), and (iii) there is a Cauchy sequence of Dedekind reals which does not converge to any Dedekind real. See [77] for a more complete treatment of these issues.

What we adopt

It now befits us to set forth the foundational system(s) we will adhere to. For the work without choice, particularly that of section 3.1, we have in mind **CZF**, since we feel the need for the Dedekind reals to form a set. For any results which require choice, under which the Dedekind and Cauchy reals coincide, we have in mind **CST**. The few simple independence results we present will be over **IZF**, possibly with other assumptions of the background universe (see section 1.3 for details). We have made an attempt to flag any use we make of choice principles beyond the axiom of unique choice.

1.3 Models of IZF

This section gives a gentle introduction to models of **IZF**, in particular to topological models. The take home message is: the move from classical logic to intuitionistic logic allows much greater freedom in the choice of forcing semantics. Throughout this section we use $\to$ for internal implication in a model, and $\Rightarrow$ for implication at the metalevel.

1.3.1 Kripke models

We first introduce a Kripke semantics for intuitionistic logic. Kripke models have their root in the possible world interpretation of modal logic. In possible world semantics we are given a lattice $\mathbf{P}$ of worlds and, for $p, q \in \mathbf{P}$, we interpret $p \leqslant q$ as 'p is visible from q'. To each world we associate a classical model of propositional or first order logic, with relation and function symbols. The modal operators are then interpreted as

- $\Box\varphi$ (normally read, necessarily φ) holds at world p if φ holds at all worlds visible from p;
- $\Diamond\varphi$ (normally read, possibly φ) holds at world p if there exists some world visible from p in which φ holds.

Long before Kripke had proved completeness results for the possible world semantics of various modal logics, Gödel gave an interpretation of intuitionistic logic in the modal logic **S4**. This presaged Kripke's possible-world-like semantics for the intuitionistic predicate calculus; subsequently Kripke models were extended to first order logic. We describe Kripke models for **IQC** and a language $\mathcal{L}$ with predicate symbols.

A *Kripke model* for **IQC** with relation symbols is a quadruple $\mathcal{K} = (K, \leqslant, D, \Vdash)$ where

(i) $(K, \leqslant)$ is an inhabited poset;

(ii) the domain function D assigns an inhabited set to each element of K and is *monotone* with respect to $\leqslant$:

$$\forall_{k,k'}(k \leqslant k' \to D(k) \subset D(k'));$$

(iii) $\Vdash$ is a relation, the *forcing relation*, between K and the set of atomic formulas over $\mathcal{L}$ with parameters in $\mathbf{D} = \bigcup\{D(k) : k \in K\}$ such that for all $k, k' \in K$ and any n-ary relation R

$$\begin{array}{lll} k \Vdash R(d_1, \dots, d_n) & \Rightarrow & d_i \in D(k) \text{ for } 1 \leqslant i \leqslant n, \\ k \Vdash R(d_1, \dots, d_n) \wedge k \leqslant k' & \Rightarrow & k' \Vdash R(d_1, \dots, d_n), \\ \neg(k \Vdash \bot). & & \end{array}$$

If $k' \leqslant k$ we say that k' *extends*, or is an *extension* of, k. We extend $\Vdash$ inductively to all formulas in $\mathcal{L}$ with parameters in $\mathbf{D} = \bigcup\{D(k) : k \in K\}$:

K1 $k \Vdash A \wedge B$ if $k \Vdash A$ and $k \Vdash B$;

K2 $k \Vdash A \vee B$ if $k \Vdash A$ or $k \Vdash B$;

K3 $k \Vdash A \to B$ if for all extensions k' of k if $k' \Vdash A$, then $k' \Vdash B$;

K4 $k \Vdash \forall_x \varphi x$ if for every extension k' of k and each $d \in D(k')$, $k' \Vdash \varphi d$;

K5 $k \Vdash \exists_x \varphi x$ if there exists $d \in D(k)$ such that $k \Vdash \varphi d$.

We read $p \Vdash \varphi$ as 'φ is *valid* or *true* at p.' A formula φ is *valid in the Kripke model* $\mathcal{K}$, written $\mathcal{K} \Vdash \varphi$, if it is valid at every node of $\mathcal{K}$. In particular, if K has a root node 0, then φ is valid in $\mathcal{K}$ if and only if it is valid at 0. If φ holds in every Kripke model, we say that φ is *valid for Kripke semantics* and write $\Vdash \varphi$.

Proposition 9 *The forcing relation is monotonic: if $p \Vdash \varphi$ and $q \leqslant p$, then $q \Vdash \varphi$.*

Proof. By induction on the construction of φ. Monotonicity for atomic formula is part of the definition of a Kripke model. If $\varphi = A \wedge B$, then

$k \Vdash A$ and $k \Vdash B$; then for any extension k' of k, $k' \Vdash A$ and $k' \Vdash B$, by the induction hypothesis, so $k' \Vdash A \wedge B$. The case $\varphi = A \vee B$ is similar. Since forcing implications and universal quantifiers are defined in terms of extensions, the case $\varphi = A \to B$ and the case $\varphi = \forall_x A$ are immediate. Finally, the case $\varphi = \exists_x A$ follows from the monotonicity of D. ■

Brouwer's intuitionism gives us the following very helpful interpretation of a Kripke model. We view a node p of our poset $\mathbb{P}$ as a 'state of knowledge' and the nodes q extending p as possible future states of knowledge. With this view

- the monotonicity of D says that once we have constructed an object it always exists;
- the monotonicity of $\Vdash$ says that once φ is proved it is always true;
- $p \Vdash \varphi \to \psi$ if whenever φ is true in some possible future state, then ψ is also true in this state;
- $p \Vdash \forall_x \varphi$ if in every future state φx holds for all objects shown to exist.

The way in which forcing is defined for disjunction and implication is reminiscent of the BHK interpretation of these logical connectives.

As a simple example of unpacking a forcing statement we consider what it means for a node to force $\neg\neg\varphi$ for some formula φ. By definition, $k \Vdash \neg\varphi$ if $k \Vdash \varphi \to \bot$; since no $k \in K$ forces $\bot$, $p \Vdash \neg\varphi$ precisely when $\neg(k' \Vdash \varphi)$ for all $k' \leqslant k$. Therefore $k \Vdash \neg\neg\varphi$ if no extension of k forces $\neg\varphi$, so every extension of k has an extension k' which forces φ. With the notion of a dense subset of a poset we can give a more succinct description of when $k \Vdash \neg\neg\varphi$. A subset S of a poset K is *dense* in K if for every $k \in K$ there exists an extension of k in S; S is *dense below* $k \in K$, if S is dense in $\{k' \in K : k' \leqslant k\}$. If S is dense in K, then S is dense below k for each $k \in K$. We then have that $k \Vdash \neg\neg\varphi$ if and only if

$$\{p \in K : p \Vdash \varphi\}$$

is dense below k.

Kripke semantics is sound and complete for IQC.

Theorem 10 *Kripke semantics is sound for IQC: if* $\vdash \varphi$*, then* $\Vdash \varphi$*.*

Proof. We verify the axioms and rules of the Hilbert style deduction system for **IQC**. Most cases, including the modus ponens and generalisation rules, are straightforward. We only consider the two cases:

> $(A \to C) \to ((B \to C) \to (A \vee B \to C))$:
> We must show that any node of any Kripke model forces the above formula for any choice of A, B, C. Unpacking the implications, we must show that if k forces $A \to C$, $B \to C$, and $A \vee B$, then it forces C: since $k \Vdash A \vee B$, either $k \Vdash A$ or $k \Vdash B$; in the first case $k \Vdash C$ since $k \Vdash A \to C$ and $k \leqslant k$, and the second case is analogous.
>
> $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$:
> Suppose that k forces $A \to (B \to C)$, $A \to B$, and A; we must show that k forces C. Since $k \leqslant k$, we can apply modus ponens twice to show that k forces $B \to C$ and B, and then one final time to see that $k \Vdash C$.

■

The standard proofs of completeness are nonconstructive.

Theorem 11 (ZF) *Kripke semantics is complete for IQC: if* $\Vdash \varphi$*, then* $\vdash \varphi$*.*

Proof. See Theorem 5.11 of [107]. ■

1.3.2 Forcing models

We are interested in developing Kripke models for set theory. We will focus our attention on forcing models; for a different approach see [61] and see [77, 78, 47] for some more involved custom examples (some of which also apply forcing). In what follows our metatheory will be **IZF** unless otherwise stated.

We restrict ourselves to giving a special type of Kripke model for set theory: intuitionistic forcing. Our goal is to define a class model $V(\mathbb{P}) = (\mathbb{P}, \leqslant, D, \Vdash)$ satisfying the axioms of **IZF**; as such we want to have a fixed universe of sets. The temporal aspect of our model will be given by how the elements of the domain D are interpreted at each node. We construct D in a manner similar to the construction of the von Neumann hierarchy, but we want to build in the dependence on our future worlds.

Before defining D, we provide a few motivating examples. Suppose we want to construct a set τ which looks like $0 = \emptyset$ at some nodes, but looks like $1 = \{0\}$ at some node p of $\mathbb{P}$ (and by monotonicity at each node q extending p). Such a set will consist of the element 0 together with a label indicating that $0 \in \tau$ at p:

$$\sigma = \{< 0, p >\};$$

this is a simple example of what is called a name. As a slightly more complicated example we set $\sigma' = \{< 0, q >, < 1, r >\}$, then $0 \in \sigma$ at any node extending q and $1 \in \sigma'$ at any node extending r. So, depending on p, q and $\mathbb{P}$, σ' may appear, in various futures, to look like $0, 1, \{1\}$, or $2 = \{0, 1\}$. We could then use σ, σ' to build a more complicated set τ, which contains σ, whatever this looks like, at some node p' (and, by monotonicity, any extension of p'), and contains σ' at nodes q' and r': this gives the following name

$$\tau = \{< \sigma, p' >, < \sigma', q' >, < \sigma', r' >\}.$$

In general, a *name* τ is a set of pairs $< \sigma, p >$ where σ is a name and p is a node of $\mathbb{P}$: formally we define the class $V(\mathbb{P})$ of names over $\mathbb{P}$ by a transfinite recursion

$$\begin{aligned} V_\alpha(\mathbb{P}) &= \bigcup \{\mathcal{P}(V_\beta(\mathbb{P}) \times \mathbb{P}) \cup V_\beta(\mathbb{P}) : \beta \in \alpha\} \\ V(\mathbb{P}) &= \bigcup_{\alpha \in \mathbf{ORD}} V_\alpha(\mathbb{P}). \end{aligned}$$

Since $V(\mathbb{P})$ is a proper class, it is officially viewed as a predicate

$$x \in V(\mathbb{P}) \Leftrightarrow \exists_\alpha(\mathbf{ORD}(\alpha) \wedge x \in V_\alpha(\mathbb{P})).$$

In several places in the proof that forcing models are sound for **IZF** we do a transfinite induction on the name construction; formally, for a name τ we define the *rank* $\mathrm{rk}(\tau)$ of τ inductively by

$$\mathrm{rk}(\tau) = \bigcup\{\mathrm{rk}(\sigma) + 1 :< \sigma, p >\in \tau\}.$$

In order to discuss forcing statements we must expand our language to include a constant for each element of $V(\mathbb{P})$, but we happily gloss over such technical details.

To complete the definition of a forcing model over $\mathbb{P}$ we must interpret the membership and equality relations. For the simplest names we have already defined the interpretation of $\in$: for $\tau = \{< x_i, p_i >: x_i \in V, p_i \in \mathbb{P}, i \in I\}$

$$p \Vdash x \in \tau \quad \Leftrightarrow \quad \exists_{i \in I}(x = x_i \wedge p \leqslant p_i).$$

For higher levels of the inductive definition, we are guided by the formula

$$u \in v \leftrightarrow \exists_{y \in v}(u = y).$$

Given that v is a set of names, for p to force $u \in v$, we want to produce a name σ and $q \in K$ such that $< \sigma, q >\in v$, $p \leqslant q$ (and hence $p \Vdash \sigma \in v$), and $p \Vdash u = \sigma$. We now want to define the interpretation of $=$; our guide here is the extensionality formula

$$u = v \leftrightarrow \forall_{x \in u}(x \in v) \wedge \forall_{y \in v}(y \in u).$$

So for p to force $u \subset v$ we want that for every $< \sigma, q >\in u$ and each r extending both q (and hence forcing $\sigma \in u$) and p we want $r \Vdash \sigma \in v$. The equality $u = v$ is forced by p, if p forces both $u \subset v$ and $v \subset u$. Formally, we define $p \Vdash u = v$ and $p \Vdash u \in v$ by a simultaneous recursion on the definition of u, v:

$$\begin{aligned}
p \Vdash u \subset v &\Leftrightarrow \forall_{<\sigma,q>\in u}\forall_{r\leqslant p,q}(r \Vdash \sigma \in v);\\
p \Vdash u = v &\Leftrightarrow p \Vdash u \subset v \text{ and } p \Vdash v \subset u;\\
p \Vdash u \in v &\Leftrightarrow \exists_\sigma \exists_{q\geqslant p}(<\sigma, q>\in v \text{ and } p \Vdash u = \sigma).
\end{aligned}$$

The next lemma, at least partially, confirms that $p \Vdash \sigma \in \tau$ behaves as we had hoped.

Lemma 12 *Let τ be a name. If $<\sigma, q>\in \tau$, then $q \Vdash \sigma \in \tau$.*

Proof. Since $q \leqslant q$ the result follows from $p \Vdash \sigma = \sigma$, which is proved by induction on the rank of σ (see Theorem 15). ■

Lemma 13 *For any poset* $\mathbf{P}$*, the forcing model* $V(\mathbb{P})$ *is a Kripke model.*

Proof. Since the domain function is constant, we only need to show that $\Vdash$ is monotonic for $\in$ and $=$ and that no $k \in K$ forces $\perp$, both of which follow readily from the definitions. ■

Since our restriction to a constant domain simplifies some of the details of Kripke semantics we give the simplified forcing relations again.

$p \Vdash A \wedge B$ if $p \Vdash A$ and $p \Vdash B$;

$p \Vdash A \vee B$ if $p \Vdash A$ or $p \Vdash B$;

$p \Vdash A \rightarrow B$ if for all extensions q of p if $q \Vdash A$, then $q \Vdash B$;

$p \Vdash \forall_x \varphi x$ if $p \Vdash \varphi\tau$ for all every name τ;

$p \Vdash \exists_x \varphi x$ if there exists a name τ such that $p \Vdash \varphi\tau$.

We write $V(\mathbb{P})$ for the model $(\mathbb{P}, \leqslant, V(\mathbb{P}), \Vdash)$ and we refer to it as a forcing model. Hereafter we will assume that our poset $\mathbb{P}$ has a maximal element $\mathbf{1}$—such an element can be appended if necessary.

As a simple exercise in producing names, we give a canonical name $\check{x}$ to each set x in the real world; this is done by transfinite induction. For a set x we define

$$\check{x} = \{< \check{y}, \mathbf{1} >: y \in x\}.$$

Clearly for any $p \in \mathbb{P}$, $p \Vdash \check{x} \in \check{y}$ if and only if $x \in y$, so $x \mapsto \check{x}$ can be seen as an embedding of the von-Nuemann hierarchy into the forcing model $V(\mathbb{P})$.

Lemma 14 *If $p \Vdash x = y$ and $p \Vdash \varphi(x)$, then $p \Vdash \varphi(y)$.*

Proof. By induction on the complexity of φ. The base case is given by E3,E4,E5 which are shown to hold in the proof of Theorem 15 below. The only nontrivial induction step is when $\varphi = \psi_1 \to \psi_2$. Supposing $p \Vdash \psi_1(x) \to \psi_2(x)$ and $p \Vdash x = y$, if $r \Vdash \psi_1(y)$ for some $r \leqslant p$, then by the induction hypothesis $r \Vdash \psi_1(x)$. Then $r \Vdash \psi_2(x)$ and, again by the induction hypothesis, $r \Vdash \psi_2(y)$; whence $p \Vdash \psi_1(y) \to \psi_2(y)$. ■

Theorem 15 (IZF) *Forcing is sound for **IZF**: **IZF** $\vdash \varphi \Rightarrow V(\mathbb{P}) \Vdash \varphi$.*

Proof. That forcing models are sound for intuitionistic logic follows from Theorem 10 and Lemma 13.

Equality Axioms:

E1 $\forall_x (x = x)$;

E2 $\forall_{x,y} (x = y \to y = x)$;

E3 $\forall_{x,y,z} (x = y \wedge y = z \to x = z)$;

E4 $\forall_{x,y,z} (x = y \wedge x \in z \to y \in z)$;

E5 $\forall_{x,y,z} (x = y \wedge z \in x \to z \in y)$.

These are all proved by a straightforward induction on the rank of a name, except for E3 and E4 which use a simultaneous induction on rank, and E5 which follows straightforwardly from E4.

E1: Fix a name τ and let $<\sigma, q> \in \tau$. By the induction hypothesis, $\mathbf{1} \Vdash \sigma = \sigma$, so $q \Vdash \sigma \in \tau$. Hence $\forall_{<\sigma,q>\in\tau}\forall_{p\leqslant \mathbf{1},q}(p \Vdash \sigma \in \tau)$; that is $\mathbf{1} \Vdash \tau \subset \tau$, from which the result follows.

E2: Follows directly from the symmetry of the definition.

E3 and **E4**: To show E3, it suffices to show that if $x \subset y$ and $y \subset z$, then $x \subset z$. Suppose first that $p \Vdash x \subset y \wedge y \subset z$ and that E4 holds for all names in $\bigcup\{V_\beta : \beta \leqslant \mathrm{rk}(z)\}$. Then

$$\forall_{<\sigma,q>\in x}\forall_{r\leqslant p,q}(r \Vdash \sigma \in y) \ \text{ and } \ \forall_{<\sigma,q>\in y}\forall_{r\leqslant p,q}(r \Vdash \sigma \in z).$$

Fix $<\sigma, q> \in x$ and $r \leqslant p, q$. Since $r \Vdash \sigma \in y$, there exists $<\sigma', q'> \in y$ such that $q' \geqslant r$ and $r \Vdash \sigma = \sigma'$ and since $r \leqslant p, q'$, $r \Vdash \sigma' \in z$. Hence, by our induction hypothesis, $r \Vdash \sigma \in z$. Since $<\sigma, q> \in x$ and $r \leqslant p, q$ are arbitrary, $p \Vdash x \subset z$.

To show E4, suppose that $p \Vdash x = y \wedge x \in z$, and that E3 holds for all names with rank less than the rank of z. Since $p \Vdash x \in z$, there exists $<\sigma, q> \in z$ such that $q \geqslant p$ and $p \Vdash x = \sigma$. Applying E3 to x, σ, y, we have that $p \Vdash \sigma = y$. Hence there exists $<\sigma, q> \in z$ such that $q \geqslant p$ and $p \Vdash y = \sigma$; that is, $p \Vdash y \in z$.

E5: Suppose that p forces that $x = y$ and $z \in x$. Since p forces $z \in x$, there exists $<\sigma, q> \in x$ with $q \geqslant p$ such that $p \Vdash \sigma = z$. Thus $p \Vdash x = y$ implies that p forces $\sigma \in y$; whence $p \Vdash \sigma = z \wedge \sigma \in y$. It now follows from E4 that $p \Vdash z \in y$.

Axioms of IZF.

Extensionality: $\forall_{x,y}(\forall_t(t \in x \leftrightarrow t \in y) \to x = y)$.
We must show that $\mathbf{1}$ forces Extensionality. This means, by definition, that for all names x, y and all $p \in K$, if $p \Vdash \forall_t(t \in x \leftrightarrow t \in y)$, then $p \Vdash x = y$. Fix x, y and suppose that $p \Vdash \forall_t(t \in x \to t \in y)$. In

order to show that $p \Vdash x \subset y$, we must show that $r \Vdash \sigma \in y$ for any $< \sigma, q > \in x$ and $r \leqslant p, q$; fix such σ, q, and r . By Lemma 12, $r \Vdash \sigma \in x$, so $r \Vdash \sigma \in y$ since $r \leqslant p$ and p forces $t \in x \to t \in y$.

Pairing: $\forall_{x,y} \exists_z \forall_t (t \in z \leftrightarrow t = x \vee t = y)$.
Given names x, y, we must define a name z such that for any t

$$p \Vdash t \in z \Leftrightarrow (p \Vdash t = x \vee p \Vdash t = y).$$

This is easily done: using pairing in the metatheory, we define

$$z = \{< x, \mathbf{1} >, < y, \mathbf{1} >\}.$$

Then $\mathbf{1} \Vdash x \in z \wedge y \in z$, and if $p \Vdash t \in z$, then there exists $< \sigma, q > \in z$ such that $q \geqslant p$ and $p \Vdash \sigma = t$. Either $< \sigma, q > = < x, \mathbf{1} >$, in which case $p \Vdash \sigma = x$, or $< \sigma, q > = < y, \mathbf{1} >$ and $p \Vdash \sigma = y$.

Union: $\forall_x \exists_y \forall_s (s \in y \leftrightarrow \exists_t (s \in t \wedge t \in y))$.
We must show that for each name x there exists a name y such that for all names s and each $p \in \mathbb{P}$

$$p \Vdash s \in A \quad \Leftrightarrow \quad \exists_t (s \in t \wedge t \in x).$$

Given a name x, define

$$\begin{aligned} y &= \{< t, r >: \exists_x (r \Vdash t \in x \wedge x \in a)\} \\ &= \Big\{< t, r > \in V_{\mathrm{rk}(x)} \times K : \\ &\qquad \exists_{q, q' \leqslant r}, \exists_{x \in V_{\mathrm{rk}(a)}} (< t, q > \in x \wedge < x, q' > \in a)\Big\} \end{aligned}$$

by (bounded) separation on $V_{\mathrm{rk}(x)} \times K$. Suppose $p \Vdash t \in y$; let $< \sigma, q > \in y$ be such that $q \geqslant p$ and $p \Vdash t = \sigma$. Since $< \sigma, q > \in y$, $q \Vdash \sigma \in s \wedge s \in x$ for some name s, so $p \Vdash t \in s \wedge s \in x$. Hence $p \Vdash \exists_s (t \in s \wedge s \in x)$. Conversely if $p \Vdash t \in s \wedge s \in x$, then $< t, p > \in y$, so $r \Vdash t \in y$.

Separation: $\forall_x \exists_y \forall_t (t \in y \leftrightarrow t \in x \wedge \varphi)$.
Given a name x define

$$y = \{< t, p >: p \Vdash t \in x \wedge p \Vdash \varphi x\}$$

by separation on $V_{\mathrm{rk}(x)}(K) \times K$. Then for all names t and each $p \in K$, p forces that $t \in z$ if and only if p forces $t \in x$ and p forces φ.

Powerset: $\forall_x \exists_y \forall_t (t \in y \leftrightarrow t \subset x)$.
Given x, define

$$y = \{< t, p >: p \Vdash t \subset x\}$$

by bounded separation on $\mathcal{P}(V_{\mathrm{rk}(x)}(K) \times K)$. Then for each $p \in K$, $p \Vdash t \in y$ if and only if $p \Vdash t \subset x$.

Set induction: $\forall_x \left(\forall_y (y \in x \to \varphi(y)) \to \varphi(x)\right) \to \forall_x \varphi(x)$.
Suppose that

$$p \Vdash \forall_y (y \in x \to \varphi(y)) \to \varphi(x)$$

for all x. We prove that $p \Vdash \forall_x \varphi(x)$ by induction on the rank of x. Fix a name a and suppose that $p \Vdash \varphi(x)$ for all x with rank less than $\mathrm{rk}(a)$. If $r \Vdash y \in x$ for some $r \leqslant p$, then there exists $< \sigma, q > \in x$ such that $q \geqslant r$ and $r \Vdash \sigma = y$. Since $\mathrm{rk}(\sigma) < \mathrm{rk}(x)$, p, and hence r, forces $\varphi(\sigma)$ by the induction hypothesis; so $r \Vdash \varphi(y)$, by Lemma 14. Thus $p \Vdash \forall_y (y \in a \to \varphi(y))$ and it follows from our initial assumption that $p \Vdash \varphi(a)$, which completes the induction.

Infinity: $\exists_x (\exists_s s \in x \wedge \forall_{s \in x} \exists_{t \in x} s \in t)$.
Let x be the canonical name $\check{\mathbb{N}}$ for the natural numbers.

Collection: $\forall_{x \in a} \exists_y \theta(x, y) \to \exists_b \forall_{x \in a} \exists_{y \in b} \theta(x, y)$.
Suppose that

$$p \Vdash \forall_{x \in a} \exists_y \theta(x, y) \qquad (*)$$

and define

$$a' = \{< x, q >: q \Vdash x \in a, q \leqslant p\}$$

by (bounded) separation on $V_{\mathrm{rk}(a)} \times K$. Unpacking $(*)$ we have that

$$\forall_{<x,q>\in a'} \exists_y (q \Vdash \theta(x,y)).$$

We can now apply collection in the metatheory to construct b' such that

$$\forall_{<x,q>\in a'} \exists_{y\in b'} (q \Vdash \theta(x,y)).$$

By (full) separation $b = b' \cap V(K)$ is a name. Then

$$\forall_{q\leqslant p} \forall_x (q \Vdash x \in a \Rightarrow \exists_{y\in b} q \Vdash \theta(x,y)),$$

so $p \Vdash \exists_b \forall_{x\in a} \exists_{y\in b} \theta(x,y)$.

■

We give a simple example of a forcing model of **IZF**. A set S is said to *admit a constant function* if there exists a function $f : S \to S$ such that $f(s_1) = f(s_2)$ for all $s_1, s_2 \in S$. We show that the statement

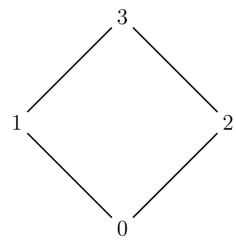
(*) Every set admits a constant function.

is independent of **IZF** plus dependent choice, and hence Bishop's constructive mathematics.[22] There are two objections to (*) as a constructive principle: (i) if we have a constructive proof of (*) restricted to a collection $\mathcal{A}$ of inhabited sets, then by examining the codomains of the constant functions $(f_a)_{a\in\mathcal{A}}$ given by (*), we are able—using unique choice on $\varphi(a,y) = \forall_{x\in a}(f_a(x) = y)$—to construct a choice function for $\mathcal{A}$; (ii) the nature of a constant function on S depends on whether S is inhabited or empty, which is not something that can be constructively decided.

Our model shows that (*) fails even for finite subsets of the natural numbers; since dependent choice holds in the model, the failure of this restriction of (*) rests with objection (ii). If we restrict (*) to Σ^0_1 sets, then it is constructively valid.

[22] Condition (*) is considered in [5] where, in particular, it is shown to imply a constructively dubious principle and a homotopy type theory taboo.

We take the model $V(\mathbb{P})$ over the poset $\mathbb{P}$ elements $\{0, 1, 2, 3\}$ and comparabilities as indicated below.



Let $\sigma = \{< \check{0}, 1 >, < \check{1}, 0 >\}$. Suppose 0 forces that a $\mathbb{P}$-name τ is a function on σ; in particular, for all $p \in \mathbb{P}$ and all $\mathbb{P}$-names x, if $p \Vdash x \in \sigma$, then there exists a $\mathbb{P}$-name η such that $p \Vdash \eta \in \sigma$ and $p \Vdash (x, \eta) \in \tau$. It follows that

$$1 \Vdash (\check{0}, \check{0}) \in \tau \quad \text{and} \quad 2 \Vdash (\check{1}, \check{1}) \in \tau.$$

Hence 3 forces that $(\check{0}, \check{0}), (\check{1}, \check{1}) \in \tau$, so that τ is not the name of a constant function on σ.

Since $\mathbb{P}$ is finite, we can code all names of natural numbers in $M(\mathbb{P})$ by natural numbers at the meta level; this allows us to prove dependent choice in $M(\mathbb{P})$ by applying dependent choice, to names for natural numbers and a predicate involving the forcing relation, at the meta-level. Hence $M(\mathbb{P})$ satisfies **DC**.

1.3.3 Topological models

Let (X, T_X) be a topological space. For a point $x \in X$, we denote by $\mathcal{N}_x$ the set of neighbourhoods of x. Let $V(X)$ be the class of names $V(T_X)$ from forcing: $V(X)$ is defined inductively by

$$\begin{aligned} V_\alpha(X) &= \bigcup \{\mathcal{P}(V_\beta(X) \times T_X) : \beta \in \alpha\}, \\ V(X) &= \bigcup_{\alpha \in \mathrm{ORD}} V_\alpha(X); \end{aligned}$$

for a name τ, the rank $\mathrm{rk}(\tau)$ of τ is $\inf\{\alpha \in \mathrm{ORD} : \tau \in V_\alpha(X)\}$. We first define, by induction, the topological forcing relation $\Vdash$ between points and formulas; an open set U is then said to force a formula A if each element of U forces A. The base cases, forcing prime formulas, are defined by simultaneous transfinite recursion on rank:

$x \Vdash u \in v$ if there exists $< \sigma, V > \in v$ such that $x \in V$ and $x \Vdash \sigma = u$:

$$\exists_{< \sigma, V > \in v}(x \in V \wedge x \Vdash \sigma = u).$$

$x \Vdash u \subset v$ if for some neighborhood W of x and all $< \sigma, U > \in u$ if any point in W forcing that $\sigma \in u$ also forces that $\sigma \in v$:

$$\exists_{W \in \mathcal{N}_x} \forall_{y \in W} \forall_{< \sigma, U > \in u}(y \in U \to y \Vdash \sigma \in v).$$

We extend $\Vdash$ to all formulas inductively by the following clauses.

$x \Vdash A \wedge B$ if $x \Vdash A$ and $x \Vdash B$.

$x \Vdash \exists_u Au$ if there exists a name τ such that $x \Vdash A\tau$.

$x \Vdash A \vee B$ if there is a neighborhood of x which either forces A or forces B:

$$\exists_{W \in \mathcal{N}_x}((\forall_{y \in W} y \Vdash A) \vee (\forall_{y \in W} y \Vdash B)).$$

$x \Vdash A \to B$ if there is a neighborhood W of x such that if $y \in W$ and $y \Vdash A$, then $y \Vdash B$:

$$\exists_{W \in \mathcal{N}_x} \forall_{y \in W}(y \Vdash A \to y \Vdash B).$$

$x \Vdash \forall_u Au$ if there is a neighborhood of x which forces $A\tau$ for any name τ:

$$\exists_{W \in \mathcal{N}_x} \forall_{y \in W} \forall_{\tau \in V(X)}(y \Vdash A\tau).$$

We further require that no node forces falsity

$$\forall_{x \in X} \neg(x \Vdash \perp)$$

and define

$$U \Vdash A \quad \text{if and only if} \quad \forall_{x \in U} x \Vdash A,$$

for $U \in T_X$. Note that monotonicity follows immediately from this definition, and a simple induction argument shows that $x \Vdash A$ if and only if there is a neighborhood W of x such that $W \Vdash A$.

The definition of the *topological model* $M_X = (X, T_X, V(X), \Vdash)$ is completed by setting A to be true in M_X if and only if A is forced by X:

$$M_X \vDash A \quad \Leftrightarrow X \Vdash A.$$

Let (X, T_X) be a topological space. By unwrapping the various definitions we have that:

$$\begin{array}{lcl}
M_X \vDash u \in v & \Leftrightarrow & \forall_{x \in X} \exists_{<\sigma, V> \in v} (x \in V \wedge x \Vdash \sigma = u); \\
M_X \vDash u \subset v & \Leftrightarrow & \forall_{<\sigma, U> \in u} \forall_{x \in U} (x \Vdash \sigma \in v); \\
M_X \vDash A \wedge B & \Leftrightarrow & M_X \vDash A \text{ and } M_X \vDash B; \\
M_X \vDash A \vee B & \Leftrightarrow & \forall_{x \in X} (x \Vdash A \vee x \Vdash B); \\
M_X \vDash A \rightarrow B & \Leftrightarrow & \forall_{x \in X} (x \Vdash A \Rightarrow x \Vdash B); \\
M_X \vDash \exists_a A & \Leftrightarrow & \forall_{x \in X} \exists_{\sigma \in V(X)} (x \Vdash A\sigma); \\
M_X \vDash \forall_a A & \Leftrightarrow & \forall_{\sigma \in V(X)} M_X \vDash A\sigma.
\end{array}$$

Comparing these with the conditions for a forcing model $V(\mathbb{P})$ to satisfy a formula, we see that $\subset, \wedge, \rightarrow$, and $\forall$ behave in essentially the same way. For example $M_X \vDash u \subset v$ if and only if each $U \in T_X$ forces $u \subset v$, which is precisely the condition for $V(T_X)$ to satisfy $u \subset v$. In contrast, the validity of membership, existential statements, and disjunctions is now local:

- for $M_X \vDash u \in v$, the $< \sigma, V > \in v$ such that $x \Vdash \sigma \in v \wedge u = \sigma$ may depend on x;
- for $M_X \vDash \exists_u A$, there need not be one name σ such that $M_X \vDash A\sigma$—points may differ on the names they force to satisfy A;
- similarly for disjunctions, points may force different parts of the disjunction, but each point forces at least one.

Despite these differences, the soundness theorem for topological models of **IZF** follows almost directly from the arguments in the soundness theorem for forcing models.

Theorem 16 (IZF) *Topological models are sound for **IZF**:*

$$\textbf{\textit{IZF}} \vdash \varphi \Rightarrow M_X \Vdash \varphi.$$

Proof. The proof is very similar to the proof that forcing is sound for **IZF**, Theorem 16. We need only check that the arguments in the proof of Theorem 16 can be carried out locally; this is left as an exercise. ■

The next theorem gives an example of a typical argument in topological forcing.[23]

Lemma 17 (LPO) *Let (X, T_X) be a locally connected topological space, and let $(X_n)_{n\in\mathbf{N}}$ be a sequence of subsets of X such that $X = X_n \cup (\sim X_n)^\circ$. Then*

$$X = \left(\bigcap_{n\in\mathbf{N}} X_n\right)^\circ \cup \bigcup_{n\in\mathbf{N}} (\sim X_n)^\circ.$$

Proof. Fix $x \in X$ and let U be an open connected neighborhood of x. Since U is connected, either $U \subset X_n$ or $U \subset (\sim X_n)^\circ$, for otherwise we could write U as the disjoint union of the nonempty open sets $U \cap X_n$ and $U \cap (\sim X_n)^\circ$. Hence it follows from **LPO** that either $U \subset X_n$ for each n or there exists n such that $U \subset (\sim X_n)^\circ$. In the former case $x \in \left(\bigcap_{n\in\mathbf{N}} X_n\right)^\circ$ and in the latter $x \in \bigcup_{n\in\mathbf{N}}(\sim X_n)^\circ$. ■

Theorem 18 (LPO) *Let (X, T_X) be a connected topological space. Then $M_X \vDash$ **LPO**.*

Proof. Suppose $U \in T_X$ forces that α is a binary sequence. For each n there exists an open subset U_n of U such that $U = U_n \cup (\sim U_n)^\circ$, and

[23] Much thanks must go to Andrew Swan for giving so much of his time to convince me that **LPO** should hold in many topological models; the proof of the next theorem is essentially his.

$U_n \Vdash \alpha(n) = 0$ and $({\sim}U_n)^\circ \Vdash \alpha(n) = 1$. Then

$$\left(\bigcap_{n\in\mathbf{N}} U_n\right)^\circ \Vdash \forall_{n\in\mathbf{N}}\alpha(n) = 0 \quad \text{and} \quad \bigcup_{n\in\mathbf{N}} ({\sim}U_n)^\circ \Vdash \exists_{n\in\mathbf{N}}\alpha(n) = 1.$$

It follows from Lemma 17 that $U = \left(\bigcap_{n\in\mathbf{N}} U_n\right)^\circ \cup \bigcup_{n\in\mathbf{N}}({\sim}U_n)^\circ$, and hence that $U \Vdash (\forall_{n\in\mathbf{N}}\alpha(n) = 0) \vee \exists_{n\in\mathbf{N}}\alpha(n) = 1$. ■

In general, since topological semantics is local, properties valid in a particular topological model only depend on the local properties of the underlying topological space.

Topological models have been well studied as sheaf models $\mathrm{Sh}(\mathcal{O}(X))$ over a topological space (X, T_X) [54], where the following very useful characterisation of the Dedekind and Cauchy reals in a topological model is given.

Proposition 19 *Let (X, T_X) be a topological space. Then the Dedekind reals in $\mathrm{Sh}(\mathcal{O}(X))$ are represented by the continuous functions from (X, T_X) to $\mathbf{R}^D$, and the Cauchy reals are represented by the locally constant functions.*

We shall have a close look at the sheaf model over $[0, 1]$, with the standard topology, in section 1.5.

1.3.4 Realizability

Topological models provide a natural semantics for Brouwer's intuitionism: in particular, every topological model satisfies Brouwer's fan theorem [47], and many natural examples satisfy Brouwer's continuity principles.[24] Moreover, Brouwer's perception of mathematics can (anachronistically) be used to motivate Kripke, forcing, and topological models for **IZF**.

There is another school of constructive mathematics, the Russian school of recursive mathematics, with a different philosophical standpoint, which also motivates a semantics for **IZF**; namely, realizability. Realizability can be

[24] In particular, Krol [71] gave a topological model satisfying Brouwer's fan theorem and continuity principles as well as Kripke's schema.

seen as an attempt to make precise the notions of proof and algorithm in the **BHK** interpretation of the logical connectives and quantifiers. We interpret both proof and algorithm by computations; partial combinatory algebra's provide us with a rich source of models of computability appropriate for this task.

Let $\mathcal{A}$ be a pca. The basic structure of realizability models is similar to that of forcing. First we inductively construct a universe of $\mathcal{A}$-*names*

$$
\begin{aligned}
V(\mathcal{A})_\alpha &= \bigcup_{\beta<\alpha} \mathcal{P}(\mathcal{A} \times V(\mathcal{A})_\beta), \\
V(\mathcal{A}) &= \bigcup_{\alpha \in \mathbf{ORD}} V(\mathcal{A})_\alpha;
\end{aligned}
$$

we interpret the labels from $\mathcal{A}$ as providing computational evidence of the properties of names. The *realizability relation* $\Vdash$, between elements of $\mathcal{A}$ and formulas of our (extended) language, is defined by induction. Let a, b be names, and let e be an element of $\mathcal{A}$, we write e_0, e_1 for the projections $\pi_0 e, \pi_1 e$ of e and $, e_0, e_1 >$ for the pair pe_0e_1. The base cases are

- $e \Vdash a \in b$ if there exists a name c such that $\langle e_0, c\rangle \in b$ and $e_1 \Vdash a = c$;
- $e \Vdash a \subset b$ if for all $\langle f, d\rangle$ in a, $e_0 f \downarrow$ and $e_0 f \Vdash d \in b$;
- $\neg e \Vdash \perp$.

The induction cases closely reflect the **BHK** interpretation:

- $e \Vdash A \wedge B$ if $e_0 \Vdash A$ and $e_1 \Vdash B$;
- $e \Vdash A \vee B$ if $e_0 = 0$ and $e_1 \Vdash A$ or $e_0 = 1$ and $e_1 \Vdash B$;
- $e \Vdash A \to B$ if for every $f \in \mathcal{A}$ which forces A, $ef \downarrow$ and $ef \Vdash B$;
- $e \Vdash \forall_x \varphi x$ if for every name a, $e \Vdash A[x/a]$;
- $e] \Vdash \exists_x \varphi x$ if there exists name a such that $e \Vdash A[x/a]$.

We call any $e \in \mathcal{A}$ such that $e \Vdash A$ a *realizer* of A. The *realizability model* $V(\mathcal{A})$ satisfies A if there is a realizer for A in $\mathcal{A}$: $V(\mathcal{A}) \models A$ if and only if $\exists_{e \in \mathcal{A}} e \Vdash A$.

Theorem 20 *Realizability is sound for **IZF**: if $\mathcal{A}$ is a pca and **IZF**$\vdash A$, then $V(\mathcal{A})$ satisfies A.*

Proof. The full proof can be found in McCarty's thesis [82]. ■

Realizability models have been put to great use in the study of the metamathematic properties of constructive set theory. For example whether constructive theories have the set existence property [90, 103], conservativity and independence [8, 38], and program extraction [13].

In the same way that Kripke models can be used to give models of classical set theory, realizability has been extended by Krivine to give models of ZFC [70] (forcing models can even be seen as a special case of Krivine's realizability).

1.4 Constructive mathematics and computability

Constructive proofs are (logically) simpler than nonconstructive proofs, and it is often the case that (the search for) a constructive proof reveals more about why a result is true—the purpose of proofs after all is to tell us why something is true (we already know what is true). So an appreciation for constructive methods may give the classical mathematician more instructive proofs. But what else might working constructively give the classical mathematician? We give a few examples.

1.4.1 Constructive mathematics and computation

The message of this brief section is

> "Algorithmic mathematics—that is, computer science—appears to be equivalent to mathematics that uses only intuitionistic logic."[25]
>
> D. Bridges and S. Reeves, [29]

The essence of our message is what is expressed by the *Curry-Howard correspondence* between proof systems and models of computation; at its most general, the correspondence is the association of propositions with types. Data types, or simply types, are the computer science equivalent of a set, each program has an associated type which describes the valid inputs and outputs of a program. More specifically, the Curry-Howard isomorphism gives a correspondence between models of computation and intuitionistic theories. For example, between natural deduction and the λ-calculus there is a correspondence between: hypothesis and free variables; implication elimination and application; and implication introduction and abstraction.

How do we get programs from proofs? The proper, or at most widely accepted, foundational system for computer science is Martin-Löf's intuitionistic type theory—the same type theory which interprets **CZF** plus relativised

[25] Although, as Martin Escardo pointed out to me, many, if not most, computer scientists use classical logic.

dependent choice (and large cardinal like assumptions)—and the Curry-Howard correspondence allows us to view mathematical proofs in **MLTT**, or **CZF**, as programs in some typed λ-calculus. This process is sketched below.

Marin-Löf Type Theory

Martin-Löf type theory can be viewed as a high level programing language,[26] or alternatively as a restricted set theory: types are roughly the completely presented sets—sets in which every element carries evidence of its membership. Types form the valid domains of quantification. Informally, in order to define a type we must describe a set in the sense of Bishop, and additionally:

(i) describe the canonical members, elements which witness there own membership;

(ii) describe how to evaluate members of A to give canonical members.

For example, the type of natural numbers $\mathbf{N}$ consists of canonical elements defined inductively by

> 0 is canonical, and if n is canonical, then the successor $s(n)$ of n is canonical,

together with elements defined from canonical elements using the operations defined by recursion. The evaluation rules for $\mathbf{N}$ are the standard rules for evaluating operations, so for example $3! + 10^2 - 3$ evaluates to 103. From a programing perspective, canonical elements correspond to data, and general elements to programs which give instructions for their own evaluation. So canonical members roughly correspond to λ terms in normal form. There is, however, no notion of canonical element in the formal theory.

MLTT is based on taking the Curry-Howard isomorphism as a philosophical principle. Above we interpreted a type as a set, but, by the 'propositions-as-types' idea, we can also interpret a type as a proposition. Under this

[26] As Beeson puts it "a sort of λ-calculus with variable types" [9].

interpretations, elements of a type correspond to 'proofs', or *realizers*, of the associated proposition, and a proposition corresponds to the collection of its realizers. Returning to the natural numbers, the type $\mathbf{N}$ can be interpreted as the proposition 'there exists a natural number', and any natural number gives a proof of this proposition. The propositions-as-types notion together with the basic type constructors, introduced below, allow us to embed intuitionistic logic into **MLTT**.

Before introducing some methods for constructing new types from old, there is one more key aspect of **MLTT** to introduce, that of judgements. Mathematics happens at the level of judgements, while realizers give only the computational content. The four basic types of *judgement* are the following.

- A is a type, written $\mathrm{Type}A$.
- a has type A, written $a : A$.
- s, t are equal elements of type A, $s = t : A$.
- A and B are equal types, $A = B$, they have the same elements and the same equality.

If A is a type, then the judgement that $a : A$ is the assertion that a is a proof of the proposition corresponding to A. So for example, if A is the type corresponding to Lagrange's four-square theorem—every natural number can be expressed as the sum of four squares—then the judgement $a : A$ corresponds to a proof of the four-square theorem, and the program a which given $n \in \mathbf{n}$ enumerates $\mathbf{N}^4$ in an appropriate order until it finds (a, b, c, d) such that $a^2 + b^2 + c^2 + d^2 = n$ is an element of A—this search is bounded by $\sqrt{n}$. Judgements can also have a dependency on a type, for example

- a is of type A given that b is of type B, written $a : A(b : B)$;
- B is a family of types over A, $\mathrm{Type}B(x)\ x : A$.

We now give the basic ways to construct *types*, and relate them to the embedding of logic via the Curry-Howard isomorphism.

- We start with the empty type $\mathbf{0}$ and the unit type $\mathbf{1}$. By the Curry-Howard isomorphism, $\mathbf{0}$ is a proposition with no proof and corresponds to false, while $\mathbf{1}$ has a single element and corresponds to true.

- **Pi-types** are analogous to indexed products of sets, they generalise function spaces by allowing the image type to vary with the input. Intuitively $f \in \Pi(A, B)$ corresponds to $f(a) \in B(a)$ $(a : A)$, where $B(a)$ is a family of types dependent on A. The logical interpretation of $\Pi(A, B)$ is $\forall_{x \in A} B(x)$; for if $f : \Pi(A, B)$, then given any realizer a of A, $f(a)$ is a realizer of $B(a)$. For types A, B, we write $A \to B$ for the type of functions from A to B. If $f : A \to B$, then f maps realizers of A to realizers of B; whence an element of $A \to B$ corresponds to a proof that A logically entails B. The Π-types can also be seen as a generalisation of λ-abstraction.

- **Sigma-types** are analogous to indexed disjoint unions; the intuitive meaning of the type $\Sigma(A, B)$ is the set $\{(a, b) : a \in A, b \in B(a)\}$ where Type$B(a)$ $(a : A)$. Thinking now in terms of propositions, (a, b) realizes $\Sigma(A, B)$ if $a : A$ and b realizes $B(a)$, so $\Sigma(A, B)$ corresponds to the existential statement $\exists_{a \in A} B(a)$. We get Cartesian products as the special case of Σ-types when the dependency of B is trivial—that is, when $B(a)$ is B for each $a : A$. An element of type $A \times B$ corresponds to a pair (a, b) such that a realizes A and b realizes B. Hence $A \times B$ gives an interpretation of $A \wedge B$.

- Given types A, B there is also a disjoint sum type $A + B$ an element of which is a member of $A \cup B$ together with a label indicating whether it is from A or B. Disjoint sums are the type equivalent of disjunction, so we must also introduce operations i, j which inject A and B respectively into $A + B$; i, j are realizers of the intuitionistic axioms $A \to A \vee B$ and $B \to A \vee B$.

- **Inductive types** are introduced as a special case of the type of well founded trees. By the Curry-Howard correspondence, primitive recursion is identified with the induction type

$$\mathbf{N} - \mathrm{elim} : P(0) \to \Pi(\mathbf{N}, P(n) \to P(sn)) \to \Pi(\mathbf{N}, P),$$

  for any P such that Type$P(n)$ $(n : \mathbf{N})$. More generally, inductive families correspond to inductively defined relations. The naturals are inductively generated by $0 : \mathbf{N}$ and $s : \mathbf{N} \to \mathbf{N}$.

Formally, type theories are usually presented as a dependently typed lambda calculus. We have basic rules for equality and substitution, and for each type constructor, rules for formation, introduction, elimination, and defining equality on the type. See [81] for a full account of **MLTT**.

Martin-Löf type theory is the first system we have seen which is clearly predicative: types can only be constructed from previously constructed types, and we avoid the appeal to a completed universe of sets which is inherent in even a predicative set theory like **CZF**. One may try to justify the constructive credentials of **CZF** directly from the axioms, but the most convincing evidence that **CZF** is constructive comes from Aczel's interpretation of **CZF** in **MLTT**. **MLTT** is intensional, Per Martin-Löf's original formulation was extensional, but Martin-Löf choose to make his type theory intensional in order to preserve the decidability of type checking. Finally we note that, since the intuitive justification for relativised dependent choice is valid for completely presented sets, relativised dependent choice is valid in **MLTT**.

We can now be a little more specific about **MLTT**'s credentials as a programing language. Given a proposition A, a mathematical proof J of A, formalized in **MLTT**, is a judgement that $a : A$ for some a. The typed λ term a is a computable function which gives a computable realisation of A, and the judgement J is a proof that a meets the programing specification given by A.

See [95] for the development of formal topology in **MLTT**.

General program extraction

The utility of **MLTT** as a programing language results from the careful assignment and preservation of computational content (realizers) as we form types. For intuitionistic logic and set theories we have given a similar systematic assignment of computational information, in the metatheory, in the form of realizability. Using these realizability models, program extraction can be achieved from constructive proofs without needing to formalise our results in the mathematically unfamiliar territory of type theory.

We only give an example, for realizability with Turing machines, of what can be proved in this direction.

Theorem 21 *There are Turing machines* $\mathbf{u}, \mathbf{v}$ *such that if*

$$e \Vdash \forall_{x \in \mathbf{N}} \exists_{y \in \mathbf{N}} \varphi(x, y),$$

then

$$(\mathbf{v}e)(n) \Vdash \varphi(\overline{n}, \mathbf{u}e(\overline{u}));$$

and if φ *is recursive, then in* ***IZF*** *we have* $\varphi(n, \{\mathbf{u}e\}(n))$ *for each* n.

Proof. See Lipton [73]. ■

Other approaches and more refined techniques have been used to extract programs from minimal logic [99], and for intuitionistic systems with choice and inductive definitions [11, 12]. Programs have also been extracted from classical proofs using these techniques [14], but in this case we no longer have a guarantee that the extracted program meets it specifications. Kohlenbach has developed very successful techniques for the extraction of computational information from classical proofs [69] using similar ideas.

Exact real number computation

Programs derived directly from constructive proofs via realizability come with a guarantee that they meet their specifications (assuming the extraction is correct!). Programs dealing with real number computation have an additional, similar, advantage over their more traditional counterparts: constructive results give exact real number calculations. When we implement a constructive proof we do not have to worry about finite computer precision producing erroneous results—since constructive mathematics deals with finite information and procedures (and only the potential infinite), constructive results will be implemented faithfully.

We borrow an example from [32, page 2]. Consider the function $f : [0, 1] \to \mathbf{R}$ given by

$$f(x) = \left(x - \frac{3}{4}\right)\left(x - \frac{1}{2}\right)^2 - 2^{-51}$$

and the classical proof of the intermediate value theorem which given reals $a < b$ with $f(a) < 0 < f(b)$ proceeds by testing the midpoint $(b - a)/2$ to see whether it is positive, negative or zero, and refines the interval $[a, b]$ to a smaller one $[x, y]$ for some $x, y \in \{a, b, (b - a)/2\}$, depending on the result of this test, which still must (classically) contain a root. If our computer implementation of this proof uses 50-bit precision, then at the first step we test $f(1/2)$ and concluding it is 0, as 2^{-51} is below the computer precision, we stop the program and output $x = 1/2$ as a root of f. But the unique root of f is greater than $3/4$! This phenomenon cannot occur with a program derived from a constructive proof: constructive proofs have interval arithmetic built in. This example is very similar to the proof that the intermediate value theorem does not admit a constructive proof.

1.4.2 Theorems for free

The message of this section is

> *The power of constructive mathematics is in its weakness: constructive proofs give us more, precisely because they require less.*

The weakness of **CZF** allows it to be interpreted into many systems of mathematics, in particular in **MLTT** and in any topos. For the classical mathematician, who—for the sake of argument—only cares for the one true system, this may seem to be of little consequence, but a constructive proof also furnishes us with additional information which may be of interest to the classical mathematician. We give two basic examples of what the classical mathematician gets for free from a (possibly inadvertent) constructive proof.

As a crude example, suppose you have have shown that f is a well defined function between metric spaces, and that you observe the proof to be fully constructive. Then you can take a proof in **IZF** and using a standard realizability interpretation find a computable realizer that f is a function. Since every computable function is continuous, f must be continuous. Thus from a constructive proof that f is a function, we get a classical proof that f is continuous and even computable—see [7] for a formal result.

A more advanced example was given by Joyal who gave a categorical extension of Kripke semantics for higher order Heyting arithmetic[27] (**HAH**), and used it to derive the following continuity result.

Theorem 22 (ZFC) *If we have a proof that for all $x \in \mathbf{R}^D$ there exists $y \in \mathbf{R}^D$ such that $A(x, y)$ in **HAH**, then there is a proof in **HAH** that for all $x \in \mathbf{R}^D$ there exists an open neighborhood U of x and a continuous function $f : U \to \mathbf{R}^D$ such that $A(x, f(x))$ for all $y \in U$.*

Proof. See [108, page 805]. The proof uses a sheaf model over the topological space with points the disjoint union $\mathbf{R}^D \dot\cup \mathbf{R}^D$ of $\mathbf{R}^D$ with itself, and opens $U_1 \dot\cup U_2$ where $U_1 \subset U_2$ and U_1 is open in the standard topology on $\mathbf{R}^D$. ■

This result extends to **CZF** or **IZF**. Thus if we have a constructive proof of $\forall_{x \in \mathbf{R}^D} \exists_{y \in \mathbf{R}^D} A(x, y)$, then classically we know that we can find a locally continuous solution for y in terms of x.

[27] Heyting arithmetic is an intuitionistic formulation of Peano arithmetic, and Peano arithmetic can be interpreted in Heyting arithmetic by a double negation interpretation. Higher order Heyting arithmetic is Heyting arithmetic with types for functions from the natural numbers to the natural numbers, functions between number theoretic functions, and so forth.

In general, we can interpret proofs in **CZF** not only to be about the familiar denizens of the mathematical universe, but also to those inhabitants of any model of **CZF**, and in particular any topos. Thus by catering the topos to your interests you can cater the interpretations of constructive theorems.

Constructive mathematics and computability

In [7], Bauer gives an introduction to a particularly striking and well developed example of deriving classical results from a model of constructive mathematics:

> "Computable mathematics is the realizability interpretation of constructive mathematics."

We only give a simple example of how results in constructive mathematics can be converted into results in computable mathematics using realizability. A *multivalued map* $g : X \rightrightarrows Y$ is a function from X to the collection of inhabited subsets of X. The multivalued map $g : X \rightrightarrows Y$ is *computably realized* if there exists a computable function $f : X \to Y$ such that $f(x) \in g(x)$ for each $x \in X$. The following is Proposition 4.30 of [7].

Proposition 23 *If we can prove*

$$\forall_{x \in X} \exists_{y \in Y} R(x, y)$$

with intuitionistic logic, then there is a computably realized multivalued map $g : X \rightrightarrows Y$ *such that* $R(x, y)$ *holds for all* $y \in g(x)$*.*

Bauer uses this Proposition together with [108, Theorem 7.2.7] to show that there is a computably realized multivalued map which takes $(f, x, n) \in \mathbf{R}^{\mathbf{R}} \times \mathbf{R} \times \mathbf{N}$ to a subset S of $\mathbf{N}$ such that $1/m$ is a $1/n$-modulus of continuity for f at x for each $m \in S$.

1.5 Constructive (reverse) mathematics

This section gives some basic constructive mathematics and constructive reverse mathematics. First we present some definitions which are indispensable to the practicing constructivist.

1.5.1 Constructive definitions

As we have emphasised in the previous section, any proof in Bishop's constructive mathematics contains an algorithm which 'implements' the result it proves. So if the conclusion of a theorem has some computational content, then so must the hypothesis: we cannot get something for nothing! To get constructively meaningful results, it is necessary to rewrite many classical definitions in a positive way; we must also at times make explicit conditions which hold trivially in classical mathematics, but which may fail in our framework. At still other times, a classical definition given a constructive reading becomes a far stronger property, so here we must adopt a computationally weaker, though classically equivalent alternative. Here we give some of the definitions which are fundamental to the study of constructive analysis. Other constructive definitions will be introduced as required.

Let X be a metric space and let S be a subset of X. If there exists $x \in S$, then S is said to be *inhabited*; constructively this is a stronger property than S being nonempty, $\neg(S = \emptyset)$. An inhabited set S is said to be *located* if for each $x \in X$ the *distance*

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

from x to S exists. Let $\varepsilon > 0$. An ε-*approximation* to S is a subset T of S such that for each $s \in S$, there exists $t \in T$ such that $\rho(s, t) < \varepsilon$. We say that S is *totally bounded* if for each $\varepsilon > 0$ there exists a finitely enumerable[28] ε-approximation to S. A metric space is said to be *compact* if it is complete

[28] A set is *finitely enumerable* if it is the image of $\{1, \ldots, n\}$ for some $n \in N^+$, and a set is *finite* if it is in bijection with $\{1, \ldots, n\}$ for some $n \in N^+$; constructively these notions are distinct.

and totally bounded. A totally bounded subset of X is located [18, page 95, Proposition (4.4)]. In a metric space X, we denote by $B(x, r)$ and $\overline{B}(x, r)$ the open and closed balls, respectively, centred on $x \in X$ with radius $r > 0$.

Continuity is a thorny issue in constructive mathematics.[29] Since the three major extensions of Bishop's constructive mathematics—classical mathematics, Brouwers intuitionism, and recursive mathematics–all disagree on which continuity principles hold, we must be very careful; in particular, in recursive mathematics there is a continuous function on the unit interval which is not uniformly continuous. In [16], Bishop elegantly sidesteps this problem by including the uniform continuity theorem in the definition of continuity: a function f with totally bounded domain is *Bishop-continuous* if f is uniformly continuous on each compact subset of its domain. A *modulus of continuity* for $f : X \to Y$ is a function which takes values[30] $(x, \varepsilon) \in X \times \mathbf{R}^+$ and outputs $\delta > 0$ such that (ε, δ) satisfies the definition of continuity for f at x. A modulus of continuity is a *uniform modulus of continuity* if δ does not depend on x.

The law of trichotomy

$$\forall_{x \in \mathbf{R}} (x < 0 \vee x = 0 \vee 0 < x)$$

implies **LLPO**, and hence is not constructively valid. The *constructive law of trichotomy*

$$\forall_{x,y,z \in \mathbf{R}} (x < y \rightarrow x < z \vee z < y)$$

is sufficient for most arguments.

1.5.2 Constructive reverse mathematics

The goal of a *reverse mathematics* is to classify theorems according to the axioms they require in addition to some basic theory, our *base system*. Results

[29] See [80] for a discussion of continuity in constructive mathematics.

[30] For a subset S of $\mathbf{R}$ we use S^+ to denote the positive elements of S; S^-, S^{0+} are defined similarly.

in reverse mathematics have two parts: first we give a standard mathematical proof of a theorem φ in some extension $Base^{+}$ of our base system $Base$, showing the sufficiency of the principles of that extension; and then we give a reversal, a proof in $Base$ that if φ holds, then so does each theorem of $Base^{+}$. Typically the extensions of our base system are represented by the addition of a single axiom, and a reversal is given by proving that a theorem implies this axiom. The theorems we study in a reverse mathematics will be some subset of those results which are consistent with our base system.

In classical reverse mathematics [100] the base system $\mathbf{RCA}_0$ is roughly the study of computable objects with classical logic, and the base system is extended by set existence axioms. In constructive reverse mathematics we extend or base system by logical axioms and by function existence axioms. The main reason for focusing on function existence axioms rather than set existence axioms is that set membership is generally not decidable in constructive mathematics: function existence axioms give us this decidability. This is related to the adoption of exponentiation, in place of the powerset axiom, in constructive set theories. Logical principles can be viewed as function existence axioms, but this is sometimes disruptive; for example **LLPO** could be viewed as a function from $\{\alpha \in 2^{\mathbf{N}} : \forall_{n,m\in\mathbf{N}}\alpha(n)\alpha(m) = 0\}$ to 2 such that

$$\begin{aligned} \mathbf{LLPO}(\alpha) = 0 &\Rightarrow \forall_{n\in\mathbf{N}}\alpha(2n) = 0; \\ \mathbf{LLPO}(\alpha) = 1 &\Rightarrow \forall_{n\in\mathbf{N}}\alpha(2n+1) = 0. \end{aligned}$$

However this removes the possibility that applying **LLPO** twice to the zero sequence produces different responses—this function form of **LLPO** will be equivalent to Weak König's lemma,

WKL: Every infinite decidable, binary tree has an infinite path,

and hence to **LLPO** plus binary choice for Π^0_1 formulas, Π^0_1-$\mathbf{AC}_{\omega,2}$ [64].

In constructive reverse mathematics we take as our base theory some system of constructive mathematics, and study theorems from the classical, intuitionistic, and recursive schools of mathematics (or indeed any result consistent with our base theory). The extensions of our base theory will be given by some combination of a logical principle and a function existence axiom. The literature on constructive reverse mathematics is somewhat confused; in particular, very little constructive reverse mathematics has been done over an explicit formal system, and there is no formal framework which has gained much attention (see [74] for a critic of constructive reverse mathematics as a programme).

There are three types of constructive reverse mathematics in the literature. The majority of reverse constructive mathematics has been done informally. In informal constructive reverse mathematics, our base theory is Bishop's constructive mathematics, and we focus our attention on fragments of **LEM**, and the semi-constructive principles of **RUSS** and **INT**. It is the expectation that this programme could be formalised in **CST** or some higher order Heyting arithmetic plus dependent choice. If we are interested in constructive mathematics without choice or its numerous models, then we can take **CZF** or **IZF** as our base system and set our focus on the fragments of the axiom of choice, logical principles, and the semi-constructive principles of **RUSS** and **INT**.

Ishihara has proposed a formal approach to constructive reverse mathematics [64] more in line with the classical reverse mathematics of Friedman and Simpson. We take as our base (a subsystem of) elementary analysis **EL**. Elementary analysis is an extension of Heyting arithmetic with variables for number theoretic functions. In addition to the axioms of Heyting arithmetic, **EL** has primitive recursion, λ-conversion, countable choice for quantifier free formulas, and the induction scheme

$$(A(0) \wedge \forall_n (A(n) \to A(n+1))) \to \forall_n A(n),$$

for any formula A, where n ranges over the natural numbers. See [107] for a detailed study of elementary analysis.

Since the focus of constructive reverse mathematics is on logical principles, it is natural to work with a strong set theoretic base theory as in the informal approach. However, we feel that the relationship between **CZF** and topos theory makes it beneficial to omit choice from our base theory. Consequently we adopt **CZF** without choice as our base theory. The next theorem gives an example of some of the results of constructive reverse mathematics. See [65] for a review of informal constructive reverse mathematics, and [64] for constructive reverse mathematics formalised in elementary analysis.

Theorem 24 (CZF) *The following are equivalent to* ***WKL****.*

1. ***LLPO*** *plus* Π^0_1-$\boldsymbol{AC}_{\omega,2}$.
2. *If a sequence of (enumerably) closed subsets in* $\mathbf{R}$ *has the finite intersection property, then it has inhabited intersection.*

The following are equivalent to ***FT*** *for decidable trees.*

1. *An infinite binary tree with at* most one infinite branch—*for all distinct* $\alpha, \beta \in 2^{\mathbf{N}}$ *there exists* n *such that one of* $\alpha(n), \beta(n)$ *is not in the tree—has an infinite branch.*
2. *Each positive valued uniformly continuous function defined on a compact metric space has positive infimum.*
3. *The Heine-Borel theorem for* Δ_0 *open covers: every* Δ_0 *open cover of* $[0, 1]$ *has a finite subcover.*

The following are equivalent to ***BD***-$\mathbf{N}$*.*

1. *Each sequentially continuous mapping from a separable metric space into a metric space is continuous.*
2. *Each bijective continuous linear mapping between Banach spaces has a continuous inverse.*
3. $\mathcal{K}(\mathbf{R})$ *is sequentially complete.*[31]

[31] $\mathcal{K}(\mathbf{R})$ is the uniformly convex space of uniformly continuous functions in $\mathbf{R}^{\mathbf{R}}$ with the seminorms

$$f \mapsto \sup_n \sup_{|x| \geqslant n} 2^{\alpha(n)} |f(x)| \qquad (\alpha \in 2^{\mathbf{N}}).$$

An important principle in reverse constructive mathematics is the minimum principle

> **MIN**: A continuous real valued function on a compact metric space attains its infimum.

In the presence of $\mathbf{AC}_{\omega,2}$, **MIN** is equivalent to **WKL**. **MIN** implies the following weakening of the law of trichotomy

> $\mathbf{LLPO_R}$: for all $x \in \mathbf{R}^D$ either $x \leqslant 0$ or $x \geqslant 0$.

If weak countable choice holds, then **LLPO** and $\mathbf{LLPO_R}$ are equivalent. We present a model of **IZF** below in which **WKL** holds and $\mathbf{LLPO_R}$ fails; it follows, in particular, that **MIN** is strictly stronger than **WKL**.

The little reverse constructive mathematics we indulge in will be done, informally, over **CZF** or **IZF**—our only countermodel is a model of **IZF**, so can be seen as a nonderivability result over any reasonable constructive base theory without choice axioms. For a theorem T, we write $T \in [A, B]$ to mean $(B \to T) \wedge (T \to A)$. For example, many of the fixed point theorems we consider are easily seen to be in $[\mathbf{LLPO_R}, \mathbf{MIN}]$. If we have Π^0_1-$\mathbf{AC}_{\omega,2}$ then $[\mathbf{LLPO}, \mathbf{MIN}]$ collapses to $[\mathbf{WKL}, \mathbf{MIN}]$; if we have $\mathbf{AC}_{\omega,2}$ in our base theory, then $[\mathbf{LLPO}, \mathbf{MIN}]$ collapses to **LLPO**; indeed without choice the implications in

$$\begin{tikzcd}
 & \boxed{\mathbf{WKL}} \arrow[dr] & \\
\boxed{\mathbf{LLPO}} \arrow[ur] & & \boxed{\mathbf{MIN}} \arrow[dl] \\
 & \boxed{\mathbf{LLPO_R}} \arrow[ul] &
\end{tikzcd}$$

are fully indicated.

Some of the principles which occur in constructive mathematics and there interrelations (over **CZF**) are given in Figure 1.

WWKL, weak weak König's lemma, is the restriction of **WKL** to trees with positive measure; $\mathbf{MP}^{\vee}$ is **LLPO** restricted to binary sequences which

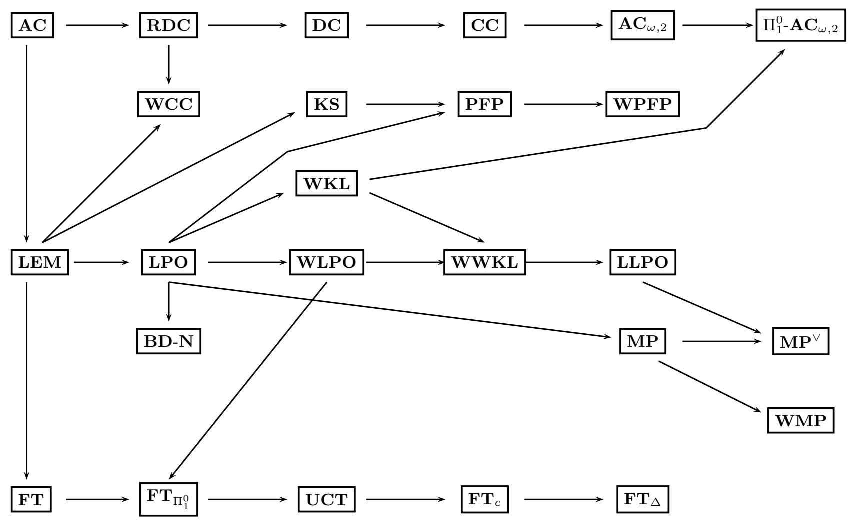


Figure 1: Some classically valid principles of constructive reverse mathematics.

also satisfy the hypothesis of **MP**; and **WMP**, weak Markov's principle, is a strange weakening of **MP** which holds in the three main schools of constructive mathematics, but is independent fo **IZF** [75]. If we add countable or dependent choice to our base system, then more arrows can be added to Figure 1. For example, under countable choice **WKL** and **LLPO** are equivalent and imply **UCT**.

One of the drawbacks of constructive reverse mathematics, which contrasts greatly with classical reverse mathematics, is the number of principles that occur and the complex relationships between them. The above diagram, which we emphasise deals only with principles that arise in mathematical practice, exemplifies this. We give a further example. In [34] it was shown that the statement "every continuous homomorphism θ from $\mathbf{R}$ onto a compact abelian group such that $\{\theta(t) : t > 0\}$ is open has a minimal period"[32] implies that **LLPO** and **WLPO** are equivalent, and Hannes Diener asked

[32]It is shown in [34] that every such homomorphism has a period; that is, there exists $\tau > 0$ such that $\theta(\tau) = \theta(0)$.

whether there is a natural principle representing this gap. We show that **WLPO** is equivalent to the conjunction of **LLPO** and a weakening of the principle of finite possibility:

> (**WPFP**) For each binary sequence $(a_n)_{n\geqslant 1}$, there exists a binary sequence $(b_n)_{n\geqslant 1}$ such that $a_n = 0$ for all n if and only if it is not the case that $b_n = 0$ for all n.

Indeed, this equivalence holds with **LLPO** replaced by $\mathbf{MP}^{\vee}$; we can isolate a tighter bounding principle on the gap between **LLPO** and **WLPO**, but it would be somewhat contrived.

Proposition 25 ***WLPO*** *is equivalent to* $\boldsymbol{MP}^{\vee}$ *+* ***WPFP****.*

On first sight it seems that **LEM** can be decomposed into separate pieces **KS** and **LPO**, but Proposition 25 shows that both **KS** and **LPO** imply **PFP** and hence they are not quite separate.

In order to prove Proposition 25 we need the next lemma, which says that if **LLPO** holds, then given any two binary sequences we can decide in which the first nonzero term does not appear. To make this precise we introduce a function $F : 2^{\mathbf{N}} \to \mathbf{N} \cup \{\infty\}$ defined by

$$F(\alpha) = \inf\{n \in \mathbf{N} : a_n = 1\}.$$

Note that F is not a constructively well defined function. For a given sequence α we may not be able to calculate $F(\alpha)$, but for each natural number n we can decide whether $F(\alpha)$ is less than, equal to, or greater than n. In particular, statements like $F(\alpha) \leqslant F(\beta)$ have a clear interpretation: namely,

- $F(\alpha) \leqslant F(\beta)$ is read as 'if a_n is the first nonzero term of α, then $b_k = 0$ for all $k < n$';
- $F(\alpha) < F(\beta)$ means that 'if a_n is the first nonzero term of α, then $b_k = 0$ for all $k \leqslant n$';

- ▶ $F(\alpha) \neq F(\beta)$ is interpreted as 'for all $x \in \mathbf{N} \cup \{\infty\}$, either $\neg F(\alpha) = x$ or $\neg F(\beta) = x$.

If $F(\alpha) \leqslant F(\beta)$ and $F(\alpha) \neq F(\beta)$, then $F(\alpha) < F(\beta)$.

Lemma 26 ***LLPO*** *is equivalent to the statement*

> *(*) For all binary sequences $\alpha = (a_n)_{n\geqslant 1}$ and $\beta = (b_n)_{n\geqslant 1}$ either $F(\alpha) \geqslant F(\beta)$ or $F(\beta) \leqslant F(\alpha)$.*

Further, if ***LLPO*** *holds and it is impossible that $a_n = 0$ for all n, then $F(\alpha) < F(\beta)$ or $F(\alpha) = F(\beta)$ or $F(\alpha) > F(\beta)$, and* ***MP***$^{\vee}$ *is equivalent to the restriction of $(*)$ to sequences such that it is impossible that $a_n = 0$ for all n.*

Proof. We define a binary sequence $(c_n)_{n\geqslant 1}$ with at most one nonzero term by

$$c_n = \begin{cases} a_{n/2} - \max\{c_k : k \leqslant n\} & n \text{ even} \\ b_{(n+1)/2} - \max\{c_k : k \leqslant n\} & n \text{ odd.} \end{cases}$$

By **LLPO**, either $c_{2n} = 0$ for all n or $c_{2n+1} = 0$ for all n. In the first case suppose that $n = F(\alpha) > F(\beta)$. Then $b_k = 0$ for all $k \leqslant n$ and $a_k = 0$ for all $k < n$, so

$$c_{2n} = a_n - \max\{c_k : k \leqslant n\} = 1$$

—a contradiction. It follows that $F(\alpha) \leqslant F(\beta)$. Similarly, in the second case $F(\alpha) \geqslant F(\beta)$.

Now suppose that it is impossible for $a_n = 0$ for all n; then $\neg F(\alpha) = \infty$. Define $(d_n)_{n\geqslant 1}$ by

$$d_n = \begin{cases} b_{n/2} - \max\{c_k : k \leqslant n\} & n \text{ even} \\ a_{(n+1)/2} - \max\{c_k : k \leqslant n\} & n \text{ odd.} \end{cases}$$

Applying **LLPO** to both $\gamma = (c_n)_{n\geqslant 1}$ and $\delta = (d_n)_{n\geqslant 1}$ we have four possible outcomes:

1. $c_{2n} = 0$ and $d_{2n} = 0$ for all n;

2. $c_{2n+1} = 0$ and $d_{2n} = 0$ for all n;

3. $c_{2n} = 0$ and $d_{2n+1} = 0$ for all n;

4. $c_{2n+1} = 0$ and $d_{2n+1} = 0$ for all n.

In the first and fourth cases we have that $F(\gamma) \leqslant F(\delta) \leqslant F(\gamma)$, so $F(\gamma) = F(\delta)$. In the second case, $F(\alpha) \geqslant F(\beta)$. If $F(\alpha) = F(\beta) = n$, then $d_{2n} = 1$, a contradiction; whence $F(\alpha) > F(\beta)$. A similar argument shows that in the third case $F(\alpha) < F(\beta)$. The last statement follows immediately from the first part of the proof. ■

We can now give the **proof of Proposition 25**:

Proof. Let $\alpha = (a_n)_{n\geqslant 1}$ be a binary sequence; without loss of generality $\alpha = (a_n)_{n\geqslant 1}$ has at most one nonzero term. Using **WPFP**, construct a sequence $\beta = (b_n)_{n\geqslant 1}$ such that

$$\forall_{n\in\mathbf{N}}\,(a_n = 0) \quad \Leftrightarrow \quad \neg\forall_{n\in\mathbf{N}}\,(b_n = 0)\,.$$

By **LLPO** and Lemma 26, either $F(\alpha) \geqslant F(\beta)$ or $F(\beta) \leqslant F(\alpha)$. If $F(\alpha) \geqslant F(\beta)$ and it is not the case that $a_n = 0$ for all n, then $F(\beta) = \infty$, which is absurd. Hence in the first case $\neg\neg(a_n = 0$ for all $n)$; that is, $a_n = 0$ for all n. In the latter case if $a_n = 0$ for all n, then $F(\alpha) = \infty$ again gives a contradiction, so in this case we have $\neg(a_n = 0$ for all $n)$.

Conversely, let $(a_n)_{n\geqslant 1}$ be a binary sequence. By **WLPO**, either $a_n = 0$ for all n or it is impossible for $a_n = 0$ for all n. In the first case we set $b_n = 1$ for each n and in the second we set $b_n = 0$ for all n. Then $(b_n)_{n\geqslant 1}$ satisfies **WPFP** for $(a_n)_{n\geqslant 1}$. If $(a_n)_{n\geqslant 1}$ has at most one nonzero term, then applying **WLPO** to the sequence $(a_{2n})_{n\geqslant 1}$ allows us to decide whether $a_{2n} = 0$ for all n or $a_{2n+1} = 0$ for all n. ■

Rathjen and Chen's extension of Lifschitz realizability to **IZF** (see [39]) gives a model of **IZF** in which **LLPO** holds and **WLPO** is false, and hence a model in which **WPFP** fails.

1.5.3 A useful model

We will explore a little, from the vantage of **ZFC**, the constructive landscape of the sheaf $\mathrm{Sh}(\mathcal{O}([0,1])$ over $[0,1]$ with the standard topology. Recall (Proposition 19) that the Dedekind reals are represented by the continuous real valued functions on $[0,1]$, and the Cauchy reals are represented by the constant functions. Therefore the Cauchy and Dedekind reals are not isomorphic, and hence $\mathbf{AC}_{\omega,2}$ fails. However, by Theorem 18 **LPO** holds in $\mathrm{Sh}(\mathcal{O}([0,1])$, so Π^0_1-$\mathbf{AC}_{\omega,2}$ is validated. We show that even **WCC** fails in $\mathrm{Sh}(\mathcal{O}([0,1])$. We first observe that

$$\mathrm{Sh}(\mathcal{O}([0,1])) \not\models \forall_{x\in\mathbf{R}^D}(x<0 \vee x=0 \vee x>0),$$

since for example $1/2$ does not force that $t \mapsto t-1/2$ satisfies the law of trichotomy. Indeed, by varying our internal Dedekind reals (continuous functions from $[0,1]$ to $\mathbf{R}$), we see that no $x \in [0,1]$ forces the law of trichotomy, and so it is even provably false in $\mathrm{Sh}(\mathcal{O}([0,1])$. If we can show that **WCC** together with **LPO** implies the law of trichotomy, then we must have that $\mathrm{Sh}(\mathcal{O}([0,1]) \not\models$ **WCC**. Fix $x \in \mathbf{R}^D$. By **WCC** there exists a binary sequence $\alpha \in 2^{\mathbf{N}}$ such that

$$\begin{aligned} \alpha(n)=0 \quad &\Rightarrow \quad |x| < \frac{1}{n}; \\ \alpha(n)=1 \quad &\Rightarrow \quad |x| > \frac{1}{n+1}. \end{aligned}$$

Applying **LPO** to α, either $\alpha(n)=0$ for all n and $x=0$, or there exists n such that $\alpha(n)=1$. In this latter case $a \neq 0$, and hence either $x>0$ or $x<0$.

Other intuitionistic non-implications we get from $\mathrm{Sh}(\mathcal{O}([0,1])$ are: **LPO** does not imply the intermediate value theorem, since the former is validated and the latter fails (see [107]); **LLPO** and even **LPO** does not imply that the Cauchy and Dedekind reals are isomorphic.[33] We can also conclude that

[33] I was asked by Hajime Ishihara whether **LLPO** implied the intermediate value theorem in the absence of choice. The possibility that **LLPO** implied $\mathbf{R}^D \cong \mathbf{R}^C$ was raised

for a Dedekind real X and a rational q, the statements $X > q$ and $X \geqslant q$ are not Π^0_1 for otherwise Π^0_1-$\mathbf{AC}_{\omega,2}$ would imply the equivalence of **LPO** and the law of trichotomy.

Since the space of continuous real valued functions on $[0,1]$ with the supremum norm is not complete, not every Cauchy sequence of Dedekind reals in $\mathrm{Sh}(\mathcal{O}([0,1])$ converges to a Dedekind real.

We finish with a sequence $(x_n)_{n\in\mathbf{N}}$ of Dedekind reals such that[34]

$$\mathrm{Sh}(\mathcal{O}([0,1]) \not\models \forall_{x\in\mathbf{R}}\exists_{n\in\mathbf{N}} x \neq x_n.$$

Let $(q_n, r_n)_{n\in\mathbf{N}}$ be an enumeration of $\mathbf{Q}^2$, and let x_n be the linear function joining $(0, q_n)$ and $(1, r_n)$—

$$x_n : t \mapsto (r-q)t + q$$

—and consider $x \in \mathbf{R}$. Let x be an internal Dedekind real. If $n \in \mathbf{N}$ such that $q_n < x(0)$ and $r_n > x(1)$, then by the (approximate) intermediate value theorem $\|x - x_n\|_\infty = 0$, so $\mathrm{Sh}(\mathcal{O}([0,1]) \not\models x \neq x_n$.

at a talk by Michael Rathjen on Lifschitz realizability [39], a model in which weak choice principles fail, but these two definitions of the reals coincide.

[34]This example came out of conversations with Andrew Swan. However, as Peter Schuster pointed out to us, it has already been presented by Bas Spitters [102].

Chapter 2

Some preliminary results

This chapter presents some results that we shall need later. The results of 2.3, on the construction of demand functions, are of independent interest in mathematical economics.

2.1 Fan Theorems and Uniform continuity

Let $2^{\mathbf{N}}$ denote the space of infinite binary sequences, Cantor's space, and let 2^* be the set of finite binary sequences. A subset S of 2^* is *decidable* if for each $a \in 2^*$ either $a \in S$ or $a \notin S$. For two elements $u = (u_1, \ldots, u_m), v = (v_1, \ldots, v_n) \in 2^*$ we denote by $u \frown v$ the *concatenation*

$$(u_1, \ldots, u_m, v_1, \ldots, v_n)$$

of u and v. For each $\alpha \in 2^{\mathbf{N}}$ and each $N \in \mathbf{N}$ we denote by $\overline{\alpha}(N)$ the finite binary sequence consisting of the first N terms of α. A set B of finite binary sequences is called a *bar* if for each $\alpha \in 2^{\mathbf{N}}$ there exists $N \in \mathbf{N}$ such that $\overline{\alpha}(N) \in B$. A bar B is said to be *uniform* if there exists $N \in \mathbf{N}$ such that for each $\alpha \in 2^{\mathbf{N}}$ there is $n \leqslant N$ with $\alpha(n) \in B$. The weakest form of Brouwer's fan theorem is:

> $\mathbf{FT}_{\Delta}$: Every decidable bar is uniform.

Stronger versions of the fan theorem are obtained by allowing more complex bars. A set S is said to be a Π_1^0-set if there exists a decidable subset D of $2^* \times \mathbf{N}$ such that for each $u \in 2^*$ and each $n \in \mathbf{N}$, if $(u, n) \in D$, then $(u \frown 0, n) \in D$ and $(u \frown 1, n) \in D$,

and $S = \{u \in 2^* : \forall_{n \in \mathbf{N}} (u, n) \in D\}$. It is easy to see that the fan theorem for bars which are also Π^0_1-sets

FT$_{\Pi^0_1}$: Every Π^0_1-bar is uniform.

implies **FT**$_\Delta$. The converse (and non-implications between other variants of Brouwer's fan theorem) have recently been shown not to hold in **IZF** [47]. Brouwer's fan theorem is not intuitionistically valid (that is, provable in **CZF**), but is accepted by some schools of constructive mathematics (see [23]).

Each fan theorem has an equivalent formulation where we allow an arbitrary finitely branching tree in the place of the binary fan $2^{\mathbf{N}}$. With this notation, Brouwer's full fan theorem can be stated as follows.

FT: Every bar is uniform.

The fan theorem for a finitely branching tree is reduced to that on $2^{\mathbf{N}}$ by replacing tree width by tree depth: a tree consisting of a root with n branches can be treated as a binary tree with depth $\lceil \log_2(n) \rceil$, possibly with some branches duplicated.

A predicate P on $S \times S \times \mathbf{R}^+$ is said to be a *pointwise continuous predicate on* S if

(i) for each $\varepsilon > 0$ and each $x \in S$, there exists $\delta > 0$ such that if $y, y' \in B(x, \delta)$, then $P(y, y', \varepsilon)$;

(ii) if $\varepsilon > 0$, $x' \in S$, and $(x_n)_{n \geqslant 1}$ is a sequence in S converging to a point x of S and such that $P(x_n, x', \varepsilon)$ for each n, then $P(x, x', \varepsilon)$;

(iii) for all $x, x' \in S$ and each $\varepsilon > 0$ either $P(x, x', \varepsilon)$ or $\neg P(x, x', \varepsilon/2)$.

If the δ in condition (i) can be chosen independently of x, then P is said to be a *uniformly continuous predicate on* S.

For example to each pointwise (resp. uniformly) continuous function $f : S \to \mathbf{R}$ we can associate a pointwise (resp. uniformly) continuous predicate given by

$$P(x, x', \varepsilon) \equiv |f(x) - f(x')| \leqslant \varepsilon.$$

The proof of the following is based on that of Theorem 2 of [46].

Lemma 27 ($\mathbf{AC}_{\omega,2}$) *Assume the fan theorem for Π^0_1-bars. Then every pointwise continuous predicate on* $[0, 1]$ *is uniformly continuous.*

Proof. Let P be a pointwise continuous predicate on $[0, 1]$ and fix $\varepsilon > 0$. Let $X = \{-1, 0, 1\}$ and let X^* be the set of finite sequences of elements of X. Define $f : X^{\mathbf{N}} \to [0, 1]$ by

$$f(\alpha) = \frac{1}{2} + \sum_{n=1}^{\infty} 2^{-(n+1)} \alpha(n);$$

then f is a uniformly continuous function which maps $X^{\mathbf{N}}$ onto $[0, 1]$. Using $\mathbf{AC}_{\omega,2}$, we construct a binary valued function λ on $X^* \times X^*$ such that

$$\begin{aligned} \lambda(u, v) = 1 \quad &\Rightarrow \quad P(f(u \frown \mathbf{0}), f(v \frown \mathbf{0}), \varepsilon); \\ \lambda(u, v) = 0 \quad &\Rightarrow \quad \neg P(f(u \frown \mathbf{0}), f(v \frown \mathbf{0}), \varepsilon/2), \end{aligned}$$

where $\mathbf{0} = (0, \ldots)$. Let D be the set of pairs (u, n) in $X^* \times \mathbf{N}$ such that

> for all $v, w \in X^*$ with lengths at most $n - |u|$, $\lambda(u \frown v, u \frown w) = 1$;

clearly D is a decidable set and if $(u, n) \in D$, then $(u \frown a, n) \in D$ for each $a \in \{-1, 0, 1\}$. Hence

$$B = \{u \in X^* : \forall_{n \in \mathbf{N}} (u, n) \in D\}$$

is a Π^0_1-set. To see that B is a bar, consider any $\alpha \in X^{\mathbf{N}}$. Since P is a pointwise continuous predicate, there exists $\delta > 0$ such that $P(y, y', \varepsilon/2)$ for all $y, y' \in B(f(\alpha), \delta)$. If $N > 0$ is such that $2^{-N} < \delta$, then for all $u, v \in X^*$

$$f(\overline{\alpha}(N) \frown u \frown \mathbf{0}), f(\overline{\alpha}(N) \frown v \frown \mathbf{0}) \in B(x, \delta),$$

so

$$P(f(\overline{\alpha}(N) \frown u \frown \mathbf{0}), f(\overline{\alpha}(N) \frown v \frown \mathbf{0}), \varepsilon/2).$$

Hence $\lambda(\overline{\alpha}(N) \frown u, \overline{\alpha}(N) \frown u) = 1$ for all $u, v \in X^*$ and therefore $\overline{\alpha}(N) \in B$.

Using $\mathbf{FT}_{\Pi^0_1}$, compute $N \in \mathbf{N}$ such that $\overline{\alpha}(N) \in B$ for each $\alpha \in X^{\mathbf{N}}$. Let $x, y \in [0,1]$ be such that $\rho(x, y) < 2^{-(N+1)}$. Then there exist $\alpha, \beta \in X^{\mathbf{N}}$ such that $f(\alpha) = x$, $f(\beta) = y$, and $\overline{\alpha}(N) = \overline{\beta}(N)$. Since $\overline{\alpha}(N) \in B$,

$$P(f(\overline{\alpha}(N) \frown u \frown \mathbf{0}), f(\overline{\alpha}(N) \frown v \frown \mathbf{0}), \varepsilon)$$

for all $u, v \in X^*$. It now follows from condition (ii) of being a pointwise continuous predicate that $P(x, y, \varepsilon)$ holds. Hence P is uniformly continuous.
■

Theorem 28 ($\mathbf{AC}_{\omega,2}$) *Assume the fan theorem for Π^0_1-bars. Then every pointwise continuous predicate on $[0,1]^n$ is uniformly continuous.*

Proof. We proceed by induction on n. The case $n = 1$ is just Lemma 27. Suppose that the result holds for predicates on $[0,1]^{n-1}$, and let P be a predicate on $[0,1]^n$. For each x in $[0,1]$ let P_x be the predicate on $[0,1]^{n-1}$ given by

$$P_x(z, z', \varepsilon) \Leftrightarrow P((z, x), (z', x), \varepsilon).$$

Since P is a pointwise continuous predicate, P_x is pointwise continuous for each $x \in [0,1]$. It follows from our induction hypothesis that each P_x is uniformly continuous. Define a predicate P' on $[0,1]$ by

$$P'(s, t, \varepsilon) \quad \Leftrightarrow \forall_{y \in [0,1]^{n-1}} P_x((s, y), (t, y), \varepsilon).$$

It is easily shown that P' is also a pointwise continuous predicate and that $P'(s, t, \delta)$ holds for all $s, t \in [0,1]$ if and only if $P(x, x', \delta)$ holds for all

$x, x' \in [0,1]^n$. By Lemma 27, P' is uniformly continuous; whence P is uniformly continuous. ■

A predicate P on $S \times \mathbf{R}^+ \times \mathbf{R}^+$ is said to be a *weakly continuous predicate on S* if

(i) for each $x \in S$ and each $\varepsilon > 0$, there exists $\delta > 0$ such that $P(x, \varepsilon, \delta)$;

(ii) if $P(x, \varepsilon, \delta)$ and $|x - y| < \delta' < \delta$, then $P(y, \varepsilon, \delta - \delta')$.

If in addition, for each $\varepsilon > 0$, there exists $\delta > 0$ such that $P(x, \varepsilon, \delta)$ for all $x \in S$, then P is a *uniformly weakly continuous predicate on S.*

Theorem 29 *The statement*

> *Every weakly continuous predicate on $[0,1]^n$ is uniformly weakly continuous.*

is equivalent to the full fan theorem.

Proof. Let P be a weakly continuous predicate on $[0,1]$ and fix $\varepsilon > 0$. Define a uniformly continuous function f from $2^{\mathbf{N}}$ onto $[0,1]$ by

$$f(\alpha) = \sum_{n=1}^{\infty} \left(\frac{2}{3}\right)^{n-1} \left(\frac{(-1)^{a_n} + 1}{2}\right),$$

where $\alpha = (a_n)_{n \geqslant 1}$, and let

$$B = \left\{ a \in 2^* : \forall_{x \in (f(a \frown \mathbf{0}), f(a \frown \mathbf{i}_1))} P\left(x, \varepsilon, (2/3)^{|a|}\right) \right\},$$

where $\frown$ denotes concatenation, $\mathbf{0} = (0, \ldots)$, and $\mathbf{i}_1 = (1, 0, \ldots)$. We show that B is a bar. Let $\alpha \in 2^{\mathbf{N}}$, and, using (i), pick $\delta > 0$ such that $P(f(\alpha), \varepsilon, \delta)$. Pick n such that $(2/3)^{n-1} < 2\delta$. Then

$$(f(\overline{\alpha}(n) \frown \mathbf{0}), f(\overline{\alpha}(n) \frown \mathbf{i_1}))_{(2/3)^n} \subset (f(\alpha) - \delta, f(\alpha + \delta)).$$

It follows from condition (ii) that $\alpha(n) \in B$; whence B is a bar.

By Brouwer's fan theorem, there exists $N > 0$ such that for all $\alpha \in 2^{\mathbf{N}}$ there is $n < N$ with $\alpha(n) \in B$. Then, by condition (ii), $P\left(x, \varepsilon, (2/3)^N\right)$ for all $x \in [0,1]$. This extends to weakly continuous predicates on $[0,1]^n$ by an induction argument similar to that in the proof of Theorem 28.

Conversely, let B be a bar that is closed under extension and define a predicate P by

$$P(x, \varepsilon, \delta) \equiv \forall_x \left(f(x) = \alpha \rightarrow \exists_{N>0}(2^{-N} > \delta \wedge \overline{\alpha}(N) \in B)\right).$$

It is easy to show that P is a pointwise continuous predicate. Hence P is uniformly continuous; in particular, there exists $\delta > 0$ such that $P(x, 1, \delta)$ holds for all $x \in [0,1]$. Pick $N > 0$ such that $2^{-N} < \delta$. Then for all $\alpha \in 2^{\mathbf{N}}$, $\overline{\alpha}(N) \in B$. ■

We present here a generalisation, along similar lines to the above, of a recent result of Hannes Diener [45].[35]

A predicate P on $S^2 \times \mathbf{R}^+$ is said to be a *strongly pointwise continuous predicate on* S if

(i) for each $\varepsilon > 0$ and each $x \in S$, there exists $\delta > 0$ such that if $y, y' \in B(x, \delta)$, then $P(y, y', \varepsilon)$;

(ii)′ if $(x_n)_{n \in \mathbf{N}}$ and $(x'_n)_{n \in \mathbf{N}}$ are sequences in S converging to x and x', respectively, such that $\neg P(x, x', \varepsilon)$, then there exists $n \in \mathbf{N}$ such that $\neg P(x_n, x'_n, \varepsilon)$;

(iii)′ for all $x, x' \in S$, each $\varepsilon > 0$, and any $r \in (0,1)$, either $P(x, x', \varepsilon)$ or $\neg P(x, x', r\varepsilon)$.

As before, if the δ in condition (i) can be chosen independent of x, then P is said to be a *uniformly continuous predicate on* S. Note that the predicates derived from a (uniformly) continuous function is in fact a (uniformly) strongly continuous predicate.

[35] There is no claim that this is an exciting generalisation, but our proof is both considerably simpler and more natural than the intricate proof in [45].

Proposition 30 (CC) ***WKL*** *implies that every strongly pointwise continuous predicate on* $[0,1]$ *is uniform.*

For each $n > 0$ let $S_n = \{0, 2^{-n}, \ldots, 1-2^{-n}\}$ and define $g : 2^{<\omega} \cup 2^{\omega} \to [0,1]$ by

$$g(a) = \sum_{i=0}^{|a|} a(i) 2^{-(i+1)}.$$

We write $\mathcal{D}$ for the set $\cup\{S_n : n \in \mathbf{N}\}$ of dyadic rationals in $[0,1)$; g is a bijection between 2^* and $\mathcal{D}$. For a pointwise continuous predicate P on a subset S of $\mathbf{R}$ we define

$$\varphi_P(\varepsilon, \delta) \equiv \forall_{x,y \in S}(|x-y| < \delta \to P(x,y,\varepsilon)),$$

and we use $\varphi_P^{\neg}(\varepsilon,\delta)$ to denote the existence statement classically equivalent to $\neg\varphi_P(\varepsilon,\delta)$. To ease notation, we write $\varphi_{P|_S}$ for $\varphi_{P|_{S\times S\times \mathbf{R}^+}}$.

Lemma 31 (CC) ***WKL*** $\vdash$ *Let* P *be pointwise continuous predicate on* $[a,b]$ *and let* δ, ε *be positive real numbers. Then there exist* $x, y \in [a,b]$ *such that*

$$\varphi_P^{\neg}(\varepsilon,\delta) \quad \longrightarrow \quad |x-y| < \delta \wedge \neg P(x,y,\varepsilon).$$

Proof. We may assume that $a = 0$ and $b = 1$. Using countable choice, construct a function $\gamma : (\mathbf{Q} \cap [a,b])^2 \times \mathbf{N}$ such that

$$\begin{aligned} \gamma(x,y,n) = 0 \quad &\Rightarrow \quad |x-y| > \delta - 2^{-n} \vee P(x,y,\varepsilon + 2^{-n}); \\ \gamma(x,y,n) = 1 \quad &\Rightarrow \quad |x-y| < \delta \wedge \neg P(x,y,\varepsilon). \end{aligned}$$

Using γ, we construct an increasing binary sequence $(\lambda_n)_{n\in\mathbf{N}}$ such that

$$\begin{aligned} \lambda_n = 0 \quad &\Rightarrow \quad \forall_{x,y\in S_n} \gamma(x,y,n) = 0; \\ \lambda_n = 1 \quad &\Rightarrow \quad \exists_{x,y\in S_n} \exists_{m \leqslant n} \gamma(x,y,m) = 1. \end{aligned}$$

Finally we construct a decidable binary tree T as follows. If $\lambda_n = 0$ we let $T_n = 2^n$, and if $\lambda_{n-1} = 1$ we set

$$T_n = \{\sigma \frown 0 : \sigma \in T_{n-1} \text{ and } |\sigma| = \mathrm{ht}(T)\}.$$

If $\lambda_n = 1 - \lambda_{n-1}$, we let x, y be the minimal elements of S_n such that $\gamma(x, y, n) = 1$ and we set $T_n = 2^n \cup \{g^{-1}(x) \oplus g^{-1}(y)\} \downarrow$—the branch $g^{-1}(x) \oplus g^{-1}(y)$ is the unique branch of T_n with length $\mathrm{ht}(T_n)$, and it codes the witnesses x, y that δ is not a modulus of uniform continuity for ε. Then

$$T = \bigcup_{n \in \mathbf{N}} T_n$$

is an infinite decidable tree.

Using **WKL** we can construct an infinite path α through T. Set $x = g(\pi_0 \alpha), y = g(\pi_1 \alpha)$. Suppose there exist $u, v \in [0, 1]$ such that $|u - v| < \delta$ and $\neg \varphi_P(\varepsilon, \delta)$. Since f is pointwise continuous, there must exist such $u, v \in \mathcal{D}$. Hence $\gamma(u, v, n) = 1$ for some $n \in \mathbf{N}$ such that $u, v \in S_n$, so $\lambda_n = 1$. It now follows from the construction of T that x, y have the desired property. ∎

We recall a lemma of Hajime Ishihara [66]: **WKL** is equivalent to the longest path principle

> **LPP**: Let T be a decidable tree. Then there exists $\alpha \in 2^\omega$ such that for all n, if $\alpha(n) \notin T$, then $T \subset 2^{<n}$.

Here is the **proof of Proposition 30**:

Proof. Let P be a strongly pointwise continuous predicate on $[0, 1]$ and fix $\varepsilon > 0$. We define a function J taking finite binary sequences to subintervals of $[0, 1]$ inductively as follows: we set $J_{<>} = [0, 1]$ and if $J_u = [p, q]$, then we set $J_{u \frown 0} = [p, (p+q)/2]$ and $J_{u \frown 1} = [(p+q)/2, q]$. By repeated application of the lemma, we can construct sequences $(x_u)_{u \in 2^*}, (y_u)_{u \in 2^*}$ such that for each u, $x_u, y_u \in J_u$ and

$$\varphi^{\neg}_{P|_{J_u}}(\varepsilon, 2^{-|u|}) \quad \longrightarrow \quad |x_u - y_u| < 2^{-|u|} \wedge \neg P(x_u, y_u, \varepsilon).$$

Using countable choice again, we construct a decidable tree T such that

$$\begin{aligned} u \in T &\Rightarrow P(x_u, y_u, \varepsilon - 2^{-|u|}); \\ u \notin T &\Rightarrow P(x_u, y_u, \varepsilon). \end{aligned}$$

Let α be a longest path of T, and let ξ be the unique element of

$$\bigcap_{n \in \mathbf{N}} J_{\overline{\alpha}(n)}.$$

Let $\delta > 0$ be such that $P(x, y, \varepsilon/2)$ for all $x, y \in (\xi - \delta, \xi + \delta)$, and let n be such that $2^{-n+1} < \max\{\delta, \varepsilon\}$. If $u = \overline{\alpha}(n) \in T$, then $|x_u - y_u| < \delta$ and $|f(x_u) - f(y_u)| > \varepsilon - 2^{-n} > \varepsilon/2$ contradicting our choice of δ. Hence $\overline{\alpha}(n) \notin T$, so $T \subset 2^n$. It follows that for all $x, y \in [0,1]$, if $|x - y| < 2^{-n}$, then $P(x, y, 2\varepsilon)$. ■

2.2 Boundary crossings

Let C be a located convex subset of a Banach space X. Then for each $\xi \in C^\circ$ and each $z \in -C$ there exists a unique point $h(\xi, z)$ in the intersection of the *interval*

$$[\xi, z] = \{t\xi + (1-t)z : t \in [0,1]\}$$

and the boundary ∂C of C; moreover, the mapping $(\xi, z) \mapsto h(\xi, z)$—the *boundary crossing map* of C—is pointwise continuous on $C^\circ \times -C$ [32, Proposition 5.1.5]. The next lemma shows that for a fixed $\xi \in C^\circ$, this mapping is uniformly continuous.

Lemma 32 *Let X be a bounded convex subset of $\mathbf{R}^N$ and let $\xi \in X^\circ$. Then the function $h : \mathbf{R}^N \to \overline{X}$ which fixes each point of $\overline{X}$ and sends $y \in \sim X$ to the unique intersection point of $[\xi, y]$ and ∂X is uniformly continuous.*

Proof. Without loss of generality we suppose $\xi = 0$. Let $N > 0$ be such that $X \subset B(0, N)$ and let $r > 0$ be such that $B(0, r) \subset X$. Since the function mapping a point $x \neq 0$ to the unique intersection point of

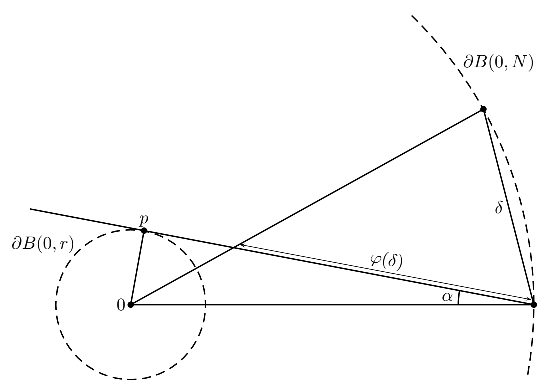
$$\mathbf{R}x = \{rx : r \in \mathbf{R}\}$$

and $\partial B(0,N)$ is uniformly continuous on $-B(0,r/2)$, it suffices to show that h is uniformly continuous on $\partial B(0,N)$.

Given $\delta > 0$, set $\theta = \cos^{-1}(1-(\delta^2/2N^2))$, $\beta = \cos^{-1}(\delta/2N)$ and $\alpha = \sin^{-1}(r/N)$. Define

$$\varphi(\delta) = \frac{\delta|\sin(\beta)|}{|\sin(\alpha+\theta)|}.$$

The function φ is constructed as a 'worst case scenario' given that X contains $B(0,r)$ and is strictly contained in $B(0,N)$; see the following diagram.



Fix $a, b \in \partial B(0,N)$ with $0 < \|a-b\| < \delta$, and let $x \in [0,a] \cap X$ and $y \in [0,b] \cap X$ such that $\|x-y\| > \varphi(\delta)$; without loss of generality, $\|x\| < \|y\|$. It suffices to show that it cannot occur that both $x, y \in \partial X$, for then the assumption that $\|h(x) - h(y)\| > \varphi(\delta)$ leads to a contradiction. By the construction of φ, the unique line passing through x and y must intersect $B(0,r)$. It follows that

$$x \in (\mathrm{conhull}\,(B(0,r) \cup \{y\}))^\circ \subset X^\circ,$$

where conhull(S) is the convex hull of S. Hence if $\|a-b\| < \delta$, then $\|h(a) - h(b)\| \leqslant \varphi(\delta)$.

It only remains to show that for each $\varepsilon > 0$ we can find a $\delta > 0$ such that $\varphi(\delta) < \varepsilon$. From elementary calculations we have that

$$\begin{aligned}
\varphi(\delta) &= \frac{\delta\sqrt{4N^2 - \delta^2}}{2r(1 - (\delta^2/2N^2)) + 2\delta\sqrt{(1 - (r^2/N^2))(1 - (\delta^2/2N^2))}} \\
&\leq \frac{\delta\sqrt{4N^2 - \delta^2}}{2r(1 - (r^2/2N^2)) + 2\delta(1 - (r^2/N^2))} \quad \longrightarrow \quad 0
\end{aligned}$$

as $\delta \to 0$. ∎

2.3 Demand functions

This section gives conditions under which the demand function of a strictly convex preference relation can be constructed, and should be seen as a continuation of the work of Douglas Bridges [22, 24, 25, 28] to examine aspects of mathematical economics in a rigorously constructive manner. In particular, Bridges considered the problem that we consider here in [25]. Corollary 43 is a generalisation of the main result of [25]; our proof, although less elegant, is also somewhat simpler.

Following Bridges we take, as our starting point, the standard configuration in microeconomics consisting of a consumer whose consumption set X is a compact, convex subset of $\mathbf{R}^n$ ordered by a strictly convex preference relation $\succ$. For a given price vector $p \in \mathbf{R}^n$ and a given initial endowment w, the consumers *budget set*

$$\beta(p, w) = \{x \in X : p \cdot x \leqslant w\}$$

is the collection of all consumption bundles available to the consumer.

As detailed in [25], it is easy to show that classically, if $\beta(p, w) \neq \emptyset$, then there exists a unique $\succ$-*maximal* point $\xi_{p,w} \in \beta(p, w)$: $\xi_{p,w} \succcurlyeq x$ for all $x \in \beta(p, w)$. Let T be the set of pairs consisting of a price vector p and an initial endowment w for which $\beta(p, w)$ is inhabited. If the preference relation $\succ$ is continuous, then a sequential compactness argument gives the sequential, and hence pointwise, continuity of the *demand function* F on T which sends (p, w) to the maximal element $\xi_{p,w}$ of $\beta(p, w)$ (see, for example, chapter 2, section D of [104]).

Bridges asked under what conditions can we

1. Compute the demand function F;

2. Compute a modulus of uniform continuity for F: given $\varepsilon > 0$, can we produce $\delta > 0$ such that if $(p, w), (p', w') \in T$ with $\|(p, w) - (p', w')\| < \delta$, then $\|F(p, w) - F(p', w')\| < \varepsilon$.

In [25] Bridges introduced the notion of a uniformly rotund preference relation and showed that if $\succ$ is uniformly rotund and you restrict F to a compact subset of T on which the consumer cannot be satiated, then F is uniformly continuous. Theorem 43 shows that we do not need the hypothesis that our consumer is nonsatiated. Theorems 33 and 41 encapsulate what we can say about strictly convex preference relations, which is more than one might think.

We direct the reader to [22, 24] for an introduction to Bridges' programme to constructivise mathematical economics.

A *preference relation* $\succ$ on a set X is a binary relation which is

- *asymmetric*: if $x \succ y$, then $\neg(y \succ x)$;

- *negatively transitive*: if $x \succ y$, then for all z either $x \succ z$ or $z \succ y$.

If $x \succ y$, we say that x is *preferable* to y. We write $x \succcurlyeq y$, x is *preferable or indifferent to* y, for $\neg(y \succ x)$. We note that $x \succ x$ is contradictory, that $\succ$ and $\succcurlyeq$ are transitive, and that if either $x \succcurlyeq y \succ z$ or $x \succ y \succcurlyeq z$, then $x \succ z$.

Let $\succ$ be a preference relation on a subset X of $\mathbf{R}^N$.

- $\succ$ is a *continuous preference relation* if the graph

$$\{(x, x') : x \succ x'\}$$

of $\succ$ is open.

- $\succ$ is *strictly convex* if X is convex and $tx + (1-t)x' \succ x$ or $tx + (1-t)x' \succ x'$ whenever $x \neq x' \in X$ and $t \in (0,1)$.

- X is *uniformly rotund* if for each $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x, x' \in X$, if $\|x - x'\| \geqslant \varepsilon$, then

$$\left\{ \frac{1}{2}\left(x + x'\right) + z : z \in B(0, \delta) \right\} \subset X,$$

  where $B(x, r)$ is the open ball of radius r centred on x. The preference relation $\succ$ is uniformly rotund if X is uniformly rotund and for each $\varepsilon > 0$ there exists $\delta > 0$ such that if $\|x - x'\| \geqslant \varepsilon$ $(x, x' \in X)$, then for each $z \in B(0, \delta)$ either $\frac{1}{2}(x + x') + z \succ x$ or $\frac{1}{2}(x + x') + z \succ x'$.

A uniformly rotund preference relation is strictly convex.

2.3.1 Constructing maxima

In this section we focus on the construction of maximally preferred elements of a consumption set X. Our main result is

Theorem 33 ($\mathbf{AC}_{\omega,2}$) *Let $\succ$ be a continuous, strictly convex preference relation on an inhabited, compact subset X of Euclidean space. Then there exists a unique $\xi \in X$ such that $\xi \succcurlyeq x$ for all $x \in X$.*

Our proof proceeds by induction. The following lemma provides the key to proving the one dimensional case.

Lemma 34 *Let $\succ$ be a strictly convex preference relation on* $[0, 1]$*. Then either* $1/2 \succcurlyeq x$ *for all* $x \in [0, 1/4)$ *or* $1/2 \succcurlyeq x$ *for all* $x \in (3/4, 1]$.

Proof. Applying the strict convexity of $\succ$ to $1/4 \in (0, 3/4)$, $1/2 \in (1/4, 3/4)$, $3/4 \in (1/2, 1)$ yields

$$\begin{aligned} 1/4 \succ 0 \quad &\text{or} \quad 1/4 \succ 3/4; \\ 1/2 \succ 1/4 \quad &\text{or} \quad 1/2 \succ 3/4; \\ 3/4 \succ 1/4 \quad &\text{or} \quad 3/4 \succ 1. \end{aligned}$$

It follows that either $1/2 \succ 1/4 \succ 0$ or $1/2 \succ 3/4 \succ 1$. In the first case suppose that $x \succcurlyeq 1/2$ for some $x \in [0, 1/4)$. Then, by the strict convexity and transitivity of $\succ$, $1/4 \succ 1/2$; this contradiction ensures that $1/2 \succcurlyeq x$ for all $x \in [0, 1/4)$. Similarly, in the second case $1/2 \succcurlyeq x$ for all $x \in (3/4, 1]$. ■

Lemma 35 ($\mathbf{AC}_{\omega,2}$) *If $\succ$ is a strictly convex, continuous preference relation on $[0,1]$, then there exists $\xi \in [0,1]$ such that $\xi \succcurlyeq x$ for all $x \in [0,1]$.*

Proof. We inductively construct intervals $[\underline{\xi}_n, \overline{\xi}_n]$ such that, for each n,

1. $[\underline{\xi}_n, \overline{\xi}_n] \subset [\underline{\xi}_{n-1}, \overline{\xi}_{n-1}]$;
2. $\overline{\xi}_n - \underline{\xi}_n = (4/5)^n$;
3. for each $x \in [0,1] \setminus [\underline{\xi}_n, \overline{\xi}_n]$,there exists $y \in [\underline{\xi}_n, \overline{\xi}_n]$ such that $y \succcurlyeq x$.

To begin the construction set $\underline{\xi}_0 = 0$ and $\overline{\xi}_0 = 1$. At stage n, rescaling for $n > 1$, we apply Lemma 34; if the first case obtains, then we set $\underline{\xi}_n = (3\underline{\xi}_{n-1} + \overline{\xi}_{n-1})/4$ and $\overline{\xi}_n = \overline{\xi}_{n-1}$. In the second case we set $\underline{\xi}_n = \underline{\xi}_{n-1}$ and $\overline{\xi}_0 = (\underline{\xi}_{n-1} + 3\overline{\xi}_{n-1})/4$. By the transitivity of $\succcurlyeq$, we need only check condition 3. for $[\underline{\xi}_{n-1}, \overline{\xi}_{n-1}] \setminus [\underline{\xi}_n, \overline{\xi}_n]$, and by Lemma 34 $y = (\underline{\xi}_{n-1} + \overline{\xi}_{n-1})/2$ suffices for each such point.

Let ξ be the unique intersection of the $[\underline{\xi}_n, \overline{\xi}_n]$. Since $\succcurlyeq$ is continuous, the maximality of ξ follows from 3. ■

Lemma 36 ($\mathbf{AC}_{\omega,2}$) *If $\succ$ is a strictly convex, continuous preference relation on $[a,b]$, where $a \leqslant b$, then there exists $\xi \in [a,b]$ such that $\xi \succcurlyeq x$ for all $x \in [a,b]$.*

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

$$\begin{aligned} \lambda_n = 0 \quad &\Rightarrow \quad b - a < 1/n; \\ \lambda_n = 1 \quad &\Rightarrow \quad b - a > 1/(n+1). \end{aligned}$$

Without loss of generality, we may assume that $\lambda_1 = 0$. If $\lambda_n = 0$, set $x_n = a$ and if $\lambda_n = 1 - \lambda_{n-1}$, then we apply Lemma 35, after some scaling, to construct a $\succ$-maximal element x in $[a, b]$, and set $x_k = x$ for all $k \geqslant n$. Then for $m > n$, $|x_n - x_m| < 2/(n-1)$, so $(x_n)_{n \geqslant 1}$ converges to some element $\xi \in [a, b]$. If there exists $x \neq \xi$ such that $x \succ \xi$, then $b - a > 0$ and we get a contradiction to Lemma 35. The result now follows from continuity. ∎

We use π_i to denote the i-th projection function, and we write $[x, y]$ for

$$\{tx + (1-t)y : t \in [0, 1]\}.$$

Here is the **proof of Theorem 33:**

Proof. We proceed by induction on the dimension n of the space containing X. Lemma 36 is just the case $n = 1$. Now suppose we have proved the result for n and consider a strictly convex preference relation $\succ$ on a compact, convex subset X of $\mathbf{R}^n$. Define a preference relation $\succ'$ on $\pi_1(X) = [a, b]$ by

$$s \succ_i t \quad \Leftrightarrow \quad \exists_{x \in X} \forall_{y \in X} \left(\pi_1(x) = s \text{ and if } \pi_1(y) = t, \text{ then } x \succ y\right).$$

Then $\succ'$ is strictly convex and sequentially continuous: let $s_1, s_2, t \in [a, b]$ with $s_1 < t < s_2$. By the induction hypothesis there exist ξ_1, ξ_2 such that $\pi_1(\xi_i) = s_i$ and $\xi_i \nsucc x$ for all $x \in X$ with $\pi_1(x) = s_i$ $(i = 1, 2)$. Let z be the unique element of $[\xi_1, \xi_2]$ such that $\pi_1(z) = t$. Then, by the strict convexity of $\succ$, either $z \succ \xi_1$ or $z \succ \xi_2$. In the first case $t \succ' s_1$ and in the second $t \succ' s_2$. Hence $\succ$ is strictly convex. That $\succ'$ is continuous is straightforward.

We can now apply Lemma 36 to construct a maximal element ξ_1 of $(\pi_1(X), \succ'$ $)$, and then the induction hypothesis to construct a maximal element of $S = \{x \in X : \pi_1(x) = \xi_1\}$ with $\succ|_S$. Clearly $\xi = \xi_1 \times \xi_2$ is a $\succ$-maximal element of X. The uniqueness of maximal elements follows directly from the strict convexity of $\succ$. ∎

We shall have need for the following simple corollary, which is of independent interest.

Corollary 37 *Under the conditions of Theorem 33, if $x \in X$ and $x \neq \xi$, then $\xi \succ x$.*

Proof. Let $y = (x+\xi)/2$. Then either $y \succ x$ or $y \succ \xi$. Since $\xi \not\prec y$ the former must attain, so $\xi \not\prec y \succ x$. ∎

If we are not interested in unique maxima, then we might suppose that $\succ$ only satisfies the weaker condition of being *convex*: for all $x, y \in X$ and each $t \in [0,1]$, either $(x+y)/2 \not\prec x$ or $(x+y)/2 \not\prec y$. We give a Brouwerian counterexample to show that this condition is not strong enough to allow the construction of a maximal point. Let $x \in (-1/4, 1/4)$ and let $f_x : [0,1] \to \mathbf{R}$ be the function given by

$$f_x(t) = \begin{cases} \operatorname{sign}(x)(t - \max\{x,0\}) & t \in [0, \max\{x,0\}] \\ 0 & t \in [\max\{x,0\}, 1-\max\{x,0\}] \\ -\operatorname{sign}(x)(t - \max\{x,0\}) & t \in [1-\max\{x,0\}, 1], \end{cases}$$

where[36]

$$\operatorname{sign}(x) = \begin{cases} -1 & x < 0 \\ 0 & x = 0 \\ 1 & x > 0. \end{cases}$$

Define a preference relation $\succ$ on $[0,1]$ by

$$t \succ s \Leftrightarrow f_x(t) > f_x(s).$$

It is easy to see that $\succ$ is continuous and convex. Further, if $x > 0$, then 0 is the unique maximal element, and if $x < 0$, then 1 is the unique maximal element. Now suppose that we can construct $\xi \in [0,1]$ such that $\xi \not\prec t$ for all $t \in [0,1]$; either $\xi > 0$ or $\xi < 1$. In the first case we have $\neg(x > 0)$ and in the second $\neg(x < 0)$, so the statement

> 'Every continuous, convex preference relation on $[0,1]$ has a maximal element'

implies $\forall_{x \in \mathbf{R}}(x \leqslant 0 \vee x \geqslant 0)$.

[36] This is just convenient notation, formally sign is not a constructively well defined function, but the function f does exist constructively. We can define $x \mapsto f_x$ for $x \neq 0$ and extend to all x by continuity.

2.3.2 Continuous demand functions

We now consider a consumer whose consumption set X is a closed convex subset of $\mathbf{R}^n$ ordered by a strictly convex preference relation $\succ$, and who has an initial endowment $w \in \mathbf{R}$. For a given price vector $p \in \mathbf{R}$, a consumers *budget set*

$$\beta(p, w) = \{x \in X : p.x \leqslant w\}$$

is the collection of commodity bundles the consumer can afford. The collection of maximal elements of $\beta(p, w)$ form the consumers *demand set* for price p and initial endowment w.

Lemma 38 *If $p > 0$ and there exists $x \in X$ such that $p{\cdot}x \leqslant w$, then $\beta(p, w)$ is compact and convex.*

Proof. Convexity is clear. See [25] for a proof that $\beta(p, w)$ is compact. ■

We use ∂S to denote the boundary of a subset S of some metric space.

Lemma 39 *The boundary of $\beta(p, w)$ is compact.*

Proof. If X is colocated, then $\rho(x, \partial X) = \max\{\rho(x, X), \rho(x, -X)\}$ and hence the boundary of X is located. Therefore it suffices to show that $-\beta(p, w)$ is located. This is similar to the proof of Lemma 38. ■

It now follows from Theorem 33 that the function F, the consumers *demand function*, that maps (p, w), where p is a price vector and w an initial endowment, to the unique maximal element of $\beta(p, w)$, is well defined. By logical considerations, see section 1.4, we have that any function which can be proven to exist within Bishop's constructive mathematics alone is classically continuous, so the consumers demand function is continuous in the classical setting.

We seek conditions under which F is constructively continuous. We study the continuity of F by looking at the map Γ, on the set T of all inhabited

$\beta(p,w)$, which maps $\beta(p,w)$ to $F(p,w)$. We give T the Hausdorff metric: for located subsets A, B of a metric space Y

$$\rho_H(A,B) = \max\{\sup\{\rho(a,B) : a \in A\}, \sup\{\rho(b,A) : b \in B\}\}.$$

Our next lemma shows how studying Γ allows us to show the continuity of F.

Lemma 40 *If Γ is continuous, then F is continuous. If Γ is uniformly continuous, then for each $p \in \mathbf{R}^n$, $w \mapsto F(p,w)$ is uniformly continuous, and for each $w \in \mathbf{R}$, $p \mapsto F(p,w)$ is Bishop continuous.*

Our next result says that adopting Brouwer's fan theorem is sufficient to prove the classical result that F is continuous when $\succ$ is continuous and strictly convex. We observe that if $\beta(p,w)$ is inhabited and every component of p is positive, then

$$\beta(p,w) = \left\{ x \in \mathbf{R}^n : \sum_{i=1}^{n} p_i x_i \leqslant w \right\}$$

is a diamond, and if the *diameter*

$$\sup\{\rho(x,y) : x, y \in \beta(p,w)\}$$

of $\beta(p,w) > 0$, then $\beta(p,w)$ has inhabited interior.

Theorem 41 *Suppose Brouwer's full fan theorem holds. If $\succ$ is continuous and strictly convex, then F is Bishop continuous.*

Proof. Since **FT** implies that every continuous function on a compact space is uniformly continuous, it suffices, by Lemma 40, to show that Γ is continuous. Fix $\varepsilon > 0$, and $(p,w) \in \mathbf{R}^{n+1}$ such that $\beta(p,w)$ is inhabited; we write $S = \beta(p,w)$ and $\xi = F(p,w)$.

Either $\rho(\xi, \partial S) > 0$ or $\rho(\xi, \partial S) < \varepsilon/2$. In the first case ξ is maximal on the entire set of consumer bundles, so it suffices to set $\delta = \varepsilon$. In the second case, let φ be the natural bijection of $[0,1]^n$ with $T \equiv \partial\beta(p,w) \setminus B_{\rho_1}(x, \varepsilon/2)$; without loss of generality, φ is nonexpansive. We define a predicate on $[0,1]$ by

$$P(x, \alpha, \delta) \quad \Leftrightarrow \quad \forall_{y \in B(\varphi(x),\delta)} \xi \succ y.$$

Then P is a weakly continuous predicate: condition (i) follows from Corollary 37 and the lower pointwise continuity of $\succ$; condition (ii) follows from elementary geometry, given that φ is nonexpansive. By Theorem 29, P is weakly uniformly continuous and hence there exists $\delta > 0$ such that every $y \in B(x, \delta)$ is strictly less preferable than ξ for all $x \in T$. If $\rho(x, S) < \min\{\delta, \varepsilon\}/2$, then $\rho(x, T) < \delta$, $x \in S$, or $x \in B(\xi, \varepsilon)$. In the first two cases $\xi \succ x$; it follows that $F(p', w') \in B(\xi, \varepsilon)$ whenever $\rho_H(\beta(p, w), \beta(p', w') < \min\{\delta, \varepsilon\}/2$. ■

It may seem a little odd that we choose to work in Bishop's constructive mathematics because we are interested in producing results with computational meaning, but that we then add an extra principle **FT** to our framework. In particular, the inconsistency of Brouwer's fan theorem with recursive analysis [23] may cause some consternation. The constructive nature of the fan theorem can be intuitively justified as follows: in order to assert that B is a bar we must have a proof that B is a bar, and a proof is a finite object; therefore an examination of the finite information used in the proof that B is a bar should reveal the uniform bound that the fan theorem gives us. Although this argument does not hold up under scrutiny, with the right formulation, versions of the fan theorem can be proved in some systems of computation [37, 106].

2.3.3 Uniformly rotund preference relations

In order to prove Theorem 41 we effectively strengthened our theory, and therefore weakened our notion of computable. The other natural approach toward proving the existence of a Bishop continuous demand function is to strengthen the conditions on $\succ$. We follow the lead of Bridges in [25] and focus on uniformly rotund preference relations.

Hereafter, we extend the domain of Γ to all inhabited compact convex subsets of X. Theorem 33 still ensures that Γ is well defined.

Theorem 42 *If $\succ$ is a uniformly rotund preference relation, then Γ is uniformly continuous.*

Proof. Let S, S' be compact, convex subsets of X and let ξ, ξ' be their $\succ$-maximal points. Fix $\varepsilon > 0$ and let $\delta' > 0$ be such that if $\|x - x'\| \geqslant \varepsilon$ $(x, x' \in X)$, then for each $z \in B(0, \delta')$ either $\frac{1}{2}(x + x') + z \succ x$ or $\frac{1}{2}(x + x') + z \succ x'$, and set $\delta = \min\{\varepsilon, \delta'\}/2$.

If $\rho_H(S, S') < \delta$, then $\|\xi - \xi'\| \leqslant \varepsilon$: Let S, S' be such that $\rho_H(S, S') < \delta$ and suppose that $\|\xi - \xi'\| > \varepsilon$. Since S, S' are convex

$$S \cap B((\xi + \xi')/2, \delta) \quad \text{and} \quad S' \cap B((\xi + \xi')/2, \delta)$$

are both inhabited; let z be an element of the former set and let z' be an element of the latter. By the maximality of $\xi \in S$ and our choice of δ, $z \succ \xi'$; similarly, $z' \succ \xi$. Therefore

$$\xi \not\succ z \succ \xi' \not\succ z' \succ \xi,$$

which is absurd. Hence $\|\xi - \xi'\| \leqslant \varepsilon$. ∎

As a corollary we have the following improvement on the main result of [25].

Corollary 43 *Let $\succ$ be a uniformly rotund preference relation on a compact, uniformly rotund subset X of $\mathbf{R}^n$, and let S be a subset of $\mathbf{R}^n \times \mathbf{R}$ such that $\beta(p, w)$ is inhabited for each $(p, w) \in S$. Then for each $p \in \mathbf{R}^n$, the function $w \mapsto F(p, w)$ is uniformly continuous, and for each $w \in \mathbf{R}$, the function $p \mapsto F(p, w)$ is Bishop continuous. In particular, F is Bishop Continuous.*

Proof. The result follows directly from Lemma 40 and Theorem 42. ∎

Not surprisingly, a less uniform version of rotundness is enough to give us the pointwise continuity of Γ. A subset X of $\mathbf{R}^n$ is *rotund* if for each $x \in X$

and each $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x' \in X$, if $\|x - x'\| \geqslant \varepsilon$, then

$$\left\{\frac{1}{2}\left(x + x'\right) + z : z \in B(0,\delta)\right\} \subset X.$$

A preference relation $\succ$ is *rotund* if X is rotund and for each $x \in X, \varepsilon > 0$ there exists $\delta > 0$ such that if $\|x - x'\| \geqslant \varepsilon$ $(x' \in X)$, then for each $z \in B(0,\delta)$ either $\frac{1}{2}(x + x') + z \succ x$ or $\frac{1}{2}(x + x') + z \succ x'$.

Theorem 44 *If $\succ$ is a rotund preference relation, then Γ is continuous.*

Proof. The proof is, of course very similar to the proof of Theorem 42. Let S be a compact, convex subset of X and let ξ be the unique maximal element of S. Fix $\varepsilon > 0$. Pick $\delta > 0$ such that if $\|\xi - x\| \geqslant \varepsilon$ $(x \in X)$, then for each $z \in B(0,\delta)$ either $\frac{1}{2}(\xi + x) + z \succ \xi$ or $\frac{1}{2}(\xi + x) + z \succ x'$. If S' is a compact, convex subset of X, with maxima ξ', such that $\rho_H(S, S') < \delta$, then the assumption that $\|\xi - \xi'\| > \varepsilon$ leads to a contradiction as in the proof of Theorem 42. ■

By the next result, Theorem 42 can be used to improve on Theorem 41.

Proposition 45 *Assume Brouwer's full fan theorem. If $\succ$ is continuous and strictly convex, then $\succ$ is uniformly rotund.*

Proof. Without loss of generality,

$$C = \{(x, y) \in X^2 : \|x - y\| \geqslant \varepsilon\}$$

is compact; moreover

$$P((x, y), \varepsilon, \delta) \equiv \|x - y\| < \varepsilon \vee \forall_{z \in B((x+y)/2,\delta)}(z \succ x \vee z \succ y)$$

defines a continuous predicate on C. Hence P is uniformly continuous by Theorem 29, but the uniformity of P says precisely that $\succ$ is uniformly rotund. ■

Corollary 46 *Suppose Brouwer's full fan theorem holds. If $\succ$ is continuous and strictly convex, then Γ is uniformly continuous.*

Chapter 3

Constructing fixed points

In this chapter we consider the problem of constructing the fixed points of the most prominent classical fixed point theorems.

Fixed-point theorems are a major tool in both functional analysis and mathematical economics and are used to prove the existence of solutions to differential equations and the existence of Nash equilibria among other things. Despite this, the constructive literature on fixed point theorems has been scant[37]. There are (at least) two reasons for this:

(i) The standard proof of the simplest and most useful of the well known fixed point theorems, the Banach fixed point theorem, is essentially constructive.

(ii) The nonconstructive nature of Brouwer's fixed point theorem, and the subsequent rejection of this theorem by Brouwer, is well known, and a constructive approximate version for simplices (via Sperner's lemma) is part of the folklore.

Only recently has a fully constructive proof of the approximate version of Brouwer's fixed point theorem, for simplices, been presented [110].

Let X be a metric space and let f be a function from X into X. If $\rho(x, f(x)) < \varepsilon$, then x is called an *ε-fixed point*. A function $f : X \to X$

[37]Although [26] gives a Bishop-style constructive treatment of Edelstein's fixed point theorem; and Kohlenbach [69, Chapter 18] examines contractive and nonexpansive fixed point theorems for computational content using tools from proof theory.

has *approximate fixed points* if for each $\varepsilon > 0$ there exists an ε-fixed point of f in X. If every uniformly continuous function from X into X has approximate fixed points, then X is said to have the *approximate fixed point property.*

This chapter is split into five sections. The first two look at the question 'when can we construct fixed points which exist classically?', the first focuses on the intermediate value theorem in choice free constructive mathematics, and the second looks at the various contraction mapping theorems. The final three sections demonstrate the construction of approximate fixed points for Brouwer's fixed point theorem, and the two most commonly applied extensions, those of Schauder and Kakutani.

3.1 The intermediate value theorem

In this section we study the intermediate value theorem as the simplest interesting case of the question

> Given a nonconstructive fixed point theorem, under what extra conditions can we construct a fixed point?

In particular we study the question: when can we construct a (non-unique) fixed point without choice axioms?

3.1.1 The intermediate value theorem in CLASS and BISH

The intermediate value theorem

> **IVT**: Let $f : [a, b] \to \mathbf{R}$ be a continuous function such that $f(a) \leqslant 0 \leqslant f(b)$. Then there exists an $x \in [a, b]$ such that $f(x) = 0$.

has a surprisingly rich history within Bishop's constructive mathematics. There are two standard classical proofs of this result. In the first we define $x = \sup\{y : f(y) \leqslant 0\}$ and show that this x is a root of f. Constructively, however, this supremum may not exist; the supremum of a subset X of $\mathbf{R}$

exists constructively if and only if X is *upper order located*: for all $\alpha, \beta \in \mathbf{Q}$, if $\alpha < \beta$, then either $x < \beta$ for all $x \in X$ or there exists $x \in X$ with $\alpha < x$.

In the second proof we 'construct' a sequence of nested intervals $[a_n, b_n]$, whose lengths tends to zero, such that $f(a_n) \leqslant 0 \leqslant f(b_n)$ for all n. This is done using interval halving: set $a_1 = a$ and $b_1 = b$. Suppose that we have constructed $(a_k)_{k=1}^{n-1}$ and $(b_k)_{k=1}^{n-1}$, and let m be the midpoint of $[a_{n-1}, b_{n-1}]$. Either $f(m) \geqslant 0$ or $f(m) \leqslant 0$. In the first case we set $a_n = a_{n-1}, b_n = m$, and in the second we set $a_n = m, b_n = b_{n-1}$. The desired root is then given by the shared limit of $(a_n)_{n \geqslant 1}$ and $(b_n)_{n \geqslant 1}$, which are nondecreasing and nonincreasing, respectively. The problem with this argument constructively is in deciding, for each n, whether the value of f at the midpoint of $[a_{n-1}, b_{n-1}]$ is greater than or equal to zero, or less than or equal to zero: the statement

$$\forall_{x \in \mathbf{R}} x \geqslant 0 \wedge x \leqslant 0$$

implies **LLPO** (and is equivalent to **LLPO** under **WCC**). So neither of these proofs is constructive, and in fact **IVT** implies **LLPO**, and in the presence of Π^0_1-$\mathbf{AC}_{\omega,2}$, **IVT** and **LLPO** are equivalent. The first constructive version of **IVT**, the approximate intermediate value theorem, was given by Bishop in [16] and showed, using choice, the existence of approximate solutions of arbitrary precision:

> **aIVT**: Let $f : [a, b] \to \mathbf{R}$ be a (uniformly) continuous function such that $f(a) \leqslant 0 \leqslant f(b)$. Then for all $\varepsilon > 0$ there exists $x \in [a, b]$ such that $|f(x)| < \varepsilon$.

Bishop's proof of **aIVT** essentially uses the classical interval halving argument above, stopping when the value of $|f|$ at the midpoint of $[a_{n-1}, b_{n-1}]$ is known to be less than ε. In [23, Chapter 3, Theorem 2.5], Bridges and Richman gave an exact constructive version of the intermediate value theorem[38] requiring only the additional hypothesis that f be *locally nonzero*: for each

[38] In the presence of countable choice, we only require f to be sequentially continuous for both **aIVT** and this exact version of the intermediate value theorem (see exercise 11, page 21 of [32]).

$x \in (a,b)$ and each $\varepsilon \in (0, \min\{x-a, b-x\})$, there exists $y \in (x-\varepsilon, x+\varepsilon)$ such that $f(y) \neq 0$. Once again the proof proceeds by an interval halving argument, this time with the midpoint of $[a_{n-1}, b_{n-1}]$ perturbed slightly, if necessary, so that the function evaluation is nonzero. The known proofs of this exact constructive intermediate value theorem use some form of countable choice, and it is shown below that some form of choice is necessary.

A root x of a function $f : \mathbf{R} \to \mathbf{R}$ is *stable* if

$$\forall_{\varepsilon>0} \exists_{y,z \in (x-\varepsilon, x+\varepsilon) \cap I} (f(y) < 0 < f(z));$$

that is, a root x of f is stable if f crosses the horizontal axis arbitrarily close to x. Both of the above classical proofs and the constructive proof of the intermediate value theorem for locally nonzero functions, $\mathbf{IVT_{loc}}$, produce stable roots. Indeed a function satisfying the hypothesis of the intermediate value theorem might only have stable roots, and constructively we could never hope to construct a root which was not stable.

3.1.2 The intermediate value theorem in CZF

So under what conditions can we construct, in **CZF**, roots of a continuous function which crosses the x- axis? The answer from Bishop's constructive mathematics, when the function is locally nonzero, requires some form of the axiom of choice, at least when working with the Dedekind reals.

Proposition 47 (IZF) *The intermediate value theorem restricted to locally nonzero functions on the Dedekind reals is independent of **IZF**.*

Proof. We give a topological space (X, T_X) such that the topological model over X satisfies **IZF** plus $\neg\mathbf{IVT_{loc}}$. The points of X are uniformly continuous, locally nonzero, real valued functions on $[0,1]$ such that $-f(0) = f(1) = 1$ for each $f \in X$. A basic open $U \in T_X$ is a pair (U_1, U_2) where U_1 is a finite subset of $(0,1) \times \mathbf{R} \setminus \{0\}$ and U_2 is a finite subset of $\mathbf{Q}^+ \times \mathbf{Q}^+$. A function $f \in X$ is in U if

(i) $U_1 \subset \mathrm{graph}(f)$;

(ii) for each $(\varepsilon, \delta) \in U_2$, there exists $\delta' < \delta$ such that δ' is an ε-modulus of uniform continuity for f.

Let f_G be the function associated to the generic. We first show that f_G is a total function; we show that if $U \vdash x \in [0,1]$, then for all $p, q \in \mathbf{Q}$, $U \vdash (g(x) < q) \vee (p < g(x))$. Set $\varepsilon = (q-p)/2$. Then the basic opens $\{(y,z), (\varepsilon, 1/n)\}$, where $n \in \mathbf{N}$, z is a nonzero real and $U \vdash |x-y| < 1/n$, cover U; whence $U \Vdash (g(x) < q) \vee (p < g(x))$. Then f_G is uniformly continuous, since for all $\varepsilon > 0$ the opens $(\emptyset, \{(\varepsilon, 1/n)\})$ $(n \in \mathbf{N})$ cover X, and f_G is locally nonzero since, given any open interval I in $[0,1]$ and any $U \in T$, if U does not already force f_G to be nonzero on I, then we can extend U by adding (x,r) to U_1 for some $x \in I$ and some sufficiently small nonzero r. To see that X forces that there does not exist x such that $f(x) = 0$, we show that $U \nVdash f_G(\sigma) = 0$ for every name σ and every nonempty $U \in T_X$. For suppose $U \Vdash \sigma \in \mathbf{R}$. Since U_1, U_2 are finite, we can pick $a, r > 0$ sufficiently small such that the opens

$$U_q = (U_1 \cup \{(q,b)\}, U_2 \cup \{(a/2, r)\}$$

are nonempty for all rational q in $[0,1]$ and some $b \geqslant a$ depending on q (such a b can be given explicitly, so this does not require choice). Then U forces σ to be within r of some rational $q \in [0,1]$, and hence U_q is an extension of U forcing that $f_G(\sigma) \neq 0$. ■

Proposition 47 shows that **IVT** for locally nonzero functions requires some form of choice. However, we show below that this version of **IVT** holds in the sheaf model $\mathrm{Sh}(\mathcal{O}[0,1])$, and hence that **IVT** for locally nonzero functions does not imply any of the choice principles which fail in that model. We take **IZF** plus **CC** as our metatheory since we will need to apply $\mathbf{IVT_{loc}}$, to produce a stable root, at the meta-level; we only distinguish the Dedekind and Cauchy reals in the model, so $\mathbf{R}^D, [0,1]^D$ will be used for the internal Dedekind reals. Let $f : [0,1]^D \to \mathbf{R}^D$ be a uniformly continuous, locally nonzero function in $\mathrm{Sh}(\mathcal{O}[0,1])$. Since f is locally nonzero, we have that for all positive $\varepsilon \in \mathbf{Q}$ and all $x \in [0,1]^D$ there exists $y \in [0,1]^D$ such that[39] $\|x-y\|_\infty < \varepsilon$ and $f(y) \neq 0$ for all $t \in [0,1]$. By $\mathbf{IVT_{loc}}$, in the metatheory, we can construct a stable root $s \in [0,1]$ of the function $r \mapsto f(\mathbf{r})(1/2)$, where $\mathbf{r}$ is the constant function on $[0,1]$ with codomain $\{r\}$. If $f(\mathbf{s})$ is not the constant zero function, then it follows from the continuity of f that there exists $r \in [0,1]$ such that $f(\mathbf{r})$ has a stable root. Since this contradicts that f is locally nonzero, $\mathbf{s}$ is a root of f in $\mathrm{Sh}(\mathcal{O}[0,1])$. This sheaf model also shows that **LPO** does not imply the intermediate value theorem.

We now turn our attention to obtaining positive results. For completeness we begin with a proof of the approximate version of the intermediate value theorem, and by giving the choice free constructive content of the classical proof of **IVT** by interval halving.

Proposition 48 *Let $f : [a,b] \to \mathbf{R}$ be a continuous function such that $f(a) \leqslant 0 \leqslant f(b)$. Then for all $\varepsilon > 0$ there exists $x \in [a,b]$ such that $|f(x)| < \varepsilon$.*

Proof. Since f is continuous and $[a,b]$ is connected, $f([a,b])$ is connected. Hence it is not the case that $f([a,b])$ is bounded away from 0; it follows that $\inf\{|f(x)| : x \in [a,b]\} = 0$. ■

The classical interval halving argument yields the following proposition within **CZF**.

Proposition 49 *Let $f : [a,b] \to \mathbf{R}$ be a function such that $f(a) < 0 < f(b)$ and $f(x) \neq 0$ for all $x \in [0,1]$. Then there exists $x \in [0,1]$ such that $f(x)$ is not sequentially continuous at x: there exists a sequence $(x_n)_{n\geqslant 1}$ converging to x such that $(f(x_n))_{n\geqslant 1}$ is bounded away from $f(x)$.*

Proof. Construct sequences $(a_n)_{n\geqslant 1}$, $(b_n)_{n\geqslant 1}$ as follows. Set $a_1 = a, b_1 = b$. Suppose we have constructed $a_1, \ldots, a_{n-1}$ and $b_1, \ldots, b_{n-1}$, and let $m = (a_{n-1} + b_{n-1})/2$. Since $f(m) \neq 0$, either $f(m) > 0$ or $f(m) < 0$. In the first case we set $a_n = a_{n-1}, b_n = m$, and in the second we set $a_n = m, b_n = b_{n-1}$. Then

[39] The distance between x and y in the model $\mathrm{Sh}(\mathcal{O}[0,1])$ is, externally, $\|x-y\|_\infty$.

- $f(a_n) < 0$ for all $n \in \mathbf{N}$;
- $f(b_n) < 0$ for all $n \in \mathbf{N}$;
- $(a_n)_{n\geqslant 1}$ and $(b_n)_{n\geqslant 1}$ converge to a shared limit x.

Either $f(x) > 0$ or $f(x) < 0$. In the first case $a_n \to x$, but $f(a_n) < 0$ for all n; therefore $(f(a_n))_{n\geqslant 1}$ is bounded away from $f(x)$. In the second case $(b_n)_{n\geqslant 1}$ defies sequential continuity. ■

Corollary 50 *Let $f : [a, b] \to \mathbf{R}$ be a sequentially continuous function such that $f(a) \leqslant 0 \leqslant f(b)$. Then $\neg\left(\forall_{x\in[0,1]}\ f(x) \neq 0\right)$.*

The above corollary has a further simple consequence which, in view of Theorem 47, is a little surprising. The previous result says that the function f cannot fail to have a root; with this in mind we define: $x \in \mathbf{R}$ is said to be *within ε of a root of f* if

$$\neg\neg\exists_{y\in\mathbf{R}}(|x - y| < \varepsilon \wedge f(y) = 0).$$

Corollary 51 *Let $f : [a, b] \to \mathbf{R}$ be a sequentially continuous function such that $f(a) \leqslant 0 \leqslant f(b)$. Then for each $\varepsilon > 0$ there exists $x \in (a, b)$ such that x is within ε of a root of f.*

Proof. Pick $n \in \mathbf{N}$ such that $\varepsilon > (a - b)/n$. Let $\xi_0 = a, \xi_{2n} = b$ and for each $0 < i < 2n$ let $\xi_i \in [a, b]$ be such that

$$\left|\xi_i - \frac{(a-b)i}{2n}\right| < \varepsilon/2 \qquad \text{and} \qquad f(\xi_i) \neq= 0$$

—this is an instance of finite choice. Then the function

$$g(x) = \begin{cases} \frac{f(\xi_0)-f(\xi_1)}{\xi_0-\xi_1}(x - \xi_0) + f(\xi_0) & x \in [\xi_0, \xi_1] \\ \vdots & \\ \frac{f(\xi_{2n-1})-f(\xi_{2n})}{\xi_{2n-1}-\xi_{2n}}(x - \xi_{2n-1}) + f(\xi_{2n-1}) & x \in [\xi_{2n-1}, \xi_{2n}] \end{cases}$$

has an inhabited and finite set of roots each element of which is, by Corollary 50, within ε of a root of f. ■

Using countable choice we can, as in the proof of this last corollary, construct a sequence of functions $(g_n)_{n\in\mathbf{N}}$ such that for each n, the set of roots of g_n is an inhabited, finite set, and each root of g_n is within 2^{-n} of a root of f. Using countable choice again (to repeatedly apply the approximate law of trichotomy), we can form an inhabited, finitely branching forest F of decidable trees with labels from $[a, b]$ such that

- the nodes in the n^{th} level of any T in F are the zeros of g_n,
- if a node labeled x is a child of a node labeled y and y is a root of g_n, then $|x - y| < 2^{-n+1}$,
- the branches of F represent all stable roots of f as Cauchy reals.

Then each T in F is a spread (that is, has no leaves), and any branch of F realizes the intermediate value theorem for locally nonzero functions. So we can recover the locally nonzero version of the intermediate value theorem.

We now turn our attention to the construction of exact roots. For the remainder of this section we use I to denote either a closed bounded interval or the entire real line. Let f be a continuous real valued function on I. Since the property of all the roots of a function f being stable only gives us positive information about f when we already have a root of f, it seems unlikely that this condition is strong enough to prove a constructive version of the intermediate value theorem, even when we allow choice. We give a strengthening of the notion of a function having only stable roots. A function $f : I \to \mathbf{R}$ is said to have *uniformly stable roots* if

$$\forall_{\varepsilon>0}\exists_{\delta>0}\forall_{x\in I}(|f(x)| < \delta \Rightarrow \exists_{y,z\in(x-\varepsilon,x+\varepsilon)\cap I} f(y) < 0 < f(z)).$$

It is easy to see that if f has uniformly stable roots, then f is locally nonzero and has only stable roots.

Since, given a function $f : [a, b] \to \mathbf{R}$ with $f(a) \leqslant 0$ and $f(b) \geqslant 0$, we need only construct the smallest root of f, we only require f to have *uniformly upper-stable roots*:

$$\forall_{\varepsilon>0}\exists_{\delta>0}\forall_{x\in(a,b)}(|f(x)| < \delta \Rightarrow \exists_{z\in(x-\varepsilon,x+\varepsilon)}f(z) > 0).$$

Our choice free constructive version of the intermediate value theorem is the following.

Proposition 52 *Let $f : [a, b] \to \mathbf{R}$ be a uniformly continuous function with uniformly upper-stable roots such that $f(a) \leqslant 0$ and $f(b) \geqslant 0$. Then there exists $x \in [a, b]$ such that $f(x) = 0$ and $f(y) < 0$ for all $y \in [0, x)$.*

Before proving Proposition 52, we give some examples of functions with uniformly stable roots. If $f : \mathbf{R} \to \mathbf{R}$ is a uniformly continuous, continuously differentiable function which only has roots of multiplicity one in the following strong sense

> There exists $r > 0$ such that $\max\{|f(x)|, |f'(x)|\} > r$ for all $x \in \mathbf{R}$.

then f has uniformly stable roots. To see this, fix $\varepsilon > 0$ and let $r > 0$ be such that $\max\{|f(x)|, |f'(x)|\} > r$ for all $x \in \mathbf{R}$; without loss of generality $\varepsilon < 1/2$ and $r < 1$. Since f is uniformly continuous, there exists $\delta_1 \in (0, \varepsilon)$ such that if $|x-y| < \delta_1$, then $|f(x)-f(y)| < r/2$. Set $\delta = \delta_1 r$ and fix $x \in \mathbf{R}$ such that $|f(x)| < \delta$. Then for all $y \in (x - \delta_1, x + \delta_1)$

$$\begin{aligned} |f(y)| &\leqslant |f(x)| + |f(y) - f(x)| \\ &< \delta + r/2 \\ &< r, \end{aligned}$$

so, by our choice of r, $|f'(y)| > r$. Suppose that there exist $y, y' \in (x-\delta_1, x+\delta_1)$ such that $f'(y) < 0 < f'(y')$. Applying **aIVT** to $f'|_{(\min\{y,y'\},\max\{y,y'\})}$ produces $z \in (\min\{y, y'\}, \max\{y, y'\})$ such that $|f'(z)| < r$, which is a contradiction. Hence either $f'(y) \geqslant r$ for all $y \in (x - \delta_1, x + \delta_1)$ or else

$f'(y) \leqslant -r$ for all $y \in (x-\delta_1, x+\delta_1)$. In the first case

$$\begin{aligned} f(x+\delta_1) &\geqslant r\delta_1 + f(x) = \delta + f(x) > 0, \text{ and} \\ f(x-\delta_1) &\leqslant -r\delta_1 + f(x) = -\delta + f(x) < 0. \end{aligned}$$

In the second case $f(x+\delta_1) < 0 < f(x-\delta_1)$. Since $x+\delta_1, x-\delta_1 \in (x-\varepsilon, x+\varepsilon)$, f has uniformly stable roots.

If f is a polynomial, then we can do better. First we require a few notions from [92]. We denote by $\pi_n(\mathbf{R})$ the set of monic polynomials of degree n with real coefficients given any of the standard metrics. An *n-multiset* is an image of $\{1, 2, \ldots, n\}$. The distance between two n-multisets $X = \{x_1, x_2, \ldots, x_n\}$ and $Y = \{y_1, y_2, \ldots, y_n\}$ is given by

$$\rho_n(X, Y) = \inf_{\sigma \in S_n} \sup\{|x_i - y_{\sigma(i)}| : 1 \leqslant i \leqslant n\},$$

where S_n denotes the group of permutations on n elements; the set $M_n(\mathbf{C})$ of n-multisets of complex numbers taken with ρ_n is a pseudometric space. In [92] it is shown that the natural map F from $M_n(\mathbf{C})$ to $\pi_n(\mathbf{R})$ is one-one and uniformly bicontinuous.

Let

$$p(x) = \sum_{i=1}^{n} (x - r_i) \qquad\qquad (r_i \in \mathbf{C})$$

be a polynomial[40] such that $r_i \neq r_j$ for all $1 \leqslant i < j \leqslant n$. Then p has uniformly stable roots. We prove that there exists $r > 0$ such that for all $x \in \mathbf{R}$ either $|p(x)| > r$ or $|p'(x)| > r$; this is done by induction on the degree of p. The case where p is linear is trivial. Suppose that the result holds for polynomials of degree $n-1$ and write

$$p(x) = (x - r_1) \sum_{i=2}^{n} (x - r_i) \equiv (x - r_1) q(x).$$

[40] Of course, the application of Proposition 52 to such polynomials is not very interesting.

Then by our induction hypothesis there exists $r' \in (0,1)$ such that for all $x \in \mathbf{R}$ either $|q(x)| > r'$ or $|q'(x)| > r'$. Fix $x \in \mathbf{R}$ and let $R \in \mathbf{R}$ be such that $p'(y) < R$ for all $y \in [r_1 - 1, r_1 + 1]$; without loss of generality $r' < R$. If $|q(x)| > r'$, then either $|x - r_1| > r'/3R$ and

$$|p(x)| = (x - r_1)q(x) > \frac{r'}{3R}r' = \frac{r'^2}{3R},$$

or $|x - r_1| < 2r'/3R$. In the latter case

$$|(x - r_1)q'(x)| < \frac{2r'}{3R}R = \frac{2r'}{3},$$

so

$$|p'(x)| = |(x - r_1)q'(x) + q(x)| > \frac{r'}{3}.$$

It remains to consider the case $|q'(x)| > r'$. Let $\varepsilon > 0$ be such that $|r_1 - r_i| > \varepsilon$ for all $i \in \{2, \ldots, n\}$. Using the uniform bicontinuity of the natural map F sending n-multisets of $\mathbf{C}$ to polynomials of degree n, there exists $\delta > 0$ such that for all $q' \in \pi_n(\mathbf{R})$ if $\rho(q, q') < \delta$, then $\rho_n(F^{-1}(q), F^{-1}(q')) < \varepsilon$. Suppose that $|q(r_1)| < \delta$. Then $\rho(q, q + q(r_1)) < \delta$, so $\rho_n(F^{-1}(q), F^{-1}(q + q(r_1))) < \varepsilon/2$. Since $q + q(r_1)$ has a root at r_1, there exists $i \in \{1, \ldots, n\}$ such that $|r_1 - r_i| < \varepsilon$, which is absurd. It follows that $|q(r_1)| \geqslant \delta > 0$. By the continuity of q at r_1, there exists $\delta' \in (0, \min\{r'/R, r'\})$ such that for all $x \in \mathbf{R}$ either

$$|x - r_1| > \delta' \text{ or } |q(x)| > \delta'.$$

In the first case either $|q(x)| > r'\delta'/3$ and

$$|p(x)| = |(x - r_1)q(x)| > \frac{r'\delta'}{3}\delta' = \frac{r'\delta'^2}{3},$$

or $|q(x)| < 2r\delta'/3$. If $|q(x)| < 2r'\delta'/3$, then $|(x - r_1)q'(x)| > r'\delta'$ and

$$|p'(x)| = |(x - r_1)q'(x) + q(x)| > \frac{r'\delta'}{3} > \frac{r'\delta'^2}{3}.$$

In the second case $|x-r_1| > \delta'/3R$ and so $|p(x)| > \delta'^2/3R$, or $|x-r_1| < \delta'/2R$ and $|p'(x)| > \delta'/2 > \delta'^2/3R$. Given that $\delta' < r' < 1$, it just remains to set

$$r = \min\left\{\frac{r'\delta'^2}{3}, \frac{\delta'^2}{3R}\right\};$$

then for all $x \in \mathbf{R}$ either $|p(x)| > r$ or $|p'(x)| > r$.

The next lemma is a (very) slight generalisation of a result in [97] which considers strictly increasing functions.

Lemma 53 *Let f be a locally nonzero nondecreasing real valued function on $[a,b]$ such that $f(a) \leqslant 0$ and $f(b) \geqslant 0$. Then there exists a unique $x \in [a,b]$ with $f(x) = 0$.*

Proof. Without loss of generality $a = 0$ and $b = 1$. Define

$$L = \{y \in [0,1] : f(y) \leqslant 0\};$$

it is easily shown that the required x is given by the supremum of L provided that this supremum exists. Since $f(0) \leqslant 0$, $0 \in L$ so L is nonempty. Let $\alpha, \beta \in \mathbf{Q}$ be such that $\alpha < \beta$; without loss of generality $\alpha, \beta \in [0,1]$. Since f is locally nonzero, there exist α', β' such that $\alpha < \alpha' < \beta' < \beta$ and $f(\alpha')$ and $f(\beta')$ are nonzero. If $f(\alpha') > 0$ and $f(\beta') < 0$, then $f(\beta') < 0 < f(\alpha')$, which contradicts that f is nondecreasing. Hence either $f(\alpha') < 0$ or $f(\beta') > 0$. If $f(\alpha') < 0$, then $f(\alpha) < 0$ and $\alpha \in L$. If $f(\beta') > 0$, then $f(y) < f(\beta')$ for each $y \in L$, so, since f is nondecreasing, $y \leqslant \beta$ for each $y \in L$ and β is an upper bound for L. Hence L is upper order located and therefore $x = \sup(L)$ exists.

It remains to show that this x is unique. To this end, let $x, y \in [0,1]$ be such that $f(x) = f(y) = 0$, and suppose that $x \neq y$. Without loss of generality $x < y$. Since f is locally nonzero, there exists $z \in (x,y)$ such that $f(z) \neq 0$; either $f(z) > 0$ or $f(z) < 0$. In the first case we have that $f(z) < f(x)$ and $x < z$, contradicting that f is nondecreasing. In the second case, $f(y) < f(z)$ and $z < y$, which again gives a contradiction. Hence $x = y$ and f has a unique root. $\blacksquare$

Here is the **proof of Proposition 52**:

Proof. We may assume that $a = 0$ and $b = 1$. Define a nondecreasing function $g : [0,1] \to \mathbf{R}$ by[41]

$$g(x) = \sup\{f(y) : y \in [0,x]\};$$

it is easily shown that g is uniformly continuous. Then any zero of g is a zero of f: let $x \in [0,1]$ be a zero of g and suppose that $f(x) \neq 0$. If $f(x) > 0$, then $g(x) > 0$, which is absurd. Therefore $f(x) < 0$. Using the continuity of f at x, pick $r > 0$ such that $f(y) < f(x)/2$ for all $y \in (x - r, x)$. Let $\delta \in (0, |f(x)|/2)$ be such that for all $x' \in [0,1]$ if $|f(x')| < \delta$, then there exists $z \in (x' - r, x' + r)$ with $f(z) > 0$. Since

$$g(x) = \sup\{f(y) : y \in [0,x]\} = 0,$$

there exists $x' \in [0,x]$ such that $|f(x')| < \delta$; moreover, since $\delta < |f(x)|/2$, $x' \leqslant x - r$. Then, by our choice of δ, there exists $z < x' + r \leqslant x$ such that $f(z) > 0$, and so $g(x) > 0$. This final contradiction ensures that $f(x) = 0$.

Thus in order to show that f has a zero we need only show that g has a zero; in turn, by Lemma 53, it suffices to show that g is locally nonzero. To this end, fix $x \in [0,1]$ and $\varepsilon > 0$; without loss of generality $x \notin \{0,1\}$ and $\varepsilon < \min\{x, 1-x\}$. Using that f has uniformly upper-stable roots, pick $\delta > 0$ such that for all $x \in [0,1]$, if $|f(x)| < \delta$, then there exists $z \in (x - \varepsilon, x + \varepsilon)$ such that $f(z) > 0$.

Either $|g(x)| > 0$ or, as we may assume, $|g(x)| < \delta$. Then there exists $y \in [0,x]$ such that $|f(y)| < \delta$; whence, by our choice of δ, there exists

[41] The supremum of $\{f(y) : y \in [0,x]\}$ is given by

$$\left(\{q \in \mathbf{Q} : \exists_{y\in[0,x]}(q < f(y))\}, \{q \in \mathbf{Q} : \forall_{y\in[0,x]}(f(y) < q)\}\right).$$

This is a Dedekind real because f is uniformly continuous and $[0,x]$ is totally bounded.

$z \in (y - \varepsilon, y + \varepsilon)$ such that $f(z) > 0$. Thus $g(x + \varepsilon) \geqslant f(z) > 0$. It now follows from the continuity of g at $x + \varepsilon$ that g is locally nonzero.

Let x be the unique root of g; consider $y \in [0, x)$ and set $\varepsilon' = y - x$. Since f has uniformly upper-stable roots, there exists $\delta' > 0$ such that for all $x' \in [0, 1]$

$$|f(x)| < \delta' \Rightarrow \exists_{z' \in (x' - \varepsilon', x' + \varepsilon') \cap [0,1]} f(z') > 0.$$

Either $f(y) < 0$ or $f(y) > -\delta'$. In the latter case, there exists $z < y + \varepsilon' = x$ such that $f(z) > 0$. This contradicts our construction of x; whence $f(y) < 0$. ■

Corollary 54 *If $f : [0, 1] \to \mathbf{R}$ is uniformly continuous, has uniformly stable roots, and $f(0)f(1) < 0$, then Z_f is located.*

Proof. Fix $x \in \mathbf{R}$ and let $s, t \in \mathbf{Q}$ with $s < t$. Let $\delta > 0$ be such that if $|f(x)| < \delta$, then there exist $y, y' \in (x - |t - s|, x + |t - s|)$ such that $f(y) < 0 < f(y')$. Either

$$\inf\{|f(z)| : z \in (x - s, x + s)\} > 0 \text{ or } \inf\{|f(z)| : z \in (x - s, x + s)\} < \delta.$$

In the first case, $\rho(x, Z_f) > 0$. In the second case, there exist $y, y' \in (x - t, x + t)$ such that $f(y) < 0 < f(y')$. By the Theorem there exists $z \in (y, y')$ such that $f(z) = 0$; whence $\rho(x, Z_f) < t$. ■

In general: the statement

> 'If f is continuous, then Z_f is located'

is equivalent to the *weak law of trichotomy*: for all $x \in \mathbf{R}$ either $x = 0$ or $\neg x = 0$. To show that Z_f is located it is sufficient to apply the weak law of trichotomy to $\inf\{|f(s)| : s \in [t_1, t_2]\}$ for appropriate $t_1, t_2 \in \mathbf{R}$. Conversely for sufficiently small $x \in \mathbf{R}$ if the zero set of

$$f : t \mapsto (t - 1)^2(t - 2)(t - 3)^2 + x$$

is located then we can decide that $x = 0$, $x \leqslant 0$ and $\neg(x = 0)$, or $x \geqslant 0$ and $\neg(x = 0)$.

In [97], Peter Schuster put forward the 'guarded thesis' that in order to construct an object in **BISH** without choice we require that object to be unique (in some strong sense). In Proposition 52 this uniqueness is given by the produced x being the minimal root of f. Finding the minimal root of a locally nonconstant uniformly continuous function $f : [a, b] \to \mathbf{R}$ with $f(a) \leqslant 0$ and $f(b) \geqslant 0$ is not possible in general, even when we allow choice principles. For example, let a be a real number close to 0 and define $f : [-1, 2] \to \mathbf{R}$ by

$$f(x) = x^2(x - 1) + a.$$

Then, for sufficiently small a, f has minimal root less than $1/2$ if and only if $a \geqslant 0$ and has minimal root greater than 0 if and only if $a < 0$. Knowing that f has only stable roots does not help, since there exists a stable root of f less than $1/2$ if and only if $a > 0$.

For a function $f : I \to \mathbf{R}$, we say that f has *roots isolated above* if for each $x \in I$ with $f(x) = 0$, there exists $\varepsilon > 0$ such that $f(y) \neq 0$ for all $y \in (x, x + \varepsilon) \cap I$. Applying Proposition 52 iteratively to a function $f : [a, b] \to \mathbf{R}$ which has roots isolated above, has uniformly stable roots and is such that $f(a) \leqslant 0 \leqslant f(b)$, we can *find all the initial roots of* f in the following sense: there exists a nondecreasing sequence $(x_i)_{i \geqslant 1}$ contained in $[a, b] \cup \{b + 1\}$ such that for each i

(i) either $f(x_i) = 0$ and $x_i < x_{i+1}$, or else $x_i = b + 1$; and

(ii) if $x \in [a, b]$, $f(x) = 0$ and there exists $i \in \mathbf{N}$ such that $x < x_i$, then $x = x_j$ for some $j < i$.

Proposition 55 *Let $f : [a, b] \to \mathbf{R}$ be a continuous function with uniformly stable roots such that $f(a) \leqslant 0 \leqslant f(b)$ and such that each root of f is isolated above. Then we can find all the initial roots of f.*

Proof. By Proposition 52, there exists $x \in [a,b]$ such that $f(x) = 0$ and $f(y) \neq 0$ for all $y \in [0,x)$; set $x_1 = x$. Suppose that we have constructed $x_1, x_2, \ldots, x_{n-1}$ such that $f(x_i) = 0$ for each i and such that

$$\forall_{y \in [0, x_{n-1}) - \{x_1, \ldots, x_{n-2}\}} f(y) \neq 0. \tag{$*$}$$

Let $\varepsilon \in (0, 1 - x_{n-1})$ be such that $f(y) \neq 0$ for all $y \in (x_{n-1}, x_{n-1} + \varepsilon)$, and let $\delta > 0$ be such that for all $x \in [0,1]$

$$|f(x)| < \delta \Rightarrow \exists_{x,y \in (x-\varepsilon, x+\varepsilon) \cap [a,b]} f(y) < 0 < f(z).$$

Suppose that there exist $y, y' \in (x_{n-1}, x_{n-1} + \varepsilon)$ such that $f(y) < 0 < f(y')$. If $y < y'$, applying Proposition 52 to $f|_{(y,y')}$ produces $z \in (y, y')$ such that $f(z) = 0$—a contradiction. If $y' < y$ we get a similar contradiction. Hence either $f(y) > 0$ for all $y \in (x, x+\varepsilon)$ or else $f(y) < 0$ for all $y \in (x_{n-1}, x_{n-1} + \varepsilon)$.

For illustration, we consider the case where $f(y) > 0$ for all $y \in (x, x + \varepsilon)$. Define

$$m = \inf \{ |f(z)| : z \in [x + \varepsilon, 1] \}.$$

Either $m > 0$ or $m < \delta$. In the first case $f(z) \neq 0$ for all $z \in [0,1] - \{x_1, \ldots, x_{n-1}\}$ and we set $x_i = b + 1$ for all $i \geqslant n$. In the second case there exists $z \in [x_{n-1} + \varepsilon, 1]$ such that $|f(z)| < \delta$; thus there exists $z' \in [x_{n-1} + \varepsilon, 1)$ with $f(z') < 0$. Applying Proposition 52 to $-f|_{[x_{n-1}+\varepsilon/2, z]}$ produces $x_n \geqslant x_{n-1} + \varepsilon$ with $f(x_n) = 0$. For each $y \in [0, x_n) - \{x_1, \ldots, x_{n-1}\}$, $f(y) \neq 0$ by ($*$), Proposition 52, and by our choice of ε. The case when $f(y) < 0$ for all $y \in (x_{n-1}, x_{n-1} + \varepsilon)$ is handled similarly.

Since each root of f is isolated above, $S_n = \{x_1, \ldots, x_n\}$ is discrete for each $n \in \mathbf{N}$. If $y < x_i$ for some $i \in \mathbf{N}$ and $\rho(y, S_i) > 0$, then $y \in [0, x_i) - \{x_1, \ldots, x_{i-1}\}$, so $f(y) \neq 0$. Hence if $y < x_i$ for some $i \in \mathbf{N}$ and $f(y) = 0$, then $\rho(y, S_i) = 0$. Since S_i is closed and discrete, $y = x_j$ for some $j \in \{1, \ldots, i-1\}$; whence we can find all the initial roots of f. ∎

3.1.3 Continuity

Joyal's derived continuity rule shows that many results about the Dedekind reals in **CZF** give algorithms which are at least locally continuous. We present a few results on the continuity of algorithms for the intermediate value theorem.

The function $f : [0,1] \to \mathbf{R} : x \mapsto 2x - 1$ shows that the interval halving proof of the exact intermediate value theorem with **LLPO** and the interval halving proof of $\mathbf{IVT_{loc}}$ are not even operations (they may return different outputs for the same input), let alone continuous. The functions $f_a : (x - 1)^2(x - 2) + a$ $(a \in \mathbf{R})$ shows that the standard classical sup argument is also discontinuous.

It was observed by Helmut Schwichtenberg that if f maps $\mathbf{Q}$ in to $\mathbf{Q}$, then, since equality on $\mathbf{Q}$ is decidable, we can use interval halving to find a root of f without choice. If we restrict this to the family of functions $\{f_a : a \in \mathbf{Q}\}$ from above, we get a function $\mathbf{Q} \to \mathbf{R} : a \mapsto x$ such that $f_a(x) = 0$, which cannot be globally continuous.

The proof of Proposition 52 is more well behaved. Let

$$S = \{f \in \mathcal{C}([0,1], \mathbf{R}) : f \text{ has uniformly upper} \\ \text{stable roots }, f(0) < 0 < f(1)\}$$

with the supremum norm.

Proposition 56 *The mapping $F : S \to (0,1)$ which sends $f \in S$ to the minimal root of f is continuous.*

Proof. Fix $f \in S$ and, using Proposition 52, construct the minimal root x of f. Let $\varepsilon > 0$ and let $z \in (x-\varepsilon, x+\varepsilon)$ be such that $f(z) > 0$. Since x is the minimal root of f and f is continuous, $z > x$. Set $\delta = \min\{f(z), -f(x-\varepsilon)\}$ and consider $g \in S$ such that $\rho(f,g) < \delta$; suppose that $|F(g)-x| > \varepsilon$. Then either $F(g) > x+\varepsilon$ or $F(g) < x-\varepsilon$. In the first case, $g(0) < 0 \leqslant f(z)-\delta <$

$g(z)$, so by Proposition 48 it is impossible for the minimal root $F(g)$ of g to be greater than $x+\varepsilon$. Suppose there exists $y \in [0, x-\varepsilon]$ with $f(y) > -\delta$. Then there exists $z' \in (0, x)$ such that $f(z') > 0$. Applying Proposition 48 once again leads to a contradiction. Hence $|F(g) - x| < \varepsilon$. ■

Having uniformly stable roots can be characterised in terms of the function $f \mapsto Z_f$. We restrict our attention to functions with located zero sets: let

$$S_l = \{f \in \mathcal{C}([0,1], \mathbf{R}) : Z_f \text{ is located}, f(0) < 0 < f(1)\}$$

with the supremum norm. Then

- ► x is a stable root of f if

$$\forall_{\varepsilon>0} \exists_{\delta>0} \forall_{g \in S} (\|f-g\|_\infty < \delta \to \rho(x, Z_g) < \varepsilon);$$

- ► f has uniformly stable roots if

$$\forall_{\varepsilon>0} \exists_{\delta>0} \forall_{g \in S} (\|f-g\|_\infty < \delta \to \sup\{\rho(x, Z_g) : x \in Z_f\} < \varepsilon).$$

A natural question to ask is: what condition on f corresponds to the continuity of the mapping $F : S_l \to \mathcal{P}[0,1]$ sending f to Z_f. Our next proposition answers this question. A function f has *strongly uniformly stable roots* if

$$\forall_{\varepsilon>0} \exists_{\delta>0} \forall_{x \in \mathbf{R}} (|f(x)| < \delta \Rightarrow \exists_{y,z \in (x-\varepsilon, x+\varepsilon)} f(y), -f(z) > \delta) \qquad (*).$$

The examples we gave above of functions with uniformly stable roots satisfy this stronger condition.

Proposition 57 *Let $f \in S_l$. Then f has strongly uniformly stable roots if and only if the function $F : S \to \mathcal{P}[0,1]$ which sends $g \in S$ to Z_g is continuous at f.*

Proof. Suppose that f has strongly uniformly stable roots, fix $\varepsilon > 0$, and let $\delta > 0$ be as in (*). We must show that, if $g \in S_l$ is such that $\|f - g\|_\infty < \delta$, then

$$\rho(Z_f, Z_g) = \max\{\sup\{\rho(x, Z_g) : x \in Z_f\}, \sup\{\rho(x, Z_f) : x \in Z_g\}\} < \varepsilon.$$

If $x \in Z_f$, then there exist $y, z \in B(x, \varepsilon)$ such that, without loss of generality, $g(y) < 0 < g(z)$. By Corollary 50, we have that $Z_g \cap B(x, \varepsilon)$ cannot be empty; since Z_g is located, $Z_g \cap B(x, \varepsilon)$ must be inhabited and hence $\rho(x, Z_g) < \varepsilon$. If $x \in Z_g$, then $|f(x)| < \delta$. It follows from $(*)$, Corollary 50, and the locatedness of Z_f that $\rho(x, Z_f) < \varepsilon$. Hence $\rho(Z_f, Z_g) < \varepsilon$.

Conversely, suppose that F is continuous at f. Fix $\varepsilon > 0$ and let $\delta > 0$ be such that if $\|f - g\|_\infty < 3\delta$, then $\rho(Z_f, Z_g) < \varepsilon/2$. If $|f(x)| < \delta$, then $f - f(x)$ has a zero at x and $\|f - (f - f(x))\| < 3\delta$; whence there exists a root x' of f in $B(x, \varepsilon/2)$. Also $g_1 = f - f(x) + 2\delta$ and $g_2 = f - f(x) - 2\delta$ satisfy $\|f - g_i\|_\infty < 3\delta$ $(i = 1, 2)$, so there exist $x_1, x_2 \in B(x', \varepsilon/2)$ such that $g_i(x_i) = 0$ $(i = 1, 2)$. Then $x_1, x_2 \in B(x, \varepsilon$ and $f(x_1), -f(x_2) > \delta$. ■

Classically a real valued function on $[0, 1]$ satisfies the *intermediate value property* if there exists $z \in (x, y)$ such that $f(z) = 0$ whenever $x, y \in [0, 1]$ and $f(x) < 0 < f(y)$. The intermediate value theorem states that every continuous real valued function on the unit interval has the intermediate value property. Given a constructive reading, the intermediate value theorem is unstable: let $a \in (0, 1)$ and define

$$f_a(x) = \begin{cases} x - 1/3 & x \in [0, 1/3 + a] \\ a & x \in [1/3 + a, 2/3 + a] \\ x - 2/3 & x \in [2/3 + a, 1]. \end{cases}$$

Then f_0 has the intermediate value property, but f_a has the intermediate value property if and only if $a \geqslant 0$ or $a \leqslant 0$. The reason for this nonconstructivity is that the intermediate value property is concerned only with the existence of roots, but the function taking f to the set Z_f of zeros of f can be discontinuous. We define a *constructive intermediate value property*:

IVP$_c$: The function on S_l which sends $f \in \mathbf{R}^{[0,1]}$ to the set

$Z_f = \{x \in [0,1] : f(x) = 0\}$ is continuous at f.

Proposition 57 can be restated as: a function satisfies the constructive intermediate value property if and only if it has strongly uniformly stable roots.

In the setting of type two effectivity [111] a new programme of reverse mathematics with emphasis on the computational content of theorems has began [20, 21]. The important questions here are of the form: how noncomputable is a particular multivalued function. For example, how noncomputable is the multivalued function which sends a continuous function $f \in \mathcal{C}[0,1]$ to its set of zeros. In the foregoing we consider this from a different angle; we ask not how noncomputable this function is, but what is the largest subset of its domain on which it is computable. Although this question seems of less general interest, we feel it gives a natural approach to the systematic study of constructive and computable analysis.

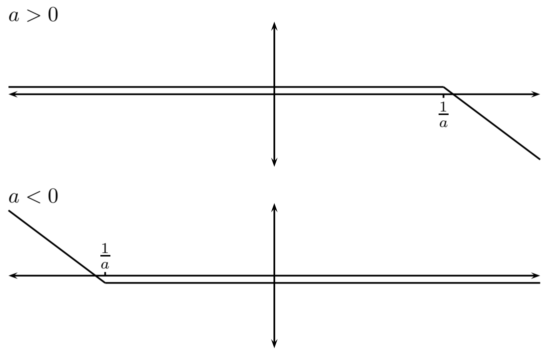
3.1.4 Relationships between notions

We finish by presenting some examples showing the relationship between the various properties on a function f.

There exists a function $f : \mathbf{R} \to \mathbf{R}$ that has only stable roots, for which it is impossible not to have a root, and for which we can, in general, only find a root if Markov's Principle holds, but which is only locally nonzero if **MP** holds. Let $a \in \mathbf{R}$ be such that $\neg(a = 0)$ and define $f : \mathbf{R} \to \mathbf{R}$ by[42]

$$f(x) = \begin{cases} -x - \frac{1}{a} + a & a \geqslant \frac{1}{x} > 0 \\ -x + \frac{1}{a} + a & a \leqslant \frac{1}{x} < 0 \\ a & \text{otherwise.} \end{cases}$$

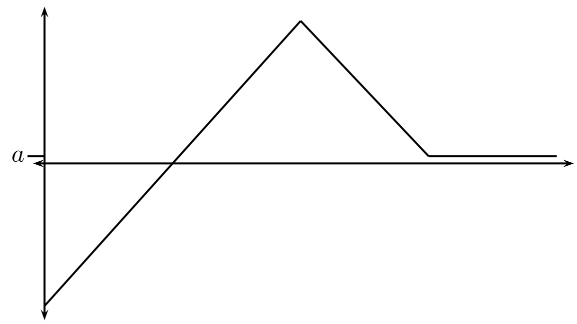
[42]The standard construction of functions on $\mathbf{R}$ by case does not require choice.



To see that f has only stable roots, let $x \in \mathbf{R}$ be a root of f. Note that if $a \neq 0$, then f is locally linear at x with gradient -1, so f has only stable roots. Pick $N > 0$ such that $|x| < N$. If $|a| < 1/N$, then $|x| > N$, which is absurd. Hence $|a| > 0$ and f has only stable roots. Now suppose that f is locally nonzero. Then there exists $x \in (-1, 1)$ with $f(x) \neq 0$ and so $|a| \geqslant \min\{1, |f(x)|\} > 0$.

This example does not have a direct impact on whether having only stable roots is sufficient to prove a constructive intermediate value theorem (without choice), since it is necessary in this example that the domain of f is unbounded. It is possible to give an example of a function $f : [0, 2] \to \mathbf{R}$ with compact domain such that f has only stable roots, f is only locally non-zero, in general, if **MP** holds, and it is impossible for f not to have a root: let a be a non-negative real number such that $\neg(a = 1)$ and define

$$f(x) = \begin{cases} 2x - 1 & x \in [0, 1] \\ 2(a-1)x + 3 - 2a & x \in [1, 3/2] \\ a & x \in [3/2, 2]. \end{cases}$$



However, in this case it is easy to find the root of f.

The condition that $f : I \to \mathbf{R}$ have *at most one root*

$$\forall_{x,y \in I}(x \neq y \Rightarrow f(x) > 0 \vee f(y) > 0)$$

is not strong enough to prove that f has uniformly upper-stable roots, even when f is continuous, has only stable roots, and there exist $a, b \in I$ such that $f(a)f(b) < 0$. To see this, recall that in Russian recursive mathematics there exists a continuous function $g : [0, 1/2] \to \mathbf{R}$ with supremum 0 such that $g(x) < 0$ for each $x \in [0, 1/2]$ (see [23, Chapter 6, Corollary 2.9]). The construction of g uses the existence of an enumeration of the set of partial functions from $\mathbf{N}$ into $\mathbf{N}$ with countable domain. Assuming the Church-Markov-Turing thesis, this enumeration is constructed using countable choice [107]. Define

$$f(x) = \begin{cases} g(x) & x \in [0, 1/2] \\ g(1/2) & x \in [1/2, 3/4] \\ 4(1 - g(1/2))x - 3 + 4g(1/2) & x \in [3/4, 1]. \end{cases}$$

Then f is a continuous real valued function on the unit interval, $f(0) = g(0) < 0$, and $f(1) = 1 > 0$. Since f is linear with gradient $4(1 - f(1/2))$ on $[3/4, 1]$ and since the only root of f is within this interval, all roots of f are stable. To see that f has at most one root, note that $f(x) \neq 0$ for all

$$x \in [0,1] - \left\{ \frac{3 - 4f(1/2)}{4(1 - f(1/2))} \right\}.$$

But f does not have uniformly upper-stable roots: let $\varepsilon = 1/4$. For all $\delta > 0$ there exists $x \in [0, 1/2]$ such that $|f(x)| < \delta$, but $f(y) \neq 0$ for all $y \in (x - \varepsilon, x + \varepsilon) \cap [0, 1]$.

If a continuous function $f : [a, b] \to \mathbf{R}$ with $f(a) < 0$ and $f(b) > 0$[43] has *uniformly at most one zero*—that is,

$$\forall_{\varepsilon>0} \exists_{\delta>0} \forall_{x,y \in I} \left(|f(x)| < \delta \wedge |f(y)| < \delta \Rightarrow |x - y| < \varepsilon\right)$$

—then f has uniformly stable roots. To see this, let $f : [a, b] \to \mathbf{R}$ be a continuous function with uniformly at most one root and fix $\varepsilon > 0$. Let $\delta \in (0, r)$ be such that for all $x, y \in [a, b]$

$$|f(x)| < \delta \wedge |f(y)| < \delta \Rightarrow |x - y| < \varepsilon,$$

and let $x \in [a, b]$ be such that $|f(x)| < \delta$. Using the continuity of f at a and b, we may assume that $|f(x)| > \delta$ for all $x \in [a, a + \varepsilon) \cup (b - \varepsilon, b]$. Then $x \in [a + \varepsilon, b - \varepsilon]$ and so, by continuity, it suffices to show that $f(x - \varepsilon) < 0 < f(x + \varepsilon)$. Suppose that $f(x + \varepsilon) < \delta$. If $f(x + \varepsilon) > -\delta$, then $|f(x + \varepsilon)| < \delta$, so $|x - (x - \varepsilon)| < \varepsilon$—a contradiction from which it follows that $f(x + \varepsilon) < 0$. Then by the approximate intermediate value theorem, there exists $z \in (x + \varepsilon, 1)$ such that $|f(z)| < \delta$, but

$$|z - x| > |x + \varepsilon - x| = \varepsilon,$$

again contradicting our choice of δ. Thus $f(x + \varepsilon) \geqslant \delta > 0$. A similar proof shows that $f(x - \varepsilon) < 0$.

With countable choice any continuous function $f : [a, b] \to \mathbf{R}$ with $f(a) < 0$

[43] We need this condition to rule out a function contained entirely above or below the x-axis (for example, $f(x) = |x - 1/2|$ on $[0, 1]$).

and $f(b) > 0$ that has at most one root has only stable roots. Let $f : [a, b] \to \mathbf{R}$ be a continuous function with $f(a) < 0$ and $f(b) > 0$ that has at most one root, let $x \in [a, b]$ be such that $f(x) = 0$, and let $\varepsilon > 0$. Then $x \notin \{a, b\}$ and without loss of generality $(x - \varepsilon, x + \varepsilon) \subset [a, b]$. Since f has at most one root, $f(x + \varepsilon/2) \neq 0$ and $f(x - \varepsilon/2) \neq 0$. Suppose that $f(x + \varepsilon/2) < 0$. Having at most one zero implies being locally nonconstant, so we can apply the standard constructive intermediate value theorem (which requires countable choice) to construct $y \in [x + \varepsilon/2, 1]$ with $f(y) = 0$—a contradiction. Hence $f(x + \varepsilon/2) > 0$. A similar proof shows that $f(x - \varepsilon/2) < 0$, so x is a stable root and f has only stable roots.

As you would expect, if $f : \mathbf{R} \to \mathbf{R}$ has uniformly stable roots and does not have two roots, then f has uniformly at most one root. Let $\varepsilon > 0$. Since f has uniformly stable roots, there exists $\delta < 0$ such that for all $x \in \mathbf{R}$ if $|f(x)| < \delta$, then there exist $y, z \in (x - \varepsilon/2, x + \varepsilon/2)$ with $f(y) < 0 < f(z)$. Let $x, x' \in \mathbf{R}$ be such that $|f(x)| < \delta$ and $|f(x')| < \delta$, and suppose that $|x - x'| > \varepsilon$. Then there exist $y, z \in (x - \varepsilon/2, x + \varepsilon/2)$ and $y', z' \in (x' - \varepsilon/2, x' + \varepsilon/2)$ such that $f(y) < 0 < f(z)$ and $f(y') < 0 < f(z')$. By Proposition 52, there exist $w \in (y, z)$ and $w' \in (y', z')$ such that $f(w) = f(w') = 0$. But, since $|x - x'| > \varepsilon$, $w \neq w'$—this contradicts that f does not have two roots. Hence $|x - x'| \leqslant \varepsilon$, and f has uniformly at most one zero.

The following diagram summarises the relationships between the various properties of a function $f : \mathbf{R} \to \mathbf{R}$; all implications are strict.

$$
\begin{tikzcd}
 & \boxed{\exists_{a,b\in\mathbf{R}} f(a)f(b) < 0 + \text{uniformly at most one}} \arrow[d, Rightarrow] & \\
\boxed{\begin{array}{c}\text{uniformly at most one}\\ +\\ \text{stable}\end{array}} \arrow[r, Rightarrow] & \boxed{\text{uniformly stable}} \arrow[r, Rightarrow] & \boxed{\begin{array}{c}\text{locally nonzero}\\ +\\ \text{stable}\end{array}}
\end{tikzcd}
$$

$$
\begin{tikzcd}
\boxed{\begin{array}{c}\text{uniformly stable}\\ +\\ \text{not 2 roots}\end{array}} \arrow[r, Rightarrow] & \boxed{\text{uniformly at most one}} \arrow[r, Rightarrow] & \boxed{\text{at most one}} \arrow[d, Rightarrow] \\
 & & \boxed{\text{locally nonzero}}
\end{tikzcd}
$$

$$
\begin{tikzcd}
\boxed{\begin{array}{c}\exists_{a,b\in\mathbf{R}} f(a)f(b) < 0\\ +\\ \text{at most one}\end{array}} \arrow[r, Rightarrow, "+\ \mathbf{CC}"] & \boxed{\text{stable}}
\end{tikzcd}
$$

3.2 Contractive and nonexpansive mappings

In this section we look at the simplest, but seemingly most general, case of classical fixed point theorems in which we can guarantee the constructive existence of exact fixed points: contractive mappings. This is the one place in this thesis where will give a little attention to the computational efficiency of the algorithms that we insist our proofs contain.[44]

Let S be a subset of a metric space X and let $f : S \to X$. There are three basic types of nonexpansive properties we may require of f:

- f is *uniformly contractive* if there exists $r \in (0, 1)$ such that

$$\|f(x) - f(y)\| \leqslant r\|x - y\|$$

  for all $x, y \in S$;

- f is *contractive* if $\|f(x) - f(y)\| < \|x - y\|$ for all $x, y \in S$ with $x \neq y$;

- f is *nonexpansive* if $\|f(x) - f(y)\| \leqslant \|x - y\|$ for all $x, y \in S$.

Nonexpansive, and hence (uniformly) contractive, mappings are uniformly continuous, and, as the following lemma shows, any fixed point of a contractive mapping is unique.

Lemma 58 *If f is a contractive mapping of X into itself and x, y are distinct points of X, then either $f(x) \neq x$ or $f(y) \neq y$.*

Proof. Since

$$\|x - f(x)\| + \|y - f(y)\| \geqslant \|x - y\| - \|f(x) - f(y)\| > 0,$$

either $\|x - f(x)\| > 0$ or $\|y - f(y)\| > 0$. ■

As we have already said, the standard classical proof of the contractive mapping theorem

[44] Constructive mathematicians are generally not interested in the efficiency of their proofs—the important question is what can be computed—and the contents of this section represent the only (and very minimal) thought I have ever given to algorithmic efficiency.

> Every uniformly contractive mapping of an inhabited closed subset of a complete metric space into itself has a fixed point,

is fully constructive: for any $x \in S$ the sequence $(f^{n+1}(x))_{n\in\mathbf{N}}$ converges at a rate determined by r and $\|x - f(x)\|$. If $1 - r$ is small (and $\|x - f(x)\|$ is large), then this sequence will converge very slowly; while this proof is mathematically very simple, it is computationally poor. So what can we do? This is really a question for the numerical analyst rather than the constructive mathematician, but we will give another proof of the contractive mapping theorem for $\mathbf{R}^n$ (as a special case of Edelstein's fixed point theorem on non-uniformly contractive mappings) which is far more efficient—see the discussion after the proof of Theorem 61. First we consider the converse of the contractive mapping theorem.

3.2.1 Uniformly contractive mappings

A metric space X is said to have the *Banach fixed point property* if every uniformly contractive mapping on an inhabited closed subspace of X has a fixed point. Classically, the contractive mapping theorem gives a characterisation of complete spaces

> (*) A metric space is complete if and only if it has the Banach fixed point property.

Proposition 59 (WCC) *(*) is equivalent to* ***LPO****.*

Proof. Suppose that (*) holds, fix $a \in \mathbf{R}$, and set $X = \{0, 1\}a$. Consider an inhabited closed subspace Y of X and a contraction mapping $f : Y \to Y$ with contraction constant $r \in (0, 1)$. Fix $y \in Y$ and let $\lambda \in \{0, 1\}$ be such that $f(y) = \lambda a$; then $\lambda a \in Y$. Suppose that $f(\lambda a) \neq \lambda a$. Then $a \neq 0$, and $f(\lambda a) = y = (1 - \lambda)a$. But

$$a = \rho(f(y), f(\lambda a)) \leqslant r\rho(y, \lambda a) = ra,$$

so $a = 0$—a contradiction. Hence λa is a fixed point of f and X has the Banach fixed point property. By (*), X is then complete. By examining the limit of the sequence $\lambda_n a$, where $(\lambda_n)_{n\in\mathbf{N}}$ is a binary sequence such that

$$\begin{aligned}\lambda_n = 0 \quad &\Rightarrow \quad a < \frac{1}{n}\\ \lambda_n = 1 \quad &\Rightarrow \quad a > \frac{1}{n+1},\end{aligned}$$

we can decide whether $a = 0$ or $a \neq 0$.

Conversely, assume **LPO**, let X be a metric space with the Banach fixed point property, and let $(x_n)_{n\geqslant 1}$ be a Cauchy sequence in X. Define

$$\theta(x) = \inf\left\{\rho(x, x_n) : x_n \neq x, n \in \mathbf{N}\right\}.$$

Using **LPO**, we may assume that the terms of $(x_n)_{n\geqslant 1}$ are all distinct and that $\theta\,(x_n) > 0$ for each $n \in \mathbf{N}$. Define $(k_n)_{n\geqslant 1}$ inductively as follows. Set $k_1 = 1$ and suppose that we have constructed $k_1, \ldots, k_m$. Let k_{m+1} be such that

$$\rho(x_s, x_t) \leqslant \theta(x_{k_m})$$

for all $s, t \geqslant k_{m+1}$. Set $S = \{x_{k_n} : n \in \mathbf{N}\}$ and define a contraction mapping f on S by setting $f\,(x_{k_n}) = x_{k_{n+1}}$ for each $n \in \mathbf{N}$. Extend f, by continuity, to a contraction mapping on $\overline{S}$. Since X has the Banach mapping property, there exists $x \in \overline{S}$ such that $f(x) = x$. Then $x \in X$ and, it is easy to see that, x is the limit of $(x_n)_{n\geqslant 1}$. ■

Given an inhabited metric space X and a function $f : X \to X$ without a fixed point, the sequence $(f^{n+1}(x))_{n\in\mathbf{N}}$, for any $x \in X$, witnesses that X is not complete. Indeed, $(f^{n+1}(x))_{n\in\mathbf{N}}$ is eventually bounded away from each point of X: extend f to a (uniformly contractive) mapping on $\hat{X}$, and let $\hat{x}$ be the unique fixed point of f in $\hat{X}$. Pick $x_0 \in X$ and let $x_n = f^n(x_0)$ for each $n \geqslant 1$; then $(x_n)_{n\geqslant 0}$ is a Cauchy sequence in X which converges to $\hat{x}$ in $\hat{X}$. By Lemma 58, $f(x) \neq x$ for each $x \in X$ and so, by continuity, $x \neq \hat{x}$ for each $x \in X$. Since $(x_n)_{n\geqslant 0}$ converges to $\hat{x}$, it follows that $(x_n)_{n\geqslant 0}$ is a

Cauchy sequence in X which is eventually bounded away from each point of X.

There is a weaker property associated with uniformly contractive mappings. A a metric space S is said to have the *contraction mapping property* if every uniformly contractive mapping from S into S has a fixed point. We present a constructive proof of a result, due to Borwein [19], which characterises when sets with a strong connectedness property have the contraction mapping property. A metric space X is said to be *uniformly Lipschitz-connected* if there exists a positive real number L such that for all $x, y \in X$, there exists a function $g : [0, 1] \to X$ such that $g(0) = x, g(1) = y$ and

$$\rho(g(s), g(t)) \leqslant L|s-t|\rho(g(0), g(1))$$

for all $s, t \in [0, 1]$.

Theorem 60 (CC) *Let S be an inhabited uniformly Lipschitz-connected metric space. Then S is complete if and only if it has the contraction mapping property.*

Proof. Suppose that S is complete. Let f be a contraction mapping on S with contraction constant $r \in (0, 1)$, and fix $x \in S$. Let $x_0 = x$ and $x_n = f^{n-1}(x)$ for each $n \in \mathbf{N}$. Then $(x_n)_{n\geqslant1}$ is a Cauchy sequence in S: if $m \geqslant n \geqslant 1$, then

$$\rho(x_n, x_m) \leqslant \sum_{k=n}^{m-1} \rho(x_k, x_{k+1}) < \frac{r^n}{1-r}\rho(x_0, x_1) \longrightarrow 0 \text{ as } n \longrightarrow \infty.$$

Hence $(x_n)_{n\geqslant1}$ converges to some point $x \in S$; clearly $f(x) = x$.

Conversely, suppose that S has the contraction mapping property, let $(x_n)_{n\geqslant1}$ be a Cauchy sequence in S, and let $\hat{x}$ be the limit of $(x_n)_{n\geqslant1}$ in the completion $(\hat{S}, \hat{\rho})$ of (S, ρ). Without loss of generality we may take

$$\rho(x_n, x_m) < 2^{-\min\{m,n\}} \qquad (m, n \in \mathbf{N}).$$

Using countable choice and the uniform Lipschitz-connectedness of S, construct $L > 0$ and functions $g_k : [0,1] \to S$ $(k \in \mathbf{N})$ such that for each k, $g_k(0) = x_{k+1}$, $g_k(1) = x_k$, and

$$\rho(g_k(s), g_k(t)) < L|s-t|\rho(x_k, x_{k+1})$$

for all $s, t \in [0,1]$. Using the gluing lemma [33], define a mapping

$$g : \{0\} \cup \bigcup_{k \geqslant 1} [2^{-(k+1)}, 2^{-k}] \cup (1, \infty) \to S \cup \{\hat{x}\}$$

by

$$g(t) \equiv \begin{cases} \hat{x} & \text{if } t = 0 \\ g_k(2^{k+1}t - 1) & \text{if } t \in [2^{-(k+1)}, 2^{-k}] \\ x_1 & \text{if } t > 1. \end{cases}$$

Suppose that there exist t_1, t_2 with $t_2 < t_1$ and $\rho(g(t_1), g(t_2)) > L|t_1 - t_2|$. By continuity we may assume, without loss of generality that $t_i \in [2^{-n_i}, 2^{-n_i+1}]$ $(i = 1, 2)$; set $s_0 = t_1, s_{n_1 - n_2} = t_2$, and $s_k = 2^{-n_2 - 1 + k}$ for $1 \leqslant k \leqslant n_1 - n_2 - 1$. Then

$$\begin{aligned} \sum_{k=1}^{n_1 - n_2} \rho(g(s_{k-1}), g(s_k)) &\geqslant \rho(g(t_1), g(t_2)) \\ &> L|t_1 - t_2| \\ &= \sum_{k=1}^{n_1 - n_2} |s_{k-1} - s_k|, \end{aligned}$$

so there exists $1 \leqslant k \leqslant n_1 - n_2 - 1$ such that $\rho(g(s_{k-1}), g(s_k)) > L|s_{k-1} - s_k|$. But

$$\begin{aligned} \rho(g(s_{k-1}), g(s_k)) &= \rho\left(g_k(2^{k+1}s_{k-1} - 1), g_k(2^{k+1}s_k - 1))\right) \\ &\leqslant L2^{k+1}|s_{k-1} - s_k|\rho(x_{k+1}, x_k) \\ &\leqslant L|s_{k-1} - s_k| \end{aligned}$$

—a contradiction. Hence

$$\rho(g(s), g(t)) \leqslant L|s-t| \tag{1}$$

for all s, t in the domain of g. It follows that g is uniformly continuous on its domain, and so extends to a uniformly continuous mapping $g : [0, \infty) \to S \cup \{\widehat{x}\}$ such that (1) holds for all $s, t \in [0, \infty)$.

Now define a uniformly continuous mapping $h : \widehat{S} \to [0, \infty)$ by

$$h(x) \equiv \frac{2}{L}\rho(x, \widehat{x}).$$

Then $g \circ h$ is a contraction mapping on $\hat{S}$, and therefore on S, and

$$\hat{x} = g(0) = g(h(\hat{x})) = g \circ h(\hat{x}).$$

Hence $\hat{x}$ is the unique fixed point of $g \circ h$ in $\hat{S}$. Since S has the contraction mapping property, $\hat{x} \in S$, so $(x_n)_{n \geqslant 1}$ converges. ■

3.2.2 Contractive mappings

Edelstein [49] gave the following generalisation of the contractive mapping theorem.

Theorem 61 *A bounded contractive mapping of* $\mathbf{R}^n$ *into itself has a unique fixed point.*

We give a constructive proof of this result which is essentially a tidier version of the proof of Bridges et al [26]. We fix a bounded, contractive mapping $f : \mathbf{R}^n \to \mathbf{R}^n$.

Lemma 62 *For all* $x, y \in \mathbf{R}^n$

$$\|z - f(z)\| > \|f(x) - y\| - \|x - y\|.$$

Proof. Since f is contractive, we have

$$\begin{aligned}\|z - f(z)\| &\geqslant \|z - f(x)\| - \|f(x) - f(z)\| \\ &> \|z - f(x)\| - \|x - z\|.\end{aligned}$$

■

Here then is the **proof of Edelstein's theorem**:

Proof. The proof is by induction on n; the case $n = 0$ is trivial. Suppose we have proved the result for functions on $\mathbf{R}^{n-1}$ and let $f : \mathbf{R}^n \to \mathbf{R}^n$. Since f is bounded there exists $n > 0$ such that $f(\mathbf{R}^n) \subset C_0 \equiv [-N, N]^n$. Let $S = \{1/3, 2/3\}^n$ and for each $1 \leqslant k \leqslant n$ and each $s \in S$ we take a hyperplane

$$H_{s,k} = \{x + s : x \perp e_k\}$$

through x and orthogonal to the k^{th} basis vector e_k. By applying our induction hypothesis $n2^n$ times to the contractive mappings

$$f_{s,k} = \pi_{s,k} \circ f$$

restricted to $H_{s,k}$, where $\pi_{s,k}$ is the projection of $\mathbf{R}^n$ onto $H_{s,k}$, we construct points $x_{s,k}$ ($s \in S, 1 \leqslant k \leqslant n$) such that $f_{s,k}$ fixes $x_{s,k}$. By Lemma 62, for each k and each s with k^{th} coordinate $1/3$, either $f(s) \neq s$ or $f(s + e_k/3) \neq s + e_k/3$. Since $f_{s,k}$ fixes $x_{s,k}$, it follows from Lemma 62 that there is $\delta > 0$ and an n-cube C_1 with each side of length $2N/3$ such that $|f(x) - x| > \delta$ for all $x \notin C_1$. Continuing in this way, we construct a sequence $(C_i)_{i \in \mathbf{N}}$ of nested cubes such that C_i has sides of length $(2/3)^i N$; by continuity, the unique point in the intersection of $\{C_n : n \in \mathbf{N}\}$ is the unique fixed point of f. ■

A few comments on the proof are in order. Edelstein's fixed point theorem is in essence what you get from the intermediate value theorem for locally nonzero functions with at most one zero (Lemma 53) by an induction on the dimension. So, although the above proof uses $AC_{\omega,2}$ the result requires no choice. The proof could be criticised for being algorithmically very inefficient: each time we construct C_{i+1} from C_i (for $n \geqslant 2$) we apply the inductive hypothesis $n2^n$ times, each application of which in turn appeals to the induction hypothesis many times, and so on. Should we find this disturbing, after all we made such a big deal about the algorithmic content of constructive mathematics in Section 1.4. Not really: we have made no effort for efficiency, like any other mathematician we have searched for a simple, possibly elegant, and illustrative proof. We would hope that we, if so inclined, could find a more efficient proof.

So pretending for a moment that we are so inclined, let us strive a *little* more for this efficiency. Assume for simplicity that we have already cornered the fixed point into the box $B_0 = [0, 1]^n$, and let ξ be the fixed point we must find. Picking two distinct points close to the centre of the box, Lemma 58 guarantees that one of them, y say, is not fixed by f. We can now use Lemma 62 to eliminate (almost) half of the box from our search, to give a set S_0 which must contains ξ. To keep things from getting too complicated, we extend S_0 to a box and rotate and translate so that the edges are parallel to the basis vectors. We then repeat this process to produce a sequence of boxes $(B_n)_{n\in\mathbf{N}}$ such that the area of B_n is bounded above by $(1/\sqrt{2})^{n-1} + \delta$ for some arbitrarily small δ chosen beforehand.

We must ensure that the maximum of the lengths of the edges also tends to zero. If we accept Markov's principle, then this is immediate. Without Markov's principle, for example if we wish to get a known rate of convergence for our approximations to the fixed point, we must work harder. If one edge is larger than the others by r, then we can use Lemma 62 and that $f(x)$ is closer to ξ for all $x \neq \xi$ to reduce this edge by (almost) $r/2$. Finally, if need be—if there are a number of sides whose lengths are near the maximum, and the previous techniques keep reducing the smaller edges—we can apply the induction hypothesis to reduce specific edges. These very simple strategies would provide a significant increase in the efficiency of the algorithm inherent in the proof of Theorem 61, although no doubt a numerical analyst could do considerably better still (and still guarantee the convergence of the algorithm).

In the special case of a uniformly contractive mapping f, we can use the contraction constant r to ensure that the relative dimensions of the box do not become too different without having to use the induction hypothesis. Suppose for simplicity that $n = 2$ and that the box is contained in the positive quadrant with a vertex at $(0,0)$, and that the edges of our box are a side of length d_1 in direction $e_1 = (0,1)$ and a side of length d_2 in direction $e_2 = (1,0)$. If

$$d_2 < \left(\frac{r(\sqrt{2}-1)}{r+\sqrt{n}} \right) d_1,$$

then, since f has contraction constant r, a point

$$a \in \{(d_1/3, d_2/2), ((2d_1)/3, d_2/3)\}$$

is moved by enough to ensure that the smallest angle between $f(a) - a$ and e_1 is less than 45 degrees. We can then apply Lemma 62 to reduce the box by a factor of $1/\sqrt{2}$ in the e_1 direction. In higher dimensions, we can use this trick to reduce the box in a direction other than that of the smallest edge, and we can ensure that the ratio of the lengths of any two edges does not exceed

$$\frac{2(r+\sqrt{n}}{r(\sqrt{2}-1)}$$

without needing to appeal to induction. This, together with the uniform bound on the area of B_n, allows us to give a computationally simple algorithm for finding the fixed point of f which converges at a quick rate dependent on r and n.

We give a simplified proof of Proposition 3.2 from [26]. A function f from a metric space X into itself has a *uniformly unique fixed point* if for all $\varepsilon > 0$ there exists $\delta > 0$ such that if $\max\{\rho(x, f(x)), \rho(y, f(y))\} < \delta$, then $\rho(x, y) < \varepsilon$. With countable choice, $\mathbf{FT}_\Delta$ implies that a function with a unique fixed point has a uniformly unique fixed point; and if f has a uniformly unique fixed point, then we can construct the fixed point of f.

Lemma 63 *If $f : \mathbf{R} \to \mathbf{R}$ is a bounded contractive mapping with fixed point ξ and $x, y < \xi$ with $|f(x) - x| < |f(y) - y|$, then $x < y$.*

Proof. By continuity $x \neq y$; suppose that $x > y$. Then

$$\begin{aligned} |x-y| &= x - y \\ &= (x - f(x)) + (f(x) - f(y) + (f(y) - y) \\ &= (f(x) - f(y)) - (|f(x) - x| - |f(y) - y|) \\ &> |f(x) - f(y)|, \end{aligned}$$

which contradicts that f is contractive. Hence $x < y$. ■

Proposition 64 *A bounded contractive mapping of $\mathbf{R}^n$ into itself has a uniformly unique fixed point.*

Proof. Fix $\varepsilon > 0$ and let ξ be the fixed point of f. We first consider the case $n = 1$: let $f : \mathbf{R} \to \mathbf{R}$ be a bounded contractive mapping. By continuity we can find $y_1 < \xi < y_2$ such that $|f(y_i) - y_i| < \varepsilon$; set

$$\delta = \min\{|f(y_1) - y_1|, |f(y_2) - y_2|\}.$$

If $|f(x) - x| < \delta$, then either $|x - \xi| < \varepsilon$ or, as we may assume, $x \neq \xi$. If $x < \xi$, then Lemma 63, with $y = y_1$, shows that $x \in (y_1, \xi)$ and hence $|x - \xi| < \varepsilon$. If $x > \xi$ we apply Lemma 63 to the function $-f$ and with $y = y_2$ to get the same conclusion.

Now consider $f : \mathbf{R}^n \to \mathbf{R}^n$ for $n > 1$, and let l_k be the line $\xi + \mathbf{R}e_k$ ($1 \leqslant k \leqslant n$). Applying the first part of the proof repeatedly construct $\delta_1, \ldots, \delta_n > 0$ such that for each $1 \leqslant k \leqslant n$ and all $x \in l_k$, if $|\pi_k \circ f(x) - x| < \delta$, then $|\xi - x| < \varepsilon/\sqrt{n}$. Then

$$\delta = \min\{\delta_1, \ldots, \delta_n\}$$

satisfies the conclusion of the uniform unique fixed point condition for ε. ■

In [26] Bridges et al proceed to give a bound on the rate of convergence of $(f^n(x))_{n\in\mathbf{N}}$ for a bounded, contractive mapping on $\mathbf{R}^2$. Kohlenbach produced other bounds on the rate of convergence of $(f^n(x))_{n\in\mathbf{N}}$ using his proof mining techniques [69].

We finish with a question: does the proof of Theorem 61 require induction? If the answer is yes, then one suspects that the major difficulty is finding a precise formulation of the question.

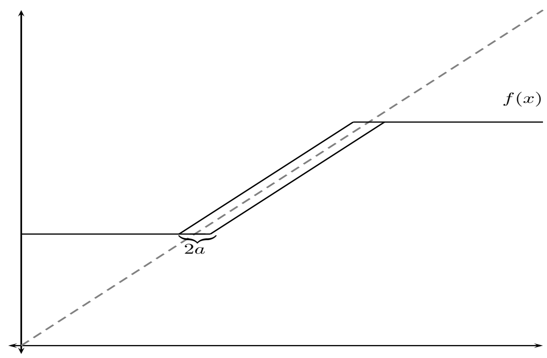
3.2.3 Nonexpansive mappings

It is with nonexpansive mappings that the classical and constructive fixed points diverge: The standard Brouwerian example showing that the intermediate value theorem implies **LLPO** is easily altered to show that we cannot prove constructively that every nonexpansive mapping on $[0,1]$ has the fixed point property.[45]

Let $a \in \mathbf{R}$ be near zero and construct a continuous mapping $f : [0,1] \to [0,1]$ such that

$$f(x) = \begin{cases} 1/3 & \text{if } x \in [0, 1/3 + a] \\ x - a & \text{if } x \in [1/3 + a, 2/3 + a] \\ 2/3 & \text{if } x \in [2/3 + a, 1]. \end{cases}$$

[45]Classically, every continuous nonexpansive mapping on a bounded closed subset of a uniformly convex Banach space has the fixed point property. The standard classical proof for Hilbert spaces (see [101]) requires the statement that 'every convex bounded closed subset of a Hilbert space is weakly compact', which implies **LPO**. As a consequence it seems likely that the most general formulation of the nonexpansive fixed point theorem does not follow from **WKL**; however, if we restrict ourselves to weakly compact subsets of a Hilbert space, then the nonexpansive fixed point theorem follows from **WKL**.



If f has a fixed point, then either it is greater than $1/3$ and $a \leqslant 0$ or it is less than $2/3$ and $a \geqslant 0$. Hence the statement 'every nonexpansive mapping of $[0, 1]$ into itself has a fixed point' implies that

$$\forall_{a \in \mathbf{R}}(a \geqslant 0 \vee a \leqslant 0),$$

which in turn implies **LLPO**.

Using a standard classical argument (see for example [101]), we can, however, show that such a mapping has approximate fixed points.

Proposition 65 *Let S be a bounded subset of a normed space X such that the closure of X is convex and let f be a nonexpansive mapping of S into itself. Then f has approximate fixed points.*

Proof. Since we are only interested in approximate fixed points we may replace X by its completion $\hat{X}$, and S by its closure in $\hat{X}$; we may also assume that $0 \in S$. Fix $\varepsilon > 0$ and let $N > 0$ be such that S is contained in the ball of radius N centered on 0. Pick $r \in (1 - \varepsilon/N, 1)$, and let x be the unique fixed point of the contraction mapping rf. Then

$$\|f(x) - x\| = \|f(x) - rf(x)\| = (1 - r)\|f(x)\| \leqslant (1 - r)N < \varepsilon.$$

Hence x is an ε-fixed point. ■

It follows from the previous proposition that **MIN** implies that every non-expansive function which maps a bounded convex subset into itself has a fixed point.

3.3 Brouwer's fixed point theorem

For completeness, we give a constructive proof of the approximate Brouwer fixed point theorem, extended to uniformly sequentially continuous functions; for novelty we give a proof based on David Gale's proof from [52]. Before we do this we require a few more definitions.

Gale's proof of Brouwer's fixed point theorem uses a generalisation of the game of Hex. An n-dimensional Hex board of size k consists of vertices $V = \{1, \ldots, k\}^n$ with edges between two vertices $x, y \in V$ if[46] $\|x - y\| = 1$ and either $x_i \leqslant y_i$ for each i or $y_i \leqslant x_i$ for each i. Then n-dimensional Hex is an n-player game where players take turns to pick unclaimed vertices. A player gains an edge of the hex board if she owns the nodes at either end; player i wins the game by connecting the two i-banks

$$\begin{aligned} i\text{-bank } 1 &= \{(v_1, \ldots, v_n) \in V : v_i = 0\}, \\ i\text{-bank } 2 &= \{(v_1, \ldots, v_n) \in V : v_i = k\}, \end{aligned}$$

with her edges. The 'Hex Theorem' of [52] (which, as finite combinatorics, is fully constructive) says that any colouring of an n-dimensional Hex board with at most n colours has a winner (for $n > 2$ there may be more than one).

Lemma 66 *Let f be a function from the unit hypercube $[0,1]^n$ into itself. Then for all $\varepsilon, \delta > 0$ either there exists $x \in [0,1]^n$ such that $\rho(x, f(x)) < \varepsilon$ or there exist $x, x' \in [0,1]^n$ such that $\rho(x, x') < \delta$ and $\rho(f(x), f(x')) > \varepsilon$.*

Proof. Write

$$f(x) = (f_1(x), \ldots, f_n(x))\,;$$

Fixing $\varepsilon, \delta > 0$, without loss of generality take $\delta < \varepsilon/3$. Pick $N > 0$ such that $1/N < \delta$, and subdivide $[0,1]^n$ into an n-dimensional Hex board of size N. We partition the set V of vertices of this Hex board into sets

[46] We use $\|\cdot\|$ throughout this section to represent the maximum norm.

$C_1^+, C_1^-, \ldots, C_n^+, C_n^-$, and B such that

$$\begin{aligned} x \in C_1^+ &\Rightarrow f_1(x) - x_1 > \frac{2\varepsilon}{3}; \\ x \in C_1^- &\Rightarrow x_1 - f_1(x) > \frac{2\varepsilon}{3}; \\ &\vdots \\ x \in C_n^+ &\Rightarrow f_n(x) - x_n > \frac{2\varepsilon}{3}; \\ x \in C_n^- &\Rightarrow x_n - f_n(x) > \frac{2\varepsilon}{3}; \\ x \in B &\Rightarrow \|f(x) - x\| < \varepsilon. \end{aligned}$$

By the Hex theorem, either B is inhabited, and there exists $x \in [0,1]$ such that $\rho(x, f(x)) < \varepsilon$, or, as we may assume, there exists an i-path from i-bank 1 to i-bank 2 for some $1 \leqslant i \leqslant n$. Since no vertex of C_i^+ has i-th coordinate 1 and no vertex of C_i^- has i-th coordinate 0, such a path contains points from each set. Hence there exist adjacent vertices x, x' such that $x \in C_i^+$ and $x' \in C_i^-$. Then $\|x - x'\| < \delta < \varepsilon/3$, $f_i(x) > f_i(x')$, and

$$\begin{aligned} f_i(x) - f_i(x') &= (f_i(x) - x) + \left(x - x'\right) + \left(x' - f_i(x')\right) \\ &> \frac{2\varepsilon}{3} - \frac{\varepsilon}{3} + \frac{2\varepsilon}{3} = \varepsilon. \end{aligned}$$

Therefore $\rho(f(x), f(x')) > |f_i(x) - f_i(x')| > \varepsilon$. ■

With this lemma at hand we can weaken the standard hypothesis of the approximate Brouwer fixed point theorem. A function $f : [0,1]^n \to [0,1]^n$ is *uniformly sequentially continuous*[47] if for all sequences $(x_n)_{n\geqslant 1}, (y_n)_{n\geqslant 1}$ in $[0,1]^n$, if $\rho(x_n, y_n)$ tends to zero as $n \to \infty$, then $\rho(f(x_n), f(y_n))$ also tends to zero as $n \to \infty$. It is easy to see that uniform continuity implies uniform sequential continuity; the converse is equivalent to **BD-N** and hence cannot be proved constructively (see [31]).

[47] Throughout this section and the next, uniformly sequentially continuous can be substituted for uniformly continuous in the definition of the approximate fixed point property.

Theorem 67 (CC) *Let f be a uniformly sequentially continuous function from the unit hypercube $[0,1]^n$ into itself. Then f has approximate fixed points.*

Proof. We construct, using countable choice, sequences $(x_n)_{n\geqslant 1}, (x'_n)_{n\geqslant 1}$ as follows. For each $n \in \mathbf{N}$, apply Lemma 66 to construct either $x \in [0,1]^n$ such that $\rho(x, f(x)) < \varepsilon$ or $x, x' \in [0,1]^n$ such that $\rho(x, x') < 1/n$ and $\rho(f(x), f(x')) > \varepsilon$. In the latter case we set $x_n = x$ and $x'_n = x'$; in the former we set $x_n = x'_n = x$. Then $(\rho(x_n, x'_n))_{n\geqslant 1}$ converges to zero. Since f is uniformly sequentially continuous, there exists $N \in \mathbf{N}$ such that $\rho(f(x_n), f(x'_n)) < \varepsilon$ for all $n \geqslant N$. It follows that $\rho(x_N, f(x_N)) < \varepsilon$. ∎

Without countable choice we seem to require f to be uniformly continuous:

Corollary 68 *Let f be a uniformly continuous function from the unit hypercube $[0,1]^n$ into itself. Then f has approximate fixed points.*

Proof. For a given $\varepsilon > 0$, let δ be as in the definition of uniform continuity. Then applying Lemma 66 must produce an ε-fixed point of f, for the other disjunct is ruled out by the choice of δ. ∎

Next we extend the approximate Brouwer fixed point theorem, for uniformly continuous functions, to compact convex subsets of $\mathbf{R}^n$. A subset S of a normed space X is *strictly convex* if for each $\varepsilon > 0$ there exists $\delta > 0$ such that for all x, y in the boundary ∂S of S, if[48] $\rho\left(\frac{1}{2}(x-y), \partial S\right) < \delta$, then $\|x - y\| < \varepsilon$. A normed space X is *uniformly convex* if its unit ball is strictly convex: for all $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x, y \in X$ with $\|x\| = \|y\| = 1$, if $\left\|\frac{1}{2}(x-y)\right\| > 1 - \delta$, then $\|x - y\| < \varepsilon$. Any inner product space is uniformly convex [32, Page 93], and the L_p spaces for $1 < p < \infty$ are uniformly convex [18, Chapter 7, (3.22)].

Let S be an inhabited subset of a metric space X, and let $x \in X$. We say that $b \in S$ is a *closest point*, or *best approximation*, to x in S if $\rho(x, b) \leq \rho(x, s)$ for all $s \in S$. The following extends Theorem 6 of [30].

[48] We do not require ∂S to be located here: for an arbitrary subset S of a metric space X we use '$\rho(x, S) < \varepsilon$' as a shorthand for 'there exists $s \in S$ with $\rho(x, s) < \varepsilon$. If S is located, then this coincides with the standard meaning.

Theorem 69 *Let S be a complete, located, convex subset of a uniformly convex normed space X. Then each point in X has a unique closest point in S. Moreover, the mapping Q from X to S sending x to the best approximation to x in S is uniformly continuous.*

Proof. The proof of Theorem 6 in [30] establishes that for each $x \in X$ and each $\varepsilon > 0$ the set

$$S_\varepsilon^x = \{y \in S : \rho(x, y) < \rho(x, S) + \varepsilon\}$$

has diameter no greater than ε, and hence that x has a unique best approximation in S. To see that Q is uniformly continuous, observe that if $\|x - y\| < \varepsilon/2$, then $S_{\varepsilon/2}^y \subset S_\varepsilon^x$. Hence $Q(x), Q(y) \in S_\varepsilon^x$, so $\|Q(x) - Q(y)\| \leqslant \varepsilon$. ∎

We call the mapping Q from the preceding theorem the *projection onto S*.

Theorem 70 *Every totally bounded set S of $\mathbf{R}^n$ with convex closure has the approximate fixed point property.*

Proof. Let S be a subset of $\mathbf{R}^n$ satisfying the conditions of the theorem; without loss of generality we may assume that S is both closed and a subset of the unit cube $[0,1]^n$. Fix $\varepsilon > 0$ and let Q be the projection mapping from $[0,1]^n$ onto S (which exists, by the preceding theorem). Applying Theorem 67 to the mapping $f \circ Q : [0,1]^n \to [0,1]^n$, construct $x \in [0,1]^n$ such that $\|x - f \circ Q(x)\| < \varepsilon/2$. Then

$$\begin{aligned} \|x - Q(x)\| &= \rho(x, S) \\ &\leqslant \|x - f \circ Q(x)\| < \frac{\varepsilon}{2}, \end{aligned}$$

so

$$\begin{aligned} \|Q(x) - f(Q(x))\| &\leqslant \|Q(x) - x\| + \|x - f \circ Q(x)\| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Hence $Q(x)$ is an ε-fixed point of f. ∎

For a subset S of a metric space X we write

$$S_\varepsilon = \{y \in X : \rho(x,y) < \varepsilon \text{ for some } x \in S\}.$$

Classically, Brouwer's fixed point theorem holds for any metric space which is homeomorphic to $[0,1]^n$; this also holds constructively. For subsets of uniformly convex normed spaces, this result is classically equivalent to, but seems constructively weaker than, the following.

Proposition 71 *Let X be a uniformly convex normed space, let S be a subset of X with the approximate fixed point property, and let T be a subset of X such that for each $\varepsilon > 0$ there exists a uniformly bicontinuous*[49] *function $f_\varepsilon : S \to T$ such that $f_\varepsilon(S)$ is convex and totally bounded and*

$$f_\varepsilon(S) \subset T \subset (f_\varepsilon(S))_\varepsilon.$$

Then T has the approximate fixed point property.

Proof. Let f be a uniformly continuous function from T into itself, and fix $\varepsilon > 0$. Let $\delta > 0$ be such that for all $x, y \in T$, if $\|x - y\| < \delta$, then

$$\left\|f_{\varepsilon/2}(x) - f_{\varepsilon/2}(y)\right\| < \varepsilon/2,$$

where $f_{\varepsilon/2}$ is as in the statement of the proposition. Let Q be the projection onto $f_{\varepsilon/2}(S)$ restricted to T, and let $I : f_{\varepsilon/2}(S) \to T$ be the inclusion mapping; note that $\|Q(t) - t\| < \varepsilon/2$ for all $t \in T$.

[49] A function f from X onto Y is *uniformly bicontinuous* if both f and its inverse are uniformly continuous.

$$\begin{tikzcd}
S \arrow[r, bend left, "f_{\varepsilon/2}"] & f_{\varepsilon/2}(S) \arrow[l, bend left, "f_{\varepsilon/2}^{-1}"] \arrow[r, bend left, "I"] & T \arrow[l, bend left, "Q"] \arrow[loop right, "f"]
\end{tikzcd}$$

Then $f_{\varepsilon/2}^{-1} \circ Q \circ f \circ I \circ f_{\varepsilon/2}$ is a uniformly continuous function from S into S. Hence there exists $x \in S$ such that

$$\|f_{\varepsilon/2}^{-1} \circ Q \circ f \circ f_{\varepsilon/2}(x) - x\| < \delta.$$

Then

$$\|Q \circ f \circ f_{\varepsilon/2}(x) - f_{\varepsilon/2}(x)\| < \varepsilon/2,$$

and so

$$\begin{aligned}
\left\|f\left(f_{\varepsilon/2}(x)\right) - f_{\varepsilon/2}(x)\right\| &\leqslant \left\|f \circ f_{\varepsilon/2}(x) - Q \circ f \circ f_{\varepsilon/2}(x)\right\| \\
&\qquad + \left\|Q \circ f \circ f_{\varepsilon/2}(x) - f_{\varepsilon/2}(x)\right\| \\
&\leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \;=\; \varepsilon.
\end{aligned}$$

Thus $Q\left(f_{\varepsilon/2}^{-1}(x)\right)$ is an ε-fixed point of f. ■

Brouwer's fixed point theorem implies the intermediate value theorem, and Edelstein's fixed point theorem, and hence implies **LLPO**. Corollary 68 together with **MIN** gives Brouwer's fixed point theorem for uniformly continuous functions; Brouwer's fixed point theorem is also, in the absence of choice, strictly stronger than **LLPO** since it implies $\mathbf{LLPO_R}$. Thus Brouwer's fixed point theorem for uniformly continuous functions is in $(\mathbf{LLPO}, \mathbf{MIN}]$. In the presence of $\mathbf{AC}_{\omega,2}$, Brouwer's fixed point theorem for sequentially continuous function also follows from **WKL**: let $f : [0,1]^n \to [0,1]^n$ be a sequentially continuous function, we first approximate f by a sequence $(f_n)_{n \geqslant 1}$ of affine, and therefore uniformly continuous, functions to which, given **WKL**, we apply Brouwer's fixed point theorem for uniformly continuous functions to produce a sequence of points $(x_n)_{n\geqslant 1}$ such that $x_n = f_n(x_n)$ for each n. By **WKL**, we may suppose that $(x_n)_{n\geqslant 1}$ converges to some point x. The sequential continuity of f then ensures that x is a fixed point of f.

Orehkov [89] constructed a continuous function f in **RUSS** which maps the unit square into itself and which does not have a fixed point: for all $x \in [0,1]^2$, $f(x) \neq x$. Consequentially, it is not likely that Brouwer's fixed point theorem holds constructively even if we impose a natural uniqueness condition. In particular, the condition that f *has at most one fixed point*

$$\text{for all distinct } x, y \in \mathrm{dom(f)}, \text{ either } f(x) \neq x \text{ or } f(y) \neq y$$

is not sufficient to give the existence of an exact fixed point constructively. On the other hand, it is easy to see, given Corollary 68, that it suffices for f to have a uniformly unique fixed point. Wim Veldman [110] has shown that Brouwer's fixed point theorem for pointwise continuous functions is equivalent to $\mathbf{FT}_\Delta$.

We finish with a question: is there a model of **IZF** which validates the intermediate value theorem and in which Brouwer's fixed point theorem fails? If the answer is yes, then it would show that any classical proof of Brouwer's fixed point theorem by induction must appeal to excluded middle in the inductive step.

3.4 Schauder's fixed point theorem

In this section we extend Brouwer's fixed point theorem by considering compact convex subsets of arbitrary Banach spaces; this gives us Schauder's fixed point theorem. We also consider Rothe's further extension.

3.4.1 Approximate fixed points

We call a located subset S of a normed space X *projective* if there exists a uniformly continuous *projection* function Q of X onto S such that $\rho(x, Q(x)) = \rho(x, S)$ for each x in X. We give an approximate version of Schauder's fixed point theorem for projective sets.

Lemma 72 *Let S be a totally bounded subset of a metric space X, fix $\beta > \alpha > 0$, and let S' be a convex set such that for each $x \in S$ there exists $x' \in S'$ such that $\rho(x, x') < \alpha/2$. Then there exists a uniformly continuous function P from S into S' such that $\|P(x) - x\| < \beta$ for all $x \in S$.*

Proof. Let $\{x_1, \ldots, x_n\}$ be an $\alpha/2$-approximation to S, and for each $1 \leqslant i \leqslant n$ pick $x_i' \in S'$ such that $\rho(x_i, x_i') < \alpha/2$. Then for each $x \in S$ there exists i such that $\rho(x, x_i') < \alpha$.

Let $f_1, \ldots, f_n$ be the uniformly continuous functions from S into $\mathbf{R}$ given by

$$f_i(x) = \max\left\{0, \gamma - \|x - x_i'\|\right\},$$

where $\gamma = (\alpha + \beta)/2$. Then for each $x \in S$, there exists i such that

$$f_i(x) > \gamma - \alpha;$$

whence

$$P(x) \equiv \frac{\sum_{i=1}^{n} f_i(x) x_i'}{\sum_{i=1}^{n} f_i(x)}$$

defines a uniformly continuous map from S into S'.

Let $r > 0$ and write $\{1, \ldots, n\}$ as the disjoint union of two sets P, Q such that

$$\begin{aligned} i \in P &\Rightarrow \|x - x'\| < \gamma + r; \\ i \in Q &\Rightarrow \|x - x'\| > \gamma. \end{aligned}$$

Then $P(x)$ is a convex combination of points in P, so

$$\|P(x) - x\| \leqslant \max\{\|x - x_i'\| : i \in P\} < \gamma + r.$$

Since $r > 0$ is arbitrary, it follows that $\|P(x) - x\| \leqslant \gamma < \beta$ for all $x \in S$. ■

Theorem 73 *Let S be an inhabited, totally bounded, projective subset of a normed space X. Then S has the approximate fixed points property.*

Proof. Let $f : S \rightarrow$ be a uniformly continuous function. Fixing $\varepsilon > 0$, let $\{x_1, \ldots, x_n\}$ be an $\varepsilon/8$-approximation to S. Using [18, Lemma 2.5, Chapter 7], construct a finite-dimensional subspace V of X, with a basis contained in S, such that

$$\rho(x_i, V) < \varepsilon/8$$

for all $i \in \{1, \ldots, n\}$. For each such i pick $x_i' \in V$ such that $\|x_i - x_i'\| < \varepsilon/8$. Then for each $x \in S$, there exists $i \in \{1, \ldots, n\}$ such that $\|x - x_i'\| < \varepsilon/4$. Let S be the closed convex hull of $\{x_1', \ldots, x_n'\}$, and let $Q : S' \rightarrow S$ be the restriction to S' of the projection onto S. If $\sum_{i=1}^n \lambda_i = 1$ and each $\lambda_i \geq 0$, then

$$\begin{aligned} \left\| Q\left(\sum_{i=1}^n \lambda_i x_i'\right) - \sum_{i=1}^n \lambda_i x_i' \right\| &\leq \sum_{i=1}^n \lambda_i \left\| Q x_i' - x_i' \right\| \\ &= \sum_{i=1}^n \lambda_i \rho\left(x_i', S\right) \\ &\leq \sum_{i=1}^n \lambda_i \left\| x_i' - x_i \right\| < \frac{\varepsilon}{8}; \end{aligned}$$

thus $\|Q(x) - x\| < \varepsilon/4$ for all $x \in S'$.

Using Lemma 72, construct a uniformly continuous function $P : S \to S'$ such that $\|P(x) - x\| < \varepsilon/3$ for all $x \in S$. Then $P \circ f \circ Q$ is a uniformly continuous map from S' into S'; by Brouwer's fixed point theorem, Theorem 67, there exists $x' \in S'$ such that

$$\|P \circ f \circ Q(x') - x'\| < 5\varepsilon/12;$$

write $x = Q(x')$. Then $x \in S$ and

$$\begin{aligned} \|f(x) - x\| &\leqslant \|f(x) - P \circ f \circ Q(x')\| + \|P \circ f \circ Q(x') - x'\| + \|x' - x\| \\ &= \|f(x) - P(f(x))\| + \|P \circ f \circ Q(x') - x'\| + \|x' - Q(x')\| \\ &< \varepsilon/3 + 5\varepsilon/12 + \varepsilon/4 = \varepsilon. \end{aligned}$$

Hence x is an ε-fixed point of f. ■

By Theorem 69, every complete, located, convex subset of a uniformly convex space is projective; this gives us the following result.

Corollary 74 *Let S be an inhabited, totally bounded subset of a uniformly convex normed space X such that the closure $\overline{S}$ of S is convex. Then S has the approximate fixed points property.*

Proof. Since we are interested in approximate fixed points, replacing X with its completion $\hat{X}$ and S with its closure in $\hat{X}$, we may assume that S is compact and convex. The result then follows from Theorems 69 and 73. ■

Strictly convex sets are also projective; the proof is similar to that of Theorem 69 (which, in turn, is based on the proof of Theorem 6 of [30]).

Theorem 75 *Let S be an inhabited, complete, located, strictly convex subset of a normed space X. Then each point in X has a unique closest point in S. Moreover, the mapping Q from X to S sending x to the best approximation to x in S is uniformly continuous.*

Proof. Let x be a point of X, fix $\varepsilon > 0$, and let $\delta \in (0, \varepsilon/2)$ be such that for all $x, y \in \partial S$, if $\rho\left(\frac{1}{2}(x-y), \partial S\right) < 2\delta$, then $\|x - y\| < \varepsilon/2$. Set

$$S_\varepsilon^x = \{y \in S : \|x - y\| < \rho(x, S) + \delta/2\};$$

and fix $y_1, y_2 \in S_\varepsilon^x$. Either $\rho(x, S) < \delta/2$ and

$$\|y_1 - y_2\| \leqslant \|y_1 - x\| + \|y_2 - x\| < \delta + \delta < \varepsilon,$$

or, as we may assume, $\rho(x, S) > 0$. Since S is located $S \cup {\sim}S$ is dense in X; whence we can apply Proposition 5.15 of [32] to construct the unique points y_1', y_2' such that y_i' is in the intersection of

$$[x, y_i] \ \equiv \ \{tx + (1-t)y : t \in [0, 1]\}$$

and ∂S. Then, for $i = 1, 2$,

$$\rho(y_i, y_i') = \rho(x, y_i) - \rho(x, y_i') < \rho(x, S) + \delta/2 - \rho(x, S) = \delta/2,$$

so

$$\begin{aligned} \left\| x - \frac{1}{2}(y_1' + y_2') \right\| &\leqslant \frac{1}{2}\|x - y_1'\| + \frac{1}{2}\|x - y_2'\| \\ &\leqslant \frac{1}{2}\left(\|x - y_1\| + \|y_1 - y_1'\| + \|x - y_2\| + \|y_2 - y_2'\|\right) \\ &< \rho(x, S) + 2\delta. \end{aligned}$$

Apply Proposition 5.15 of [32] again to construct the unique point z in the intersection of $[x, \dfrac{1}{2}(y_1', y_2')]$ and the boundary of S. Then

$$\begin{aligned} \rho(z, \frac{1}{2}(y_1' + y_2')) &= \rho(x, \frac{1}{2}(y_1' + y_2')) - \rho(x, z) \\ &< \rho(x, S) + 2\delta - \rho(x, S) \ = \ 2\delta. \end{aligned}$$

Therefore, by our choice of δ, $\|y_1' - y_2\prime\| < \varepsilon/2$, and so

$$\begin{aligned}\rho(y_1, y_2) &\leqslant \rho(y_1, y_1') + \rho(y_1', y_2') + \rho(y_2', y_2) \\ &< \delta/2 + \varepsilon/2 + \delta/2 \ < \ \varepsilon.\end{aligned}$$

Hence the diameter of S_ε^x is no greater than ε. The proof then proceeds as in Theorem 69. ■

Corollary 76 *Let S be an inhabited, totally bounded, strictly convex subset of a normed space X such that the closure of S is strictly convex. Then S has the approximate fixed points property.*

Proof. As for Corollary 8. ■

In the proof of Theorem 73, we begin by approximating the convex, totally bounded set S by a set S' contained in a finite dimensional subspace of X. We then use f to define a uniformly continuous map from S' into itself to which we can apply Brouwer's fixed point theorem. In particular, this requires us to construct a uniformly continuous map from S' into S which is close to the identity map; it is in order to construct this mapping that we require S to be projective. In the following result we circumvent this requirement: by considering only open sets, we can ensure that S' is contained in S; we can then produce a uniformly continuous function from S' into S' by restricting the domain of f, rather than composing f with a mapping from S' into S, as in Theorem 73.

Theorem 77 *Every inhabited, open, totally bounded, convex subset of a normed space has the approximate fixed points property.*

Proof. The proof is similar to that of the preceding theorem. Let S be an inhabited, open, totally bounded, convex subset of a normed space X, and let $f : S \to S$ be a uniformly continuous function. Let $\{x_1, \ldots, x_n\}$ be an $\varepsilon/6$-approximation to S. We construct, as follows, a finite-dimensional subspace V of X such that V contains an $\varepsilon/3$-approximation $\{x_1', \ldots, x_n'\}$ to S. Let $V_1 = \operatorname{span}\{x_1\}$ and $x_1' = x_1$. Suppose that we have constructed

V_{k-1} and $x'_1, \ldots, x'_k$ and let $r \in (0, \varepsilon/6)$ be such that $B(x_k, r) \subset S$. Either $\rho(x_k, V_{k-1}) > 0$ or $\rho(x_k, V_{k-1}) < r$. In the first case we set

$$V_k = \operatorname{span}\{V_{k-1}, x_k\}$$

and $x_k = x'_k$. In the second case, pick $x'_k \in V$ such that $\|x_k - x'_k\| < r$ and set $V_k = V_{k-1}$. Set $V = V_n$; it is easy to see that $\{x'_1, \ldots, x'_n\}$ is an $\varepsilon/3$-approximation to S.

Let S' be the convex hull of $\{x'_1, \ldots, x'_n\}$. Then $S' \subset S$ and, by Lemma 72, there exists a uniformly continuous function $P : S \to S'$ such that $\|P(x) - x\| < \varepsilon/2$ for all $x \in S$. Using Brouwer's fixed point theorem, applied to $P \circ f|_{S'} : S' \to S'$, construct $x \in S'$ such that $\|P \circ f(x) - x\| < \varepsilon/2$. Then

$$\begin{aligned} \|f(x) - x\| &\leqslant \|f(x) - P \circ f(x)\| + \|P \circ f(x) - x\| \\ &< \varepsilon/2 + \varepsilon/2 \;=\; \varepsilon. \end{aligned}$$

Hence x is an ε-fixed point of f. ■

We can extend Theorem 73 to give an approximate version of Rothe's theorem [94, 101] for projective sets.

Theorem 78 *Let S be a compact, convex, projective subset of a normed space X, and let f be a uniformly continuous function from S into X which maps the boundary of S into S. Then f has approximate fixed points.*

Proof. Fix $\varepsilon > 0$, let Q be the projection onto S, and let $\delta \in (0, \varepsilon/4)$ be such that if $\|x - y\| < \delta$, then $\|f(x) - f(y)\| < \varepsilon$. Since $Q \circ f$ is a uniformly continuous function from S into S, it follows from Schauder's fixed point theorem for projective sets that there exists $x \in S$ such that $\|Q \circ f(x) - x\| < \delta$. Suppose that $\rho(f(x), S) > \varepsilon/4$. Then$f(x) \notin S$ and $\rho(x, y) \geqslant \delta$ for all $y \in \partial S$—this contradicts that $\|Q \circ f(x) - x\| < \delta$.

Therefore $\rho(f(x),S) \leqslant \varepsilon/4$, so

$$\begin{aligned}\|f\circ Q(x)-Q(x)\| &\leqslant \|f\circ Q(x)-f(x)\|+\|f(x)-Q\circ f(x)\| \\ &\qquad +\|Q\circ f(x)-x\|+\|x-Q(x)\| \\ &< \varepsilon/4+\varepsilon/4+\delta+\delta<\varepsilon.\end{aligned}$$

Hence $Q(x)$ is an ε-fixed point of f. ■

Since Schauder's fixed point theorem implies Brouwer's fixed point theorem and follows from **MIN**, it is in $(\mathbf{LLPO}, \mathbf{MIN}]$.

3.4.2 An application: Peano's existence theorem

We give an application of the approximate Schauder fixed point theorem for uniformly convex spaces (Corollary 74). A standard application of Schauder's fixed point theorem is in proving Peano's Theorem asserting the existence of solutions to particular differential equations:

> **Peano's Theorem** Let A be a closed subset of $\mathbf{R}$, let $(x_0, y_0) \in A$, and let $r > 0$ be such that if $|x - x_0| \leqslant r$ and $|y - y_0| \leqslant r$, then $(x, y) \in A$. Let $f : A \to \mathbf{R}$ be continuous, let
>
> $$M \geqslant \sup\{|f(x,y)| : |x-x_0| \leqslant r, |y-y_0| \leqslant r\},$$
>
> and set $h = \min\{r, r/M\}$. Then the differential equation
>
> $$y' = f(x,y), y(x_0) = y_0 \qquad (2)$$
>
> has a solution y on the interval $[x_0 - h, x_0 + h]$.

However, since the exact version of Peano's Theorem implies **LLPO** (see [35], which also gives an alternative constructive proof of an approximate Peano's Theorem, [9] gives a proof that Peano's Theorem implies **LLPO**), we can only hope to prove an approximate version of Peano's Theorem.

There is another, more pressing, problem: Peano's Theorem asserts the existence of solutions to particular differential equations in the normed space $\mathcal{C}(I)$, for some interval I, with the supremum norm, but this normed space is not uniformly convex. To overcome this difficulty, we first approximate the sup norm with a uniformly convex norm—this relies on being able to restrict the possible solutions of (2) to a sufficiently friendly subset of $\mathcal{C}(I)$.

A solution to the differential equation (2) on an interval I in $\mathbf{R}$ is precisely a fixed point of the mapping $U : \mathcal{C}(I) \to \mathcal{C}(I)$ given by

$$U(y) = y_0 + \int_{x_0}^{x} f(t, y(t))dt.$$

The differential equation (2) is said to have approximate solutions on an interval I if for all $\varepsilon > 0$ there exists a continuous function $y : I \to \mathbf{R}$ such that $\|U(y) - y\| < \varepsilon$.

To prove a constructive version of Peano's theorem, we need the following lemma. A subset S of $\mathcal{C}(I)$ is *Lipschitz* if there exists $M > 0$ such that for all $y \in S$ and all $x_1, x_2 \in I$ we have

$$|y(x_1) - y(x_2)| \leqslant M|x_1 - x_2|;$$

that is, M is a Lipschitz constant for each $y \in S$. We call M a *Lipschitz constant* for S.

Lemma 79 *If S is a bounded Lipschitz subset of $\mathcal{C}(I)$, then for each $\varepsilon > 0$ there exists $p > 1$ such that $|\ \|y\| - \|y\|_p| < \varepsilon$ for all $y \in S$.*

Proof. Fix $\varepsilon > 0$ and let N be a bound for S and M be a Lipschitz constant for S. It suffices to choose $p > 1$ such that $|\|y\| - \|y\|_p| < \varepsilon$, where y is given by

$$y(x) = \max\left\{M, \frac{4N}{b-a}\right\}\left(1 - \frac{2}{b-a}\left|x - \frac{a+b}{2}\right|\right) - N.$$

This is possible because, in $\mathcal{C}([-1,1])$,

$$\begin{aligned}
\|1-|x|\|_p &= \left(\int_{-1}^{1} |1-|x||^p dt\right)^{1/p} \\
&= \left(2\int_{0}^{1} |1-x|^p dt\right)^{1/p} \\
&= \left(\frac{2}{p+1}\right)^{1/p} \longrightarrow 1,
\end{aligned}$$

as $p \to \infty$. ■

Theorem 80 *Let $A \subset \mathbf{R}^2$ be closed, $(x_0, y_0) \in A^\circ$, and $r > 0$ be such that if $|x - x_0| \leqslant r$, then $(x, y) \in A$. Let $f : A \to \mathbf{R}$ be uniformly continuous, let*

$$M \geqslant \sup \{|f(x,y)| : |x-x_0| \leqslant r, |y-y_0| \leqslant r\},$$

and let $h = \min\{r, r/M\}$. Then the differential equation

$$y' = f(x,y), \quad y(x_0) = y_0$$

has approximate solutions.

Proof. Fix $\varepsilon > 0$, let $I = [x_0 - h, x_0 + h]$, and set

$$\mathcal{M} = \{y \in \mathcal{C}(I) : |y(t) - y_0| \leqslant r \text{ for all } t \in I\}.$$

Since f is uniformly continuous, U is also uniformly continuous. Define

$$\begin{aligned}
S = \{ & y \in \mathcal{M} : (\|Uy\| \leqslant |y_0| + Mh) \wedge \\
& (\forall_{x_1,x_2 \in I} |y(x_1) - y(x_2)| \leqslant M|x_1 - x_2|)\}.
\end{aligned}$$

Let $y \in \mathcal{M}$ and $t \in I$. Then

$$\begin{aligned}
|Uy(t) - y_0| &= \left| \int_{x_0}^{x} f(t, y(t)) dt \right| \leqslant Mh \leqslant r, \\
\|Uy\| &= \left| y_0 + \int_{x_0}^{x} f(t, y(t)) dt \right| \leqslant |y_0| + Mh, \text{ and} \\
|Uy(x_1) - Uy(x_2)| &\leqslant \left| \int_{x_1}^{x_2} f(t, y(t)) dt \right| \\
&\leqslant M|x_1 - x_2|.
\end{aligned}$$

Hence U maps $\mathcal{M}$ into S. By (a slight variation of) [18, (5.6) pg. 102], S is compact; and, by Lemma 79, there exists $p > 1$ such that

$$|\|y\| - \|y\|_p| < \varepsilon/2,$$

for all $y \in S$. We can now apply the Schauder fixed point theorem to $U|_S$ to construct a $y \in S$ such that $\|Uy - y\|_p < \varepsilon/2$. Then

$$\begin{aligned}
\|Uy - y\| &< \|Uy - y\|_p + \varepsilon/2 \\
&< \varepsilon/2 + \varepsilon/2 = \varepsilon,
\end{aligned}$$

so y is an ε-fixed point of U. ■

The above proof readily extends to a system

$$\begin{aligned}
y_1' &= f_1(y_1, \ldots, y_n, x) \\
y_2' &= f_2(y_1, \ldots, y_n, x) \\
&\vdots \\
y_n' &= f_n(y_1, \ldots, y_n, x)
\end{aligned}$$

of linear ordinary differential equations.

3.5 Kakutani's fixed point theorem

In this section we give a constructive treatment of Kakutani's extension of Brouwer's fixed point theorem [68].

Let U be a function from a metric space X into the class $\mathcal{P}^*(X)$ of nonempty subsets of X; U is said to be a *set valued mapping on* X. We say that U is *convex* (*compact*, *closed*, etc.) if $U(x)$ is convex (compact, closed, etc.) for each $x \in X$. A mapping $U : X \to \mathcal{P}^*(X)$ is said to be *sequentially upper hemi-continuous* if for each pair of sequences $(x_n)_{n\geqslant 1}$, $(y_n)_{n\geqslant 1}$ in X converging to points x, y in X respectively, if $y_n \in U(x_n)$ for each n, then $\rho(y, U(x)) = 0$; in particular, if U is closed, then $y \in U(x)$. If U is closed, then U is sequentially upper hemi-continuous if and only if the *graph*

$$G(U) = \bigcup_{x\in X} \{x\} \times U(x)$$

of U is closed. A point $x \in S$ such that $x \in U(x)$ is called a *fixed point of* U. Kakutani's fixed point theorem is the following.

> **Kakutani's fixed point theorem** Let S be a compact, convex subset of $\mathbf{R}^n$ and let $U : S \to \mathcal{P}^*(S)$ be a closed, convex, sequentially upper hemi-continuous mapping. Then U has a fixed point.

If f is sequentially continuous, then $U(x) = \{f(x)\}$ is sequentially upper hemi-continuous; this shows that Kakutani's fixed point theorem is a generalisation of Brouwer's fixed point theorem, and hence Kakutani's fixed point theorem is not constructively valid.

Classically, the Kakutani fixed point theorem is equivalent to the Brouwer fixed point theorem (in the sense that it is straightforward to prove one given the other; in reverse mathematics, both theorems are equivalent to $\mathbf{WKL}_0$). The constructive proof of (an approximate) Brouwer fixed point theorem, for a uniformly continuous function f from $[0, 1]^n$ into itself, uses a combinatorial argument to show that for all $\delta, \varepsilon > 0$ either there exists $x, y \in [0,1]^n$ such that $\rho(x,y) < \delta$ and $\rho(f(x), f(y)) > \varepsilon$, or we can construct $x \in [0,1]^n$ such that $\rho(x, f(x)) < \varepsilon$. Given $\varepsilon > 0$, the former possibility is then ruled out by using the uniform continuity of f to choose an appropriate δ. It is clear that f must satisfy some uniform form of continuity for this approach to work; indeed Wim Veldman has shown that the approximate Brouwer fixed point theorem for pointwise continuous functions is equivalent to $\mathbf{FT}_\Delta$ [110]. If we are to prove a constructive version of Kakutani's fixed point theorem using the constructive approximate Brouwer fixed point theorem, we must therefore have our set valued mappings satisfy some form of uniform continuity.

This section is broken up into three parts. The first examines Kakutani's original proof from a constructive perspective, and in the second part we discuss the difficulties of formulating an appropriate notion of uniform continuity for set valued mappings, and hence a constructive version of the Kakutani fixed point theorem. We then give our constructive version of Kakutani's fixed point theorem in the final part; this result has, classically, both weaker hypothesis and a weaker conclusion—we only construct approximate fixed points—than the classical version, but is classically equivalent to the standard formulation.

3.5.1 Kakutani's proof

Our first question is: what is the constructive content of the standard classical proofs of Kakutani's fixed point theorem?

As we saw before, Kakutani's fixed point theorem implies **LLPO**; thus any classical proof of Kakutani's fixed point theorem must be non-constructive. On the other hand, Kakutani's original proof of his theorem only requires (several applications of) weak König's lemma, in addition to intuitionistic logic.

Theorem 81 (WCC) *Kakutani's fixed point theorem follows from **WKL**.*

Proof. Suppose **WKL** holds. We give Kakutani's original proof, adapted to the unit hypercube. Let U be a sequentially upper hemi-continuous set valued mapping on $[0,1]^n$. For each $k \in \mathbf{N}$ let f_k be the affine extension of a function on

$$\{0, 1/k, \ldots, 1\}^n$$

which takes values in $U(x)$ for each x in its domain. Since **WKL** implies Brouwer's fixed point theorem, we can construct a sequence $(x_k)_{k \geqslant 1}$ such that $x_k = f_k(x_k)$ for each $k \in \mathbf{N}$; by **WKL** we may suppose that $(x_k)_{k \geqslant 1}$ converges to some point $x_0 \in [0,1]^n$. Since **LLPO** allows us to decide whether $a \leqslant 0$ or $a \geqslant 0$ for each $a \in \mathbf{R}$, for each $k \in \mathbf{N}$ there exists $s \in S_k$ such that

$$x_k \in \{x \in [0,1]^n : 0 \leqslant x_i - s_i \leqslant 1/k, 1 \leqslant i \leqslant n\} \equiv \{x_1^k, \ldots, x_{2^n}^k\}.$$

Let $\lambda_1^k, \ldots, \lambda_{2^n}^k$ $(k \in \mathbf{N})$ be such that $\lambda_i^k \geqslant 0$ for each i, $\sum_{i=1}^{2^n} \lambda_i^k = 1$, and

$$x_k = \sum_{i=1}^{2^k} \lambda_i^k x_i^k.$$

Set $y_i^k = f_k(x_i^k)$ for each $k \in \mathbf{N}, 1 \leqslant i \leqslant 2^k$. Then $y_i^k \in U(x_i^k)$ for all i, k and

$$x_k = f_k(x_k) = \sum_{i=1}^{2^k} \lambda_i^k y_i^k.$$

Applying **WKL** repeatedly, we may assume that for each $1 \leqslant i \leqslant 2^n$ there exist sequences $\left(\lambda_i^k\right)_{k \geqslant 1}$, $\left(y_i^k\right)_{k \geqslant 1}$ such that $\left(\lambda_i^k\right)_{k \geqslant 1}$ converges to λ_i^0 in $\mathbf{R}$ and $\left(y_i^k\right)_{k \geqslant 1}$ converges to y_i^0 in $[0,1]^n$. Then $\lambda_i^0 \geqslant 0$ for each i, $\sum_{i=1}^{2^n} \lambda_i^0 = 1$, and

$$x_0 = \sum_{i=0}^{2^n} \lambda_i^0 y_i^0.$$

Moreover, $x_i^k \to x_0$, $y_i^k \to y_i^0$, and $y_i^k \in U(x_i^k)$ for each i; whence, since U is closed and upper hemi-continuous, $y_i^0 \in U(x_0)$ for each i. It now follows from the convexity of U that

$$x_0 = \sum_{i=0}^{2^n} \lambda_i^0 y_i^0 \in U(x_0);$$

that is, x_0 is a fixed point of U.

We extend this to a closed, convex, sequentially upper hemi-continuous mapping U on an arbitrary convex compact subset S of $\mathbf{R}^n$ as follows. Let Q be the uniformly continuous function from $\mathbf{R}^n$ into S which takes a point x of $\mathbf{R}^n$ to the (unique) closest point to x in S; this function exists and is uniformly continuous (Lemma 32). We may suppose, without loss of generality, that S is contained in the unit hypercube. Define a set valued mapping U' on the unit hypercube by setting

$$U'(x) = U(Q(x)).$$

It is easy to see that U' is closed, convex, and sequentially upper hemi-continuous; whence, since S is closed and U' maps into S, there exists $x \in S$ such that $x \in U'(x)$. Since $x = Q(x)$, x is also a fixed point of U. ■

Intuition may suggest that the functions $(f_k)_{k \geqslant 1}$ become closer and closer to U in some sense, and hence that Kakutani's proof contains a proof of the existence of approximate fixed points—a set valued mapping U on a metric space X has *approximate fixed points* if for each $\varepsilon > 0$ there exists $x \in X$ such that $\rho(x, U(x)) < \varepsilon$; such an x is called an *ε-fixed point of* U. However, in order to construct approximate fixed points, we must be able to quantify the 'convergence' of these affine approximations. Our eventual solution to finding a constructive Kakutani fixed point theorem is to restrict our mapping U in such a way as to ensure that for each $\varepsilon > 0$ there exists an affine function contained in an ε-expansion of the graph of U.

3.5.2 Continuity for set valued mappings

We are ready to begin our journey toward a constructive Kakutani fixed point theorem. The constructive treatment of Brouwer's fixed point theorem suggests that we take the following route:

(i) we should recast upper sequential hemi-continuity as a pointwise property;

(ii) we should further consider a uniform notion of upper sequential hemi-continuity;

(iii) we should focus on approximate fixed points.

It is also natural to insist that the image of each point be a located set; but, working constructively, this severely restricts the set valued mappings we can define. For example, in order to prove one direction of Proposition 83, we have in mind the function $U : [0,1] \to \mathcal{P}^*([0,1])$ given by

$$U(x) = \begin{cases} \{0\} & x < 1/2 \\ [0,1] & x = 1/2 \\ \{1\} & x > 1/2. \end{cases}$$

However, U is only defined on the subset $[0,1/2) \cup \{1/2\} \cup (1/2,1]$ of $[0,1]$, which equals $[0,1]$ only in the presence of **LPO**—in fact, this equality is equivalent to the law of trichotomy. We can overcome this problem by defining a set valued mapping from its graph: set G to be the subset of $[0,1]^2$ given by $[(0,0),(1/2,0)] \cup [(1/2,0),(1/2,1)] \cup [(1/2,1),1,1)]$—where $[x,y] = \{tx + (1-t)y : t \in [0,1]\}$—and let $U(x) = \{y \in [0,1] : (x,y) \in G\}$.

Note that in general $U(x)$ need not be located or even inhabited, but it is nonempty.

We say that a mapping $U : X \to \mathcal{P}^*(X)$ is *pointwise upper hemi-continuous* if for each $x \in X$ and each $\varepsilon > 0$ there exists $\delta > 0$ such that for all $y \in X$, if $\|x - y\| < \delta$, then

$$U(y) \subset (U(x))_\varepsilon ,$$

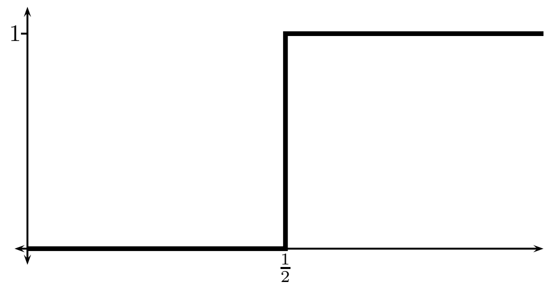


Figure 2: The graph G of $U : [0,1] \to \mathcal{P}^*([0,1])$.

where, for a subset S of a metric space X and a positive real number ε,

$$S_\varepsilon = \{x \in X : \exists_{s \in S} \rho(x,s) < \varepsilon\}.$$

If U is pointwise upper hemi-continuous, then U is sequentially upper hemi-continuous. Suppose that U is a pointwise upper hemi-continuous function and let $(x_n)_{n\geqslant 1}$, $(y_n)_{n\geqslant 1}$ be sequences in X converging to points x and y of X, respectively, such that $y_n \in U(x_n)$ for each n. Fix $\varepsilon > 0$ and let $\delta > 0$ be such that $U(z) \subset (U(x))_{\varepsilon/2}$ for all $z \in B(x,\delta)$. Pick $N > 0$ such that $\rho(x_n, x) < \delta$ and $\rho(y_n, y) < \varepsilon/2$ for all $n \geqslant N$. Then for each $n \geqslant N$ we have

$$y_n \in U(x_n) \subset (U(x))_{\varepsilon/2},$$

so

$$\begin{aligned} \rho(y, U(x)) &\leqslant \rho(y, y_n) + \rho(y_n, U(x)) \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, $\rho(y, U(x)) = 0$.

Lemma 82 (WCC+LPO) *If U is an upper hemi-continuous mapping on the unit hypercube $[0,1]^n$ and $\mathrm{G}(U)$ is separable, then $U(x)$ is located for* *each x.*

Proof. Let $(\mathbf{z}_n)_{n\in\mathbf{N}}$ be a dense sequence in the graph of U, and fix some $\mathbf{z} \in [0,1]^{2n}$ and some $r > 0$. Using the law of trichotomy, which follows from **LPO** plus **WCC**, we can construct a binary sequence $(\lambda_N)_{n\in\mathbf{N}}$ such that

$$\begin{aligned}\lambda_n = 0 \quad &\Rightarrow \quad \exists_{m\in\mathbf{N}}(\mathbf{z}_m \in \overline{B}(\mathbf{z},r) \text{ and } \rho(x,\pi_1(\mathbf{z}_m)) < \frac{1}{m}\\ \lambda_n = 1 \quad &\Rightarrow \quad \forall_{m\in\mathbf{N}}(\mathbf{z}_m \notin \overline{B}(\mathbf{z},r) \text{ or } \rho(x,\pi_1(\mathbf{z}_m)) \geqslant \frac{1}{m}.\end{aligned}$$

Applying **LPO** to $(\lambda_N)_{n\in\mathbf{N}}$ we have that either $\lambda_n = 0$ for all n in which case $\rho(\mathbf{z}, U(x)) \leqslant r$, or there exists n such that $\lambda_n = 1$ and hence $\rho(\mathbf{z}, U(x)) \geqslant r$. Since $\mathbf{z}, r$ are arbitrary, it follows that $U(x)$ is located. ∎

Proposition 83 (WCC) *The statement*

> *Every sequentially upper hemi-continuous mapping with a separable graph is pointwise upper hemi-continuous.*

is equivalent to ***LPO****.*

Proof. Let U be a sequentially upper hemi-continuous mapping on X, let $((x_n, y_n))_{n\geqslant 1}$ be a dense sequence in the graph of U, and fix $x \in X$ and $\varepsilon > 0$. Using **LPO**, construct a binary double sequence $(\lambda_{k,n})_{k,n\geqslant 1}$ such that

$$\lambda_{k,n} = 1 \quad \Leftrightarrow \quad \rho\left(x, x_n\right) < \frac{1}{k} \wedge \rho\left(y_n, U(x)\right) \geqslant \varepsilon.$$

By **LPO**, either for all k there exists an $n \in \mathbf{N}$ such that $\lambda_{n,k} = 1$, or else there exists $k \in \mathbf{N}$ such that $\lambda_{k,n} = 0$ for all $n \in \mathbf{N}$. If there exists $k \in \mathbf{N}$ such that $\lambda_{k,n} = 0$ for each $n \in \mathbf{N}$, then $\delta = 1/k$ satisfies the definition of pointwise upper hemi-continuity and we are done. Therefore it suffices to rule out the former case: if for each k there exists n such that $\lambda_{k,n} = 1$, then there exist sequences $(x_n)_{n\geqslant 1}, (y_n)_{n\geqslant 1}$ in X such that $(x_n)_{n\geqslant 1}$ converges to x, and for each n, $y_n \in U(x_n)$ and $\rho(y_n, U(x)) \geqslant \varepsilon$—a contradiction.

To show the converse consider the function $U : [0,1] \to \mathcal{P}^*([0,1])$ pictured in Figure 2. It is straightforward to show that $G(U)$ is closed and hence that U is sequentially upper hemi-continuous. Suppose that U is pointwise upper hemi-continuous and let a be a number close to 0. Let $\delta > 0$ be such that if $y \in B(|a| + 1/2, \delta)$, then

$$U(y) \subset (U(|a| + 1/2))_{1/2} .$$

Either $a \neq 0$ or $|a| < \delta$. In the latter case, if $a \neq 0$, then

$$[0, 1] = U(1/2) \subset U(|a| + 1/2)_{1/2} = [1/2, 1],$$

which is absurd. Hence a is in fact equal to 0. Thus the sequential upper hemi-continuity of f implies

$$\forall_{a \in \mathbf{R}} \, (a = 0 \vee a \neq 0) \, ,$$

which in turn implies **LPO**. ■

The natural notion of uniform pointwise upper hemi-continuity seems to be too strong to be of much interest. In particular, if U, in addition to satisfying the uniform notion of pointwise upper hemi-continuity, is located, then U is uniformly continuous , however, is equivalent to the uniform continuity of U as a function from X to $\mathcal{P}^*(X)$ endowed with the standard Hausdorff metric.

The uniform version of pointwise upper hemi-continuity is not classically equivalent to the non-uniform version because, for a fixed ε, the $\delta(x)$ satisfying sequential upper hemi-continuity at x need not vary continuously with x and may fail to be bounded below by a positive valued continuous function. Another result of this is that few functions are constructively pointwise upper hemi-continuous; for instance the (benign) mapping given in the proof of Proposition 83.

The uniform continuity of U is a significantly stronger property than that of sequential upper hemi-continuity, and easily leads to an approximate fixed point theorem of relatively little interest.

In order to find a more satisfactory constructive version of Kakutani's fixed point theorem (preferably classically equivalent to the classical one), we need to find a notion similar to pointwise sequential upper hemi-continuity, and with more computational content than sequential upper hemi-continuity, for which the uniform version is classically equivalent to the non-uniform version. To that end, we say that a mapping $U : X \to \mathcal{P}^*(X)$ is *locally approximable* if for each $x \in X$ and each $\varepsilon > 0$, there exists $\delta > 0$ such that if $y, y' \in B(x, \delta)$, $u \in U(y)$, $u' \in U(y')$, and $t \in [0, 1]$, then

$$\rho\left(\left(z_t, u_t\right), G(U)\right) < \varepsilon,$$

where $z_t = ty + (1-t)y'$ and $u_t = tu + (1-t)u'$; note that we do not require $G(U)$ to be located here: we use '$\rho(x, S) < \varepsilon$' as a shorthand for 'there exists $s \in S$ such that $\rho(x, s) < \varepsilon$'.

Proposition 84 *Every convex, pointwise upper hemi-continuous set valued mapping on a linear metric space is locally approximable.*

Proof. Let X be a linear metric space and let U be a convex, pointwise upper hemi-continuous set valued mapping on X. Fix $x \in X$ and $\varepsilon > 0$, and let $\delta \in (0, \varepsilon/2)$ be such that $U(y) \subset (U(x))_{\varepsilon/2}$ for all $y \in B(x, \delta)$. Let $y, y' \in B(x, \delta)$, $u \in U(x)$, $u' \in U(y)$, and $t \in [0, 1]$; since $(U(x))_{\varepsilon/2}$ is convex, $u_t \in (U(x))_{\varepsilon/2}$. Then

$$\rho\left(z_t, x\right) \leqslant \max\{\rho(x, y), \rho(x, y')\} < \delta < \varepsilon/2,$$

so

$$\begin{aligned}
\rho\left(\left(z_t, u_t\right), G(U)\right) &\leqslant \rho\left(\left(z_t, u_t\right), \{x\} \times U(x)\right) \\
&< \rho\left(\left(z_t, u_t\right), \left(x, u_t\right)\right) + \varepsilon/2 \\
&= \rho\left(z_t, x\right) + \varepsilon/2 \\
&< \varepsilon/2 + \varepsilon/2 = \varepsilon.
\end{aligned}$$

Hence U is locally approximable. ■

Above we have negotiated from the sequential upper hemi-continuity of a set valued mapping to its being locally approximable via the property of pointwise upper hemi-continuity. By Proposition 83 this requires **LPO**. It seems likely that we can prove this (classical) equivalence directly under the assumption of weaker nonconstructive principles; our next result is a first attempt at this.

In [62], Ishihara showed that **BD-N** holds in the classical, intuitionistic, and recursive models of **BISH**, and that the statement 'every sequentially continuous mapping from a separable metric space is pointwise continuous' is equivalent, over **BISH**, to **BD-N**. The proof of the following uses ideas from [62].

Proposition 85 ($\mathbf{AC}_{\omega,2}$) *Suppose **WKL** and **BD**-N hold. Let U be a convex, sequentially upper hemi-continuous set valued mapping on a separable metric space. Then U is locally approximable.*

Proof. Let $(x_n)_{n\geqslant 1}$ be a dense sequence in S, and fix $x \in S$ and $\varepsilon > 0$. Define

$$\begin{aligned} A &= \{0\} \cup \big\{k > 0 : \exists_{m,n} \big(x_m, x_n \in B(x, k^{-1}) \wedge \\ &\qquad \exists_{t\in[0,1]} \exists_{u\in U(x_m)} \exists_{u'\in U(x_n)} (\rho((z_t, u_t), G(U)) > \varepsilon/2)\big)\big\}, \end{aligned}$$

where $z_t = tx_m + (1-t)x_n$ and $u_t = tu + (1-t)u'$. We show that the set A is pseudobounded. It then follows from **BD-N** that there exists $M > 0$ such that $a < M$ for all $a \in A$. The definition of A then ensures that $\delta = 1/M$ satisfies the definition of local approximability.

Let $(a_n)_{n\geqslant 1}$ be a nondecreasing sequence in A. Using $\mathbf{AC}_{\omega,2}$, we construct a binary sequence $(\lambda_n)_{i\geqslant 1}$ such that

$$\begin{aligned} \lambda_n = 0 &\;\Rightarrow\; a_n/i < 1; \\ \lambda_n = 1 &\;\Rightarrow\; a_n/i > 1/2. \end{aligned}$$

Let $u \in U(x)$, and construct sequences $(x_n)_{n\geqslant 1}, (x'_n)_{n\geqslant 1}, (y_n)_{n\geqslant 1}, (y'_n)_{n\geqslant 1}$ in S and a sequence $(t_n)_{n\geqslant 1}$ in $[0,1]$ as follows. If $\lambda_n = 0$, set $x_n = x'_n = x$, $y_n = y'_n = u$, and $t_i = 0$. If $\lambda_n = 1$, then pick $x_k, x_l \in B(x, i^{-1}), u \in U(x_k), u' \in U(x_l)$, and $t \in [0,1]$ such that $\rho((z_t, u_t), G(U)) > \varepsilon$, and set $x_n = x_k$, $x'_n = x_l$, $y_n = u$, $y'_n = u'$, and $t_n = t$. Then $(x_n)_{n\geqslant 1}$ and $(x'_n)_{n\geqslant 1}$ converge to x, and, since $(a_n)_{n\geqslant 1}$ is nondecreasing, we may assume by **WKL** that there exist $y, y' \in S$ and $t \in [0,1]$ such that $(y_n)_{n\geqslant 1}$ converges to y, $(y'_n)_{n\geqslant 1}$ converges to y', and $(t_n)_{n\geqslant 1}$ converges to t. Then, by the sequential upper hemi-continuity of U, $\rho(y, U(x)) = \rho(y', U(x)) = 0$; whence, since $U(x)$ is convex, $\rho(ty + (1-t)y', U(x)) = 0$. Let $N > 0$ be such that

$$\max\left\{\rho((x_n, y_n), (x, y)), \rho((x'_n, y'_n), (x, y))\right\} < \varepsilon$$

for all $n \geqslant N$, and suppose that there exists $n \geqslant N$ with $\lambda_n = 1$. Then

$$\rho((t_n x_n + (1-t_n)x'_n, t_n y_n + (1-t_n)y'_n), G(U)) > \varepsilon,$$

but

$$\rho((t_n x_n + (1-t_n)x'_n, t_n y_n + (1-t_n)y'_n), U(x)) \;<\; \varepsilon + \rho((x,z), U(x)) \;=\; \varepsilon.$$

This contradiction ensures that $\lambda_n = 0$ for all $n \geqslant N$; thus A is pseudobounded. ∎

It is unknown whether **LLPO** implies **BD-N** over **IZF** plus **DC**.

We say that $U : X \to \mathcal{P}^*(X)$ is *approximable* if it satisfies the uniform version of local approximability: for each $\varepsilon > 0$, there exists $\delta > 0$ such that if $x, x' \in X$, $\|x - x'\| < \delta$, $u \in U(x)$, $u' \in U(x')$, and $t \in [0,1]$, then

$$\rho\left((z_t, u_t), G(U)\right) < \varepsilon,$$

where $z_t = tx + (1-t)x'$ and $u_t = tu + (1-t)u'$. Given a function $f : X \to X$, define $U_f : X \to \mathcal{P}^*(X)$ by $U_f(x) = \{f(x)\}$. If f is continuous then U_f is locally approximable, and if f is uniformly continuous then U_f

is approximable.

To each located, locally approximable (resp. approximable) function we can associate a strongly pointwise (resp. uniformly) continuous predicate P such that $P(x, x', \delta)$ if and only if for all $u \in U(x)$, $u' \in U(x')$, and $t \in [0,1]$,

$$\rho\left((z_t, u_t), G(U)\right) < \varepsilon.$$

This gives the following result.

Proposition 86 ($\mathbf{AC}_{\omega,2}$) *If the fan theorem for Π^0_1 bars holds, then every locally approximable, located mapping on $[0,1]^n$ is approximable.*

Proof. Follows directly from the above characterisation of (local) approximable by continuous predicates and Theorem 28. ■

Classically every closed, convex, sequentially upper hemi-continuous mapping on a convex subset of $\mathbf{R}^n$ is approximable. It seems unlikely that every upper hemi-continuous mapping with a located graph is constructively approximable, but the one dimensional case is straightforward.

Proposition 87 (CC) *If $U : [0,1] \to \mathcal{P}^*[0,1]$ is a closed, convex, sequentially upper hemi-continuous mapping and $\mathrm{G}(U)$ is located, then U is approximable.*

Lemma 88 (CC) *Let $U : [0,1] \to \mathcal{P}^*[0,1]$ be a closed, convex, sequentially upper hemi-continuous mapping with a located graph. Then $U([x,y])$ is convex, for all $x, y \in [0,1]$ with $x < y$.*

Proof. Let $S = U([x,y])$ and suppose that there exists that there exists $r > 0$ and $t \in \mathrm{conhull}(S)$ such that $|s - t| > r$ for all $s \in S$; since $\mathrm{G}(U)$ is located we may assume, without loss of generality, that there exist $z_x \in U(x), z_y \in U(y)$ such that $z_x < t < z_y$. We inductively construct sequences $(x_n)_{n\in\mathbf{N}}, (z_{x_n})_{n\in\mathbf{N}}, (y_n)_{n\in\mathbf{N}}, (z_{y_n})_{n\in\mathbf{N}}$ in $[0,1]$ such that

1. $(x_n, z_{x_n}), (y_n, z_{y_n}) \in \mathrm{G}(U)$ for each n;

2. $x_n \leqslant x_{n+1} < y_{n+1} \leqslant x_n$;

3. $|x_n - y_n| < (1/2)^n |x - y|$;

4. $z_{x_n} < t - r$ and $z_{y_n} > t + r$.

We begin the construction by setting $x_0 = x, y_0 = y, z_{x_0} = z_x$, and $z_{y_0} = z_y$. Having constructed the first n terms of each sequence, we take a $(1/2)^{n+2}$ approximation to $\mathrm{G}(U)$ which contains both (x_n, z_{x_n}) and (y_n, z_{y_n}). By the one-dimensional case of Sperner's lemma, which can be proved by a simple contradiction argument, we can find points $\left(x_{n+1}, z_{x_{n+1}}\right), \left(y_{n+1}, z_{y_{n+1}}\right)$ in this approximation which satisfy conditions 1 to 4. This completes the induction. It follows that the shared limit of $(x_n)_{n\in\mathbf{N}}, (y_n)_{n\in\mathbf{N}}$ is not convex—a contradiction. Hence $t \in \overline{S}$.

Now if $t \in$ conhull$S$, then, since $t \in \overline{S}$, we can use an argument almost identical to the first part of the proof to produce sequences of points $(x_n)_{n\in\mathbf{N}}$, $(z_{x_n})_{n\in\mathbf{N}}$, $(y_n)_{n\in\mathbf{N}}$, $(z_{y_n})_{n\in\mathbf{N}}$ in $[0, 1]$ satisfying 1 to 3 and such that $z_{x_n} < t + 1/n$ and $z_{y_n} > t - 1/n$. Then $t \in U(x_\infty)$ where x_∞ is the shared limit of $(x_n)_{n\in\mathbf{N}}, (y_n)_{n\in\mathbf{N}}$. ■

Here then is the **proof of Proposition 87**:

Proof. For $\varepsilon > 0$, set $\delta = \varepsilon$. Let $(x, z_x), (y, z_y) \in G(U)$ with $x < y$, and let $t \in (0, 1)$; set $r = \varepsilon - \min\{t, 1 - t\}|x - y|$. If $B(\mathbf{z}_t, r) \cap G(U)$ is empty, then $U([x, y])$ is not convex. This contradiction, and the locatedness of the graph of U, ensures that $B(\mathbf{z}_t, \varepsilon) \cap G(U)$ is inhabited. ■

3.5.3 Constructing approximate fixed points

A mapping U is approximable if for each positive ε there exists a positive δ such that the convex hull of any two points in the graph of U which are separated by less than δ never strays more than ε from the graph of U; our next lemma shows that if U is approximable, then we can generalise this from any two points of $G(U)$ to any finite subset of $G(U)$. This will allow us to give a constructive version of Kakutani's fixed point theorem which is classically equivalent to the classical version.

Lemma 89 *Let $U : X \to \mathcal{P}^*(X)$ be an approximable function. Then for each $n > 0$ and each $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x_1, \ldots, x_n$, $u_1, \ldots, u_n \in X$ and all $\mathbf{t} \in [0,1]^n$, if $u_i \in U(x_i)$ for each i, $\sum_{i=1}^n t_i = 1$, and*

$$\max\{\|x_i - x_j\| : 1 \leqslant i, j \leqslant n\} < \delta,$$

then

$$\rho((z_{\mathbf{t}}, u_{\mathbf{t}}), G(U)) < \varepsilon,$$

where $z_{\mathbf{t}} = \sum_{i=1}^n t_i x_i$ and $u_{\mathbf{t}} = \sum_{i=1}^n t_i u_i$.

Proof. We proceed by induction; the case $n = 1$ is trivial. Suppose that we have shown the result for $n = k-1$. Let $\mathbf{t} \in [0,1]^k$ and $u_1, \ldots, u_k$ be as in the statement of the lemma, and let $\delta > 0$ be such that for all $x_1, \ldots, x_{k-1} \in X^n$ and all $\mathbf{t} \in [0,1]^{k-1}$, if $u_i \in U(x_i)$ for each i, $\sum_{i=1}^{k-1} t_i = 1$, and

$$\max\{\|x_i - x_j\| : 1 \leqslant i, j \leqslant k-1\} < \delta,$$

then $\rho((z_{\mathbf{t}}, u_{\mathbf{t}}), G(U)) < \varepsilon/2$. Let $\mathbf{t}'$ be the $k-1$ dimensional vector with i-th component $t_i / \sum_{j=1}^{k-1} t_j$. Then

$$\rho\left((z_{\mathbf{t}'}, u_{\mathbf{t}'}), G(U)\right) < \varepsilon/2.$$

Picking $(x, u) \in G(U)$ with $\rho\left((z_{\mathbf{t}'}, u_{\mathbf{t}'}), (x, u)\right) < \varepsilon/2$ and $t \in [0,1]$ such that $\rho((z_{\mathbf{t}}, u_{\mathbf{t}}), (tx + (1-t)x_n, tu + (1-t)u_n)) < \varepsilon/2$, we have that

$$\begin{aligned}
\rho\left((z_{\mathbf{t}}, u_{\mathbf{t}}), G(U)\right) &< \rho((z_{\mathbf{t}}, u_{\mathbf{t}}), (tx + (1-t)x_n, tu + (1-t)u_n)) + \varepsilon/2 \\
&< \varepsilon/2 + \varepsilon/2 = \varepsilon.
\end{aligned}$$

This completes the induction. ∎

Theorem 90 *Let S be a totally bounded subset of $\mathbf{R}^n$ with convex closure and let U be an approximable set valued mapping on S. Then for each $\varepsilon > 0$ there exists $x \in S$ such that $\rho(x, U(x)) < \varepsilon$.*

Proof. Fix $\varepsilon > 0$ and let $\delta > 0$ be such that for all $x_1, \ldots, x_k \in X^n$ and all $\mathbf{t} \in [0,1]^n$, if $u_i \in U(x_i)$ for each i, $\sum_{i=1}^n t_i = 1$, and $\max\{\|x_i - x_j\| : 1 \leqslant i, j \leqslant k\} < \delta$, then

$$\rho((z_{\mathbf{t}}, u_{\mathbf{t}}), G(U)) < \varepsilon/3.$$

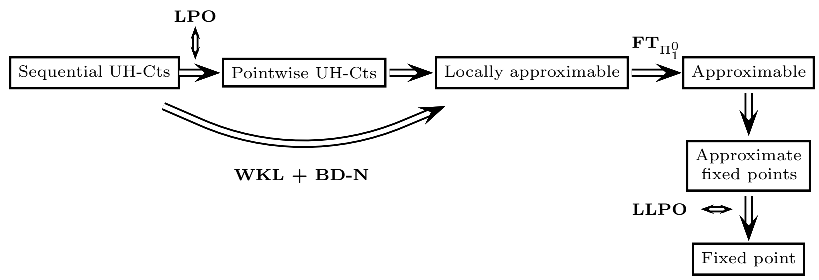
Let $S' = \{x_1, \ldots, x_l\}$ be a discrete δ-approximation to S. For each $x_i \in S'$, pick $u_i \in U(x_i)$; let g be the uniformly continuous affine function on S that takes the value u_i at x_i for each $1 \leqslant i \leqslant l$. By the approximate Brouwer fixed point theorem, there exists $y \in S$ such that $\rho(y, g(y)) < \varepsilon/3$. By our choice of δ, there exists $(x, u) \in G(U)$ such that $\rho((y, g(y)), (x, u)) < \varepsilon/3$. Therefore

$$\begin{aligned} \rho(x, u) &\leqslant \rho(x, y) + \rho(y, g(y)) + \rho(g(y), u) \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon, \end{aligned}$$

so $\rho(x, U(x)) < \varepsilon$. ∎

To see that Theorem 90 is classically equivalent to the classical theorem let $U : S \to \mathcal{P}^*(S)$ be as in the classical Kakutani fixed point theorem; then, as previously shown, U is approximable under classical logic. Using the above theorem (and countable choice), construct a sequence $(x_n)_{n \geqslant 1}$ in X such that $\rho(x_n, U(x_n)) < 1/n$ for each n. Since X is compact, we may assume, by **WKL**, that $(x_n)_{n \geqslant 1}$ converges to some $x \in X$. For each n, pick $y_n \in U(x_n)$ such that $\rho(x_n, y_n) < 1/n$. Then $y_n \to x$ and so, since U is closed and sequential upper hemi-continuous, $x \in U(x)$.

The following diagram summarises this classical proof; we suppress the use of $\mathbf{AC}_{\omega,2}$.



In particular, we feel that this gives a conceptually more straightforward classical proof of the Kakutani fixed point theorem than the standard classical proofs: that sequential upper hemi-continuity classically implies approximability is quite intuitive, and that approximable mappings have approximate fixed points is very natural in light of Brouwer's fixed point theorem; it then just remains to apply **MIN**, which roughly says that any 'continuous problem' on a compact space which has approximate solutions has an exact solution.

Our constructive version of Kakutani's fixed point theorem is clearly motivated by trying to reduce to Brouwer's fixed point theorem. In turn, the approximate Brouwer's fixed point theorem is normally proved using finite combinatorics. An alternative approach to giving an approximate Kakutani fixed point theorem would be to apply finite combinatorics to data coming directly from our set valued mapping.

Classically any ε approximation to the graph of a closed, convex, sequentially upper hemi-continuous mapping must contain an ε-fixed point. Can we prove an approximate version of Kakutani's fixed point theorem for convex, sequentially upper hemi-continuous mapping U? perhaps by taking an ε-approximation to the graph of U and applying combinatorics in a similar manner to the proof of Brouwer's fixed point theorem. The one dimensional case of this version of Kakutani's fixed point theorem follows directly from Proposition 87 and Theorem 90.

Proposition 91 *Let U be convex, sequentially upper hemi-continuous set valued mapping on $[0,1]$. Then for each $\varepsilon > 0$, there exists $x \in [0,1]$ such that $\rho(x, U(x)) < \varepsilon$.*

3.5.4 An extension

Theorem 90 gives a very simple and intuitive constructive version of Kakutani's fixed point theorem. An examination of the proof shows that we in fact only require our set valued mapping U to satisfy the following condition, weaker than approximability and often much easier to verify. A set valued mapping U on a metric space X is said to be *weakly approximable* if for each $\varepsilon > 0$, there exist

- a positive real number $\delta < \epsilon$,
- a $\delta/2$-approximation S of X, and
- a function V from S into $\mathcal{P}^*X$ with $G(V) \subset G(U)$,

such that if $x, x' \in S$, $\|x - x'\| < \delta$, $u \in V(x)$, $u' \in V(x')$, and $t \in [0,1]$, then

$$\rho\left((z_t, u_t), G(U)\right) < \varepsilon.$$

If V can be chosen independent of ε, in which case S is a dense subset of X, then U is said to be *weakly approximable with respect to V*.

The proofs of Lemma 89 and Theorem 90 readily extend to give the following result.

Theorem 92 *Let S be a totally bounded subset of $\mathbf{R}^n$ with convex closure and let U be a weakly approximable set valued mapping on S. Then for each $\varepsilon > 0$ there exists $x \in S$ such that $\rho(x, U(x)) < \varepsilon$.*

Chapter 4

Constructing economic equilibria

4.1 The minimax theorem

In [68], Kakutani presented his fixed point theorem and used it to give a simple proof of von Neumann's minimax theorem, which guarantees the existence of saddle points for particular functions:

Theorem 93 *Let* $f : [0,1]^n \times [0,1]^m \to \mathbf{R}$ *be a continuous function such that for each* $x_0, y_0 \in [0,1]$ *and each real number* r *the sets*

$$\begin{aligned} &\{y \in L : f(x_0, y) \leqslant r\} \quad \textit{and} \\ &\{x \in L : f(x, y_0) \geqslant r\} \end{aligned}$$

are convex. Then

$$\sup_{x\in[0,1]} \inf_{y\in[0,1]} f(x,y) = \inf_{y\in[0,1]} \sup_{x\in[0,1]} f(x,y).$$

Classically the suprema and infima are attained, and so can be replaced by minimum and maximum—hence the name 'minimax' theorem. We will give a constructive proof of this 'infisup' theorem using Theorem 92.

Throughout this section we fix a uniformly continuous function $f : [0,1]^n \times [0,1]^m \to \mathbf{R}$ satisfying the conditions of Theorem 93, and for each $\varepsilon > 0$ we set

$$
\begin{aligned}
V_\varepsilon &= \left\{ (x_0, y_0) \in [0,1]^n \times [0,1]^m : f(x_0, y_0) \leqslant \inf_{y \in [0,1]} f(x_0, y) + \varepsilon \right\}; \\
W_\varepsilon &= \left\{ (x_0, y_0) \in [0,1]^n \times [0,1]^m : f(x_0, y_0) \geqslant \sup_{x \in [0,1]} f(x, y_0) - \varepsilon \right\}.
\end{aligned}
$$

In order to prove the minimax theorem, we extend, in the obvious way, the definition of approximable, weakly approximable with respect to, and weakly approximable to functions which take points from a metric space X to subsets of a second metric space Y; we call such a function a *set valued mapping from X into Y*. We associate V_ε and W_ε with the set valued mappings given by

$$
V_\varepsilon(x) = \{y \in [0,1]^n : (x, y) \in V\} \text{ and } W_\varepsilon(y) = \{x \in [0,1]^n : (y, x) \in W\};
$$

note that $V_\varepsilon, W_\varepsilon$ are convex valued. Let U_i be a set valued mapping from X_i into Y_i $(i = 1, 2)$. The *product* of U_1 and U_2, written $U_1 \times U_2$, is the set valued mapping from $X_1 \times X_2$ to $Y_1 \times Y_2$ given by

$$
U_1 \times U_2(x_1, x_2) = U_1(x_1) \times U_2(x_2).
$$

We omit the straightforward proof of the next lemma.

Lemma 94 *Let U_i be a set valued mapping from X_i into Y_i ($i = 1, 2$). If U_1, U_2 are (weakly) approximable, then $U_1 \times U_2$ is (weakly) approximable, and if U_1, U_2 are weakly approximable with respect to V_1, V_2 respectively, then $U_1 \times U_2$ is weakly approximable with respect to $V_1 \times V_2$.*

Lemma 95 *For each $\varepsilon > 0$, V_ε is weakly approximable with respect to $V_{\varepsilon/2}$ and W_ε is weakly approximable with respect to $W_{\varepsilon/2}$.*

Proof. We only give the proof for V_ε; the proof for W_ε is entirely analogous. Since f is uniformly continuous, there exists $\delta > 0$ such that $(V_{\varepsilon/2})_\delta$ is contained in V_ε. Let x, x' be points of $[0,1]^n$ such that $\|x - x'\| < \delta$ and fix y, y' such that $(x, y), (x', y') \in V_{\varepsilon/2}$. Then $(\overline{x}, y), (\overline{x}, y) \in V_\varepsilon$, where

$\overline{x} = (x + x')/2$. Since $V_\varepsilon(\overline{x})$ is convex valued, $ty + (1-t)y' \in V_\varepsilon$ for each $t \in [0,1]$; whence

$$\begin{aligned} \rho((\mathbf{z}_t, \mathbf{u}_t), G(u)) &\leqslant \rho((\mathbf{z}_t, \mathbf{u}_t), \{\overline{x}\} \times V_\varepsilon(\overline{x})) \\ &= \rho(tx + (1-t)x', \overline{x}) \ < \ \delta. \end{aligned}$$

Since δ can be chosen to be arbitrarily small, this completes the proof. ■

We now have the **proof of Theorem 93**:

Proof. Let $f : [0,1] \times [0,1] \to \mathbf{R}$ be as in the statement of the theorem. It is easy to see that

$$\sup_{x\in[0,1]} \inf_{y\in[0,1]} f(x,y) \leqslant \inf_{y\in[0,1]} \sup_{x\in[0,1]} f(x,y).$$

Fix $\varepsilon > 0$ and let $\delta > 0$ be such that $\|f(x,y) - f(x',y')\| < \varepsilon/4$ whenever $\|(x,y) - (x',y')\| < \delta$. By Lemmas 94 and 95 the set valued mapping U on $[0,1]^{n+m}$ given by

$$U = W_{\varepsilon/2} \times V_{\varepsilon/2}$$

is approximable with respect to $W_{\varepsilon/4} \times V_{\varepsilon/4}$. By Theorem 92, there exists $(x_0, y_0) \in [0,1]^{n+m}$ such that $\rho((x_0,y_0), U(x_0,y_0)) < \delta$. It follows from the definition of U and our choice of δ, that

$$\begin{aligned} f(x_0,y_0) &< \inf_{y\in[0,1]} f(x_0,y) + \varepsilon, \text{ and} \\ f(x_0,y_0) &> \sup_{x\in[0,1]} f(x,y_0) - \varepsilon. \end{aligned}$$

Hence

$$\begin{aligned} \inf_{y\in[0,1]} \sup_{x\in[0,1]} f(x,y) &\leqslant \sup_{x\in[0,1]} f(x_0,y_0) \\ &< f(x_0,y_0) + \varepsilon \\ &< \inf_{y\in[0,1]} f(x_0,y) + 2\varepsilon \\ &\leqslant \sup_{x\in[0,1]} \inf_{y\in[0,1]} f(x,y) + 2\varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, this completes the proof. $\blacksquare$

4.2 Walras's existence theorem

In [105] Yasuhito Tanaka shows that Walras's existence theorem implies **LLPO**—he essentially gives a translation of the standard Brouwerian counterexample for Brouwer's fixed point theorem—it follows that Walras's existence theorem does not admit a constructive proof. We give a constructive proof of an approximate version of Walras's existence theorem from which the full theorem can be recovered with an application of **MIN**. We then push Uzawa's equivalence theorem to the level of approximate solutions, before considering economies with at most one equilibrium.

Consider an economy E with $n+1$ commodities, a set of price vectors P, and excess demand functions $f_0, \ldots, f_n$ from P into $\mathbf{R}$; we associate E with the set $\{f_0, \ldots, f_n\}$. For $p \in P$, we interpret $f_i(p)$ as the excess demand for commodity i when the price p prevails. If $f_i(p)$ is negative, then there is a surplus of commodity i, and if $f_i(p)$ is positive, there is a deficit of commodity i. An *equilibrium of* E is a price vector $p \in P$ such that

$$f_i(p) \leqslant 0 \qquad\qquad (0 \leqslant i \leqslant n).$$

Given an economy E, we are interested in finding an equilibrium of E. A fundamental result of equilibrium theory is Walras's existence theorem, which gives a sufficient condition for the existence of an equilibrium. For simplicity, we set P to be the regular n-simplex

$$\Delta^n = \left\{(p_0, \ldots, p_n) : \sum p_k = 1\right\}$$

with the metric ρ inherited from $\mathbf{R}^n$. Walras's existence theorem can be stated as follows.

> **Theorem** Let P be the regular n-simplex and let $f_0, \ldots, f_n$ be continuous real valued functions on P which satisfy *Walras's Law*: for all $p \in P$
>
> $$p_0 f_0(p) + \cdots + p_n f_n(p) = 0.$$

Then there exists $p \in P$ such that

1. $f_i(p) \leqslant 0$ for each $0 \leqslant i \leqslant n$;
2. if $p_i > 0$, then $f_i(p) = 0$.

The standard proof of Walras's existence theorem involves an application of Brouwer's fixed point theorem to a cleverly defined function φ from P into itself. The fixed point of φ is then shown to be an equilibrium of the economy. Since Brouwer's fixed point theorem implies **LLPO**, the standard proof of Walras's existence theorem is nonconstructive. It is then natural to ask: Is the Walras existence theorem nonconstructive? The answer, suggested by Uzawa's equivalence theorem—which states that Brouwer's fixed point theorem and Walras's equivalence theorem are classically equivalent—and proven by Yasuhito Tanaka in [105], is that Walras's existence theorem also implies **LLPO** and therefore does not admit a constructive proof.

The body of this section is broken up into two parts. In the first we prove, under the assumptions of Walras's existence theorem, the existence of *approximate equilibriums*: an economy E is said to have approximate equilibriums if for each $\varepsilon > 0$, there exists $p \in P$ such that $f_i(p) \leqslant \varepsilon$ for each i. The second section gives a version of Uzawa's equivalence theorem at the level of computational content, and then considers economies with at most one equilibrium.

4.2.1 A constructive Walras existence theorem

Our constructive version of the Walras existence theorem is

Proposition 96 *Let P be the regular n-simplex and let $f_0, \ldots, f_n$ be uniformly continuous real valued functions on P which satisfy Walras's Law. Then for each $R, \varepsilon > 0$, there exists $p \in P$ such that*

1. *$f_i(p) \leqslant \varepsilon$ for each $0 \leqslant i \leqslant n$;*

2. *if* $p_i > R$*, then* $|f_i(p)| < \varepsilon$.

We call this the approximate Walras existence axiom.

For the remainder of this section, let P be the regular n-simplex, let $f_0, \ldots, f_n$ be uniformly continuous real valued functions on P which satisfy Walras's Law, and fix $\varepsilon > 0$. Since $P \times \{0, \ldots, n\}$ is compact and each f_i is uniformly continuous,

$$M = \sup\{f_i(p) : 0 \leqslant i \leqslant n, p \in P\} + 1/4$$

exists. For each i let

$$v_i(p) = p_i + \max\{0, f_i(p) - \varepsilon/2n\};$$

then $v_i(p) > p_i$ if and only if $f_i(p) > \varepsilon/2n$. Define uniformly continuous real valued functions $\varphi_0, \ldots, \varphi_n$ on P by

$$\varphi_i = \lambda^{-1} v_i(p),$$

where[50]

$$\lambda = \sum v_k(p).$$

Then $\varphi(p) = (\varphi_0, \ldots, \varphi_n)$ is a uniformly continuous function from P into itself: for each p, $\varphi_i(p) = \lambda^{-1} v_i(p) \geqslant 0$ and

$$\sum \varphi_k(p) = \sum \lambda^{-1} v_k(p) = \lambda^{-1} \sum v_k(p) = \lambda^{-1}\lambda = 1.$$

Before proving Proposition 96 we require a number of lemmas.

Lemma 97 *If there exists i such that $f_i(p) > R > \varepsilon/2n$, then $\lambda > 1 + (R - \varepsilon/2n)$.*

[50] Throughout this section summands are over $\{0, \ldots, n\}$ unless otherwise stated, and have index k.

Proof. Suppose that $f_i(p) > R > \varepsilon/2n$. Then

$$\begin{array}{rcl} \lambda & = & \sum v_k(p) \\ & = & \sum p_k + \sum \max\{0, f_k(p) - \varepsilon/2n\} \\ & \geqslant & 1 + (f_i(p) - \varepsilon/2n) \quad > \quad 1 + (R - \varepsilon/2n). \end{array}$$

■

The next lemma and its successor are intimately related to the first and second part, respectively, of the conclusion of Proposition 96.

Lemma 98 *Let $p \in P$. If $\varepsilon \in (0,1)$ and*

$$\rho(p, \varphi(p)) < r = \frac{\varepsilon^2}{2(\varepsilon + 2)(2Mn + \varepsilon)},$$

then $f_i(p) \leqslant \varepsilon$ for each i.

Proof. We have that $\left|p_i - \lambda^{-1} v_i(p)\right| < r$ for each i, so

$$v_i(p) < \lambda(p_i + r).$$

Suppose that there exists $k \in \{0, \ldots, n\}$ such that $f_k(p) > \varepsilon$. Then $\lambda > 1 + \varepsilon/2$. Write $\{0, \ldots, n\}$ as the disjoint union of two sets P and Q such that

$$\begin{array}{rcl} i \in P & \Rightarrow & p_i > \left(\dfrac{\varepsilon + 2}{\varepsilon}\right) r; \\ i \in Q & \Rightarrow & p_i < 2\left(\dfrac{\varepsilon + 2}{\varepsilon}\right) r. \end{array}$$

Since $\varepsilon < 1$ and $M \geqslant 1/4$, we have that $\varepsilon^{-1}(\varepsilon + 2) r < 1/(n+1)$, so P is inhabited. If $i \in P$, then

$$v_i(p) > \lambda(p_i - r) > (1 + \varepsilon/2)(p_i - r) > p_i,$$

by our choice of r. Hence

$$\begin{aligned}
\sum p_i f_i(p) &= \sum_{i\in P} p_i f_i(p) + \sum_{i\in Q} p_i f_i(p) \\
&> \left(\sum_{i\in P} p_i\right) \varepsilon/2n - 2\left(\frac{\varepsilon+2}{\varepsilon}\right) rnM \\
&> \left(1 - 2n\left(\frac{\varepsilon+2}{\varepsilon}\right) r\right) \varepsilon/2n - 2\left(\frac{\varepsilon+2}{\varepsilon}\right) rnM \quad = \quad 0,
\end{aligned}$$

by our choice of r. This contradiction of Walras's law ensures that $f_i(p) \leqslant \varepsilon$ for each i. ■

Lemma 99 *Let $p \in P$. If*

$$\rho(p, \varphi(p)) < r = \frac{\varepsilon^2}{M(\varepsilon + 2n)},$$

then $|p_i f_i(p)| \leqslant \varepsilon$ for each i.

Proof. Suppose that $p_i f_i(p) > \varepsilon/n$ for some i. Then $f_i(p) > \varepsilon/n$, so $\lambda > 1 + \varepsilon/2n$. If $p_k > \varepsilon^{-1}(\varepsilon + 2n)r$, then

$$v_i(p) > \lambda\,(p_i - r) \;>\; (1 + \varepsilon/2n)(p_k - r) \;>\; p_k,$$

by our choice of r; whence $f_k(p) > 0$. Then

$$\begin{aligned}
\sum p_k f_k(p) &= p_i f_i(p) + \sum_{k\neq i} p_k f_k(p) \\
&\geqslant p_i f_i(p) - Mn\left(\frac{\varepsilon+2n}{\varepsilon}\right) r \\
&> \varepsilon/n - Mn\left(\frac{\varepsilon+n}{\varepsilon}\right) r \quad = \quad 0,
\end{aligned}$$

again by our choice of r. It follows from this contradiction that $p_i f_i(p) \leqslant \varepsilon/n$ for each i.

Now suppose that $p_i f_i(p) < -\varepsilon$. Then

$$\sum p_i f_i(p) = p_i f_i(p) + \sum_{k \neq i} p_k f_k(p) < -\varepsilon + n(\varepsilon/n) = 0,$$

which is absurd. Hence $p_i f_i(p) \geqslant -\varepsilon$ for each i, so $|p_i f_i(p)| \leqslant \varepsilon$. ■

We now have the **proof of Proposition 96:**

Proof. Fix $R, \varepsilon \in (0,1)$ and let

$$r = \min \left\{ \frac{\varepsilon^2 R^2}{M(\varepsilon R + 2n)}, \frac{\varepsilon^2}{2(\varepsilon+4)(4Mn+\varepsilon)} \right\}.$$

Using the approximate Brouwer fixed point theorem, construct $p \in P$ such that $\rho(p, \varphi(p)) < r$. Then for each i we have that $f_i(p) \leqslant \varepsilon/2 < \varepsilon$ and $|p_i f_i(p)| < \varepsilon R$, by Lemma 98 and Lemma 99. Suppose that $p_i > R$. Then

$$f_i(p) > -\frac{\varepsilon R}{p_i} > -\frac{\varepsilon R}{R} = -\varepsilon;$$

whence $|f_i(p)| < \varepsilon$. ■

With an application of **MIN** we can recover the full classical version of Walras's existence theorem for economies with uniformly continuous demand functions.

Note that proving, using **MIN**, the full classical version of Walras's existence theorem from Proposition 96 only requires the first condition of that proposition. However, condition 2 of Proposition 96 is vital for proving, in the next section, a version of Uzawa's equivalence theorem for the approximate versions of Brouwer's fixed point theorem and Walras's existence theorem.

4.2.2 Uzawa's equivalence theorem

In [105] Yasuhito Tanaka states that the Uzawa equivalence theorem is nonconstructive, when what he actually shows is that Walras's existence theorem is nonconstructive. In fact, the standard proof of the equivalence of Walras's existence theorem and Brouwer's fixed point theorem is essentially constructive; see [109]. Uzawa's equivalence theorem can be more informatively stated as follows; we use $\mathcal{E}$ to denote the set of economies with uniformly continuous excess demand functions which satisfy Walras's law, and $\mathcal{C}(P)$ to denote the set of uniformly continuous functions from P into itself.

Theorem 100 *There exist mappings $U : \mathcal{E} \to \mathcal{C}(P)$ and $V : \mathcal{C}(P) \to \mathcal{E}$ such that*

1. *for each E in $\mathcal{E}$, $p \in P$ is a fixed point of $U(E)$ if and only if p is an equilibrium of E;*

2. *for each φ in $\mathcal{C}(P)$, $p \in P$ is an equilibrium of $V(\varphi)$ if and only if p is a fixed point of φ.*

Proof. We sketch Uzawa's original proof [109]. For an economy E and a price vector p we set

$$U(E)(p) = \frac{v(p)}{\sum v_k(p)},$$

where $v_i(p) = p_i + \max\{0, f_i(p)\}$. If p is a fixed point of $U(E)$, then $v_i(p) = \lambda p_i$ for each i, where

$$\lambda = \sum v_k(p).$$

The assumption that $f_i(p) > 0$ for some i implies that $\lambda > 1$, which leads to a contradiction of Walras's law.

Now let φ be a continuous function from P into itself and define an economy $V(\varphi)$ by setting

$$f_i(p) = \varphi_i(p) - p_i \mu(p),$$

where

$$\mu(p) = \frac{\sum p_k \varphi_k(p)}{\sum p_i^2};$$

it is easily checked that $f_0, \ldots, f_n$ satisfy Walras's law. Let p be an equilibrium of $V(\varphi)$: for each i

$$\varphi_i(p) \leqslant p_i \mu(p)$$

with equality if $p_i > 0$. Suppose that $\varphi_i(p) \neq p_i \mu(p)$. Then $\neg (p_i > 0)$ implies that $p_i = 0$, so

$$0 \leqslant \varphi_i(p) \leqslant p_i \mu(p) = 0;$$

whence $\varphi_i(p) = p_i \mu(p)$, which is absurd. It follows that $\varphi_i(p) = p_i \mu(p)$ for each i. Summing over all i gives that $\mu(p) = 1$, and therefore $\varphi_i(p) = p_i$ for each i. ∎

Corollary 101 *Brouwer's fixed point theorem and Walras's existence theorem are equivalent.*

Corollary 102 *Walras's existence theorem is in* $(\mathbf{LLPO}, \mathbf{MIN}]$.

Does Uzawa's equivalence theorem also hold with Walras's existence theorem and Brouwer's fixed point theorem replaced by their approximate (and constructively valid) forms? The next theorem answers this question in the affirmative.

Theorem 103 *The approximate Brouwer fixed point theorem and the approximate Walras existence theorem (Proposition 96) are equivalent.*

Proof. In the proof of Proposition 96 it is shown that the approximate Brouwer fixed point theorem implies the approximate Walras existence theorem. For the converse, let φ be a uniformly continuous function from P into P, let $\varepsilon \in (0,1)$, and let $f_0, \ldots, f_n$ be the excess demand functions of $V(\varphi)$:

$$f_i(p) = \varphi_i(p) - p_i \mu(p),$$

where

$$\mu(p) = \frac{\sum p_i \varphi_i(p)}{\sum p_i^2}.$$

Let

$$M = \sup\{\mu(p) : p \in P\},$$

and set $r = \varepsilon^2/2n$. By the approximate Walras existence theorem there exists $p \in P$ such that $f_i(p) \leqslant r/(n+1)$ for each i, and if $p_i > r/(2M(n+1))$, then

$$0 \leqslant \varphi_i(p) \leqslant p_i\mu(p) + r/(n+1).$$

Either $p_i < r/(M(n+1))$ or $p_i > r/(2M(n+1))$. In the the former case $p_i\mu(p) + r(n+1) < \varepsilon$, so

$$\begin{aligned} |p_i - \varphi_i(p)| &< \varepsilon, \\ p_i\mu(p) &< r/(n+1), \\ |\varphi_i(p) - p_i\mu(p)| &< r/(n+1). \end{aligned}$$

On the other hand, if $p_i > r/(2M(n+1))$, then $|f_i(p)| < r/(n+1)$ and once again $|\varphi_i(p) - p_i\mu(p)| < r/(n+1)$. Summing over all i we have

$$\begin{aligned} |1 - \mu(p)| &= \left|\sum \varphi_k(p) - \sum p_k\mu(p)\right| \\ &\leqslant \sum |\varphi_k(p) - p_k\mu(p)| \quad < \quad r. \end{aligned}$$

Hence

$$(1-r)\sum p_i^2 < \sum p_k\varphi_k(p) < (1+r)\sum p_k^2$$

and

$$\varphi_i(p) \leqslant p_i\mu(p) + r/(n+1) < p_i(1+r) + r/(n+1) < p_i + 2r \leqslant p_i + \varepsilon.$$

Now suppose that there exists i such that $\varphi_i(p) < p_i - \varepsilon$. Then $p_i > \varepsilon$ and

$$\begin{aligned}
\sum p_k \varphi_k(p) &= p_i \varphi_i(p) + \sum_{k \neq i} p_k \varphi_k(p) \\
&< p_i(p_i - \varepsilon) + \sum_{k \neq i} p_k \\
&< p_i^2 - \varepsilon^2 + \sum_{k \neq i} p_k \left(p_k + \delta\right) \\
&= \left(\sum p_k^2\right) + 2nr - \varepsilon^2 \quad \leqslant \quad (1-r) \sum p_k^2
\end{aligned}$$

—a contradiction. Thus $\varphi_i(p) \geqslant p_i - \varepsilon$ for each i, so $|p_i - \varphi_i(p)| \leqslant \varepsilon$. Since all norms on $\mathbf{R}^n$ are equivalent, this completes the proof. ■

Suppose that $\mathbf{AC}_{\omega,2}$ holds. Then, since Brouwer's fixed point theorem and Walras's existence theorem both

a) imply **LLPO**, and

b) are equivalent to their approximate versions in the presence of **LLPO**,

the Uzawa equivalence theorem (Corollary 101) follows directly[51] from Theorem 103.

When trying to produce an exact constructive version of a nonconstructive result a natural additional hypothesis to add is that any exact solution is unique in a computationally meaningful way. A function φ from P into P is said to have *at most one fixed point* if for all distinct $p, p' \in P$ either

$$\rho\left(p, \varphi(p)\right) > 0 \text{ or } \rho\left(p', \varphi\left(p'\right)\right) > 0.$$

We say that an economy E has *at most one equilibrium* if for all distinct $p, p' \in P$ either

$$\max\left\{f_i(p) : 0 \leqslant i \leqslant n\right\} > 0 \text{ or } \max\left\{f_i\left(p'\right) : 0 \leqslant i \leqslant n\right\} > 0.$$

[51] Though, in order to avoid a circular argument we must appeal to Tanaka's direct proof that Walras's existence theorem implies **LLPO**, rather than to Corollary 102.

In [110], Veldman showed that Brouwer's fixed point theorem for functions with at most one fixed point is equivalent, under countable choice, to Brouwer's fan theorem for decidable bars. Using Theorems 100 and 103 we can extend this result to economies with at most one equilibrium.

In order to do this we must recall the function U from Theorem 100 and the approximations to U used in the proof of Proposition 96: for an economy E, a price vector p, and a positive real number δ we set

$$\begin{aligned} U(E)(p) &= \frac{v(p)}{\sum v_k(p)}, \\ U_\delta(E)(p) &= \frac{v_\delta(p)}{\sum v_{\delta,k}(p)}, \end{aligned}$$

where $v_i(p) = p_i + \max\{0, f_i(p)\}$ and $v_{\delta,i}(p) = \max\{0, f_i(p) - \delta\}$.

Lemma 104 *The functions U_δ converge pointwise to U as δ tends to zero.*

Proof. Fix $\varepsilon > 0$, let $M = \sup\{v_i(p) : 0 \leqslant i \leqslant n, p \in P\}$, and let

$$\delta = \frac{\varepsilon}{2M(n+1)+1}.$$

By the definitions of v_i and $v_{\delta,i}$ we have

$$\begin{aligned} |v_i(p) - v_{\delta,i}(p)| &< \delta, \\ \left|\lambda - \lambda'\right| &< (n+1)\delta, \end{aligned}$$

where

$$\lambda = \sum v_k(p) \text{ and } \tilde{\lambda} = \sum v_{\delta,k}.$$

Note that $\lambda \leqslant 1 + M(n+1)$. Then, for each $E \in \mathcal{E}$ and each $p \in P$,

$$\begin{aligned}
|U_i(E)(p) - U_{\delta,i}(E)(p)| &= \left|\lambda^{-1}v_i(p) - \tilde{\lambda}^{-1}v_{\delta,i}(p)\right| \\
&= \left(\lambda\tilde{\lambda}\right)^{-1}\left|\tilde{\lambda}v_i(p) - \lambda v_{\delta,i}(p)\right| \\
&\leqslant \left|\tilde{\lambda}v_i(p) - \lambda v_i(p)\right| + |\lambda v_i(p) - \lambda v_{\delta,i}(p)| \\
&= \left|\lambda - \tilde{\lambda}\right| |v_i(p)| + \lambda\,|v_i(p) - v_{\delta,i}(p)| \\
&< (n+1)\delta M + (1 + M(n+1))\delta \quad = \quad \varepsilon,
\end{aligned}$$

by our choice of δ. ■

Theorem 105 (CC) *The statement*

> *Every economy on P with uniformly continuous excess demand functions which satisfies Walras's law and which has at most one equilibrium has an equilibrium.*

is equivalent to $\boldsymbol{FT}_\Delta$.

Proof. Let E be an economy with uniformly continuous excess demand functions $f_0, \dots, f_n$, let φ be a uniformly continuous function from P into P, and let U, V be the functions from Theorem 100. The proof of Theorem 103 taken together with Lemma 104 establishes the following.

a) For each $\varepsilon > 0$ there exists $\delta > 0$ such that if $f_i(p) \leqslant \delta$ for each i, then $\rho(p, U(E)(p)) < \varepsilon$.

b) For each $\delta > 0$ there exists $\varepsilon > 0$ such that if $\rho(p, \varphi(p)) < \varepsilon$, then $V(\varphi)_i(p) \leqslant \delta$ for each i.

The contrapositives of a) and b) together with Theorem 100 show that to each economy with at most one equilibrium we can associate a function from P into itself with at most one fixed point, and to each function from P into P with at most one fixed point we can associate an economy with at most one equilibrium. The result now follows from the equivalence of $\mathbf{FT}_\Delta$ with Brouwer's fixed point theorem for functions with at most one zero (see [110]). ■

4.3 McKenzie's theorem

A *competitive equilibrium* of an economy consists of a price vector $\mathbf{p} \in \mathbf{R}^N$, points $\xi_1, \ldots, \xi_i \in \mathbf{R}^N$, and a vector η in the *aggregate production set*

$$Y = Y_1 + \cdots + Y_n,$$

satisfying

E1 $\xi_i \in D_i(\mathbf{p})$ for each $1 \leqslant i \leqslant m$.

E2 $\mathbf{p} \cdot \mathbf{y} \leqslant \mathbf{p} \cdot \eta = 0$ for all $\mathbf{y} \in Y$.

E3 $\sum_{i=1}^{m} \xi_i = \eta$.

An economy is said to have *approximate competitive equilibria* if for all $\varepsilon > 0$ there exist a price vector $\mathbf{p} \in \mathbf{R}^N$, points $\xi_1, \ldots, \xi_i \in \mathbf{R}^N$, and a vector η satisfying **E1**,**E3**, and

AE $\mathbf{p} \cdot \eta > -\varepsilon$.

Let F_i denote the demand function on $(X_i, \succcurlyeq_i)$. A subset Y of a normed space is said to be a *convex cone* if $\lambda y \in Y$ and $y + y' \in Y$ whenever $y, y' \in Y$ and $\lambda \geqslant 0$. The *convex conic closure* $\mathrm{cone}(Y)$ of Y is the smallest convex cone containing Y; that is,

$$\mathrm{cone}(S) = \{r(tx + (1-t)y) : r > 0, t \in [0,1], x, y \in S\}.$$

We use S° to denote the interior of a subset S of a metric space.

We can now state McKenzie's theorem on the existence of competitive equilibria.

McKenzie's Theorem
Suppose that

(i) each X_i is compact and convex;

(ii) *each* $\succcurlyeq_i$ *is continuous and strictly convex;*

(iii) $(X_i \cap Y)^\circ$ *is nonempty for each* i*;*

(iv) Y *is a closed convex cone;*

(v) $Y \cap \{(x_1, \ldots, x_N) : x_i \geqslant 0$ *for each* $i\} = \{0\}$*; and*

(vi) *for each* $\mathbf{p} \in \mathbf{R}^N$ *and each* i*, if* $\sum_{i=1}^m F_i(\mathbf{p}) \in Y$*, then there exists* $\boldsymbol{x}_i \in X_i$ *such that* $\mathbf{x}_i \succ_i F_i(\mathbf{p})$*.*

Then there exists a competitive equilibrium.

The standard proofs of McKenzie's theorem all contain seemingly necessary applications of Brouwer's fixed point (often in the guise of Kakutani's fixed point theorem). Since the construction of exact fixed points is not possible—even with strong assumptions on the function, like the uniqueness of any fixed point [89, 110]—it seems unlikely that a constructive proof of the existence of exact competitive equilibria is possible under any economically reasonable assumptions.

It may seem that given Theorem 43 and the approximate version of Brouwer's fixed point theorem, that all the hard work for giving a computational version of McKenzie's theorem has already been done. This is not the case: much care and attention must be given to the construction of the, family, of set valued mappings to which we will apply Kakutani's fixed point theorem.

4.3.1 Constructing competitive equilibria

Our constructive version of McKenzie's theorem is the following.

Theorem 106 *Suppose that*

(i) *each* X_i *is compact and convex;*

(ii) *each* $\succcurlyeq_i$ *is continuous and uniformly rotund;*

(iii) $(X_i \cap Y)^\circ$ *is inhabited for each* i*;*

(iv) Y *is a located closed convex cone;*

(v) $Y \cap \{(x_1, \ldots, x_N) : x_i \geqslant 0$ *for each* $i\} = \{0\}$*; and*

(vi) for each $\mathbf{p} \in \mathbf{R}^N$ *and each* i*, if* $\sum_{i=1}^{m} F_i(\mathbf{p}) \in Y$*, then there exists* $\boldsymbol{x}_i \in X_i$ *such that* $\mathbf{x}_i \succ_i F_i(\mathbf{p})$*.*

Then there are approximate competitive equilibria.

Our proof follows the standard classical proof via Kakutani's fixed point theorem (see [86]) as closely as possible; typical of constructive mathematics, it has a distinctly geometric character. The *polar* of a subset S of $\mathbf{R}^N$ is the set

$$S^{\mathrm{pol}} = \left\{\mathbf{p} \in \mathbf{R}^N : \mathbf{p} \cdot \mathbf{x} \leqslant 0 \text{ for all } \mathbf{x} \in S\right\}.$$

It follows directly from the definition that two sets $S, T \subset \mathbf{R}^N$ have the same polar if and only if they have the same convex conic closure. Classically, a little further work shows that

> (*) the polar of the polar of a set is equal to its convex conic closure,

but this is not the case in our intuitionistic setting as the following Brouwerian counterexample shows.

Given a proposition P, let S be the set

$$S = \{(0,1)\} \cup \{x : x = (1/2, 0) \wedge P\} \cup \{x : x = (1,0) \wedge \neg P\}.$$

Then $(1,1) \in (S^{\mathrm{pol}})^{\mathrm{pol}}$. Suppose that $(1,1) \in \mathrm{cone}(S)$; that is, suppose there exist $r > 0$, $t \in [0,1]$, and $x, y \in S$ such that

$$(1,1) = r(tx + (1-t)y).$$

Without loss of generality, we may suppose $x = (0,1)$. Then either $r > 2$, in which case y must be $(1/2, 0)$ and so P holds, or $r < 3$ and, similarly, $\neg P$ must hold. Hence (*) implies the law of excluded middle.

The above counterexample is rather contrived and seems to have little to do with real mathematics or economics, but seems merely to indicate how one would show (*) to be independent of some formalisation of constructive mathematics. It is, however, relevant: since our framework encapsulates what is computable in a strict, though ill-defined, way, this example shows that we cannot compute the information implicit in "$x \in \mathrm{cone}(S)$"—that there exists $r > 0, t \in [0,1], y, y' \in S$ such that $x = r(ty + (1-t)y')$—given only the information that for all z, if $z \cdot p \leqslant 0$ for all $p \in S$, then $x \cdot z = 0$. Succinctly, belonging to the conic closure of a set gives more computational information than belonging to the polar of the polar of that set, and when it comes to computational information we cannot get something for nothing!

The above failure of (*) results from us having a poor handle on the set S. The sets we deal with in practice are generally more well behaved, and (*) can be proved, constructively, for a large class of sets. The following result meets our needs.

Proposition 107 *Let S be a located closed convex cone in $\mathbf{R}^N$. Then the polar of the polar of S equals S.*

Proof. By definition $x \in (S^{\mathrm{pol}})^{\mathrm{pol}}$ if and only if

$$\bigcap_{s \in S} \{z \in \mathbf{R}^N : z \cdot s \leqslant 0\} \;=\; S^{\mathrm{pol}} \;\subset\; \{x\}^{\mathrm{pol}}.$$

The assumption that $\rho(x, S) > 0$ would contradict the above equation: let y be the closest point to x in S, this exists by Theorem 6 of [30]. Since S is a closed convex cone, $y - x \in S^{\mathrm{pol}}$, but $x - y \notin \{x\}^{\mathrm{pol}}$. Hence $\rho(x, S) = 0$, and, since S is closed, $x \in S$. The converse is straightforward. ■

For each i we fix $\overline{\xi}_i \in (X_i \cap Y)^\circ$ and let $\overline{\xi} = \sum_{i=1}^m \overline{\xi}_i$; without loss of generality, each term of $\overline{\xi}$ is nonzero. The proof of Theorem 106 proceeds by an application of a constructively valid version of Kakutani's fixed point theorem to the set

$$P = \left\{ \mathbf{p} \in Y^{\mathrm{pol}} : \mathbf{p} \cdot \overline{\xi} = -1 \right\}$$

of normalised price vectors. First, however, we require a number of lemmas. For the remainder of the section we assume that the hypothesis of Theorem 106 hold.

Lemma 108 *If* $\mathbf{y} \in Y^{\circ}$*, then* $\mathbf{p} \cdot \mathbf{y} < 0$ *for all nonzero* $\mathbf{p} \in Y^{\mathrm{pol}}$*. Moreover,* $\sup\{\mathbf{p} \cdot \mathbf{y} : \mathbf{p} \in Y^{\mathrm{pol}}, \|p\| = 1\} < 0$.

Proof. Let $\mathbf{p}$ be a nonzero element of Y^{pol}; pick $1 \leqslant i \leqslant N$ such that $p_i \neq 0$, and fix $r > 0$ such that $\overline{B}(\mathbf{y}, r) \subset Y$. Then

$$\mathbf{y}' \equiv \mathbf{y} + (\mathrm{sign}(p_i)r)\mathbf{e}_i \in Y,$$

where $\mathbf{e}_i$ is the ith basis vector. Hence

$$\mathbf{p} \cdot \mathbf{y} < \mathbf{p} \cdot \mathbf{y} + |rp_i| = \mathbf{p} \cdot \mathbf{y}' \leqslant 0.$$

If $\|p\| = 1$, then we may suppose that $|p_i| > 1/2\sqrt{N}$; thus $\mathbf{p} \cdot \mathbf{y} < -|rp_i| < -r/2\sqrt{N}$. ■

Let S be a subset of a metric space X. The *complement* of S is

$$\sim S = \{x \in X : x \neq s \text{ for all } s \in S\}.$$

If S is located, then the *apartness complement* of S is the set

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

Lemma 109 *For each* i *the demand function* F_i *for* X_i *maps into* $\sim(Y^{\circ})$.

Proof. Suppose that $F(\mathbf{p}) \in Y^{\circ}$. Then, by Lemma 108, $\mathbf{p} \cdot F(\mathbf{p}) < 0$, which contradicts Theorem 43. ■

We cannot assert that the (inhabited) intersection of two totally bounded sets is totally bounded—in general this is equivalent to the law of excluded middle.

Lemma 110 *Let X, Y be convex subsets of a normed space such that X, Y are both totally bounded, and $(X \cap Y)^\circ$ is inhabited. Then $X \cap Y$ is totally bounded.*

Proof. Let $\xi \in (X \cap Y)^\circ$ and let $R > 0$; without loss of generality $\xi \in B(0, R)$. Let $Y' = Y \cap B(0, R)$ and let h be the uniformly continuous function which fixes X and maps each point y in $-X$ to the unique point in $[\xi, z] \cap \partial X$. Fix $\varepsilon > 0$ and let $\delta \in (0, \varepsilon/4)$ be such that if $\|y - y'\| < \delta$, then $\|h(y) - h(y')\| < \varepsilon/4$. Let $\{y_1, \ldots, y_k\}$ be a $\delta/2$-approximation of Y and partition $\{1, \ldots, k\}$ into disjoint sets P, Q such that

$$\begin{aligned} i \in P &\Rightarrow \rho(y_i, X) < \delta; \\ i \in Q &\Rightarrow \rho(y_i, X) > \delta/2. \end{aligned}$$

If $i \in P$, then there exists $x \in X$ such that $\rho(x, y_i) < \delta$. Then

$$\|y_i - h(y_i)\| \leqslant \|y_i - x_i\| + \|x_i - h(y_i)\| < \varepsilon/4 + \varepsilon/4 = \varepsilon/2$$

and, since Y is convex, $h(y_i) \in X \cap Y$. The set

$$S = \{h(y_i) : i \in P\}$$

is an ε-approximation of $X \cap Y \cap B(0, R) = X \cap Y'$: fix $z \in X \cap Y$ and pick $1 \leqslant i \leqslant k$ such that

$$\|z - y_i\| < \delta/2.$$

Then $i \in P$, so $h(y_i) \in S$ and

$$\begin{aligned} \|z - h(y_i)\| &\leqslant \|z - y_i\| + \|y_i - h(y_i)\| \\ &< \delta/2 + \varepsilon/2 < \varepsilon. \end{aligned}$$

■

In order to apply Kakutani's fixed point theorem we must have a weakly approximable mapping on a metric space with a compact convex closure.

We can now show that the set of price vectors is a suitable domain for our weakly approximable mapping.

Lemma 111 *P is compact and convex.*

Proof. It is straightforward to show that P is closed and convex; it just remains to show that P is totally bounded. By the bilinearity of the mapping $(\mathbf{p}, \mathbf{x}) \mapsto \mathbf{p} \cdot \mathbf{x}$, both Y^{pol} and

$$\{\mathbf{p} \in \mathbf{R}^N : \mathbf{p} \cdot \overline{\xi} = -1\}$$

are locally totally bounded. Since P is the intersection of these two sets, P is locally totally bounded by Lemma 110. It remains to show that P is bounded: by Lemma 108

$$M = \sup\{\mathbf{p} \cdot \mathbf{y} : \mathbf{p} \in Y^{\text{pol}}, \|p\| = 1\} < 0.$$

Suppose that there exists $\mathbf{p} \in P$ such that $\|\mathbf{p}\| > -1/M$. Then $\mathbf{p}/\|\mathbf{p}\| \in Y^{\text{pol}}$ and $(\mathbf{p}/\|\mathbf{p}\|) \cdot \overline{\xi} = -1/\|\mathbf{p}\| > M$—a contradiction. ■

It remains to produce a weakly approximable mapping on P.

Lemma 112 *For each $\mathbf{y} \in Y$ and each $r > 0$ there exists $\mathbf{p} \in P$ such that $\mathbf{p} \cdot \mathbf{y} > -r$.*

Proof. Fix $\mathbf{y} \in \partial Y$. Suppose that

$$\sup\{\mathbf{p} \cdot \mathbf{y} : \mathbf{p} \in P\} < 0;$$

this supremum exists since P is totally bounded and $(\mathbf{p}, \mathbf{x}) \mapsto \mathbf{p} \cdot \mathbf{x}$ is uniformly continuous. Then there exists $\mathbf{z} \in -Y$ such that $\mathbf{p} \cdot \mathbf{z} < 0$ for all $\mathbf{p} \in P$. But

$$\mathbf{z} \in P^{\text{pol}} = (Y^{\text{pol}})^{\text{pol}} = Y.$$

This contradiction ensures that $\sup\{\mathbf{p} \cdot \mathbf{y} : \mathbf{p} \in P\} = 0$, from which the result follows. ■

The proof of the following simple lemma is left to the reader.

Lemma 113 *The composition of a weakly approximable mapping with a uniformly continuous function is weakly approximable.*

Here is the last piece of the puzzle:

Lemma 114 *For a fixed $r > 0$ and for each $\mathbf{z} \in \partial Y$, define*

$$g_r(\mathbf{z}) = \{\mathbf{p} \in P : \mathbf{p} \cdot \mathbf{z} > -r\}.$$

Then $g_r(\mathbf{z})$ is inhabited and located for each $r > 0$, and g_r is weakly approximable.

Proof. That $g_r(\mathbf{z})$ is inhabited for each $\mathbf{z}$ follows from Lemma 112. Fix $\varepsilon > 0$ and let $\delta > 0$ be such that for all $\mathbf{z}, \mathbf{z}' \in \mathbf{R}^N$, if $\|\mathbf{z} - \mathbf{z}'\| < \delta$, then $\|\mathbf{p} \cdot \mathbf{z} - \mathbf{p} \cdot \mathbf{z}'\| < r/2$ for all $\mathbf{p} \in P$—such a δ exists since the mapping $(\mathbf{p}, \mathbf{x}) \mapsto \mathbf{p} \cdot \mathbf{x}$ is uniformly continuous and P is totally bounded. Let $\mathbf{z}, \mathbf{z}' \in \mathbf{R}^N$ be such that $\|\mathbf{z} - \mathbf{z}'\| < \delta$ and let $\mathbf{p} \in g_{r/2}(\mathbf{z})$ and $\mathbf{p}' \in g_{r/2}(\mathbf{z}')$. For each $t \in [0, 1]$, let $\mathbf{p}_t = t\mathbf{p} + (1 - t)\mathbf{p}'$ and $\mathbf{z}_t = t\mathbf{z} + (1 - t)\mathbf{z}'$. Then for all $t \in [0, 1]$ we have

$$\begin{aligned}
\mathbf{p}_t \cdot \mathbf{z}_t &= (t\mathbf{p} + (1 - t)\mathbf{p}') \cdot (t\mathbf{z} + (1 - t)\mathbf{z}') \\
&= t^2\mathbf{p} \cdot \mathbf{z} + t(1 - t)(\mathbf{p} \cdot \mathbf{z}' + \mathbf{p}' \cdot \mathbf{z}) + (1 - t)^2\mathbf{p}' \cdot \mathbf{z}' \\
&> -t^2 r/2 - 2t(1 - t)r - (1 - t)^2 r/2 \ = \ -r.
\end{aligned}$$

Hence g_r is weakly approximable with respect to $g_{r/2}$. That $g_r(\mathbf{z})$ is located for each $\mathbf{z} \in \partial Y$ follows from Theorem (4.9) on page 98 of [18], and the uniform continuity of the mapping $\mathbf{p} \mapsto \mathbf{p} \cdot \mathbf{y}$ on P. ■

We now have the **proof of Theorem 106:**

Proof. Let F_i be the demand function for the ith consumer and let

$$F = \sum_{i=1}^{m} F_i.$$

Fix $\varepsilon > 0$ and let $\delta > 0$ be such that for all $\mathbf{p} \in P$, if $\|\mathbf{x} - \mathbf{x}'\| < \delta$, then $\|\mathbf{p} \cdot \mathbf{x} - \mathbf{p} \cdot \mathbf{x}'\| < \varepsilon/2$. Set

$$m = \min\left\{\frac{\varepsilon}{2}, \frac{\delta}{\sup\left\{\left\|\overline{\xi} - \eta\right\| : \eta \in F(P)\right\}}\right\}.$$

For each $r > 0$, define a set valued mapping Φ_r on P by

$$\Phi_r = g_r \circ h \circ F,$$

where h, g_r are as in Lemma 32 and Lemma 114 respectively; Φ_r is well defined by Lemma 109. By Lemmas 32,111,113,114 and Corollary 43, Φ_r is approximable for each $r > 0$ and P is compact and convex. Using Theorem 90, construct $\mathbf{p} \in P$ such that

$$\mathbf{p} \in g_m \circ h \circ F(\mathbf{p}).$$

Set $\xi_i = F_i(\mathbf{p})$ for each i, and set $\eta = F(\mathbf{p})$. Then, by definition, the ξ_i satisfy condition **E1**, and η satisfies **E3**. Pick $t \in [0, 1)$ such that

$$\zeta \equiv h(F(\mathbf{p})) = t\overline{\xi} + (1 - t)\eta.$$

Since $\mathbf{p} \in g_m(\zeta)$ and $\eta \in Y$,

$$-m < \mathbf{p} \cdot \zeta = t\mathbf{p} \cdot \overline{\xi} + (1 - t)\mathbf{p} \cdot \eta \leqslant -t,$$

so $t < m$; whence $\|\zeta - \eta\| < \delta$. By our choice of δ, it follows that $\|\mathbf{p} \cdot \eta - \mathbf{p} \cdot \zeta\| < \varepsilon/2$. Thus $\mathbf{p} \cdot \eta > -\varepsilon$, so **AE** is satisfied. ■

With the help of **MIN** we can recover the existence of an exact competitive equilibrium in the conclusion of Theorem 106. With **WKL** and $\mathbf{AC}_{\omega,2}$ we have: repeatedly apply Theorem 106 to construct sequences $(\mathbf{p}_n)_{n\geqslant 1}$, $(\xi_{1,n})_{n\geqslant 1}, \ldots, (\xi_{m,n})_{n\geqslant 1}$, $(\eta_n)_{n\geqslant 1}$ in $\mathbf{R}^N$ such that $\mathbf{p}_n, \xi_{1,n}, \cdots, \xi_{m,n}, \eta_n$ satisfy **E1,E3** and $\mathbf{p}_n \cdot, \eta_n > -1/n$ for each n. With $m+2$ applications of **WKL** we can construct an increasing sequence $(k_n)_{n\geqslant 1}$ and points $\mathbf{p}, \xi_1, \cdots, \xi_m, \eta \in$

$\mathbf{R}^N$ such that $\mathbf{p}_n \to \mathbf{p}, \xi_{i,n} \to \xi_i$ $(1 \leqslant i \leqslant m)$, and $\mathbf{eta}_n \to \eta$ as $n \to \infty$. The continuity of the demand functions, the dot product, and summation ensure that $\mathbf{p}, \xi_1, \cdots, \xi_m, \eta \in \mathbf{R}^N$ is a competitive equilibrium.

The work above together with Theorem 41 gives the next result, which in the presence of Brouwer's fan theorem improves on Theorem 106.

Theorem 115 *Assume that Brouwer's full fan theorem holds. Suppose that*

(i) each X_i is compact and convex;

(ii) each $\succ_i$ is continuous and strictly convex;

(iii) $(X_i \cap Y)^\circ$ is inhabited for each i;

(iv) Y is a located closed convex cone;

(v) $Y \cap \{(x_1, \ldots, x_N) : x_i \geqslant 0$ for each $i\} = \{0\}$; and

(vi) for each $\mathbf{p} \in \mathbf{R}^N$ and each i, if $\sum_{i=1}^m F_i(\mathbf{p}) \in Y$, then there exists $\boldsymbol{x}_i \in X_i$ such that $\mathbf{x}_i \succ_i F_i(\mathbf{p})$.

Then there are approximate competitive equilibria.

Proof. Using Theorem 41 in place of Corollary 43 in the proof of Theorem 106, we only require each $\succ_i$ to be strictly convex. ■

References

[1] P. Aczel, 'The type theoretic interpretation of constructive set theory', in A. MacIntypre, L. Pacholski, and J. Paris (eds.), Logic Colloquium 77, North-Holland, Amsterdam, p. 55–66, 1978.

[2] P. Aczel, 'The type theoretic interpretation of constructive set theory: choice principles', in A.S. Troelstra and D. van Dalen (eds.), The L.E.J. Brouwer Centenary Symposium,North-Holland, Amsterdam, p. 1–40, 1982.

[3] P. Aczel, 'The type theoretic interpretation of constructive set theory: inductive definitions', in R.B. Marcus et al. (eds.), Logic, Mathodology and Philosophy of Science VII, North-Holland, Amsterdam, p. 17–49, 1986.

[4] P.H.G. Aczel and M. Rathjen, *Notes on Constructive Set Theory*, Report No. 40, Institut Mittag-Leffler, Royal Swedish Academy of Sciences, 2001.

[5] T. Alternkirch, T. Coquand, M. Escardo, and N. Kraus, 'Generalizations of Hedberg's Theorem', TLCA 2013, LNCS 7941, p. 173-188. Springer, Heidelberg, 2013.

[6] H.P. Barendregt, 'The Lambda Calculus, Its Syntax and Semantics', Studies in Logic and the Foundations of Mathematics, Volume 103, North-Holland, 1984.

[7] A. Bauer, 'Realizability as the Connection between Computable and Constructive Mathematics', Lecture notes for CCA 2004.

[8] M. Beeson, 'Goodman's theorem and beyond', Pacific J. Math. **84**(1), pp 1-16, 1979.

[9] M. Beeson, 'Foundations of constructive mathematics', Metamathematical studies, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), Springer-Verlag, Berlin, 1985.

[10] J. Bell, 'Toposes and local set theories', Oxford Logic Guides 14, 1988.

[11] U. Berger, 'A computational interpretation of open induction', Proceedings of the Nineteenth Annual IEEE Symposium on Logic in Computer Science, p. 326–334, IEEE Computer Society Press, 2004.

[12] U. Berger, 'Realisability for Induction and Coinduction with Applications to Constructive Analysis', Journal of Universal Computer Science **16**(18), p. 2535–2555, 2010.

[13] U. Berger, 'From coinductive proofs to exact real arithmetic: theory and applications', Log. Methods Comput. Sci. **7**(1), 2011.

[14] U. Berger, W. Buchholz, H. Schwichtenberg, 'Refined program extraction from classical proofs', Commemorative Symposium Dedicated to Anne S. Troelstra, Ann. Pure Appl. Logic **114**(1-3), p. 3-25, 2002.

[15] H. Billinge, 'Did Bishop have a philosophy of mathematics?', Philos. Math. **11**(2), p. 176-194, 2003.

[16] E.A. Bishop, *Foundations of Constructive Analysis*, McGraw-Hill, New York, 1967.

[17] E. Bishop, 'The crisis in contemporary mathematics', Proceedings of the American Academy Workshop on the Evolution of Modern Mathematics, Historia Math. **2**(4), p. 507-517, 1974.

[18] E.A. Bishop and D.S. Bridges, *Constructive Analysis*, Grundlehren der Math. Wiss. **279**, Springer-Verlag, Heidelberg, 1985.

[19] J. M. Borwein, 'Completeness and the Contraction Principle', Proceedings of the American Mathematical Society **87**(2), p. 246–250, 1983.

[20] V. Brattka and G. Gherardi, 'Weihrauch degrees, omniscience principles and weak computability', J. Symbolic Logic **76**, 2011.

[21] V. Brattka, G. Gherardi and A. Marcone, 'The Bolzano-Weierstrass theorem is the jump of weak Konig's lemma' Ann. Pure Appl. Logic **163**(6), p. 623-655, 2012.

[22] D. Bridges, 'Preference and utility: a constructive development', J. Math. Econom. **9**, p. 165–185, 1982.

[23] D.S. Bridges and F. Richman, *Varieties of Constructive Mathematics*, London Math. Soc. Lecture Notes **97**, Cambridge Univ. Press, 1987.

[24] D. Bridges, 'The constructive theory of preference relations on a locally compact space', Proc. Koninklijke Nederlandse Akad. van Wetenschappen, Ser. A **92**(2), p. 141–165, 1989.

[25] D. Bridges, 'The construction of a continuous demand function for uniformly rotund preferences', Journal of Mathematical Economics **21**, p. 217–227, 1992.

[26] D.S. Bridges, W.H. Julian, R. Mines, and F. Richman, 'Extensions and fixed points of contractive maps in $\mathbf{R}^n$', J. Math. Anal. and Applications **165**, p. 438–456, 1992.

[27] D. Bridges, 'Constructive notions of strict convexity', Math. Logic Quarterly **39**, p. 295-300, 1993.

[28] D. Bridges, 'The constructive theory of preference relations on a locally compact space—II', Mathematical Social Sciences **21**, p. l–9, 1994.

[29] D.S. Bridges and S. Reeves, Constructive mathematics, in theory and programming practice, Philosophia Mathematica **7**(1), p. 65-104, 1999.

[30] D.S. Bridges, F. Richman, and P. Schuster, 'Linear independence without choice', Ann. Pure Appl. Logic **101**, p. 95–102, 2000.

[31] D.S. Bridges, H. Ishihara, P. Schuster, and L. Vita, 'Strong continuity implies uniform sequential continuity', Arch. Math. Logic **44**, p. 887–895, 2005.

[32] D.S. Bridges and L.S. Vîţă, *Techniques of Constructive Analysis*, Universitext, Springer-New-York, 2006.

[33] D. Bridges and I. Loeb, 'Glueing continuous functions constructively', Arch. Math. Logic **49**(5), 2010.

[34] D. Bridges and M. Hendtlass, 'Continuous homomorphisms of R onto a compact group', Math. Log. Q. **56**(2), p. 191-197, 2010.

[35] D.S. Bridges, 'Constructive solutions of differential equations', preprint.

[36] L.E.J. Brouwer, "Collected Works, Vol. I", Amsterdam, North-Holland, 1975.

[37] J. Cederquist, S. Negri, 'A constructive proof of the Heine-Borel covering theorem for formal reals', Types for proofs and programs (Torino, 1995), Lecture Notes in Comput. Sci. 1158, p. 62-75, Springer, Berlin, 1996.

[38] R.M. Chen, 'Independence and conservativity results for intuitionistic set theory', PhD thesis, University of Leeds, 2010.

[39] R.M. Chen and M. Rathjen, 'Lifschitz realizability for intuitionistic ZermeloFraenkel set theory', Archive for Mathematical Logic **51**(7-8), p. 789-818, 2012.

[40] P. Cohen, 'Set theory and the continuum hypothesis', W. A. Benjamin, Inc., New York-Amsterdam, 1966.

[41] T. Coquand, H. Lombardi, 'A logical approach to abstract algebra', Math. Structures Comput. Sci. **16**(5), p. 885-900, 2006.

[42] L. Crosilla and P. Schuster, 'Finite Methods in Mathematical Practice', In: M. Detlefsen, G. Link, eds., Formalism and Beyond. Ontos, Heusenstamm.

[43] H. Curry and R. Feys, *Combinatory Logic* Vol. I, Amsterdam: North Holland, 1958.

[44] D. van Dalen, 'Mystic, Geometer, and Intuitionist: The Life of L.E.J. Brouwer' volume 1 (1999) and 2 (2005), Oxford University Press, Oxford.

[45] H. Diener, 'Weak König's lemma implies the uniform continuity theorem', Computability: the Journal of CiE, to appear.

[46] H. Diener, I. Loeb, 'Sequences of real functions on $[0,1]$ in constructive reverse mathematics', Ann. Pure Appl. Logic **157**(1), p. 50-61, 2009.

[47] R. Lubarsky and H. Diener, 'Separating the Fan Theorem and Its Weakenings', In: Proceedings of LFCS '13, Lecture Notes in Computer Science 7734, p. 280–295, 2013.

[48] R. Diaconescu, 'Axiom of choice and completion', Proc. Amer. Math. Soc. **51**, p. 176–178, 1975.

[49] M. Edelstein, 'On fixed and periodic points under contractive mappings', J. London Math. Soc. **37**, p. 74–79, 1962.

[50] M. Escardó, 'Infinite sets that satisfy the principle of omniscience in all varieties of constructive mathematics', J. Symbolic Log., to appear.

[51] H.M. Friedman, 'Set Theoretic Foundations for Constructive Analysis', Ann. Math. **105**(1), p. 1–28, 1977.

[52] D. Gale, 'The game of hex and the Brouwer fixed-point theorem', American Mathematical Monthly, p 818–827, 1979.

[53] N.D. Goodman and J. Myhill, 'Choice implies excluded middle', Zeit. Math. Logik and Grunderlagen der Math. 24, p. 461, 1975.

[54] R.J. Grayson, 'Heyting-valued semantics', Logic colloquium '82, Stud. Logic Found. Math., 112, North-Holland, Amsterdam, p. 181-208, 1984.

[55] M. Hendtlass, 'The intermediate value theorem in constructive mathematics without choice', Annals Pure and Applied Logic **163**(8), p. 1050-1056, 2012.

[56] M. Hendtlass, 'Fixed point theorems in constructive mathematics', Journal of Logic and Analysis, Volume 4, Paper 10.

[57] M. Hendtlass, 'Kakutani's fixed point theorem in constructive mathematics', Logic Journal of the IGPL, to appear.

[58] M. Hendtlass, 'The computational content of Walras' existence theorem' Applied Mathematics and Computation **217**(13), p. 6185–6191, 2011.

[59] M. Hendtlass and N. Miheisi, 'On the construction of a general equilibrium in a competitive economy' (with Nazar Miheisi), Journal of Mathematical Economics, under revision.

[60] C. Heunen, N.P. Landsman, B. Spitters, 'A topos for algebraic quantum theory', Comm. Math. Phys. **291**(1), p. 63-110, 2009.

[61] R. Iemhoff, 'Kripke models for subtheories of **CZF**', Arch. Math. Logic **49**(2), p. 147-167, 2010.

[62] H. Ishihara, Continuity properties in constructive mathematics, J. Sym. Logic **57**(2), p. 557-565, 1992.

[63] H. Ishihara, 'A note on Goedel-Gentzen translation', Math. Log. Q. **46**, p. 135–137, 2000.

[64] H. Ishihara, 'Constructive reverse mathematics: compactness properties', In: L. Crosilla and P. Schuster eds., From Sets and Types to Topology and Analysis: Towards Practicable Foundations for Constructive Mathematics, Oxford University Press, Oxford, p. 245–267, 2005.

[65] H. Ishihara, 'Reverse mathematics in Bishop's constructive mathematics', Philosophia Scientiae, Cahier special **6**, p. 43–59, 2006.

[66] H. Ishihara, 'Weak König's Lemma Implies Brouwer's Fan Theorem: A Direct Proof', Notre Dame Journal of Formal Logic **47**(2), p. 249–252, 2006.

[67] P.T. Johnstone, *Topos Theory*, L. M. S. Monographs no. 10. Academic Press, 1977.

[68] S. Kakutani, 'A generalization of Brouwer's fixed point theorem', Duke Math. J. **8**(3), p. 457–459, 1941.

[69] U. Kohlenbach, *Applied proof theory: proof interpretations and their use in mathematics*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2008.

[70] J.L. Krivine,‘ Realizability in classical logic’, Panoramas et synthses, Socit Mathmatique de France **27**, p. 197-229, 2009.

[71] M.D. Krol, ‘A topological model for intuitionistic analysis with Kripke’s schema’, Z. Math. Logik Grundlagen Math. **24**, p. 427–436, 1978.

[72] P. Lietz, From Constructive Mathematics to Computable Analysis via the Realizability Interpretation, Ph.D. thesis, Technische Universitt, Darmstadt, 2004.

[73] J. Lipton, ‘Realizability, set theory and term extraction’, The Curry-Howard isomor- phism, Cahiers du Centre de Logique de l’Universite Catholique de Louvain vol. 8, p. 257–364, 1995.

[74] I. Loeb, ‘Questioning Constructive Reverse Mathematics’, Constructivist Foundations **7**(2), p. 131-140, 2012.

[75] H. Lombardi, C. Quitte, Alg‘ebre Commutative. Methodes constructives (Modules projectifs de type fini), Calvage et Mounet, 2011.

[76] R.S. Lubarsky, ‘Independence results around constructive ZF’, Ann. Pure Appl. Logic **132**(2-3), p. 209–225, 2005.

[77] R. Lubarsky, ‘On the Cauchy Completeness of the Constructive Cauchy Reals’, Math. Logic Q. **53**(4-5), p. 396–414, 2007.

[78] R. Lubarsky and M. Rathjen, On the constructive Dedekind reals, in Sergei N. Artemov and Anil Nerode (eds.), Proceedings of LFCS 07, Lecture Notes in Computer Science Vol. 4514, Springer, p. 349–362, 2007.

[79] R. Lubarsky, F. Richman, and P. Schuster, ‘ The Kripke Schema in Metric Topology’, Mathematical Logic Quarterly **58**(6), p. 498–501, 2012.

[80] M. Mandelkern, 'Constructive Continuity', Mem. Amer. Math. Soc. **42**(277), 1983.

[81] P. Martin-Löf, 'Intuitionistic type theory' Notes by Giovanni Sambin, Studies in Proof Theory., Lecture Notes, 1. Bibliopolis, Naples, 1984.

[82] D.C. McCarty, 'Realizability and recursive mathematics', PhD thesis, Oxford University, 1984.

[83] D.C. McCarty, 'Proofs and constructions', Foundational theories of classical and constructive mathematics, p. 209-225, West. Ont. Ser. Philos. Sci. **76**, Springer, Dordrecht, 2011.

[84] L.W. McKenzie, 'On the existence of general equilibrium for a competitive market', Econometrica **27**, p. 54–71, 1959.

[85] L.W. McKenzie, 'On the existence of general equilibrium for a competitive market: some corrections', Econometrica **29**, p. 247–248, 1961.

[86] L.W. McKenzie, 'The Classical Theorem on Existence of Competitive Equilibrium', Econometrica **49**(4), p. 819–841, 1981.

[87] R. Mines, F. Richman, and W. Ruitenburg, 'A Course in Constructive Algebra', Universitext, Springer-Verlag, New York, 1988.

[88] J. Myhill, 'Constructive set theory', J. Symbolic Logic **40**(3), p. 347–382, 1975.

[89] V. P. Orevkov, 'A constructive map of the square into itself, which moves every constructive point', Dokl. Akad. Nauk SSSR **152**, p. 55-58, 1963.

[90] M. Rathjen,'The disjunction and related properties for constructive Zermelo-Fraenkel set theory', J. Symbolic Logic **70**(4), p. 1232-1254, 2005.

[91] F. Richman, 'Interview with a constructive mathematician', Modern Logic **6**, p. 247–271 1996.

[92] F. Richman, ‘The fundamental theorem of algebra: a constructive development without choice’, Pacific Journal of Mathematics **196**, p. 213–230, 2000.

[93] F. Richman, ‘Spreads and choice in constructive mathematics’, Indag. Math. **13**(2), p. 259-267, 2002.

[94] E. Rothe, ‘Zur Theorie der topologischen Ordnung und der Vektorfelder in Banachschen Räumen’, Compos. Math. **5**, p. 177–196, 1937.

[95] G. Sambin, ‘The Basic Picture Structures for Constructive Topology’, Oxford University press, to appear.

[96] H. Scarf, ‘On the approximation of fixed points of continuous mappings’, SIAM Journal of Applied Mathematics **15**, p. 1328–1343, 1967.

[97] P. Schuster, ‘Unique solutions’, Math. Log. Q. **52**(6), p. 534-539, 2006.

[98] P. Schuster and J. Zappe, ‘Über das Kripke-Schema und abzählbare Teilmengen’, Logique et Anal. **51**, no. 204, p. 317-329, 2008.

[99] H. Schwichtenberg, ‘Realizability interpretation of proofs in constructive analysis’, Theory Comput. Syst. **43**(3-4), p. 583–602, 2008.

[100] S. Simpson, *Subsystems of second order arithmetic*, Perspectives in Logic (2nd ed.), Cambridge University Press, 2009.

[101] D.R. Smart, *Fixed point theorems*, Cambridge Univ. Press, 1974.

[102] B. Spitters, ‘Constructive algebraic integration theory’, Ann. Pure Appl. Logic **137**(1-3), p. 380-390, 2006.

[103] A. Swan, ‘**CZF** does not have the existence property’, Ann. Pure Appl. Logic, to appear.

[104] A. Takayama, *Mathematical Economics*, The Dryden Press, Hinsdale, Illinois, 1974.

[105] Y. Tanaka, 'Undecidability of Uzawa equivalence theorem and LLPO', App. math. and computation **201**, p. 378–383, 2008.

[106] P. Taylor, 'A lambda calculus for real analysis', J. Log. Anal. 2, Paper 5, 2010.

[107] A.S. Troelstra, D. van Dalen, 'Constructivism in mathematics: An introduction, Vol. I', Studies in Logic and the Foundations of Mathematics, 121. North-Holland Publishing Co., Amsterdam, 1988.

[108] A.S. Troelstra, D. van Dalen, 'Constructivism in mathematics: An introduction, Vol. II', Studies in Logic and the Foundations of Mathematics, 123.North-Holland Publishing Co., Amsterdam, 1988.

[109] H. Uzawa, Walras' Existence Theorem and Brouwer's Fixed Point Theorem, Economic Studies Quarterly **8**, p. 59–62, 1962.

[110] W. Veldman, 'Brouwer's approximate fixed-point theorem is equivalent to Brouwer's fan theorem', *Logicism, intuitionism, and formalism*, Synth. Libr. **341**, p. 277–299, Springer, Dordrecht, 2009.

[111] K. Weihrauch, 'Computable analysis: An introduction', Texts in Theoretical Computer Science, Springer-Verlag, Berlin, 2000.