Comparison and Analysis of Constructive Set Theories

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Logic and Computation

eingereicht von

Robert Freiman
Matrikelnummer 1225784

an der Fakultät für Informatik
der Technischen Universität Wien

Betreuung
Betreuer/in: Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Christian Fermüller
Mitwirkung: Univ.Prof.i.R. Dr.phil. Alexander Leitsch

Wien, 09.12.2019

Unterschrift Verfasser/in

Unterschrift Betreuer/in

Erklärung

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst, die verwendeten Quellen und Hilfsmittel vollständig angegeben und Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe

Ort, Datum

Unterschrift

Acknowledgements

Ich möchte mich hiermit herzlich bei meinen Betreuern, Prof. Christian Fermüller und Prof. Alexander Leitsch, bedanken, die mir während der gesamten Zeit mit positivem Input, konstruktiver Kritik und aufmerksamen Korrekturhinweisen zur Seite standen.

Kurzfassung

Seit der Entdeckung der Curry-Howard-Korrespondenz kennen wir die tiefliegenden Zusammenhänge zwischen Berechenbarkeit und Beweisbarkeit in intuitionistischer Logik. Die konstruktive Mathematik, die auf intuitionistischer Logik basiert, ist daher ein vielversprechender Ansatz in der Untersuchung, welche Teile der Mathematik computational fassbar sind. Potentielle Anwendungen liegen dabei im Bereich der automatischen Deduktion und der automatischen Theorembeweiser.

Das Ziel der vorliegenden Arbeit ist es, verschiedene Ansätze der konstruktiven Mengenlehre in einer verständlichen und in sich geschlossenen Art und Weise zu präsentieren und zu vergleichen und dadurch das Potential der Anwendbarkeit in automatischer Deduktion und automatischen Theorembeweisern offenzulegen. Das spezielle Augenmerk auf Mengenlehre ist ihrer historischen Rolle als Grundlage der gesamten Mathematik geschuldet.

Die wichtigsten Theorien, die in dieser Arbeit behandelt werden, ist Brouwers Mengenlehre, die durch die Zermelo-Fraenkelsche Mengenlehre inspirierten Mengenlehren **IZF** und **CZF**, sowie Martin-Löfs Mengenlehre **ML**. In einem ersten Schritt werden die Theorien und ihre Axiomatisierungen vom konstruktivistischen Standpunkt aus motiviert. Einige grundlegende Resultate werden abgeleitet, um einen Eindruck vom Arbeiten in den Theorien und deren Grenzen zu erhalten. Der Schwerpunkt der Untersuchung liegt allerdings in der metamathematischen Analyse.

Die metamathematische Analyse der konstruktiven Zermelo-Fraenkel Mengenlehren wird durch zwei semantische Methoden durchgeführt: Realisierbarkeit und topologische Semantik. Erstere baut direkt auf Konzepten der Berechenbarkeit auf und ist deshalb geeignet, konstruktivistisch wünschenswerte metamathematische Eigenschaften zu untersuchen. Ein Beweis wird präsentiert, der sich auf die Verwendung topologischer Semantik stützt und zeigt, dass das Prinzip der entscheidbaren Bar-Induktion unabhängig von einer auf **IZF** basierten Variante der Brouwer'schen Mathematik ist. Schlussendlich wird eine sinnbewahrende Interpretation von **CZF** in **ML** – einer Theorie, mit konstruktiv sehr klarem und wohlbegründetem Mengenbegriff – diskutiert. Dadurch stellt sich **CZF** nicht nur besonders geeignet für den mathematischen Alltag, sondern auch als konstruktiv gerechtfertigte Theorie und dadurch als vielversprechende Ausgangbasis für Anwendungen heraus.

Abstract

Since the discovery of the Curry-Howard-correspondence we know of the deep-lying connections between computability and provability in intuitionistic logic. Hence, constructive mathematics, being based on intuitionistic logic, promises to be a fruitful tool in investigating the computational content of classical mathematics with potential applications in the areas of automated deduction and automated theorem proving.

The objective of this thesis is to present and compare different approaches to constructive set theories in a comprehensible and self-contained fashion and thereby demonstrate its potential for applications in automated deduction and automated theorem proving. The particular significance placed here on set theory is due to its historically proven relevance in providing the very foundation of mathematics.

The main theories discussed in the thesis are Brouwerian set theory, the axiomatic Zermelo-Fraenkel-style set theories **IZF** and **CZF** and Martin-Löf's set theory **ML**. In a first step, the theories and their axiomatizations are motivated from the constructive standpoint. Some basic results are inferred to get used to reasoning and limitations within the respective system. The broadest investigation, however, is conducted by means of metamathematical analysis.

Metamathematical analysis of constructive Zermelo-Fraenkel set theories is executed by two semantical tools: Realizability and topological semantics. The former builds directly on notions from computability theory and thus allows for an investigation of metamathematical properties that are constructively desirable. A proof using topological semantics is presented to obtain an independence proof of the principle of decidable bar induction from a variant of Brouwer's mathematics formalized within **IZF**. Finally, a meaning-persevering interpretation of **CZF** into **ML** – a theory that is considered to give a constructively clear and well-justified notions of sets – is discussed. This makes **CZF** not only especially well-suited for mathematical practice, but also vindicates its constructive nature and makes it a promising starting point for applications.

Contents

Erklärung 1

Acknowledgements 2

Kurzfassung 3

Abstract 4

Contents 5

1 Introduction 7

2 Set Theory and Constructivism 11

2.1 Informal constructive set theory 11

2.2 Brouwer's set theory 16

2.3 Martin-Löf's set theory 20

2.4 **IZF** and **CZF** 23

3 Axiomatic constructive set theories – **IZF** and **CZF** 34

3.1 Setting the stage 34

3.2 Basic concepts in **BST** 36

3.3 **ZF** vs **IZF** or the problem with trichotomy 39

3.4 **IZF** vs **CZF** or how to live without the powerset operation 42

3.5 Relation between the theories 47

4 Metamathematical properties of constructive axiomatic set theories 49

4.1 Some aspects of computability theory 49

4.2 Metamathematical properties 50

4.3 Realizability of **CZF** 53

4.4 A simple Completeness Theorem 56

4.5 Soundness Theorem 57

4.6 Proof of metamathematical properties 65

4.7 Further results 67

5 Martin-Löf's set theory 70

5.1 Formulating **ML** 70

5.2 Additional rules 78

5.3 Interpreting **CZF** in **ML** 81

6 Topological Semantics and Independence of Bar induction 87
6.1 Bar induction in Brouwer's mathematics 87
6.2 Heyting-valued semantics 88
6.3 Heyting-valued interpretation of **IZF** 96
6.4 Independence of Bar induction 111
6.5 Compatibility with the Fan theorem 114
6.6 Compatibility with the Weak Continuity Principle 117
Conclusion 120
References 123
Index 129

1 Introduction

Since the discovery of the Curry-Howard-correspondence we know of the deep connections between computability and provability in intuitionistic logic. Hence, constructive mathematics, being based on intuitionistic logic, promises to be a fruitful tool in investigating the computational content of classical mathematics with potential applications in the areas of automated deduction and automated theorem proving. For example, Martin-Löf's intuitionistic type theory was intended as a formalization of the constructive set theory **ML**. Explicitly based on paradigms from programming, **ML** has inspired programming languages like Agda and Idris and was able to provide concepts for automated proof assistants like NuPRL and Coq (see, for example, [*43*], [*2*], [*33*], [*39*], [*BO*]). Famously, Gonthier gave a proof of the four-color theorem in Coq (see [*27*]) testifying to the efficacy of **ML**-based proof assistants.

The objective of this thesis is to present and compare different approaches to constructive set theories in a comprehensible and self-contained fashion and thereby demonstrate its potential for applications in automated deduction and automated theorem proving. The particular significance placed here on set theory is due to its historically proven relevance in providing the very foundation of mathematics.

Set theory

In [*14*], Georg Cantor defined sets as

> *... a gathering together into a whole of definite, distinct objects of our perception or of our thought—which are called elements of the set.*

Cantor was the first to start a systematic mathematical analysis of the notion of infinity, that had been used informally up to that date. He proved that the set of natural numbers $\mathbb{N}$ and the set of rational numbers $\mathbb{Q}$ is equipotent and that the set of real numbers $\mathbb{R}$ is uncountable. In 1878 he formulated the *Continuum Hypothesis*, which asserts that every infinite set of real numbers is either countable, i.e., it has the same cardinality as $\mathbb{N}$, or has the same cardinality as $\mathbb{R}$. Already with the 1890ies it turned out that set theory is capable of providing a foundation of contemporary mathematics. With the discovery of Russel's paradox in 1901, however, it soon became evident that mathematics in general and set theory in particular were in desperate need of solid foundations. Mathematics plunged into the *Grundlagenkrise*.

As the Grundlagenkrise dissipated, Cantor's set theory – now formalized by Zermelo and Fraenkel as **ZFC** – witnessed its revival. Nowadays, **ZFC** is considered as the "classical set theory" and as such the foundation of nearly all mathematical activity and is for the most part widely undisputed in this role. As Džamonja writes [22]:

> *Anecdotal evidence from working mathematicians suggests that this axiom system is viewed as more than sufficient for what mathematics needs. While the axiom of choice or the continuum hypothesis might still excite some occasional discussion,*

*mathematicians in most areas seem to happily accept the ''sufficiency'' of **ZFC** and with somewhat less assurance, also its ''necessity''.*

The discipline of metamathematics, founded in the course of *Hilbert's program*, provided us with some celebrated results. For example, Cohen showed in [*16*], that the *continuum hypothesis* independent from **ZFC**. Hence, Platonists like Gödel tend towards extending this system.

Constructivism

Besides Frege's logicism, also formalism, represented by David Hilbert, and intuitionism competed at the *Grundlagenkrise*. The intuitionistic school of thought, championed by Luitzen Brouwer, criticized Cantor's set theory sharply. Besides a fundamental critique of axiomatic formalizations of mathematics, intuitionism blamed the acceptance of infinite collections of objects as complete mathematical entities for the paradoxes. Due to the central role of set theory, however, it was clear to Brouwer that a rethinking of mathematics could not circumvent the development of an intuitionistic notion of set. Before a discussion how this might be accomplished and why it could be of interest, allow first an illustration of a few aspects of the constructivist standpoint.

Nowadays, concerning *mathematical constructivism* we understand a summary of several currents of mathematical practice, like Russian constructivism, Martin-Löf's type theory and others (see [*65*], [*41*], [*43*]). All these currents, however, can be traced back to Brouwer's intuitionism and rely on one major standpoint: Mathematics is a function of human intellect. Particularly, constructivism criticizes (see [*34*], [*9*]):

- Logicism, for favoring the idea of building mathematics on purely logical terms. Constructivism claims that logic is part of mathematics and not the other way around.
- Formalism, for defending the idea that mathematics may be reduced to nothing more but mere manipulations with strings according to predefined rules. Admittedly, many constructivists see formal language as useful tool of communicating mathematical subjects. But the formalism should never be mistaken for mathematics itself, which is assumed to be free of any (formal) language.
- Platonism, for holding that mathematical objects exist in an ideal world, independent of man and time. The mathematician's job is to advance to truths of this world and to describe them. Obviously, intuitionism rejects any kind of transcendental existence of mathematical objects. Instead, mathematical objects are created by mental constructions of an (ideal) mathematician.

The last point makes particularly clear, why constructivism rejects the common *law of excluded middle* (**LEM**). According to this rule, every mathematical statement is either true or false ($\phi \vee \neg\phi$). Indeed, this could hold in a perfect, Platonist world, but it is incompatible with the spirit of constructivism. Let us consider a simple example to illustrate the different roles of time and actual infinity in classical and constructive mathematics. We start with a proposition ψ about natural numbers, such that for each n, a finite procedure is known to check whether or not $\psi(n)$ holds. Our statement ϕ is the assertion that there is some natural number n satisfying $\psi(n)$: $\exists n. \psi(n)$. How would a constructive proof of $\phi \vee \neg\phi$ look like? If ψ is sufficiently non-trivial, all we can do is to calling the procedure to check $\psi(n)$ for each number n

after the other. Having found such an n, we conclude ϕ. On the other hand, we could at no point stop the process and conclude $\neg\phi$. Hence, in the second case, this method is not sufficient for constructively proving $\phi \vee \neg\phi$.

To put it positively: If a constructive proof of a statement of the form $\phi \vee \neg\phi$ is given, this will always give rise to a *decision procedure* for checking which one of ϕ or $\neg\phi$ holds in finitely many steps.

Classically, the existence of a natural number n with property $\psi(n)$ could be proved by leading the assumption of $\neg\exists n.\psi(n)$ into a contradiction and concluding with transcendental existence of an n with $\psi(n)$. This kind of reasoning by contradiction is constructively not justifiable. We give a further investigation of provability in the constructive sense in section 2.1.

Constructive set theory

The first attempt at building a set theory as foundation for constructive mathematics is credited to Brouwer. We will discuss some of his ideas in sections 2.2 and 6.1. In his book [*6*], Bishop showed that it is possible to justify large parts of mathematics constructively.

Being based on paradigms from programming, Martin-Löf's set theory **ML** seems particularly interesting to the computer scientist. We have listed some successful implementations of **ML** at the beginning of this introduction. We give a introduction to the main ideas of **ML** in section 2.3 and a whole description of the theory in chapter 5.

Another approach to constructive set theory is to restrict the successful system **ZFC**, with the hope of obtaining a theory that is justifiable from the constructive viewpoint. We will motivate the most prominent such theory, **IZF**, and its further restriction **CZF** in section 2.4 and further investigate their connections in chapter 3.

The ultimate confirmation that these theories are constructive are obtained only metamathematically. It is, for example, not clear a priori, whether **IZF** implies unwanted instances of **LEM**. Therefore, in chapter 4, we will discuss realizability-semantics referring directly to the notions of computability. We can thus show that **IZF** and **CZF** are suitable starting points for mathematics building on different constructivist currents.

In chapter 5, we show how **CZF** may be interpreted in **ML** in a sense-preserving fashion. Thus, we obtain a metamathematical justification of **CZF** as a restriction of **IZF**.

Finally, we introduce topological semantics of **IZF** in chapter 6. With this semantics we can show, that **IZF** is compatible with Brouwerian mathematics. It turns out, that in this context, the schema of bar induction is too strong an assumption for Brouwer's proof of the Fan-theorem (and in further consequence his theorem that every total function $[0,1] \to \mathbb{R}$ is uniformly continuous).

Throughout the thesis, we infer classical and constructive results as well as results of metamathematical nature. To prevent confusion, we mark classical results with ⓒ. Metamatheorems usually rely on classical reasoning and are marked with Ⓜ.

2 Set Theory and Constructivism

In the following, we will discuss some versions of constructive set theory. We start with an informal version of set theory, that has not been advocated by any subcurrent of constructivism in particular, but can be seen as least common denominator and a good starting point to get used to concepts and operating principles of the field. We will continue with a brief introduction to Brouwer's ideas on set theory as well as Martin-Löf's set theory **ML** and the Hilbert-style systems **IZF** and **CZF**.

2.1 Informal constructive set theory

However different the schools of constructivism may seem, they all converge on the following: A proof of a statement of the form "there is x such that $A(x)$" must be given by a method (construction of the mind, algorithm, etc.) constructing x together with a proof of the fact $A(x)$. To get used to this idea, let us informally discuss some classical proofs that would be rejected by constructivists.

2.1.1 Some non-constructive proofs and notions

The two most important classes of classical proofs that are problematic in a constructive setting are proofs by contradiction and a special form of proof by cases.

Proof by contradiction

$$\frac{\neg\neg\phi}{\phi}$$

Proof by cases

$$\frac{\psi \rightarrow \phi \quad \neg\psi \rightarrow \phi}{\phi}$$

Of course, both rules rely on the law of excluded middle and are thus not accepted by constructivists. We will explain the motivation behind this rejection in the following by examples.

Proof by contradiction

Let us consider König's lemma:

©**König's lemma 2.1 (KL)**: Each infinite, finitely branching tree contains an infinite path.

Proof: Let T be such a tree. For a node d of T denote by T_d the subtree rooted at d. We "construct" an infinite path $\alpha\colon \mathbb{N} \rightarrow T$ as follows: For $\alpha(0)$ take the root. Having defined $\alpha(n)$ suppose, $T_{\alpha(n)}$ is infinite. If for all descendants d of $\alpha(n)$, the trees T_d were finite, then so would be $T_{\alpha(n)}$, contradiction. Thus, there is a descendant d of $\alpha(n)$ such that T_d is infinite. We set $\alpha(n+1) = d$. Clearly, we can repeat this process indefinitely. ∎

Why is the word "construct" written in quotation marks? In the proof we seem to construct an infinite path, but the word cannot be understood in the sense of constructivism: The crucial point is of course the moment when we choose $\alpha(n+1)$: It is impossible, in general decide, which of the T_d is indeed infinite, hence we cannot always explicitly construct the next step of our path. To put it bluntly: The absurdity of non-existence of an object alone does not guarantee its existence. What we mean is of course: The absurdity of non-existence of an object **does not provide us with a construction**.

In first-order theories, constructivist mathematics can be interpreted as subtheory of a classical one (for examples the theories **IZF** and **CZF** are subtheories of **ZFC** – we will discuss these theories later). Indeed, usually every classical mathematician will accept a constructive proof. This means that all constructive theorems and notions can be preserved when passing to the classical setting. The converse, however, cannot be expected to be true. Indeed, we will observe the phenomenon of classically equivalent notions and theorems to become different in constructivism. To put it positively: Often, constructive mathematics distinguishes classically equivalent notions. One example is the Fan theorem, which is classically equivalent to König's lemma.

Definition 2.2: A subset of nodes B of a tree T is called a *bar of* T iff each infinite path of T finally passes through B. A bar is called a *uniform bar*, if there is a number z such that each path α of length $\geq z$ passes through B at or before $\alpha(z)$.

Fan Theorem 2.3 (FT): If B is a bar of a finitely branching tree T, then it is a uniform bar.

ⓒ**Theorem 2.4**: König's lemma is equivalent to the fan theorem.

Proof: FT→KL: Let T be an infinite, finitely branching tree with no infinite path. Let B be the set of nodes v such that T_v is finite. Trivially, B is a bar and thus, by the Fan theorem, there is some $z \in \mathbb{N}$ such that all paths α of length $\geq z$ pass through B at or before $\alpha(z)$. Without loss of generality, suppose this z is minimal. If $z \neq 0$ and v is a node of level $z - 1$, then all its successors d have the property that T_d is finite and hence T_v must be finite. As v was arbitrary in level $z - 1$ this shows that actually z must be 0, i.e. T is finite, contradiction.

KL→FT: Let T be a finitely branching tree and let B be a bar, but not a uniform bar. Let T' be the subtree of nodes reachable from the root without passing B. B not being uniform means that there are unbounded paths not passing through B, i.e. T' is infinite. By König's lemma, T' must have an infinite path, which shows that B cannot be a bar after all. ■

In both directions, the proof is by contradiction and thus rejected by constructivists. As an example of classically equivalent but constructively distinguishable notions we give the following definitions: Let $(V, <)$ be a poset. We say that u is an *upper bound of* $A \subseteq V$ if for all $x \in A$ we have $x \leq u$.

Definition 2.5 of least upper bound, version 1: For $A \subseteq V$ we say that l is the least upper bound of A iff for each upper bound u, we have $u \leq l$.

Definition 2.6 of least upper bound, version 2: For $A \subseteq V$ we say that l is the least upper bound of A iff for each $x \in A$ either $x = l$ or there is some $y \in A$ such that $x < y \leq l$.

Again, classically both notions are equivalent, but every classical proof of this equivalence will not work in a constructive setting. When formulating the least upper bound principle in constructive analysis (each bounded subset of $\mathbb{R}$ has a least upper bound), one usually prefers version 2. An objection to version 1 could be that this definition is *impredicative*: The object l is defined referring to a collection containing l – this is problematic according to some constructivists, in the case when this collection is infinite, but more on this later.

Proof by cases (decidability)

We consider the following simple example:

©**Proposition 2.7**: There are irrational numbers a, b such that a^b is rational.

Proof: If $\sqrt{2}^{\sqrt{2}}$ is rational, put $a = b = \sqrt{2}$. Else, let $a = \sqrt{2}^{\sqrt{2}}$ and $b = \sqrt{2}$. In both cases, a^b is rational. ■

The problem with this proof is that it does not declare a method how to decide whether or not $\sqrt{2}^{\sqrt{2}}$ is rational[1]. In conducting these kinds of case distinctions, one must refer to the kind of transcendental mathematical truth that is rejected in constructivism.

Next, we consider an example from set theory. We call two sets A and B *(extensionally) equal* and write $A = B$ if they contain the same elements, i.e. $\forall x(x \in A \leftrightarrow x \in B)$. We define the set-theoretic (Kuratowski) ordered pair $\langle a, b\rangle \coloneqq \{\{a\}, \{a, b\}\}$. We have the following simple result:

Proposition 2.8: If $\langle a, b\rangle = \langle c, d\rangle$, then $a = c$ and $b = d$.

Classical proof: If $a = b$, then $\langle a, b\rangle = \{\{a\}\}$ and thus $\{a\}$ is the only element of $\langle c, d\rangle$. Hence $a = b = c = d$. On the other hand, for $a \neq b$, the only one-element set of both $\langle a, b\rangle$ and $\langle c, d\rangle$ must be $\{a\}$. Hence, $a = c$. The only two-element set both $\langle a, b\rangle$ and $\langle c, d\rangle$ is $\{a, b\}$ – we conclude $b = d$. ■

The problem with this proof is that we assume that we can decide whether or not $a = b$. But this is not the case in general: Consider two infinite sets of natural numbers that do not follow any apparent law. The only possible procedure of checking if the two sets are equal is to check each element one by one. But such a procedure will never tell with certainty if sets are equal indeed, even if they seem to coincide after any arbitrarily large number of steps. However, we are able to prove the proposition constructively:

Constructive proof: $\{a\}$ is an element of $\langle a, b\rangle$ and $\langle c, d\rangle$ and thus, $\{a\} = \{c\}$ or $\{a\} = \{c, d\}$. In either case, $a = c$. Again, $\{a, b\}$ is contained in both sides and hence, $\{a, b\} = \{c\}$ or $\{a, b\} = \{c, d\}$. In either case, $b = c$ or $b = d$. If $b = c$, then $a = b = c = d$ and both $\langle a, b\rangle$ and $\langle c, d\rangle$ are equal. If $b = d$, then $\langle a, b\rangle = \{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\} = \langle a, b\rangle$. ■

Note that in this proof, we did not make any assumption on the decidability of set equality. We continue another pair of classically equivalent definitions that turn out to be different in constructivism:

[1] Gelfond and Schneider proved independently that $\sqrt{2}^{\sqrt{2}}$ is irrational, see [26]

Definition 2.9: A set A is called *(Kuratowski) finite* iff there is a number n together with a bijective function $f: \{0, \ldots, n-1\} \to A$. The set A is called *subfinite* iff it is a subset of a finite set.

Clearly, classically both notions coincide. Constructively however, they are different. For example, let P be any unsolved mathematical problem, like Goldbach's conjecture. Consider the set

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

Clearly, A is a subset of the finite set $\{0,1\}$ and thus subfinite. But to show that A is finite we need to know whether to define a function with domain $\{0\}$ or $\{0,1\}$, i.e. we would need to solve or reject P.

2.1.2 The notion of set

Following Beeson's analysis [4], in order to define a set S, one must

(i) say what has to be done in order to construct canonical members of S,
(ii) say what has to be done to prove two canonical members of S equal,
(iii) and prove that the equality defined in (ii) is an equivalence relation.

Although merely informal, we cannot leave this definition without remarks. Concerning (i) it seems that there is the important objection that given a set A, the set $B = \{x \in A: \phi(x)\}$, i.e. the subset of elements x of A satisfying some property $\phi(x)$ is acceptable in constructivist set theories such as Martin-Löf's or **IZF** and **CZF** (if required with regard to suitable restrictions on ϕ)[2]. In constructing B however, one would not so much talk about "constructing" the set B but rather of "separating the elements of A" with respect to ϕ.

Beeson writes that sets coming equipped with an equality is a unique feature of constructive set theory distinguishing it from classical set theory, as there is a global relation of equality defined on the universe V of all sets. This statement seems imprecise for two reasons: Firstly, in the theories **IZF** and **CZF**, equality will be a global relation as well (the constructive justification of this, however, is debatable). Secondly, although demanding (ii) is certainly true for **informal constructive** mathematics – so it is for **informal classical** mathematics. For example, in defining the set of real numbers, intuitively, two Cauchy-sequences represent the same real if their difference approaches 0. **Informally**, what one does is to **define equality** on the set of real numbers in this way. Of course, **technically**, one considers equivalence classes (defined according to the "equality" given by intuition) and inherits the global equality from V – the definition of equality is "hidden" in the equivalence classes. We see that (ii) does not seem to be that a unique feature of (informal) constructive set theories after all.

One example of a set is the set $\mathbb{N}$ of natural numbers. Its canonical members are 0 and for each canonical member n, the object $s(n)$ is canonical too. Two canonical members are equal if their "numbers of s's" coincides. How would we interpret expressions like 10^{10} or $3 + 5$? Although 10^{10} is not a canonical

[2] Brouwer would reject this, see [62]

member of $\mathbb{N}$ in the just defined sense, we would like to say that it belongs to the set $\mathbb{N}$ and write $10^{10} \in \mathbb{N}$. Intuitively, the expression 10^{10} stands for a **method computing a canonical member of** $\mathbb{N}$.

We thus say that a *is an element of the set* A and write $a \in A$ if either a is a canonical member of A, according to (i), or a is a method giving instructions for its own evaluation and of which we can prove that it yields some canonical member of A. This distinction between canonical and non-canonical elements is crucial in Martin-Löf's set theory – this is why we included the word "canonical" in (i) and (ii).

As another important example we may consider, given sets A and B with equalities $=_A$ and $=_B$ respectively, the set B^A of functions $f: A \to B$. To give an element f of B^A means to give a way to construct for each $a \in A$ a member $b \in B$ in such a way that equalities are respected, i.e. if $a_1 =_A a_2$, then $f(a_1) =_B f(a_2)$. If this is the case, we write $f(a)$ for the element b as usual. To prove two elements $f, g \in B^A$ equal, one needs to prove that for each $a \in A$, $f(a) =_B g(a)$. Clearly, this equality is an equivalence relation.

2.1.3 BHK-interpretation

In constructive mathematics, truth coincides with the notion of provability. The most important comprehensible semantics in the spirit of constructivism is the "Brouwer-Heyting-Kolmogorov"-semantics of intuitionistic logic (from here on: "BHK", [31], [29], [37], [30], for a discussion see [4] or [64]).

We will here give a variant of BHK-semantics tailored for the language of set-theory. We say that p proves

$a \in A$, if p is a pair $\langle x, q \rangle$, where x is a canonical member of A and q proves that a reduces to x.
$\phi \wedge \psi$, if p is a pair $\langle q, r \rangle$, where q proves ϕ and r proves ψ
$\phi \vee \psi$, if p is a pair $\langle n, q \rangle$, where n is a natural number and q proves ϕ if $n = 0$ and ψ, else.
$\phi \to \psi$, if p is a method transforming proofs of ϕ into proofs of ψ.
$\perp$, never
$\neg\phi$, if p proves $\phi \to \perp$
$\forall x. \phi(x)$, if p is a method giving for each object x a proof $p(x)$ of $\phi(x)$
$\exists x. \phi(x)$, if p is a pair $\langle x, q \rangle$, where x is a (construction of an) object and q proves $\phi(x)$
$\forall x \in A. \phi(x)$, if p is a method transforming a proof of $a \in A$ into a proof of $\phi(a)$
$\exists x \in A. \phi(x)$, if p is a triple $\langle a, q, r \rangle$, where q is a proof of $a \in A$ and r is a proof of $\phi(a)$

Note that the clauses for $\exists x\ \phi(x)$ and especially $\forall x\ \phi(x)$ make only sense, if we have specified a certain domain of discourse. Only then can we talk about "each object" and "being an object". Such a domain could be the collection of all natural numbers or more abstract notions such as the universe of all sets or all constructions of the mind.

In the following examples, we will use the λ-calculus informally to make our argumentation clearer. This choice is not by chance: Howard proved in 1969 that programs of the simply-typed λ-calculus correspond exactly proofs of natural deduction.

Examples:
We can give a proof the logical truth $(\phi \wedge \psi) \rightarrow \phi$ as follows: Suppose, p proves $\phi \wedge \psi$, then p is a pair $\langle q, r \rangle$, where q proves ϕ and r proves ψ. Thus, $\mathrm{pr}_1(p)$ proves ϕ, where pr_1 extracts the first component of a pair. Hence $\lambda x. \mathrm{pr}_1(x)$ proves $(\phi \wedge \psi) \rightarrow \phi$.

We prove $(x \in a \wedge a = b) \rightarrow x \in b$, where $a = b$ means extensional equality, i.e. $\forall x(x \in a \leftrightarrow x \in b)$. Here, as usual, "$\phi \leftrightarrow \psi$" is an abbreviation for $\phi \rightarrow \psi \wedge \psi \rightarrow \phi$. If $p = \langle q, r \rangle$ is a proof of $x \in a \wedge a = b$, i.e. q proves $x \in a$ and r proves $a = b$, then $\mathrm{pr}_1(r)(x)$ proves $x \in a \rightarrow x \in b$. Thus, $\mathrm{pr}_1(r)(x)(q)$ proves $x \in b$. Altogether, $\lambda p. \mathrm{pr}_1\big(\mathrm{pr}_1(p)\big)(x)\big(\mathrm{pr}_2(p)\big)$proves $(x \in a \wedge a = b) \rightarrow x \in b$.

2.2 Brouwer's set theory

Luitzen Brouwer can be seen as the founding father of the intuitionistic school of thought. As the name suggests, in his point, mathematics should be based upon intuition rather than language or any kind of formalism. The existence of mathematical objects is justified by mental constructions. In [*8*], Brouwer describes the first act of intuitionism as

> *Completely separating mathematics from mathematical language and hence from the phenomena of language described by theoretical logic, recognizing that intuitionistic mathematics is an essentially languageless activity of the mind having its origin in the perception of a move of time. This perception of move of time may be described as the falling apart of a life moment into two distinct things, one of which gives way to the other, but is retained by memory. If the twoity thus born is divested of all quality, it passes into the empty form of the common substratum of all twoities. And it is this common substratum, this empty form, which is the basic intuition of mathematics.*

In particular, Brouwer rejects Cantor's set theoretic construction of the continuum. According to his early standpoint [*10*], the continuum is given as an intuitive notion. It is impossible to conceive "all" its points, as intuition allows us to construct only denumerably many elements. Later however, he recognized the need to give a construction of the continuum. In the second act of intuitionism, he accepted two ways of constructing new mathematical objects: Choice sequences – infinite sequences whose elements are created more or less freely from the preceding ones and species – properties of objects previously acquired. Note that in his early writings, he referred to species as "Mengen" (germ. "set"); only later he changed this notion to emphasize the difference to classical sets.

2.2.1 Species and Spreads

A *species* is a (constructive) property P of already constructed objects. As usual, we write $a \in A$ if the object a is an element of the spread A. Note that the clause "of already defined objects" is essential to block paradoxes such as Russel's. We can thus construct species one step at a time, hence the notion of *order of a species* defined as follows is meaningful:

(i) Concrete mathematical objects (natural numbers, sequences of natural numbers, rational numbers, choice sequences etc.) are species of order 0.

(ii) If the already constructed objects have order n, then P applied to them has order $n + 1$.

For example, the spread of all natural numbers, denoted ω, has order 1. A recursive sequence of natural numbers $\left(\alpha(n)\right)_{n\in\omega}$ may be represented as species of order 2, namely the as species of pairs $\left\{\left(n, \alpha(n)\right): n \in \omega\right\}$. Note that many set-theoretic notions can be developed within this framework. It is straightforward to define union, intersection or the empty species. A function between two species A and B can be defined as a species given by a method constructing from each element of A an element of B (although Brouwer reserves these notions for the special case of functions from species generated by a spread). In [7], a theory of constructible cardinal and ordinal numbers is developed. However powerful this may seem, it does not permit a satisfying construction of the continuum.

This is the reason, Brouwer developed the notion of spread. A *spread* $M = (\Lambda, \Gamma)$ is determined by the two laws Λ and Γ:

(A) The spread law Λ decides if a finite sequence of natural numbers is accepted or not under the following restrictions:

  i. It decides which sequences of length one are accepted.
  ii. If $(n_1, n_2, \dots, n_k, n_{k+1})$ is accepted, also $(n_1, n_2, \dots, n_k)$ is accepted.
  iii. If $(n_1, n_2, \dots, n_k)$ is accepted, it decides for each m, if some sequence $(n_1, n_2, \dots, n_k, m)$ is accepted or not
  iv. If $(n_1, n_2, \dots, n_k)$ is accepted, then there is a natural number m, such that the successor sequence $(n_1, n_2, \dots, n_k, m)$ is accepted

(B) The complementary spread law Γ assigns to any Λ-accepted sequence an already constructed mathematical object:

$$\begin{aligned} (n_1) &\mapsto \alpha_1 \\ (n_1, n_2) &\mapsto \alpha_2 \\ &\vdots \\ (n_1, n_2, \dots, n_k) &\mapsto \alpha_k \\ &\vdots \end{aligned}$$

We can think of a spread as tree where each node is labelled by a mathematical object. *M-sequences* are sequences $(n_1, n_2, \dots)$, where each initial segment is accepted. The corresponding sequence $(\alpha_1, \alpha_2, \dots)$ is called a *choice sequence* – Brouwer considers them to be perfectly justified mathematical objects. To each spread M we can thus define the species $[M]$ of choice sequences of M. Note that for a choice sequence β

constructed by another spread it is in general undecidable whether or not $\beta \in [M]$, since we would need to check each initial segment of β.

Examples: We give some important examples of spreads:

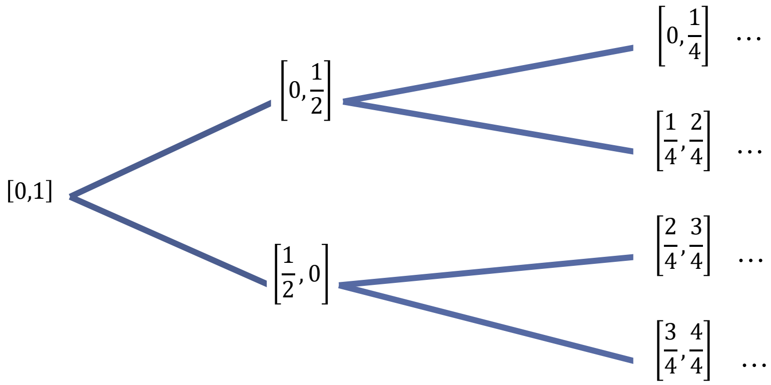
- If the spread law permits no choices of natural numbers, we end up with the empty spread.
- The species of all functions $\omega \to \omega$, denoted ω^ω can be defined via a spread: The spread law permits in each step for every natural numbers to be chosen. The complementary spread law assigns each finite sequence to itself, i.e. $(n_1, n_2, \dots, n_k) \mapsto (n_1, n_2, \dots, n_k)$.
- The binary spread allows in each step only choices of 0 or 1. Again, finite sequences are mapped to themselves.
- Given an enumeration of the rational numbers $q_1, q_2, \dots$ we define the following spread S by giving its spread law: We accept all sequences of length 1. If $(n_1, \dots, n_k)$ is accepted, then $(n_1, \dots, n_k, m)$ is accepted iff
$$|q_{n_k} - q_m| < \frac{1}{2^{k+1}}.$$
Clearly, this law will always permit some successor sequences. The complementary spread law is given by the correspondence $(n_1, \dots, n_k) \mapsto (q_{n_1}, \dots, q_{n_k})$.

We can think of the elements of $[S]$ as Cauchy-sequences of rational numbers. We define equality on $[S]$ as

$$\alpha \approx \beta \Leftrightarrow |\alpha_n - \beta_n| < \frac{1}{2^n} \text{ for all } n \in \omega.$$

Finally, the continuum $\mathbb{R}$ is given as the spread of all equivalence classes of $[S]$ under $\approx$.

Similarly, we assign to finite binary choices of the binary spread instead intervals, as indicated:


$[0,1]$
$\left[0, \frac{1}{2}\right]$
$\left[\frac{1}{2}, 0\right]$
$\left[0, \frac{1}{4}\right]$ …
$\left[\frac{1}{4}, \frac{2}{4}\right]$ …
$\left[\frac{2}{4}, \frac{3}{4}\right]$ …
$\left[\frac{3}{4}, \frac{4}{4}\right]$ …

We can thus define the spread of the real unit interval [0,1]. To conclude, we discuss Brouwer's notion of functions:

2.2.2 Intuitionistic functions

Given a spread M, an *intuitionistic function* ϕ *from* $[M]$ *to* ω, denoted $\phi\colon [M] \to \omega$ is a law ϕ that computes for each M-sequence α a natural number N_α and a natural number $\phi(\alpha)$ based on the initial segment $\left(\alpha_1, \dots, \alpha_{N_\alpha}\right)$.

From this definition we immediately see the following continuity principle: If $\phi\colon [M] \to \omega$ is an intuitionistic function, then it is continuous in the sense that for $\alpha, \beta \in [M]$ there exists some N such that $\phi(\alpha) = \phi(\beta)$, whenever $(\alpha_1, \dots \alpha_N) = (\beta_1, \dots \beta_N)$ – simply put $N = \max\left(N_\alpha, N_\beta\right)$).

Recall the definitions of continuity from analysis:

Definition 2.10: A function $f\colon D \to W$ with $D, W \subseteq \mathbb{R}$ is *continuous* at $x \in D$, if for each $n \in \mathbb{N}$ there is some $m \in \mathbb{N}$ such that for all $y \in D\colon |x - y| < 2^{-m} \to |f(x) - f(y)| < 2^{-n}$.

If we put $[M] = \mathbb{R}$ in the above uniformity principle, we can thus say

Theorem 2.11: Every function $f\colon \mathbb{R} \to \mathbb{N}$ is continuous.

Clearly, this theorem is classically not valid, since the function $f\colon \mathbb{R} \to \mathbb{N}$ with

$$f(x) = \begin{cases} 1, & \text{if } x = 0 \\ 0, & \text{if } x \neq 0 \end{cases}$$

is not continuous. Note however, that intuitionistically speaking, this function is not total, as we cannot in general, decide by a finite procedure whether $x = 0$ or $x \neq 0$. In textbooks, such as [*66*], the definition of an intuitionistic function is formulated as the axiom

$$(\mathbf{WCN}) \qquad \forall \alpha \in \omega^\omega . \exists x \in \omega . A(\alpha, x) \to \forall \alpha \in \omega^\omega . \exists n, b \in \omega . \forall \gamma \in \omega^\omega \left(\bar{\alpha}_n = \bar{\gamma}_n \to A(\gamma, b)\right),$$

where we abbreviate $\bar{\alpha}_n = (\alpha_1, \dots, \alpha_n)$. Although clearly incompatible with classical mathematics, this axiom may be added to any constructive formalism whenever one wishes to do mathematics Brouwer-style ([*59*]).

Unfortunately, in defining functions $[M] \to [N]$ between two spreads, Brouwer did not stick to the internal fashion and defined instead a function $[0,1] \to \mathbb{R}$ to be "a law that, with each of certain point cores of the unit continuum, […] and form the "domain of definition" of the function, associates one point core of the linear continuum" [*11*].

The following approach of defining these functions is more in the spirit of Brouwer's definition of functions $[M] \to \omega$ and can be found in [*50*]: A function $[M] \to [N]$ is a law which corresponds to each N-sequence an M-sequence based on the law Φ^* wich correlates finite sequences of naturals such that

(i) If $n < m$, then $\Phi^*(\alpha_1, \alpha_2, \dots, \alpha_n) \prec \Phi^*(\alpha_1, \alpha_2, \dots, \alpha_m)$, i.e. $\Phi^*(\alpha_1, \alpha_2, \dots, \alpha_n)$ is an initial segment of $\Phi^*(\alpha_1, \alpha_2, \dots, \alpha_m)$.

(ii) Φ^* is not finally constant

(iii) $\Phi(\alpha) = \sup \Phi^*(\bar{\alpha}_n)$, i.e. $\Phi(\alpha)$ is approximated by the segments $\Phi^*(\bar{\alpha}_n)$. We say that Φ^* computes Φ.

We have the following generalized continuity principle:

$$\forall n \in \omega . \exists m \in \omega . \left(\alpha_m = \beta_m \to \left(\Phi(\alpha)\right)_n = \left(\Phi(\beta)\right)_n\right)$$

This gives us the following result:

Theorem 2.12: Every real function $\mathbb{R} \to \mathbb{R}$ is continuous.

In [*11*], Brouwer, using his more general notion of real function, proved only a negative formulation of this result, namely, that a hypothesis of discontinuity of a real function leads to contradiction.

2.3 Martin-Löf's set theory

Axiomatized mathematical theories usually come in two layers: Firstly, a deductive system (some background logic, be it intuitionistic or classical) and secondly, the particular axioms of the respective theory (for example, axioms of set theory). In this sense, there is a divide between the elements of discourse (elements of the second layer) and their interplay (propositions, elements of the first layer). In contrast, in Martin-Löf's set theory (**ML**), both layers come to the same basic notion: sets. In this "logic-free" approach, **ML** aims to better model the constructive activity of a mathematician:

The task of proving a theorem is identified with a special case of the mathematical activity of constructing an object – we speak of this correspondence as *propositions-as-types* (or *Curry-Howard correspondence*). Following this intuition, mathematical objects such as sets, elements and functions are explained in terms of concepts from programming such as data structures, data types and programs. Hence, it will come as no surprise that various automated proof assistants like NuPRL and Coq and programming languages like Agda and Idris have been based on concepts featured in **ML** (see for example [2], [33], [*39*], [*BO*]).

To get used to the ideas involved, let us investigate the division theorem for natural numbers,

$$\forall a, b \in \mathbb{N}\, [b > 0 \to (\exists q, r \in \mathbb{N}\, r < b \wedge a = qb + r)].$$

Following the BHK-interpretation, this proposition corresponds to an algorithm Φ operating on pairs (a, b) of natural numbers and converting proofs of $b > 0$ into triples (q, r, p), where q and r are natural numbers and p is a proof that $r < b$ and $a = qb + r$. In symbols, we could write:

$$\Phi : \Pi\big((a, b) \colon \mathbb{N} \times \mathbb{N}\big)\big[G_0(b) \to (\Sigma(q, r \colon \mathbb{N} \times \mathbb{N})\big(\mathrm{I}(a, qb + r) \times G_0(b - r)\big)\big].$$

Here, "Π" denotes the product indexed by pairs of natural numbers, i.e. the set of functions from $\mathbb{N} \times \mathbb{N}$ into the specified codomain. It gives a natural interpretation for the universal quantifier. $G_0(b)$ is the set of proofs of "$b > 0$"; interpreting implications, "→", as usually denotes function spaces, "Σ" stands for the (generalized) disjoint sum, giving constructive meaning to existential quantification: its index set $\mathbb{N} \times \mathbb{N}$ are witnesses to the formula on the right: $\mathrm{I}(a, qb + r)$ are programs executing both a and $qb + r$ and finding that both computations converge to the same canonical element of $\mathbb{N}$.

Note that in our example, proving the division theorem comes down to constructing a member Φ of a particular set $\Pi\big((a, b){:}\, \mathbb{N} \times \mathbb{N}\big)[\ldots]$. Following the constructivist tradition, we have thus equated (the truth of) the division theorem with its proof. Φ itself becomes a mathematical object and no longer restricted to the metamathematical level. In **ML** we even go a step futher and **identify** the theorem with the set of its proofs. Following Martin-Löf's early nomenclature, we may thus speak of this correspondence as *propositions-as-sets* (instead of propositions-as-types). Implementing this idea, the assertions or *judgements* (we will discuss this important notion in Martin-Löf's theory more rigorously in a moment) "A set" and "$a \in A$" can have the following interpretations from [43]:

***A* set**	***a* ∈ *A***	
A is a set	a is an element of the set A	A is nonempty
A is a proposition	a is a proof (construction) of the proposition A	A is true
A is an intention (expectation)	a is a method of fulfilling (realizing) the intention (expectation) A	A is fulfillable (realizable)
A is a problem (task)	a is a method of solving the problem (doing the task) A	A is solvable

In a nutshell, the theory **ML** postulates the existence of certain sets, like $\mathbb{N}$ (set of natural number) and $\mathbb{N}_k$ (set with k elements) and gives rules for constructing new sets and their canonical elements from given ones (there are rules for "→", "Π", "Σ" and so on).

2.3.1 Proposition vs judgement

From Frege to the Principia the distinction of propositions and judgements has been vital: Although not as central in first-order theories, this distinction is still present: Formulas play the role of propositions – they stand for statements that can be made *within* the theory. Theorems on the other hand, representing judgements, are formulas that have been identified as true via some (possibly) *external* formalism or semantics. The only kind of judgements in these theories is therefore of the form "ϕ is true".

In **ML**, there are four kinds of judgements:

1. A set (A is a well-formed set)
2. $A = B$ (A and B are equal sets)
3. $a \in A$ (a is an element of A)

4. $a = b \in A$ (a and b are equal elements of the set A)

ML follows our intuitive definition quite closely. An element a of a set A is a method (or program) which, when executed, yields a canonical element of the set A as a result. Two arbitrary elements a, b of the set A are equal if, when executed, a and b yield equal canonical elements of the set A as results. Thus, actually, judgement 3. should be read as "a is a program yielding a canonical element of A" and judgement 4. as "a and b are programs yielding equal canonical elements of A).

2.3.2 *The rules of* **ML**

The rules of **ML** postulate the existence of certain sets and describe how new sets may be constructed from existing ones. We consider the following example of the Cartesian product $A \times B$ of two sets A and B.

Example: We define the set $A \times B$ via the following rules:

×-formation

$$\frac{A\ \text{set} \qquad\qquad B\ \text{set}}{A \times B\ \text{set}} \qquad\qquad\qquad \frac{A = C \qquad\qquad B = D}{A \times B = C \times D}$$

The rules of ×-formation explain how to construct the set $A \times B$ and when two such constructions lead to the same result.

×-introduction

$$\frac{a \in A \qquad\qquad b \in B}{(a, b) \in A \times B} \qquad\qquad\qquad \frac{a = c \in A \qquad\qquad b = d \in B}{(a, b) = (c, d) \in A \times B}$$

The introduction rule declares how canonical elements of $A \times B$ look like and when they are equal.

Remark: In formulating this rule, we should have written down assumptions A set and B set – without these assumptions the expression "$A \times B$" does not make any sense. However, assumptions like these are obvious and we will avoid writing down any such trivial assumptions in the future.

×-elimination

$$\frac{c \in A \times B \qquad\qquad \begin{matrix}(x \in A, y \in B)\\ d(x, y) \in C(x, y)\end{matrix}}{\mathrm{E}(c, d) \in C(c)} \qquad\qquad \frac{c_1 = c_2 \in A \times B \qquad\qquad \begin{matrix}(x \in A, y \in B)\\ d_1(x, y) = d_2(x, y) \in C(x, y)\end{matrix}}{\mathrm{E}(c_1, d_1) = \mathrm{E}(c_2, d_2) \in C(c)}$$

Here, the expressions

$$\begin{matrix}(x \in A, y \in B)\\ d(x, y) \in C(x, y)\end{matrix} \quad \text{and} \quad \begin{matrix}(x \in A, y \in B)\\ d_1(x, y) = d_2(x, y) \in C(x, y)\end{matrix}$$

should be read as follows: "Provided $x \in A$ and $y \in B$, the object $d(x, y)$ is an element of $C(x, y)$." and "Provided $x \in A$ and $y \in B$, the objects $d_1(x, y)$ and $d_2(x, y)$ are equal elements of $C(x, y)$.", respectively.

The elimination rules explain the behavior of elements in $A \times B$. It introduces the E-operator which, intuitively, works as follows on an input $c \in A \times B$: It executes the method c to find a canonical element (a, b) of $A \times B$. It then returns $d\big((a, b)\big) \in C(a, b)$.

×-equality

$$\frac{a \in A \qquad\qquad b \in B \qquad\qquad \begin{matrix}(x \in A, y \in B)\\ d(x, y) \in C\big((x, y)\big)\end{matrix}}{\mathrm{E}\big((a, b), d\big) = d(a, b) \in C((a, b))}$$

Finally, the rule of ×-equality declares how the E-operator works on canonical elements.

We can now give our first "proposition-as-set"-interpretation. Namely, a canonical proof of $A \wedge B$ should be a pair (a, b) of proofs a of A and b of B. It is therefore natural to identify the proposition $A \wedge B$ with its set of proofs $A \times B$. We therefore *define* $A \wedge B \equiv A \times B$. When interpreting proposition as sets, we often write A prop instead of A set and A true to indicate that there is some $a \in A$ (some proof of A). We can thus translate the ×-rules into rules familiar form logic:

∧-formation

$$\frac{A \text{ prop} \qquad\qquad B \text{ prop}}{A \wedge B \text{ prop}}$$

∧-introduction

$$\frac{A \text{ true} \qquad\qquad B \text{ true}}{A \wedge B \text{ true}}$$

∧-elimination

$$\frac{A \wedge B \text{ true} \qquad\qquad \begin{matrix}(A \text{ true}, B \text{ true})\\ C \text{ true}\end{matrix}}{C \text{ true}}$$

We can now, for example, derive the logical rule

$$\frac{A \wedge B \text{ true}}{A \text{ true}}$$

if we substitute A for C in the rule of ∧-elimination. We will continue discussing the rules of **ML** in chapter 5.

2.4 **IZF** and **CZF**

The idea behind the theories **IZF** and **CZF** is simple: We start with the Zermelo-Fraenkel set theory with choice (**ZFC**) – the Hilbert-style theory based on classical logic that most today's mathematics is encoded in. Clearly, **ZFC** is constructively not acceptable, so we try to carefully restrict it. Of course, **LEM** will be first to fall victim to these restriction – however, even after restricting the original classical background logic to intuitionistic logic, it will turn out that there are other axioms of **ZFC** that will still get us **LEM** back. We will replace these axioms by (classically) equivalent ones – this way we can always put **LEM** back to restore the initial theory. Having replaced all these axioms, we end up with the Intuitionistic Zermelo-Fraenkel-style set theory **IZF**. However, some remaining axioms of **IZF** may still seem problematic form the constructive point of view. This is especially true for impredicative axioms. We further restrict our theory and end up with the Constructive Zermelo-Fraenkel-style set theory **CZF**. We first start with a brief discussion of **ZFC**.

2.4.1 Zermelo-Fraenkel set theory **ZFC**

The theory **ZFC** is based on a classical Hilbert-style logic with equality, no function symbols and the only relational symbol being "∈". We expect axiom systems like this to be known and focus on the set-theoretic axioms of our theory:

Extensionality

The *axiom of extensionality* reads

$$\forall x, y\, [\forall z(z \in x \leftrightarrow z \in y) \rightarrow x = y].$$

It means that two sets are equal if they contain the same elements. Usually, the background logic contains axioms for equality like $(x \in a \wedge a = b) \rightarrow x \in b$ and $(x \in b \wedge a = b) \rightarrow x \in a$. We could replace these axioms if we instead defined equality via the ∈-relation by

$$x = y \leftrightarrow \forall z(z \in x \leftrightarrow z \in y).$$

This means, that unlike in **ML**, we have a global equality on all sets. Another important application of the axiom of extensionality is that we are able to define notions by the ∈-relation – we will see an example shortly, when discussion the axiom of pair.

It turns out that this seemingly innocent axiom already implies some unwanted instances of **LEM**.

Pair

The *axiom of pair* or *pairing axiom* reads

$$\forall x, y\, \big[\exists z \forall w\big(w \in z \leftrightarrow (w = x \vee w = y)\big)\big].$$

It says that for each two sets x and y there is a set containing exactly x and y as elements. Now, by extensionality such an element does not only exist but is also unique – this justifies writing $\{x, y\}$ for the pair. The singleton $\{x\}$ may be defined as $\{x, x\}$. Finite sets of more than two elements are defined similarly. The *(Kuratowski) ordered pair of* x *and* y is defined as $\langle x, y\rangle = \big\{\{x\}, \{x, y\}\big\}$. A set of ordered pairs R is called a *relation*. We often write aRb instead of $(a, b) \in R$. A relation f is called a *function* iff $x_1 = x_2 \wedge (x_1, y_1) \in f \wedge (x_2, y_2) \in f \rightarrow y_1 = y_2$. We write $f(x)$ for the thus unique y with $(x, y) \in f$.

Union

The *axiom of union* allows us to form unions of sets, i.e.

$$\forall A\, \exists u\, \forall x\, [x \in u \leftrightarrow \exists w(x \in w \land w \in A)].$$

Here, A is a collection of sets u. The union U of all these u should contain all elements x of each u. In the simplest case we are given two sets a and b. By pairing, we may form $\{a, b\}$, now the union of the sets a and b contains exactly the elements of a and b. Again, extensionality justifies writing $\bigcup A$ for the union of A and in our example, $a \cup b$ instead of $\bigcup\{a, b\}$.

Empty set

The *axiom of empty set* postulates the existence of the empty set,

$$\exists x \forall y \neg(y \in x).$$

Extensionality justifies writing $\emptyset$ for the witness of this formula. This axiom is often also formulated as the set existence axiom $\exists x(x = x)$ and the empty set constructed as a corollary.

Infinity

Before formulating the axiom, we introduce the notation $s(x) = x \cup \{x\}$. This is justified by the axioms of union, pair and extensionality (this can be seen by writing $s(x) = \bigcup\big\{x, \{x, x\}\big\}$). Now we can construct the set of natural numbers as follows:

$$\begin{aligned} &0 = \emptyset \\ &1 = s(0) = \emptyset \cup \{\emptyset\} = \{\emptyset\} \\ &2 = s(1) = \big\{\emptyset, \{\emptyset\}\big\} \\ &3 = s(2) = \Big\{\emptyset, \{\emptyset\}, \big\{\emptyset, \{\emptyset\}\big\}\Big\} \\ &\quad \vdots \end{aligned}$$

The *axiom of infinity* postulates that there is a set containing these numbers (and these numbers only):

$$\exists x\big[(\forall n \in x.\, s(n) \in x) \land \forall n \in x.\, \big(n = \emptyset \lor \exists m \in x.\, n = s(m)\big)\big].$$

It is clear that this set must be unique by extensionality. We will refer to it as ω. Note that this axiom does not suffice for internal induction or recursion on the natural numbers.

Separation schema

Given a set a and a formula ϕ such that a is not free in ϕ, we can form the (unique) set of all elements of a that we can prove to possess the property ϕ:

$$\forall a \exists x \forall y\big(y \in x \leftrightarrow y \in a \land \phi(y)\big).$$

Again, by extensionality, we may write $\{x \in a : \phi(x)\}$ for y. As an example, given two sets a and b we define their intersection as $a \cap b = \{x \in a : x \in b\}$. Another example is the definition of range and domain of a function: For a relation R and xRy note that, because of $(x, y) = \big\{\{x\}, \{x, y\}\big\}$, we have $x, y \in \bigcup \bigcup R$. Hence, we can set

$$\mathrm{dom}(R) = \left\{x \in \bigcup\bigcup R : \exists y \in \bigcup\bigcup R\,.\,(x,y) \in R\right\}$$
$$\mathrm{range}(R) = \left\{y \in \bigcup\bigcup R : \exists x \in \bigcup\bigcup R\,.\,(x,y) \in R\right\}.$$

We often write $f: A \to B$ to mean that f is a function with $A = \mathrm{dom}(f)$ and $B = \mathrm{range}(f)$.

Later in this chapter, we will discuss the fact that this axiom is constructively not unproblematic.

Replacement

Given a formula ϕ, such that B is not free in ϕ, we have

$$\forall A\big[\big(\forall x \in A.\, \exists!\, y.\, \phi(x,y)\big) \to \exists B.\, \forall x \in A.\, \exists y \in B.\, \phi(x,y)\big].$$

This axiom means intuitively, that given a function with domain A, we can form its range B. As a simple example, replacement allows us, given any set A, to "replace" each $x \in A$ by $\{x\}$, thus forming the set $A^* = \big\{\{x\}: x \in A\big\}$. Indeed, we can apply replacement since $\forall x \in A\ \exists!\, y\ y = \{x\}$. We obtain some B containing all these ys. Using separation, we can form $\big\{y \in B: \exists x \in A\ y = \{x\}\big\} = A^*$.

We can use replacement to give another definition of domain and range of a relation: Since $\forall z \in R.\, \exists!\, x.\, \big(\exists y.\, z = (x,y)\big)$, we obtain B containing all of these x. By separation, we can form $\mathrm{dom}(R) = \{y \in B: \exists x \in A\ (x,y) \in R\}$ and similarly $\mathrm{range}(R)$.

Powerset

We define the subset-relation as $x \subseteq y \leftrightarrow \forall z(z \in x \to z \in y)$. The *Powerset-axiom* states that, given a set A, there is a set containing all subsets of A:

$$\forall A.\, \exists P.\, \forall x.\, (x \subseteq A \leftrightarrow x \in P).$$

We write $\mathcal{P}(A)$ for P. It will turn out that that this axiom is not at all unproblematic from a constructive viewpoint. For example, for every formula ϕ, we can form by separation the set $\{x \in \{\emptyset\}: \phi\}$. By the Powerset axiom it must be contained in $\mathcal{P}(\{\emptyset\})$. Of course, classically, $\{x \in \{\emptyset\}: \phi\} = \emptyset$ or $\{x \in \{\emptyset\}: \phi\} = \{\emptyset\}$, but this reasoning fails with an intuitionistic background logic. Hence, already the (classically finite and easy) set $\mathcal{P}(1)$ turns out to be equivalent to the set of all formulas and hence to be highly non-trivial.

Foundation

This axiom says that every set must have an $\in$-minial element:

$$\forall A[A \to \exists m(m \in A \wedge m \cap A = \emptyset)].$$

In other words: there are no infinite $\in$-chains. This axiom allows us to prove properties by induction within the theory. This is usually done as follows: Let $\phi(x)$ be any property such that we can prove the induction step $\forall a\left((\forall x \in a.\phi(x)) \to \phi(a)\right)$. Then ϕ must hold for all sets (if it holds for at least one): Because if it does not hold for all sets, then there is (classical reasoning) a least $\in$-least element a not satisfying ϕ. By minimality, all its elements satisfy ϕ and by the inductive step, so must a, contradiction.

These lines already show that we must find a different formulation of this axiom if we want to conduct inductive proofs in systems based on intuitionistic logic. As a simple example of the foundation axiom in action, let us show that no set can contain itself:

ⓒ**Lemma 2.13**: No set can contain itself.

Proof: For each set a, the set $\{a\}$ must have an element disjoint from itself, but the only possible candidate is a, i.e. $a \cap \{a\} = \emptyset$, i.e. $\neg(a \in a)$. ∎

Choice

Let $sing(x)$ mean that x contains only one element, i.e. $\exists y[(x \in y) \wedge \forall z(z \in x \leftrightarrow z = y)]$. The *axiom of choice* (**AC**) reads:

$$\forall S\left[\left((\forall s \in S.s \neq \emptyset) \wedge (\forall s_1, s_2 \in S.s_1 \cap s_2 = \emptyset)\right) \to \exists C.\forall s \in S.sing(C \cap s)\right].$$

It says that for all collection S of nonempty sets s, there is a choice set C containing exactly one element of each s. Alternatively, it is easily seen that the axiom of choice has the equivalent formulation

$$\forall A\left[(\forall a \in A.a \neq \emptyset) \to \exists f: A \to \bigcup A.f(a) \in a\right],$$

saying that for each collection A of nonempty sets a there is a choice function f picking one element of each of the sets a. **AC** is often considered problematic, even from a classical point of view. (We will show that it is constructively inacceptable in the next section). Therefore, one often likes to consider the theory **ZF** = **ZFC** − **AC**. Kurt Gödel and Paul Cohen showed that the axiom of choice is independent from **ZF**, see for example [*40*].

2.4.2 From **ZFC** to **IZF** – weak counterexamples

We start restricting **ZFC** to obtain a theory compatible with intuitionistic logic. As noted before, the first step is to replace our classical background logic with an intuitionistic one. For now, we want to have a (more or less) informal discussion, so we skip writing down all the axioms explicitly. Unfortunately however it does not suffice to simply remove **LEM** from the background logic as we may observe what is called a weak counterexample: We say that a formula ϕ in the language of set theory is a *weak counterexample* iff ϕ implies unwanted instances of **LEM** over some simple set theory. As the background theory we take **BST** – basic set theory – over the constructively unproblematic axioms extensionality, pairing, union, empty set and infinity and a restricted version of the separation schema:

$$\forall a \exists x \forall y \big(y \in x \leftrightarrow y \in a \wedge \phi(y)\big),$$

where a is not free in ϕ and ϕ is a *restricted formula* (or *bounded formula)*, which means that quantifiers appear in their bounded forms only, i.e $\forall x \in a$ or $\exists y \in b$. **BST** will play the role of an easy to see least common divisor of the theories **ZFC**, **IZF** and **CZF**. The unwanted instances of **LEM** are usually full **LEM** or the weaker version **LEM**$'$, which is **LEM** but only for restricted formulas.

Foundation as weak counterexample

Let **LEM**$'$ be the law of excluded middle for bounded formulas. We will show that foundation implies **LEM**$'$. Assuming the full separation axiom, it even implies **LEM**:

Ⓜ**Proposition 2.14**: Let **LEM**$'$ be the law of excluded middle for bounded formulas. We then have

$$\mathbf{BST} \vdash \text{Foundation} \rightarrow \mathbf{LEM}',$$
$$\mathbf{BST} + \text{Separation} \vdash \text{Foundation} \rightarrow \mathbf{LEM}.$$

Proof: Let ϕ be any (bounded) formula and set $S = \{x \in \{0,1\}: (x = 0 \wedge \phi) \vee x = 1\}$. Since $1 \in S$, by Foundation, there is some $\in$-minimal element $s \in S$. If $s = 0$, then ϕ. If $s = 1$, then $\emptyset = 1 \cap S = \{0\} \cap S$, i.e. $0 \notin S$, which means $\neg\phi$. Altogether, $\phi \vee \neg\phi$. ■

We have already remarked that using the Foundation axiom to conduct induction proofs will be problematic in a constructive setting. We will therefore replace it with the more suitable *axiom of set-induction*:

$$\Big(\forall a \big(\forall y \in a. \phi(y)\big) \rightarrow \phi(a)\Big) \rightarrow \forall a. \phi(a).$$

Taking this axiom is also justified by the fact that both axioms are classically equivalent:

Ⓜ**Proposition 2.15**: **BST** + **LEM** $\vdash$ Foundation $\leftrightarrow$ Set-induction

"$\rightarrow$": Let ϕ be any formula and assume $\forall a\big(\forall y \in a. \phi(y)\big) \rightarrow \phi(a)$. Towards a contradiction, suppose, that there is some set A such that $\neg\phi(A)$. We want to assume that this A is $\in$-minimal with this property, i.e. $\forall y \in A. \phi(y)$. This immediately leads to the contradiction $\phi(A)$.

Let us show that this assumption is justified: If A is not $\in$-minimal, then $A' = \{b \in A: \neg\phi(A)\} \neq \emptyset$. By Foundation, let $m \in A'$ be such that $m \cap A' = \emptyset$. Then, for all $y \in m \cap A$: $\phi(y)$. We can thus use $m \cap A$ instead of A.

"$\leftarrow$": Suppose, there is some set x without $\in$-minimal element. Let $\phi(t)$ be the formula "$t \notin x$" and assume, that for all $y \in a$ $\phi(y)$. If $a \in x$, then there is some $y \in a \cap x$, hence $\neg\phi(y)$. This shows that $a \in x$ is impossible and thus $\phi(a)$. By Set-induction, $\forall a\ (a \notin x)$, i.e. $x = \emptyset$. ■

As a simple example of set-induction, let us give an alternative proof in **BST** + Set induction of the fact that no set contains itself (Lemma 2.13):

Lemma 2.16: No set can contain itself.

Proof: Suppose, $\forall y \in a\, \neg(y \in y)$. Assume, $a \in a$, then by hypothesis, $\neg(a = a)$, contradiction. ■

Axiom of choice as weak counterexample

Diaconescu showed that that the axiom of choice provides us with another weak counterexample [*21*]**:**

Ⓜ**Theorem 2.17**: We have:

$$\mathbf{BST} \vdash \mathbf{AC} \to \mathbf{LEM}',$$
$$\mathbf{BST} + \text{Separation} \vdash \mathbf{AC} \to \mathbf{LEM}.$$

Proof: Let ϕ be any (bounded) formula and set $S = \{a, b\}$, where

$$a = \{x \in 2: (x = 0 \wedge \phi) \vee x = 1\} \text{ and } b = \{x \in 2: x = 0 \vee (x = 1 \wedge \phi)\}.$$

Let $f: \{a, b\} \to 2$ be a choice function. Then one of the following cases hold:

- $f(a) = 0$, then $0 \in a$, we have ϕ.
- $f(b) = 1$, then $1 \in b$, we have ϕ.
- $f(a) = 1$ and $f(b) = 0$. Then $f(a) \neq f(b)$, and, as f is a function, $a \neq b$. If ϕ, then $0 \in a$ and $1 \in b$, so $a = 2 = b$, so $\neg\phi$. ■

The theory **IZF**

Although there are other versions of choice (not necessarily equivalent to **AC**) compatible with intuitionistic logic, we stop at this point and define **IZF** to be all axioms of **ZF** but for foundation replaced with set induction and separation replaced with the following stronger axiom, called *the collection schema*:

$$\forall A\big[\big(\forall x \in A.\, \exists y.\, \phi(x, y)\big) \to \exists B.\, \forall x \in A.\, \exists y \in B.\, \phi(x, y)\big].$$

Showing that this axiom does not go beyond **ZF**, requires some deeper results and will be postponed to another chapter. Also, it is not clear at the moment whether or not one can come up with other, more elaborate weak counterexamples, not following directly from one of the axioms – we will answer this question in chapter 6 using semantic tools.

2.4.3 *From* **IZF** *to* **CZF** *– Predicativity*

The notion of predicativity first emerged in the beginning of the 20th century in writings of Poincaré and Russel in an attempt to analyze the newly found paradoxes in Cantor's naïve set theory. Most famously, Russel's paradox [*58*] arises when one tries to form, according to Cantor's notion, the set R of all sets not containing themselves, $R = \{x : x \notin x\}$. We immediately arrive at the absurd equivalence

$$R \in R \Leftrightarrow R \notin R.$$

Other mentionable such paradoxes were Cantor's paradox (in modern terms: the class of cardinals is not a set, [*15*], [*3*] Burali-Forti paradox (the class of ordinals is not a set, [*12*], [*17*]), König's paradox [*38*] or Richard's paradox (Define, by diagonalization, a real number different from all definable real numbers. But this number has just been defined in the last sentence. [*56*] Analyzing these kinds of paradoxes, Poincaré [*51*], [*52*] found that they arise because

(i) an object O is defined by referring to a totality of objects containing O, and
(ii) in each case infinite collections are regarded as "actual" or "completed".

In response to Poincaré's diagnosis, Russel, convinced to hold on to the actual infinite, formulated the vicious circle principle (VCP) in order to pick out definitions as in (ii)

> *"Whatever contains an apparent variable must not be a possible value of that variable" (Russel, in [57])*

Definitions like these are called *predicative*, and *impredicative* if they contain such a "vicious circle". Besides the paradoxes, another impredicative definition is the one of greatest lower bound that we know from analysis: Given a bounded subset S of $\mathbb{R}$, we say that y is its greatest lower bound, $y = \text{glb}(S)$ iff $\text{lb}(y, S) \wedge \forall x\ (\text{lb}(x, S) \rightarrow x \leq y)$, where $\text{lb}(x, S) \equiv \forall s \in S\ x \leq s$ says that x is a lower bound of S. This definition is impredicative because y is defined in reference to the set of lower bounds – a set that y is itself a member of.

Russel's theory of ramified types was an attempt to give a foundation to mathematics respecting the VCP – an attempt that can be regarded today as failed [*49*]. Despite various other contributions to the question, most notably Weyl's *Das Kontinuum* [72], the discussion about predicativity abated with the development of axiomatic set theory as developed by Zermelo, Skolem and Fraenkel, which gave mathematics a solid and coherent foundation.

So why should a constructivist care about predicativity when formulating his own kind of Zermelo-Frankel-style set theory? Myhill writes in [*48*]

> *"… because in order to explain what it is to be an element of a certain set, we have to explain what it is to satisfy the defining condition of that set; that defining condition must only refer to sets which were or might have been defined previously, otherwise (on the constructive view that sets only come into being as we define them, and were not there "all along") a vicious circle might result."*

Thus, constructive set theory can be seen as bottom-up attempt at building sets. Two axioms that seem especially problematic in light of this paradigm are the axioms of Separation and Powerset. For a detailed explanation of predicativity and its history, see [*23*].

Impredicativity of Separation

Let us make precise our concerns with separation: Given a set B and a formula $\psi(x, y)$, we can define according to this axiom the set $C = \{x \in B: \forall y\, \psi(x, y)\}$. Notice, that one of the instances of ψ we have to check for x to be in C is $\psi(x, C)$. To avoid this problem, we restrict this axiom scheme to formulas with bounded quantification only – this is called *the axiom scheme of restriced separation*:

$$\forall a \exists x \forall y \big(y \in x \leftrightarrow y \in a \land \phi(y)\big),$$

for any **bounded** formula ϕ, where x is not free in ϕ. Let us show, that both bounded and unbounded separation are equivalent in the classical context:

Ⓜ**Proposition 2.18** (**BST**): The axiom of separation is equivalent to the scheme $\exists x\ (\phi \leftrightarrow \emptyset \in x)$, where x is not free in ϕ.

Proof: Given separation and a formula ϕ, let $x = \{y \in \{\emptyset\}: \phi\}$, where y is not free in ϕ. Then $\emptyset \in x \leftrightarrow \phi$. For the converse implication, let A be a set and $\phi(y)$ a formula. By assumption, there is a set x_y such that $\phi(y) \leftrightarrow \emptyset \in x_y$ and we may assume that $x \subseteq \{\emptyset\}$. Then x_y is uniquely determined by y; hence, by replacement, there is a function f defined on A with $f(y) = x_y$ and thus $\forall y \in A\ \big(\phi(y) \leftrightarrow \emptyset \in f(y)\big)$. By restricted separation, form $\{y \in A: \emptyset \in f(y)\} = \{y \in A: \phi(y)\}$. ■

Ⓜ**Proposition 2.19**: **BST** + **LEM** ⊢ Separation

Proof: Let ϕ be any formula. Because of **LEM**, we may define $x = \{\emptyset\}$ if ϕ and $x = \emptyset$ if $\neg\phi$. In either case we have $\emptyset \in x \leftrightarrow \phi$. By Proposition 2.18, this shows Separation. ■

Impredicativity of Powerset

Again, our concerns with the powerset axiom are similar to what we had before: Given a set C, form its power set $\mathcal{P}(C)$. According to bounded separation, we may form the set $B = \{n \in \omega: \forall x \in \mathcal{P}(C).\, \phi(n, x)\}$ for any restricted formula ϕ. Note that in this case in order to check whether $n \in B$ we will have to verify $\phi(n, B)$. Myhill writes in [48]:

> *“Power set seems especially nonconstructive and impredicative compared with the other axioms: it does not involve, as the others do, putting together or taking apart sets that one has already constructed but rather selecting out of the totality of all sets those that stand in the relation of inclusion to a given set.”*

Myhill notes that, given the set of natural numbers, it is sufficient for Bishop-style mathematics to have sets of functions from one set to another. This is formulated in the exponentiation axiom: The *axiom of exponentiation* states that for two sets a and b, the set b^a of functions $f: a \to b$ is a set too.

In **CZF** however, we require a generalization of Exponentiation to hold, which itself is a weakening of Powerset. This axiom is called *the subset collection schema*:

$$\forall a, b\, \exists C\, \forall u$$
$$\left[\forall x \in a. \exists y \in b. \phi(x,y,u) \to \exists d \in C\big(\forall x \in a. \exists y \in d. \phi(x,y,u) \wedge \forall y \in d. \exists x \in a. \phi(x,y,u)\big)\right],$$

The justification for taking this axiom instead of Exponentiation is given by a proof-theoretical interpretation of the notion of predicativity (see for example [24], [61], [60]), which we do not have the space to discuss in this thesis. We will make this axiom more comprehensible giving the following definition:

Definition 2.20: For two sets a and b, we call a relation R between a and b *full* iff $\forall x \in a\, \exists y \in b\, xRy$. A set C of subsets of b is a-*full* iff for each full relation R between a and b, there is a $d \in D$ such that the inverse relation R^- is full between d and a, i.e.

$$\forall x \in a. \exists y \in b\, . xRy \to \exists d \in C\ (\forall x \in a. \exists y \in d. xRy \wedge \forall y \in d. \exists x \in a. xRy).$$

The *axiom of fullness* is the assertion that for each pair of sets a and b there is a a-full set C.

Ⓜ**Proposition 2.21**: **BST** ⊢ Subset-collection ↔ Fullness

Proof: "→": Just use $\phi(x,y,R) \equiv xRy$.

"←": Let C be a-full and suppose, that for each u, $\forall x \in a. \exists y \in b. \phi(x,y,u)$. Then

$$xR_u y \leftrightarrow x \in a \wedge y \in b \wedge \phi(x,y,u)$$

is a set by $R_u \subseteq a \times b$ and bounded separation. Moreover, it is full; hence, there is $d \in C$ with

$$\forall x \in a. \exists y \in d. xR_u y \wedge \forall y \in d. \exists x \in a. xR_u y.$$

■

Clearly, the powerset of b is a-full:

Proposition 2.22: BST + Powerset ⊢ Fullness.

The theory **CZF**

All other axioms seem acceptable form the constructive viewpoint. We thus set the axioms of **CZF** to be all axioms of **IZF** but separation replaced by bounded separation, powerset replaced by the subset collection scheme and the collection schema strengthened to the *strong collection schema*

$$\forall A\big[\big(\forall x \in A. \exists y. \phi(x,y)\big) \to$$
$$\exists B\ (\forall x \in A. \exists y \in B. \phi(x,y) \wedge \forall y \in B. \exists x \in A. \phi(x,y)\big],$$

for all formulas ϕ, where B is not free in ϕ. Again, we still need to show that **ZF** proves this axiom. Clearly, constructivist analysis of **CZF** does not end with checking all the axioms. In one of the later chapters we will give a meaning-preserving interpretation of **CZF** into Martin-Löf's theory **ML** which seems to be constructively very well justified.

3 Axiomatic constructive set theories – IZF and CZF

After having motivated **IZF** anf **CZF** in section 2.4, in this chapter, we will discuss these axiomatic set theories in more detail. We start with spelling out the whole axiomatization of the theories in section 0. Before discussing differences between the theories in 3.3 and 3.4, we give some common concepts in section 3.2. Finally, in 3.5 we clear the relationship between **ZFC**, **IZF** and **CZF**.

3.1 Setting the stage

For reasons of clarity, we write down all the axioms and inference rules of logical system **HPL** (Heyting's predicate logic) as well as all the axioms and theories we would like to consider.

3.1.1 Heyting's predicate logic HPL

Axioms

(HPL1) $\phi \to (\psi \to \phi)$

(HPL2) $\big(\phi \to (\psi \to \chi)\big) \to \big((\phi \to \psi) \to (\phi \to \chi)\big)$

(HPL3) $\phi \to \big(\psi \to (\phi \wedge \chi)\big)$

(HPL4) $(\phi \wedge \psi) \to \phi$

(HPL5) $(\phi \wedge \psi) \to \psi$

(HPL6) $\phi \to (\phi \vee \psi)$

(HPL7) $\psi \to (\phi \vee \psi)$

(HPL8) $(\phi \vee \psi) \to \Big((\phi \to \chi) \to \big((\psi \to \chi) \to \chi\big)\Big)$

(HPL9) $(\phi \to \psi) \to \big((\phi \to \neg\psi) \to \neg\phi\big)$

(HPL10) $\phi \to (\neg\phi \to \psi)$

(HPL11) $\forall x\, \phi(x) \to \phi(c)$, where c is free for x in ϕ.

(HPL12) $\phi(c) \to \exists x\, \phi(x)$, where c is free for x in ϕ.

(HPL13) $\forall u \in a.\, \phi(u) \leftrightarrow \forall u[u \in a \to \phi(u)]$

(HPL14) $\exists u \in a.\, \phi(u) \leftrightarrow \exists u[u \in a \wedge \phi(u)]$

Inference rules

(DET) $\frac{\phi,\phi\to\psi}{\psi}$

(UG) $\frac{\phi\to\psi(c)}{\phi\to\forall x\,\psi(x)}$ where c is free for x in ϕ and occurs free in neither ϕ nor ψ.

(EI) $\frac{\phi(c)\to\psi}{\exists x\,\phi(x)\to\psi}$ where c is free for x in ϕ and occurs free in neither ϕ nor ψ.

Axioms of identity

(ID1) $x = x$

(ID2) $x = y \to y = x$

(ID3) $(x = y \wedge y = z) \to x = z$

(ID4) $(x = y \wedge y \in z) \to x \in z$

(ID5) $(x = y \wedge z \in x) \to z \in y$

3.1.2 Set axioms

(1) **Extensionality**: $\forall x, y\,[\forall z(z\in x\leftrightarrow z\in y)\to x=y]$,

(2) **Pair**: $\forall x, y\,\exists z\,[\forall w.\, w\in z\leftrightarrow(w=x\vee w=y)]$,
We write $\{x,y\}$ for the pair obtained from x and y. This is well-defined by extensionality.

(3) **Union**: $\forall A\,\exists u\,\forall x\,[x\in u\leftrightarrow\exists w(x\in w\wedge w\in A)]$,
We write $\bigcup A$ for the union of A and $x\cup y$ for $\bigcup\{x,y\}$. Both definitions are justified by extensionality.

(4) **Empty set**: $\exists x\forall y\,\neg(y\in x)$,
We denote the (unique) witness of this formula by $\emptyset$.

(5) **Infinity**: $\exists\omega\left[\left(\forall n\in\omega\,(s(n)\in\omega)\right)\wedge\left(\forall n\in\omega\,\left(n=\emptyset\vee\exists m\in\omega\,(n=s(m))\right)\right)\right]$,
Here $s(u)\coloneqq u\cup\{u\}$. This exists by union and pair and is well-defined by extensionality. We will refer to the witness of this formula by ω (again, this is justified by extensionality).

(6) **Separation schema**: $\forall a\exists x\forall y\big(y\in x\leftrightarrow y\in a\wedge\phi(y)\big)$,
for any formula ϕ, where x is not free in ϕ. We write $\{y\in a\colon\phi(x)\}$ for x. This is well-defined by extensionality.

(6′′) **Bounded separation schema**: $\forall a\exists x\forall y\big(y\in x\leftrightarrow y\in a\wedge\phi(y)\big)$,
for any **bounded** formula ϕ, where x is not free in ϕ. This means all quantifiers are bounded, i.e. of the form $\forall x\in A$ or $\exists x\in B$.

(7) **Replacement schema**: $\forall A\big[\big(\forall x\in A\,\exists!\,y\,\phi(x,y)\big)\to\exists B\,\forall x\in A\,\exists y\in B\,\phi(x,y)\big]$,
for all formulas ϕ, where B is not free in ϕ.

(7′) **Collection schema**: $\forall A\big[\big(\forall x\in A\,\exists y\,\phi(x,y)\big)\to\exists B\,\forall x\in A\,\exists y\in B\,\phi(x,y)\big]$,
for all formulas ϕ, where B is not free in ϕ.

(7′′) **Strong collection schema**: $\forall A[\big(\forall x\in A\,\exists y\,\phi(x,y)\big)\to$
$$\exists B\,(\forall x\in A\,\exists y\in B\,\phi(x,y)\wedge\forall y\in B\,\exists x\in A\,\phi(x,y)]$$
for all formulas ϕ, where B is not free in ϕ.

(8) **Powerset**: $\forall A\,\exists P\,\forall x\,(x\subseteq A\leftrightarrow x\in P)$,
where we define the relation $x\subseteq y\leftrightarrow\forall z(z\in x\to z\in y)$. We write $\mathcal{P}(A)$ for the powerset-operation.

(8′′) **Subset collection**: $\forall a, b\,\exists C\,\forall u$
$$\big[\forall x\in a.\,\exists y\in b.\,\phi(x,y,u)\to\exists d\in C\big(\forall x\in a.\,\exists y\in d.\,\phi(x,y,u)\wedge\forall y\in d.\,\exists x\in a.\,\phi(x,y,u)\big)\big],$$

(9) **Foundation**: $\forall A[A\neq\emptyset\to\exists m(m\in A\wedge m\cap A=\emptyset)]$,
where $x\cap y\coloneqq\{z\in x\cup y\colon z\in x\wedge z\in y\}$. This set is formed by (bounded) separation.

(9′) **Set induction**: $\Big(\forall a\big(\forall y\in a.\,\phi(y)\big)\to\phi(a)\Big)\to\forall a\,\phi(a)$,
where ϕ is any formula.

Definition 3.1: Let **LEM** (*law of excluded middle*) be the axiom schema $\phi\vee\neg\phi$. We define the theories **ZF** (*Zermelo-Fraenkel set theory*), **IZF** (*intuitionistic Zermelo-Fraenkel set theory*), **CZF** (*constructive Zermelo-Fraenkel set theory*) and **BST** (*basic set theory*) by

$$\begin{aligned}
\mathbf{ZF} &= (1)\text{-}(5) + (6) + (7) + (8) + (9) + \mathbf{LEM}\\
\mathbf{IZF} &= (1)\text{-}(5) + (6) + (7') + (8) + (9')\\
\mathbf{CZF} &= (1)\text{-}(5) + (6'') + (7'') + (8'') + (9')\\
\mathbf{BST} &= (1)\text{-}(5) + (6'') + (7)
\end{aligned}$$

It is easily seen that **BST** is the weakest of the four theories.

3.2 Basic concepts in BST

Definition 3.2: The following definition are justified within **BST** and are thus meaningful for all the set theories just defined:

- Pairs, unions and the empty set have already been defined above.
- We define the *ordered pair of* x *and* y by $(x,y) = \langle x,y\rangle = \{\{x\},\{x,y\}\}$.
- A set R is called *relation* iff is a set of ordered pairs. Instead of $(x,y) \in R$ we usually write xRy.
- For any set R, define $\mathrm{dom}(R) = \{x : \exists y\, [(x,y) \in R]\}$ and $\mathrm{range}(R) = \{y : \exists x[(x,y) \in R]\}$: Note that if $(x,y) = \{x,\{x,y\}\} \in R$, then $\{x\},\{x,y\} \in \bigcup R$ and $x,y \in \bigcup\bigcup R$. So, we justify $\mathrm{dom}(R)$ and $\mathrm{range}(R)$ using Union and Bounded Separation.
- We say that a relation R is a *relation between* A *and* B iff $\mathrm{dom}(R) \subseteq A$ and $\mathrm{range}(R) \subseteq B$.
- We say that a relation f is a *function* iff $(x,y_1) \in f \wedge (x,y_2) \in f \rightarrow y_1 = y_2$. We usually write $f(x) = y$ instead of $(x,y) \in f$. We write $f : A \rightarrow B$ and say that f is a function from A to B iff $\mathrm{dom}(f) = A$ and $\mathrm{range}(f) \subseteq B$.
- Given sets R and d, we denote by $R|_d$ *the restriction of* R *to* d by $R|_d = \{(x,y) \in R : x \in d\}$. Note that $\mathrm{dom}(R|_d) = \mathrm{dom}(R) \cap d$. We will use this notation mainly in the case when R is a function.
- Given two sets A and B, we define its *Cartesian product* $A \times B = \{(a,b) : a \in A, b \in B\}$: First use Replacement and bounded Separation to show $\{a\} \times B$ is a set for each $a \in A$. Use them again, to show $S = \{A \times \{b\} : b \in B\}$ is a set and finally $A \times B = \bigcup S$ by the Union axiom.

3.2.1 Classes

Cantor defined a set as "a gathering together into a whole of definite, distinct objects of our perception or of our thought – which are called elements of the set.", [14]. The modern pendent of this notion in axiomatic set theory is the notion of a class: A *class* A is given by a formula $\phi_A(x)$ and can be thought of as containing all sets satisfying $\phi_A(x)$, i.e. $A = \{x : \phi_A(x)\}$.

Before we go on and convey the expression that we are dealing with some type of higher order logic, let us be pedantic for a moment and insist on the fact that a class is actually a strictly metamathematical notion, as each class is given by a formula and in fact it **is nothing more** than a formula. For classes A and B, we often use terms like "$\forall x \in A\ \psi(x)$", "$\exists x \in A\ \psi(x)$", "$A \subseteq B$", "$A = B$" mimicking the language of set theory. However, what we mean by these expressions is $\forall x\big(\phi_A(x) \rightarrow \psi(x)\big)$, $\exists x\big(\phi_A(x) \wedge \psi(x)\big)$, $\forall x\big(\phi_A(x) \rightarrow \phi_B(x)\big)$ and $\forall x\big(\phi_A(x) \leftrightarrow \phi_B(x)\big)$ resp.

Sometimes we like to think about whether a given "class is a set" or not. How can that be after we insisted on classes existing strictly on the metalevel? The statement "A is a set" can be thought of as a metalinguistic abbreviation for $\exists x\, \forall y(\phi_A(y) \leftrightarrow y \in x)$. Classes that cannot are not sets are called *proper*. For example, for a set a, we can define the formula $\phi(y) \equiv y \in a$. Then the class $\mathcal{P}(a) = \{y : y \subseteq a\}$ is a set in **ZF** and **IZF** by the powerset-axiom, but the question whether or not this holds true in **CZF** is a nontrivial question (it turns out that such “powerclasses” $\mathcal{P}(a)$ are never sets in **CZF**).

Note that paradoxes such as Russel’s cannot arise, when talking about classes, since there is a strict hierarchy between sets and classes: Proper classes contain only sets, never classes, hence we cannot form objects like the class of all classes.

3.2.2 Ordinals

The notion of ordinals is central in axiomatic set theory. It generalizes the notion of natural number and allows us to talk about ordering of sets in the post-countable setting.

Definition 3.3: A set T is called *transitive* iff every element is a subset of T, i.e. $a \in T \to a \subseteq T$. A transitive set of transitive sets is called an *ordinal*. The class of all ordinals is referred to as $\mathbb{ON}$. As usual, we call $(A, \prec)$ a *well-order* iff it is well-founded, i.e. each subset has a $\prec$-minimal element and $\prec$ is irreflexive, transitive and satisfies the trichotomy law $\forall a, b \in A\ a \prec b \vee a = b \vee b \prec a$. We call two well-ordered sets $(A, \prec_A)$ and $(A, \prec_B)$ isomorphic iff there is a bijective function $f\colon A \to B$ such that $a_1 \prec_A a_2 \Leftrightarrow f(a_1) \prec_B f(a_2)$.

As noted, there are many similarities between ordinals and natural numbers. We will prove some basic results that resemble properties of ω. For an ordinal α, we define $\alpha + 1 \coloneqq s(\alpha) = \alpha \cup \{\alpha\}$.

Proposition 3.4 (BST):

1) All elements of an ordinal are ordinals.
2) $\alpha + 1$ is an ordinal for each $\alpha \in \mathbb{ON}$.
3) $\bigcup U \in \mathbb{ON}$, for each set of ordinals U.
4) $\alpha \cap \beta$ is an ordinal for $\alpha, \beta \in \mathbb{ON}$
5) ω is an ordinal.
6) All the natural numbers are ordinals.
7) $\mathbb{ON}$ is not a set.

Proof:

1) Let $\alpha \in \mathbb{ON}$ and $\beta \in \alpha$. By definition, β is a subset of α, hence each element of β must be transitive.
2) This trivial.
3) Let $x \in \bigcup U$. This means, there is some ordinal $\alpha \in U$ such that $x \in \alpha$. But α is transitive, hence $x \subseteq \alpha \subseteq \bigcup U$. This shows that $\bigcup U$ is transitive. But because of 1), and the fact that all its elements are elements of ordinals, all elements of $\bigcup U$ are transitive too.
4) This is clear.
5) Let $n \in \omega$. Then $n \in n + 1 \subseteq \omega$.
6) This follows by 1).
7) Suppose, $\exists x\ (x = \mathbb{ON})$. By 3), $\alpha = \bigcup x \in \mathbb{ON}$. But then $\alpha \in \alpha$ is a contradiction to Lemma 2.16. ∎

3.2.3 Set recursion and the natural numbers

In this section we justify the fact that ordinals can be seen as generalization of the natural numbers. A set I is called *inductive* iff $\emptyset \in I$ and $x \in I \to s(x) \in I$. Usually, in **ZF** and **IZF**, the infinity axiom is formulated without the second conjunct, thus stating "there is an inductive set". One then usually proceeds to define ω as the least inductive set. However, this requires unbounded Separation or the Powerset axiom which is problematic in **CZF**. Thus, our formulation of the infinity axiom states "there is a least inductive set":

Proposition 3.5: **BST** + Set induction $\vdash I$ inductive $\to \omega \subseteq I$.

Proof: We show by $\in$-Induction on the formula $\phi(x) \equiv x \in \omega \to x \in I$:

Suppose, $\forall y \in a. \phi(y)$ and $a \in \omega$. If $a = \emptyset$ and we are done. If there is some $n \in \omega$ with $a = s(n)$, then by inductive hypothesis, $a \in I$. By inductiveness of I, $a = s(m) \in I$. ■

We therefore obtain our usual principle of natural induction: Let ϕ be any property and suppose, we can show $\forall n \in \omega. \phi(n) \to \phi(s(n))$ and $\phi(0)$. In other words, $I = \{n \in \omega : \phi(n)\}$ is an inductive subset of ω. By the Proposition it must be equal to the whole of ω. Thus,

Ⓜ**Corollary 3.6** (**BST** + Set induction, natural induction schema): For any formula ϕ, we have the rule

$$\frac{\phi(0) \qquad \forall n \in \omega. \phi(n) \to \phi(s(n))}{\forall n \in \omega\ \phi(n).}$$

What makes the natural numbers different to the ordinals in the intuitionistic setting is that the ordering on ω is trichotomous:

Lemma 3.7 (**BST** + Set induction): $\forall n \in \omega\ \forall m \in n\ (s(m) \in n \lor s(m) = n)$.

Proof: We use natural induction on n. Let $m \in n$, then by inductive hypothesis one of the two holds:

- $s(m) \in n$, then $s(m) \in n \cup \{n\} = s(n)$.
- $s(m) = n$, then $s(m) = n \in n \cup \{n\} = s(n)$. ■

Let us adopt the more usual notations $n = m + 1$ for $n = s(m)$ and $m < n$ for $m \in n$. Also, we often write "$\forall m < n$" to signify "$\forall m \in n$". Using the lemma, we can show the trichotomy law for the relation "$<$":

Proposition 3.8 (**BST** + Set induction): $\forall m, n \in \omega : n < m \lor n = m \lor m < n$.

Proof: By natural induction on n. Let $m \in \omega$. By inductive hypothesis, we have one of the following cases:

- $n < m$, then by the last lemma,
  - $n + 1 < m$, or
  - $n + 1 = m$,
- $n = m$, then $m < n + 1$,
- $m < n$, then $m < n + 1$.

Altogether, $n+1<m$ or $n+1=m$ or $m<n+1$. ■

Corollary 3.9 (**BST** + Set induction): Equality on ω is decidable.

We will often write 0,1,2, ... instead of $0, s(0), s(s(0)), \ldots$. We will refer to a set S as *(Kuratowski) finite* iff there is a numeration of the elements of $s_0, s_1, \ldots, s_{n-1}$ of the elements of S, or to be more precise: A set is finite iff there is a natural number $n \in \omega$ and a bijective function $f\colon n \to S$. In **BST**, we can define the set of finite sequences of natural numbers, using the axioms of infinity, pairing, separation, collection and union:

Definition 3.10: We define for each $n \in \omega$, the set of *sequences of natural numbers of length* n as $\omega^n = \{((0, a_0), \ldots (n-1, a_{n-1})) \colon a_i \in \omega\}$. We define the set of *finite sequences of natural numbers* as $\omega^{<n} = \bigcup_{n\in\omega} \omega^n$.

3.3 ZF vs IZF or the problem with trichotomy

In this section we will briefly discuss some similarities and differences of **ZF** and **IZF** by the example of ordinals. Many of the important applications of ordinals in **ZF**, like induction and recursion pertain to work in **IZF**, but (the class of) ordinals fail to be well-orders in **IZF**.

3.3.1 Applications of Foundation – Ordinals in ZF

The axiom of Foundation says that every set has an $\in$-minimal element. Together with **LEM** this will make sure the class $\mathbb{ON}$ is well-ordered by $\in$. Furthermore, the ordinals in **ZF** form paradigms for well-ordered sets – a role that will be lost in **IZF**.

ⓒ**Lemma 3.11** (**BST** + **LEM** + Foundation): $\mathbb{ON}$ is well-founded, i.e. each nonempty subclass $C \subseteq \mathbb{ON}$ has an $\in$-minimal element.

Proof: Let $\alpha \in C$. Either α is already $\in$-minimal, or $\alpha \cap C \neq \emptyset$. By Foundation, there is an $\in$-minimal ordinal $\beta \in \alpha \cap C$, which must be $\in$-minimal in C too. ■

ⓒ**Proposition 3.12** (**BST** + **LEM** + Foundation): $(\mathbb{ON}, \in)$ is a well-order and so is $(\alpha, \in)$ for each ordinal α.

Proof: Irreflexivity and transitivity also hold in the intuitionistic setting and will be shown in Lemma 3.16, well-foundedness is exactly the Foundation axiom, so it remains to show the trichotomy law: We say that two ordinals η and γ are incomparable iff they are a counterexample to the trichotomy law, i.e. $\neg(\eta \in \gamma) \wedge \eta \neq \gamma \wedge \neg(\gamma \in \eta)$. Towards a contradiction, suppose that there is such an ordinal α incomparable with at least one other ordinal. By Lemma 3.11, we may assume that α is $\in$-least with this property and that β is $\in$-least among the ordinals α is incomparable with.

All $\gamma \in \beta$ must by minimality of β be comparable to α, thus $\gamma \in \alpha$, $\gamma = \alpha$ or $\alpha \in \gamma$. In the last two cases, we would have $\alpha \in \beta$, which is impossible by incomparability. Thus, $\beta \subseteq \alpha$.

The same argumentation applies to: $\gamma \in \alpha$. By minimality, it must be comparable to β. In the cases $\beta \in \gamma$ and $\beta = \gamma$, we would have the impossible $\beta \in \alpha$. Thus, $\gamma \in \beta$, i.e. $\alpha \subseteq \beta$. Altogether $\alpha = \beta$, contradiction. ■

ⓒ**Corollary 3.13** (**BST** + **LEM** + Foundation): $\in$-least elements are unique in $\mathbb{ON}$ (and thus in each ordinal α).

Lemma 3.14: Let $f: (\alpha, \in) \to (\beta, \in)$ be an isomorphism. Then f is the identity mapping.

Proof: We first show that for all $\gamma \in \alpha$, $f(\gamma)$ is an ordinal. It is clear, that $f(\gamma)$ consists of transitive elements, so it remains to show that it is transitive itself: Let $\delta \in f(\gamma)$ and let $\epsilon \in \delta$. Since f is an isomorphism, we may assume that $\delta = f(\rho)$, $\epsilon = f(\eta)$ and $\eta \in \rho \in \alpha$. By transitivity, $\eta \in \alpha$ and hence $\epsilon = f(\eta) \in f(\alpha)$.

Now, suppose, that f is not the identity and let $\gamma \in \alpha$ be the least ordinal such that $f(\gamma) \neq \gamma$. By the isomorphism property and our assumption on γ,

$$\delta \in \gamma \leftrightarrow f(\delta) \in f(\gamma) \leftrightarrow \delta \in f(\gamma).$$

We know from above, that $f(\gamma)$ consists of ordinals only, hence $f(\gamma)$ must be equal to γ after all. ■

We can now prove the following fundamental theorem of **ZF** (details can be found in [*40*]):

Theorem 3.15: Every well-ordered set $(A, \prec)$ is isomorphic to a unique ordinal α.

Proof: Uniqueness follows from Lemma 3.14. We denote by $a\downarrow = \{x \in A: x \prec a\}$ the initial segment of A given by a Note that $a\downarrow$ is well-ordered by $\prec$ as well. Let G be the set of elements $a \in A$ such that $(a\downarrow, \prec)$ is isomorphic to a (unique) ordinal ξ_a and form the function $f(a) = \xi_a$ with domain G by replacement. Note that f is in fact an isomorphism between $(G, \prec)$ and $\operatorname{ran}(f)$ and $\operatorname{ran}(f) \in \mathbb{ON}$.

If $G = A$, then $(A, \prec)$ is isomorphic to $\bigcup_{a \in A} \xi_a$. If $G \neq A$, let $a \in A$ be the least element not in G. Since $a\downarrow = G$, we have that $(a\downarrow, \prec)$ is isomorphic to $\operatorname{ran}(f) \in \mathbb{ON}$, a contradiction. ■

3.3.2 Ordinals in IZF

Lemma 3.16 (**BST** + Set Induction): $(\mathbb{ON}, \in)$ is irreflexive and transitive and so is $(\alpha, \in)$ for each ordinal α.

Proof: Irreflexivity is Lemma 2.16 and transitivity is guaranteed by definition. ■

The problem with well-ordering ordinals is the trichotomy on $\mathbb{ON}$, since it would provide us with a weak counterexample:

Ⓜ**Proposition 3.17**: Trichotomy implies forms of **LEM**:

$$\textbf{BST} + \text{Separation} \vdash [\forall \alpha, \beta \in \mathbb{ON}\ (\alpha \in \beta \lor \alpha = \beta \lor \beta \in \alpha)] \to \textbf{LEM}.$$
$$\textbf{BST} \vdash [\forall \alpha, \beta \in \mathbb{ON}\ (\alpha \in \beta \lor \alpha = \beta \lor \beta \in \alpha)] \to \textbf{LEM}'.$$

Proof: Let ϕ be any (bounded) formula and let $\beta = \{x \in 1: \phi\} \subseteq 1$. Clearly, β is an ordinal; hence, by trichotomy we have $1 \in \beta \vee 1 = \beta \vee \beta \in 1$. The first case is impossible, for $1 \in \beta \subseteq 1$ would contradict Lemma 2.16. In the second case we would have ϕ and in the third $\neg\phi$. ■

3.3.3 Recursive definitions on the ordinals

We discuss the important applications of defining (class-)functions and classes via recursion on the ordinals. A similar proof can be found in [40].

Ⓜ**Theorem 3.18** (**BST** + Replacement + Set Induction): Assume, $\forall \alpha \in \mathbb{ON}\, \forall s\, \exists!\, y\, \phi(\alpha, s, y)$ and define $G(\alpha, s)$ to be the unique y such that $\phi(\alpha, s, y)$. Then there is a formula ψ such that

1. $\forall \alpha\, \exists!\, y\, \psi(x, y)$, so ψ defines a function F, where $F(\alpha)$ is the unique y with $\psi(\alpha, y)$.
2. $\forall \alpha \in \mathbb{ON}\, F(\alpha) = G(\alpha, F|_\alpha)$ and F is uniquely by this property.

Proof: First, let us give such a ψ. Let us abbreviate

$$\mathrm{App}(d, h) \equiv h \text{ is a function} \wedge \mathrm{dom}(h) = d \subseteq \mathbb{ON} \wedge \forall \alpha \in d\; h(\alpha) = G(\alpha, h|_\alpha).$$

The functions h can be thought as local approximations to F. Our candidate for ψ is therefore

$$\psi(\alpha, y) \equiv \big[\alpha \in \mathbb{ON} \wedge \exists d, h\, [\mathrm{App}(d, h) \wedge \alpha \in d \wedge h(\alpha) = y]\big].$$

To verify that this works we have to show that the local approximations h agree on their common domains and that for each α, there is some h defined taking a value on α. For the first part, we need to show

$$\mathrm{App}(d, h) \wedge \mathrm{App}(d', h') \to \mathrm{App}(d \cap d', h \cap h').$$

We show this by ordinal induction that $h(\alpha) = h'(\alpha)$ for all $\alpha \in d \cap d'$. Suppose, we have $h(\beta) = h'(\beta)$ for all $\beta \in d \cap d' \cap \alpha$. Then $h(\alpha) = G(\alpha, h|_\alpha) = G(\alpha, h'|_\alpha) = h'(\alpha)$. So $h \cap h'$ indeed forms a function with domain $d \cap d'$ and $\mathrm{App}(d \cap d', h \cap h')$.

For the second part, let $\alpha \in \mathbb{ON}$. We need to define a function h_α defined on d_α with $\alpha \in d_\alpha$ and $\mathrm{App}(d_\alpha, h_\alpha)$. Again, we do this by Set induction: Suppose, we have defined such function h_β with $d_\beta = \beta + 1$ for all $\beta \in \alpha$ Fixing the domain lets us assume that these functions are unique by what we have shown so far. So, by Replacement and Union, we may define $\tilde{h} = \bigcup_{\beta \in \alpha} h_\beta$ and $\tilde{d} = \alpha$. Clearly, $\mathrm{App}(\tilde{h}, \tilde{d})$. Finally, we can set $h_\alpha = \tilde{h} \cup \left\{G\left(\alpha, \tilde{h}\big|_\alpha\right)\right\}$ and $d_\alpha = \alpha + 1$ and conclude $\mathrm{App}(h_\alpha, d_\alpha)$.

For the uniqueness of F, suppose, we are given two such functions F and F' and say they agree on all $\beta \in \alpha$. But then $F(\alpha) = G(\alpha, F|_\alpha) = G(\alpha, F'|_\alpha) = F'(\alpha)$. We may thus infer by ordinal induction that F and F' are identical. ■

Example (**IZF** or **ZF**): As an important example we can define the *von Neuman universe* V, also called *the universe of all sets*:

$$V_\alpha = \bigcup_{\beta\in\alpha} \mathcal{P}(V_\beta),$$
$$V = \bigcup_{\alpha\in\mathbb{ON}} V_\alpha.$$

This definition is justified by Theorem 3.18 as follows: Let $\phi(\alpha, s, y)$ be the formula $y = \bigcup_{y\in\mathrm{range}(s)} \mathcal{P}(y)$. We can now set $V_\alpha = F(\alpha)$ and observe

$$V_\alpha = F(\alpha) = G(\alpha, F|_\alpha) = \bigcup_{y\in\mathrm{range}(F|_\alpha)} \mathcal{P}(y) = \bigcup_{\beta\in\alpha} \mathcal{P}(F(\beta)) = \bigcup_{\beta\in\alpha} \mathcal{P}(V_\beta).$$

Let us consider the first few stages of the V_α:

$$V_0 = \emptyset$$
$$V_1 = \{\emptyset, \{\emptyset\}\}$$
$$V_2 = \Big\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\Big\}$$

This simple construction gives us a “bottom up” construction. V is then the class of all sets that can be obtained is this way starting from the empty set. It turns out that this class is the class of all sets justifying the second name of V:

Theorem 3.19 (**IZF** or **ZF**): $\forall x. x \in V$, or written out, $\forall x. \exists\alpha \in \mathbb{ON}. x \in V_\alpha$.

Proof: By set induction. Suppose, $\forall y \in x. \exists\alpha_y \in \mathbb{ON}. x \in V_{\alpha_y}$. In the case of **IZF**, form a set Z of such α_y by Collection. In the case of **ZF**, we may assume that the α_y are $\in$-minimal to satisfy the uniqueness requirement of the Replacement axiom. We now have $x \subseteq V_{\bigcup Z}$ and so $x \in V_{\bigcup Z+1}$. ∎

Usually, when giving inductive definitions on the ordinals in **ZF**, the definition is divided into the cases $\alpha = 0$, α is a successor ordinal (i.e. of the form $\alpha = \beta + 1$) and α a limit ordinal (α cannot be written in the form $\alpha = \beta + 1$). The definition of V would thus look like $V_0 = \emptyset$, $V_{\beta+1} = \mathcal{P}(V_\beta)$, $V_\lambda = \bigcup_{\alpha\in\lambda} V_\alpha$ and $V = \bigcup_{\alpha\in\mathbb{ON}} V_\alpha$. Although more readable, this definition like this would make sense in the intuitionistic setting as the distinction between successor and limit ordinals requires **LEM**.

3.4 IZF vs CZF or how to live without the powerset operation

In section 2.4.3 we have discussed that the axiom of Power set may be criticized for being impredicative and non-constructive in the sense, that it does not describe any procedure to construct all subsets of a given set. Myhill notes [*48*], that the weaker Exponentiation axiom, thus replacing Powerset in the formulation of **CZF**, is enough to do mathematics Bishop-style. The argument, that the Powerset axiom is necessary for a solid foundation of mathematics may therefore be rebutted by thorough investigation of Bishop’s book.

One specific application of the Powerset to mathematics in general and set theory in particular is that the class V of all sets forms a hierarchy of sets – definitions and proofs of properties on that and similar structures may thus be conducted in a recursive fashion. It turns out that – apart from the stages V_α of the hierarchy being sets – most of the recursive character of V may be preserved when passing from **IZF** to **CZF**.

3.4.1 Powerset, Subset Collection, Exponentiation

We start with formally showing that the Powerset-axiom is indeed stronger than Subset collection, which in turn is stronger than Exponentiation. Remember, that we have shown (Proposition 2.21) that the axioms of Subset collection and Fullness are in fact equivalent. For the sake of convenience, we therefore use the axiom of Fullness in the following discussion.

We start with showing that **LEM′** may be reduced to the statement that all sets either contain $\emptyset$ or do not contain $\emptyset$:

Ⓜ**Proposition 3.20**: **BST** ⊢ **LEM′** $\leftrightarrow \forall x(\emptyset \in x \vee \emptyset \notin x)$.

Proof: The direction from left to right is clear. For the other direction, let ϕ be any restricted formula. We define $x = \{y \in \{\emptyset\}: \phi\}$ with y not free in ϕ. If $\emptyset \in x$, then ϕ, if $\emptyset \notin x$, then $\neg\phi$. ■

Proposition 3.21: **BST** + Fullness ⊢ Exponentiation

Proof: Let C be an a-full set of subsets of $a \times b$. For $f: a \to b$ define $f': a \to a \times b$ by $f'(x) = \big(x, f(x)\big)$ for $x \in a$. Then f' is a full relation between a and $a \times b$, hence there is some $d \in C$ according to the fullness axiom. But then $d = \{f'(x): x \in a\} = \big\{\big(x, f(x)\big): x \in a\big\} = f$. Thus, b^a can be identified as a subset of C using restricted separation. ■

It turns out, that the seemingly weak assumption of $\mathcal{P}(\{\emptyset\})$ to be a set is enough to to deduce the full Powerset axiom from the Exponentiation axiom. By "$\mathcal{P}(a)$ is a set" we mean the formula $\exists z. z = \mathcal{P}(a)$, which of course stands for $\exists z. \forall b. (b \subseteq a \leftrightarrow b \in z)$.

Proposition 3.22: **BST** ⊢ Powerset ↔ Exponentiation + "$\mathcal{P}(\{\emptyset\})$ is a set".

Proof: One direction is clear. For the other, let A be any set and define by exponentiation and replacement the set $C = \big\{\{x \in A: \emptyset \in f(x)\}: f \in \mathcal{P}(\{\emptyset\})^A\big\}$. Clearly, $C \subseteq \mathcal{P}(A)$. For the other inclusion, let $z \subseteq A$ and let $f(x) = \{y \in \{\emptyset\}: x \in z\}$ for $x \in A$. Then $f \in \mathcal{P}(\{\emptyset\})^A$ and $z = \{x \in A: \emptyset \in f(x)\}$. This shows $C = \mathcal{P}(A)$. ■

Proposition 3.23: BST + Exponentiation + **LEM′** ⊢ Powerset.

Proof: By It turns out, that the seemingly weak assumption of $\mathcal{P}(\{\emptyset\})$ to be a set is enough to to deduce the full Powerset axiom from the Exponentiation axiom. By "$\mathcal{P}(a)$ is a set" we mean the formula $\exists z. z = \mathcal{P}(a)$, which of course stands for $\exists z. \forall b. (b \subseteq a \leftrightarrow b \in z)$.

Proposition 3.22, it suffices to show that $\mathcal{P}(\{\emptyset\})$ is a set. In fact, we show $\mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$. Let $x \subseteq \{\emptyset\}$. By **LEM**$'$, $\emptyset \in x \vee \emptyset \notin x$. In the first case, $x = \{\emptyset\}$, in the second case, $x = \emptyset$. ■

The General Uniformity Principle (GUP) is the following schema:

$$\forall a\big(\forall x. \exists y \in a. \phi(x, y) \to \exists y \in a. \forall x. \phi(x, y)\big).$$

Intuitively, GUP says that every mapping from the universe of all sets into a given set a, must in fact be constant. From the fact that **CZF** is consistent with a weaker version of this axiom (the case $a = \omega$) and a result from [*71*], one concludes that GUP is consistent with **CZF**. Given this result, one can easily show the following:

Ⓜ**Theorem 3.24**: For $a \neq \emptyset$, $\mathcal{P}(\emptyset)$ is not a set in **CZF**:

$$\mathbf{CZF} \nvdash a \neq \emptyset \to \exists z. z = \mathcal{P}(a).$$

Proof: Towards a contradiction, suppose that **CZF** proves that $\mathcal{P}(a)$ is a set for $a \neq \emptyset$. Then so does **CZF** + GUP. We will argue inside the latter theory for the rest of the proof.

By bounded separation, $\{w \in a : \emptyset \in x\}$ is a set too, for every set x. Hence,

$$\forall x. \exists y \in \mathcal{P}(a). y = \{w \in a : \emptyset \in x\},$$

and by GUP,

$$\forall y \in \mathcal{P}(a). \exists x. y = \{w \in a : \emptyset \in x\}.$$

Let this y be fixed.

- For $x = \emptyset$, $y = \bigcup\{w \in a : \emptyset \in \emptyset\} = \emptyset$.
- For $x = \{\emptyset\}$, $y = \bigcup\big\{w \in a : \emptyset \in \{\emptyset\}\big\} = a$.

Altogether, $a = \emptyset$, a contradiction. ■

This is a very strong result: It turns out, that in restricting **IZF** to **CZF** one not only loses the unwanted instances of the Powerset axiom, where the $\mathcal{P}$-operator is applied to infinite sets. For example, it seems reasonable to doubt that $\mathcal{P}(\omega)$ is consecutively justifiable. The peculiar side is that we also lose innocent instances where the resulting power-set is classically finite. Intuitionistically – and we have discussed this in the section about the Power set axiom in Zermelo-Fraenkel set theory **ZFC**2.4.1 – already the set $\mathcal{P}(1)$ is equivalent to the set of all formulas of the underlying language of set theory.

3.4.2 Inductive definitions

Note that in defining the sets V_α we used the Powerset-axiom in a decisive way. In order to give a definition of this (and other important classes) in **CZF** as well, we need to consider inductive definitions.

Definition 3.25: An *inductive definition* is a class of ordered pairs. If $\langle x, a\rangle \in \Phi$, we call x a *premise* and a a *conclusion under* Φ. For any class Y, we define the class of Φ-conclusions from premises Y,

$$\Gamma_\Phi(Y) = \{a\colon \exists X\ (X \subseteq Y \land \langle X, a\rangle \in \Phi\}.$$

We call a class Y *Φ-closed* iff $\Gamma_\Phi(Y) \subseteq Y$.

The familiar terminology is not accidental: For example, we could consider a logic with some inference rule Φ: This rule can be represented as a class of pairs $\langle X, a\rangle$, where X is a finite set of premises and a is the conclusion of x under Φ. Then, for any set M of formulas, $\Gamma_\Phi(M)$ is the class of all formulas that can be inferred from M by Φ in one step. We will be interested in the smallest Φ-closed class: In our example, it is the deductive closure of M with respect to Φ.

Ⓜ**Theorem 3.26 (inductive definition Theorem)**: For any inductive definition Φ, there is a smallest Φ-closed class I. It can be written as

$$I = \bigcup_{a\in V} I^a,$$

where the I^as satisfy

$$I^a = \Gamma^\Phi\left(\bigcup_{b\in a} I^b\right).$$

Furthermore, we have the following induction principle:

$$\frac{\forall a\left(\forall b \in a.\forall x \in I^b.\phi(x) \to \forall x \in I^a.\phi(x)\right)}{\forall x \in I.\phi(x)}$$

Before we prove the existence-part in a separate lemma, let us introduce some notation: We define for a set X and a class (or set) M,

$$M^X = \{a\colon \langle X, a\rangle \in M\},$$
$$M^{\in X} = \{a\colon \exists Y \in X\ \langle Y, a\rangle \in M\} = \bigcup_{Y\in X} M^Y\ .$$

Ⓜ**Lemma 3.27** (**BST** + Collection): There is a class J such that

$$J^a = \Gamma_\Phi\left(\bigcup_{b\in a} J^b\right).$$

Proof: We call a set G of ordered pairs *good* iff

$$\langle X, a\rangle \in G \to a \in \Gamma_\Phi(G^{\in X}),$$

Let $J = \bigcup\{G : G \text{ is good}\}$. The lemma in our new notation reads $J^X = \Gamma_\Phi(J^{\in X})$. We show this as follows:

Let $a \in J^X$. Then $\langle X, a\rangle \in G$ for some good set G, which means $a \in \Gamma_\Phi(G^{\in X})$. Since $G^{\in X} \subseteq J^{\in X}$ it follows that $a \in \Gamma_\Phi(J^{\in X})$. Thus, $J^X \subseteq \Gamma_\Phi(J^{\in X})$.

For the other inclusion, let $a \in \Gamma_\Phi(J^{\in X})$. Then $\langle Y, a\rangle \in \Phi$ for some $Y \subseteq J^{\in X}$, i.e. $\forall y \in Y\ \exists x \in X\ y \in J^x$, so

$$\forall y \in Y. \exists G. G \text{ is good and } y \in G^{\in X}.$$

By collection, there is a set Z such that

$$\forall y \in Y. \exists G \in Z. G \text{ is good and } y \in G^{\in X}.$$

Let $G = \{\langle X, a\rangle\} \cup \bigcup Z$. Then $\bigcup Z$ is good and because of $\langle Y, a\rangle \in \Phi$ and $Y \subseteq G^{\in X}$ also $a \in \Gamma_\Phi(G^{\in X})$; hence G is good. As $\langle X, a\rangle \in G$, we have $a \in J^X$. Thus $\Gamma_\Phi(J^{\in X}) \subseteq J^X$. ■

Proof of the theorem: Let the J^as be as in the lemma. Of course, we set $I^a = J^a$ and $I = \bigcup_{a\in V} I^a$ and claim that this does the job. Indeed, let $\langle X, a\rangle \in \Phi$ and $X \subseteq I$. Then for each $x \in X$ there is some y such that $x \in I^y$. So, by collection, there is a set Y, such that

$$\forall x \in X. \exists y \in Y. x \in I^y,$$

which shows that $X \subseteq \bigcup_{y\in Y} I^y$ and hence $a \in \Gamma_\Phi\left(\bigcup_{y\in Y} I^y\right) = I^Y \subseteq I$.

To show minimality, let I' be another Φ-closed class. We show $I^X \subseteq I'$ by set induction (the other inclusion is trivial). Suppose, $I^x \subseteq I'$ for all $x \in X$. But by monotonicity of Γ_Φ,

$$I^X = \Gamma_\Phi\left(\bigcup_{x\in X} I^y\right) \subseteq \Gamma_\Phi(I) \subseteq I'.$$

The induction principle is nothing more but set induction: Assume, $\forall a\left(\forall b \in a. \forall y \in J^b. \phi(y) \to \forall x \in J^a. \phi(x)\right)$. By set induction, $\forall a. \forall x \in J^a. \phi(x)$ and thus $\forall x \in J. \phi(x)$. ■

We give another formulation of Theorem 3.26 in terms of ordinals:

Ⓜ**Theorem 3.28 (inductive definition on ordinals)**: For any inductive definition Φ, there is a smallest Φ-closed class J. Furthermore,

$$I = \bigcup_{\alpha\in\mathbb{ON}} I^\alpha,$$

where the I^{α}s satisfy

$$I_{\alpha} = \Gamma_{\Phi}\left(\bigcup_{\beta\in\alpha} I_{\beta}\right).$$

Furthermore, we have the following induction principle

$$\frac{\forall\alpha \in \mathbb{ON}\left(\forall\beta \in \alpha. \forall x \in I^{\beta}. \phi(x) \to \forall x \in I^{\alpha}. \phi(x)\right)}{\forall x \in I. \phi(x)}$$

Proof: Effectively, all we need to show is that the smallest Φ-closed class I from Theorem 3.26 can be written as $\bigcup_{\alpha\in\mathbb{ON}} I^{\alpha}$. So, let $I' = \bigcup_{\alpha\in\mathbb{ON}} I^{\alpha}$, we show that indeed $I = I'$. Obviously, $I' \subseteq I$. For the other inclusion we show that I' is Φ-closed. Let $\langle X, a\rangle \in \Phi$, where X is a subset of J. We have $\forall x \in X\ \exists\alpha \in \mathbb{ON}\ x \in I^{\alpha}$. By collection, there is some B containing all these αs. Hence, by Proposition 3.4 it holds that $X \subseteq J^{\gamma}$, where $\gamma = \bigcup B + 1$. By definition, $a \in \Gamma_{\Phi}(I^{\gamma}) = I^{\gamma+1}$.

The induction principle follows easily. ■

Example: Let Φ be the class of pairs $\langle X, a\rangle$, where X is any set and $a \subseteq X$. We set $V_{\alpha} = \bigcup_{\beta\in\alpha} I^{\beta}$ and show $V_{\alpha} = \bigcup_{\beta\in\alpha} \mathcal{P}(V_{\beta})$ for all α. Suppose, this holds true for all $\beta \in \alpha$, then

$$V_{\alpha} = \bigcup_{\beta\in\alpha} J^{\beta} = \bigcup_{\beta\in\alpha} \Gamma_{\Phi}\left(\bigcup_{\gamma\in\beta} J^{\gamma}\right) = \bigcup_{\beta\in\alpha} \mathcal{P}\left(\bigcup_{\gamma\in\beta} J^{\gamma}\right) = \bigcup_{\beta\in\alpha} \mathcal{P}(V_{\beta}).$$

In later chapters, we will often prove statements by *double recursion* for an inductive class I. This is the scheme

$$\frac{\forall\alpha\left(\left(\forall\beta \in \alpha\ \forall x \in I^{\beta}\ \phi(x)\right) \to \forall x \in I^{\alpha}\ \psi(x)\right) \quad \forall\alpha\left(\left(\forall x \in I^{\alpha}\ \psi(x)\right) \to \forall x \in I^{\alpha}\ \phi(x)\right)}{\forall x \in I. \phi(x) \land \psi(x)}$$

and it is easily entangled into two separate inductions for ϕ and ψ. However, it will be convenient to choose double induction over two single inductions.

3.5 Relation between the theories

For the sake of completeness, we would like to establish the looming relationship between the theories **CZF**, **IZF** and **ZF**. First, we need some results about the interplay between the axioms of replacement, collection and strong collection – the only axioms that are strengthened when passing from **ZF** to the weaker theories.

3.5.1 Replacement, Collection, Strong Collection

Ⓜⓒ**Proposition 3.29**: **ZF** ⊢ Collection

Proof: Suppose, $\forall x \in A. \exists y. \phi(x, y)$. Then

$$\forall x \in A. \exists \alpha_x \in \mathbb{ON}. \exists y \in V_{\alpha_x}. \phi(x, y).$$

And we may assume that α_x is the smallest such β. By replacement, form the set B of all such α_x. Then $\alpha = \bigcup B$ is an ordinal and $V_\alpha = \bigcup_{x \in A} V_{\alpha_x}$ and therefore

$$\forall x \in A. \exists y \in V_\alpha. \phi(x, y).$$

■

Ⓜ**Proposition 3.30**: **IZF** ⊢ Strong collection

Proof: Suppose, $\forall x \in A. \exists y. \phi(x, y)$. By collection, we get a set B' such that $\forall x \in A. \exists y \in B'. \phi(x, y)$. To obtain a set B such as required in the strong collection schema, set $B = \{y \in B: \exists x \in A. \phi(x, y)\}$ by separation. ■

3.5.2 Proofs of inclusion

We can now gather together our results from the previous sections to establish:

Ⓜ**Theorem 3.31**: We have the strict inclusions

$$\mathbf{CZF} \subsetneq \mathbf{IZF} \subsetneq \mathbf{ZF}$$

and the equations

$$\mathbf{CZF} + \mathbf{LEM} = \mathbf{IZF} + \mathbf{LEM} = \mathbf{ZF}.$$

Proof: Ⓜ**CZF** ⊊ **IZF**: **IZF** proves Strong Collection by Proposition 3.30 and Subset Collection by Proposition 2.22 and Proposition 2.21. By Theorem 3.24, the Powerset axiom does not hold in its full generality in **CZF**, which shows that the inclusion is strict.

Ⓜ**IZF** ⊊ **ZF**: **ZF** proves Collection by Proposition 3.29 and Set induction by Proposition 2.15. We will show in Corollary 4.27, that **LEM** is indeed not derivable in **IZF** (and hence so are all weak counterexamples of this chapter and the previous one).

ⓒ**IZF** + **LEM** = **ZF**: Foundation follows from Proposition 2.15.

ⓒ**CZF** + **LEM** = **ZF**: Separation follows from Proposition 2.19 and Power set from Proposition 3.23. ■

4 Metamathematical properties of constructive axiomatic set theories

In this chapter we will discuss some metamathematical properties that we expect an axiomatic constructive set theory to possess. Many of these properties arise naturally from the BHK-semantics, others (such as Markov's or Church's principles) can be seen as starting point of a subbranch of constructivism. We will be able to show that **CZF** and **IZF** indeed satisfy some of these properties. For this purpose, we will use the concept of realizability – which itself can be seen as a specification of the BHK-interpretation (see based on effective computability. We will start the chapter with a short section recalling some notions from computability theory needed later on.

4.1 Some aspects of computability theory

In this section we will recall some notion and concepts from computability theory. We assume that the reader is familiar with the basic concepts and will therefore be relatively brief. The most prominent models of computability are Turing-machines, the lambda-calculus and recursive functions. It has been shown in [*36*] and [*67*], that the three models turn out to be equivalent, i.e. they describe the same class of functions. Additionally, as all three concepts make minimal assumptions and seem to model the general idea of computability pretty well, it is reasonable to believe that all effectively computable functions can be given in terms of one of these concepts (this assumption is known as *Church-Turing-thesis*). In the further discussion we will hence, when referring to *programs*, *algorithms* or *effectively computable functions*, actually mean recursive functions, the lambda-calculus or Turing-machines and use a lambda-calculus-style notation.

We will write Kl (after Kleene) for the structure being able of forming lambda-terms and generously containing as constants all natural numbers and standard operators such as **s**, **k**, **i**, **p** (pairing) **l** and **r** (left and right projections), logical operators etc. We will write $\tau \simeq n$ if the term τ converges to the natural number n. For terms τ, θ, we write $\tau \simeq \theta$ if the terms converge to the same term (if they do converge). The expression $\tau \downarrow$ means that τ converges (to any term) and for a formula $\phi(x)$ on natural numbers, we write $\phi(\tau)$ if τ converges to n and $\phi(n)$ holds. As usually, the set of all terms is countable, and we may assume that there is a Gödel-numbering of these terms. We write $\{e\}$ for the term with number e. Often, we will not clearly distinguish between a term and its Gödel-number and write ef instead of $\{e\}f$, where this is not problematic. Let us recall some basic results:

Theorem 4.1 (Recursion): There is a term τ^{fix} (the *fixed-point combinator*) such that for all terms σ,

$$Kl \vDash \tau^{\text{fix}}\sigma \simeq \sigma\left(\tau^{\text{fix}}\sigma\right).$$

Proof: Let $\tau^{\text{fix}} \equiv \lambda z.\left(\lambda y. z(yy)\right)\left(\lambda y. z(yy)\right)$. We verify the property:

$$\begin{aligned}&\tau^{\text{fix}}\sigma\\ &\simeq \big(\lambda y.\sigma(yy)\big)\big(\lambda y.\sigma(yy)\big)\\ &\simeq \sigma\Big(\big(\lambda y.\sigma(yy)\big)\big(\lambda y.\sigma(yy)\big)\Big)\\ &\simeq \sigma\big(\tau^{\text{fix}}\sigma\big).\end{aligned}$$

■

Theorem 4.2 (Fixed-point): All terms $\tau(x)$ with x free have a fixed-point i, this means $Kl \vDash \tau(i) \simeq i$.

Proof: Let $\sigma \equiv \lambda x.\tau(x)$ and $i \equiv \tau^{\text{fix}}\sigma$. Indeed, by the recursion theorem,

$$i \equiv \tau^{\text{fix}}\sigma \simeq \sigma\big(\tau^{\text{fix}}\sigma\big) \equiv \sigma i \equiv \big(\lambda x.\tau(x)\big)i \simeq \tau(i).$$

■

Lemma 4.3 (Double recursion): For terms $\sigma_1(x,y)$ and $\sigma_2(x,y)$ there are τ_1 and τ_2 such that

$$Kl \vDash \tau_1 \simeq \sigma_1(\tau_1,\tau_2) \wedge \tau_2 \simeq \sigma_2(\tau_1,\tau_2).$$

Proof: Given σ_1 and σ_2, apply the Fixed-point theorem to find a fixed point i of

$$\mathbf{p}\sigma_1(\mathbf{l}x,\mathbf{r}x)\sigma_2(\mathbf{l}x,\mathbf{r}x).$$

Set $\tau_1 \equiv \mathbf{l}i$ and $\tau_2 \equiv \mathbf{r}i$. We check the first property: Using the fixed-point property, we have

$$\tau_1 \equiv \mathbf{l}i \simeq \mathbf{l}\big(\mathbf{p}\sigma_1(\mathbf{l}i,\mathbf{r}i)\sigma_2(\mathbf{l}i,\mathbf{r}i)\big) \simeq \sigma_1(\mathbf{l}i,\mathbf{r}i) \equiv \sigma_1(\tau_1,\tau_2).$$

■

Definition 4.4: The lest number operator μ is known from recursive functions. For a term τ, we define by $\mu m.\tau$ the least number m such that $\tau(m) \simeq 0$.

4.2 Metamathematical properties

In this section we will discuss some metamathematical properties that we expect proper constructive set theory to possess. In the following discussion, let **T** be a theory over a language containing the relation "$\in$", some constant ω denoting the set of natural numbers and constants 0,1,2, ... denoting its elements. Furthermore, we assume that **T** decides equality on ω in the usual way and that some form of computability may be formalized within **T**.

4.2.1 Disjunction property

The disjunction property is the property that whenever the theory **T** proves $\phi \vee \psi$, it must actually prove one of ϕ or ψ:

$$\text{If } \mathbf{T} \vdash \phi \vee \psi, \text{ then } \mathbf{T} \vdash \phi \text{ or } \mathbf{T} \vdash \psi.$$

Following the BHK-interpretation, a proof p of $\phi \vee \psi$ must indeed be a pair $\langle n, q\rangle$, where n is a natural number and q proves ϕ if $n = 0$ and ψ, else. Hence, it is reasonable to expect for a constructive theory to possess the disjunction property (indeed, we will show, that both **CZF** and **IZF** have the disjunction property).

For the classical set theory **ZFC**, this property does not hold: Take any statement independent of **ZFC**, like the continuum hypothesis **CH**. Then, by **LEM**, $\mathbf{ZFC} \vdash \mathbf{CH} \vee \neg\mathbf{CH}$, but famously neither $\mathbf{ZFC} \vdash \mathbf{CH}$ nor $\mathbf{ZFC} \vdash \neg\mathbf{CH}$.

4.2.2 Existence property and numerical existence property

A theory $\mathbf{T}$ has the *existence property* if whenever $\mathbf{T} \vdash \exists x\, \phi(x)$, then there is a formula $\theta(x)$ with exactly x free such that $\mathbf{T} \vdash \exists!\, x\, [\theta(x) \wedge \phi(x)]$. This means that the witness of the formula ϕ can be constructed by the formula θ. The *numerical existence property* is a weakening of this: Whenever $\mathbf{T} \vdash \exists x \in \omega.\, \phi(x)$, then there has to be a natural number n such that $\mathbf{T} \vdash \phi(n)$. Again, both properties can be expected of a constructive theory. For example, consider the BHK-interpretation: A proof p of $\exists x\, \phi(x)$ should provide us with a construction of a witness y along with a proof of $\phi(y)$.

Note that the disjunction property is actually a special case of the numerical existence property: Let $\theta(n)$ be the formula $(n = 0 \rightarrow \phi) \vee (n \neq 0 \rightarrow \psi)$. If $\mathbf{T}$ has the numerical existence property and proves $\phi \vee \psi$ then it also proves $\exists n \in \omega.\, \theta(n)$. Now one can decide whether $\mathbf{T}$ proves ϕ or ψ by inspecting the witness n of θ, given by the numerical existence property. In particular, this shows, that **ZFC** does not possess the numerical existence property.

Surprisingly, the existence property does not hold neither for **CZF** nor for **IZF** ([25], [63]), which may seem unsatisfactory. However, the theory **CZF** may still be defended by the fact that there is a natural interpretation of **CZF** in Martin-Löf's type theory, which does allow us to extract witnesses from proofs. We will discuss this interpretation in chapter 5.

But both **CZF** and **IZF** have the numerical existence property – which we will show in sections 4.6 and 4.7.

4.2.3 Unzerlegbarkeits-rule and variants

The *Unzerlegbarkeits-rule* (UzR) states that the universe of sets is unzerlegbar[3] by a property: Whenever $\mathbf{T} \vdash \forall x[\psi(x) \vee \neg\psi(x)]$, then $\mathbf{T} \vdash \forall x\, \psi(x) \vee \forall x\, \neg\psi(x)$. In the words of McCarty in [44], this rule says that if you have two colors, "the only way to color all the sets in the [...] universe is to make everything the same color!"

A generalization is the case where you have countably many colors, formulated in the *Uniformity-rule* (UR): If whenever $\mathbf{T} \vdash \forall x\, \exists y \in \omega.\, \psi(x, y)$, then $\mathbf{T} \vdash \exists y \in \omega.\, \forall x\, \psi(x, y)$. This is indeed a generalization: If

[3] germ, indecomposable

a theory **T** has the Uniformity rule and proves $\forall x[\psi(x) \vee \neg\psi(x)]$, then it must also prove $\forall x. \exists y \in \omega. \theta(x, y)$, where $\theta(x, y) \equiv (\psi(x) \wedge y = 0) \vee (\neg\psi(x) \wedge y \neq 0)$. By Uniformity, we either have $\forall x\, \psi(x)$ or $\forall x\, \neg\psi(x)$, depending on y.

Clearly, **ZFC** does not enjoy any of the two properties: Take for example the formula $\psi(x) \equiv x = \emptyset$. Of course, **ZFC** $\vdash \forall x[x = \emptyset \vee x \neq \emptyset]$, but both $\forall x. x = \emptyset$ and $\forall x. x \neq \emptyset$ are absurd in **ZFC**[4].

4.2.4 Church's rule

Church's rule (CR, do not confuse with Church-Turing-thesis) says that all total rules on ω are given by effectively computable functions: whenever **T** $\vdash \forall x \in \omega. \exists y \in \omega. \phi(x, y)$, then there is some number e with **T** $\vdash \forall x \in \omega. \phi\big(x, \{e\}(x)\big)$.

This seems as a reasonable assumption in the intuitionistic case, but classically this is clearly not true. For example, the Halting-problem inspires the following function in **ZFC**:

$$f(n, m) = \begin{cases} 1, & \text{if } \{n\}(m) \text{ converges,} \\ 0, & \text{otherwise.} \end{cases}$$

Due to **LEM**, this function is total in **ZFC**. From Turing's famous argument ([*68*]), this function is not computable. Hence, there is no way to effectively compute witnesses for the formula $\forall \mathbf{x} \in \omega. \exists z \in \{0,1\}. f(\mathbf{x}) = z$, (we may assume that $\mathbf{x}$ codes the pair (x, y)).

This reasoning shows a clear incompatibility of Church's rule with the axiom **LEM**.

4.2.5 Markov's rule

Markov's rule (MR), central to Russian Constructivism, says that if we can prove that it is impossible that a program is never terminating, then it does terminate: If **T** $\vdash \forall n \in \omega\big(\phi(n) \vee \neg\phi(n)\big) \wedge \neg\forall n \in \omega. \neg\phi(n)$, then **T** $\vdash \exists n \in \omega. \phi(n)$.

This is the first rule we discussed, that holds (trivially) in the classical setting.

4.2.6 Rules vs. Principles

Note that all rules so far (Unzerlegbarkeit, Church, Markov) came in the form

$$\text{If } \mathbf{T} \vdash A, \text{ then } \mathbf{T} \vdash B.$$

We can associate to every such rule a *principle*, of the form

$$\mathbf{T} \vdash A \to B.$$

For example, we can formulate the Unzerlegbarkeits-principle as

[4] This counterexample does not work in **CZF** or **IZF**, as we cannot, in general, decide for each set x, whether $x = \emptyset$ or $x \neq \emptyset$.

$$\mathbf{T} \vdash \forall x[\psi(x) \lor \neg\psi(x)] \to [\forall x\, \psi(x) \lor \forall x\, \neg\psi(x)].$$

Both formulations may seem similar in description, but differ significantly in content: To show that a theory follows a rule, we must show how to transfer **T**-proofs of A into **T**-proofs of B. For the corresponding principle, however, we must give **T**-proofs of $A \to B$. Clearly, the principles imply their corresponding rules, but the converse is not the case:

For example, as both **CZF** and **IZF** are subtheories of **ZFC**, where Church's principle conflicts with **LEM**, both theories cannot possibly follow Church's principle. We will show, however, that they follow the corresponding rule. We can think of this situation as the fact that **CZF** and **IZF** do not "know" that each rule is effectively computable. Metamathematically, however, we can transform every proof of $\forall x \in \omega. \exists y \in \omega. \phi(x, y)$ into a procedure e computing for each $x \in \omega$ some $\{e\}(x) \in \omega$ such that $\phi\big(x, \{e\}(x)\big)$. Knowing about the Curry-Howard correspondence, this fact is perhaps not too big a surprise.

Although **CZF** and **IZF** cannot follow Church's principle, will show that both theories are are **compatible** with it, i.e. we can consistently extend **T** to a theory **T**′ following this principle. Of course, **T**′ $\nsubseteq$ **ZFC** (and also **ZFC** $\nsubseteq$ **T**′).

We denote the Unzerlegbarkeits principle, Uniformity principle, Church's principle and Markov's principle as UzP, UP, CP and MP respectively.

4.3 Realizability of **CZF**

Realizability is a semantical method inspired by the BHK-interpretation. While the latter leaves open the precise notion of proof, realizability interprets proofs as elements of a certain "proof structure". In our discussion this proof structure will V^* and proofs will be interpreted as effectively computable functions. Not further specifying our domain of discourse (let it be an abstract $\mathcal{V}$ for now), and leaving out the atomic case and bounded quantification, we can transfer the rules in the BHK-semantics:

$$\begin{array}{rcl}
e \Vdash \phi \land \psi & \text{iff} & (e)_0 \Vdash \phi \land (e)_1 \Vdash \psi \\
e \Vdash \phi \lor \psi & \text{iff} & [(e)_0 = 0 \land (e)_1 \Vdash \phi] \lor [(e)_0 \neq 0 \land (e)_1 \Vdash \psi] \\
e \Vdash \neg\phi & \text{iff} & \forall f[\neg(f \Vdash \phi)] \\
e \Vdash \phi \to \psi & \text{iff} & \forall f[(f \Vdash \phi) \to (\{e\}(f) \Vdash \psi)] \\
e \Vdash \forall x\, \phi & \text{iff} & \forall \mathfrak{a} \in \mathcal{V}\, [e \Vdash \phi(\mathfrak{a})] \\
e \Vdash \exists x\, \phi & \text{iff} & \exists \mathfrak{a} \in \mathcal{V}\, [e \Vdash \phi(\mathfrak{a})] \\
e \Vdash \perp & \text{iff} & \perp
\end{array}$$

If $e \Vdash \phi$, we say that e *realizes* ϕ or that e *is a realizer for* ϕ.

4.3.1 The proof structure V^*

Let us now motivate the atomic case. As in section 2.1.2, we imagine a set $\mathfrak{a}$ to be given not only by its elements x, but also by (coding of) proofs n of the fact that $x \in \mathfrak{a}$. Thus, what we can do is to **identify** the set $\mathfrak{a}$ with all such pairs $\langle n, x\rangle$. The problem, however, is that if we want sets to be foundational, the element x itself must be of this shape. To follow this idea, one will thus have to define inductively the universe $V(Kl)$ as:

$$V(Kl) = \{\langle n, \mathfrak{a}\rangle : n \in \omega \wedge \mathfrak{a} \in V(Kl)\}.$$

Note that this is actually an inductive definition and justified by Theorem 3.28. This structure is studied in [44] and [54]. For our purposes, however, we will use a similar, but slightly more complicated structure, presented in [55]. The reason for this is, that in order to verify the rules of section 4.2, we want to refer to truth inside **CZF** in our definition of realizability. Define for a pair $\langle a, m\rangle$ the expressions $\langle a, m\rangle^\circ = a$ and $\langle a, m\rangle^* = m$. The definition of the universe V^* looks as follows:

$$\begin{aligned} V_\alpha^* &= \bigcup_{\beta\in\alpha} \left\{\langle a, m\rangle : a \in V_\beta \wedge m \subseteq \omega \times V_\beta^* \wedge \forall x \in m.\, (x^*)^\circ \in a\right\}, \\ V^* &= \bigcup_{\alpha\in\mathbb{ON}} V_\alpha^*. \end{aligned}$$

We can now give the definitions of realizability for atomic formulas $\mathfrak{a} \in \mathfrak{b}$ and $\mathfrak{a} = \mathfrak{b}$. This definition is by simultaneous recursion and incorporates the extensionality axiom:

$$\begin{array}{lll} e \Vdash \mathfrak{a} \in \mathfrak{b} & \text{iff} & \mathfrak{a}^\circ \in \mathfrak{b}^\circ \wedge \exists \mathfrak{d}[\langle (e)_0, \mathfrak{d}\rangle \in \mathfrak{b} \wedge (e)_1 \Vdash \mathfrak{a} = \mathfrak{d}] \\ e \Vdash \mathfrak{a} = \mathfrak{b} & \text{iff} & \mathfrak{a}^\circ = \mathfrak{b}^\circ \wedge \\ & & \quad \forall f, \mathfrak{d}[(\langle f, \mathfrak{d}\rangle \in \mathfrak{a} \to \{(e)_0\}(f) \Vdash \mathfrak{d} \in \mathfrak{b}) \wedge (\langle f, \mathfrak{d}\rangle \in \mathfrak{b} \to \{(e)_1\}(f) \Vdash \mathfrak{d} \in \mathfrak{a})] \end{array}$$

We will want to carry out these constructions inside **CZF** (definition inside **IZF** is less problematic). We do this in the following lemma:

Lemma 4.5: The classes V_α^* and V^* are definable in **CZF.**

Proof: For each set or class X, define $X' = \{a : \exists m.\, \langle a, m\rangle \in X\}$. We use Theorem 3.28 to define the V_α^* and at the same time show inductively that $(V_\alpha^*)' = V_\alpha$.

Let Φ be the inductive definition with

$$\langle X, \langle a, m\rangle\rangle \in \Phi \quad \text{iff} \quad a \in X' \text{ and } m \subseteq \omega \times X, \text{where } \forall x \in m.\, (x^*)^\circ \in a,$$

and suppose, $\left(V_\beta^*\right)' = V_\beta$ for all $\beta \in \alpha$. By Theorem 3.28, we have $J = \bigcup_{\alpha\in\mathbb{ON}} J_\alpha$ and for each α, $J_\alpha = \Gamma_\Phi\left(\bigcup_{\beta\in\alpha} J_\beta\right)$. Let $V_\alpha^* \coloneqq \bigcup_{\beta\in\alpha} J_\beta$. Then,

$$\begin{aligned} V_\alpha^* &= \bigcup_{\beta\in\alpha} J_\beta \\ &= \bigcup_{\beta\in\alpha} \Gamma_\Phi\left(\bigcup_{\gamma\in\beta} J_\gamma\right) \end{aligned}$$

$$
\begin{aligned}
&= \bigcup_{\beta\in\alpha}\left\{\langle a,m\rangle : a \in \left(\bigcup_{\gamma\in\beta} J_\gamma\right)' \wedge m \subseteq \omega \times \bigcup_{\gamma\in\beta} J_\gamma \wedge \forall x \in m.\, (x^*)^\circ \in a\right\}\\
&= \bigcup_{\beta\in\alpha}\left\{\langle a,m\rangle : a \in \left(V_\beta^*\right)' \wedge m \subseteq \omega \times V_\beta^* \wedge \forall x \in m.\, (x^*)^\circ \in a\right\}\\
&= \bigcup_{\beta\in\alpha}\{\langle a,m\rangle : a \in V_\beta \wedge m \subseteq \omega \times V_\beta^* \wedge \forall x \in m.\, (x^*)^\circ \in a\}\\
&= V_\alpha^*.
\end{aligned}
$$

The inductive step follows from

$$
\begin{aligned}
(V_\alpha)' &= \left(\bigcup_{\beta\in\alpha} J_\beta\right)'\\
&= \left\{a : \exists m.\, \langle a,m\rangle \in \left(\bigcup_{\beta\in\alpha} J_\beta\right)'\right\}\\
&= \left\{a : \exists m.\, \langle a,m\rangle \in \bigcup_{\beta\in\alpha}\{\langle a,m\rangle : a \in V_\beta \wedge m \subseteq \omega \times V_\beta^* \wedge \forall x \in m.\, (x^*)^\circ \in a\}'\right\}\\
&\supseteq \left\{a : \langle a,\emptyset\rangle \in \bigcup_{\beta\in\alpha}\{\langle a,m\rangle : a \in V_\beta \wedge m \subseteq \omega \times V_\beta^* \wedge \forall x \in m.\, (x^*)^\circ \in a\}\right\}\\
&= \left\{a : a \in \bigcup_{\beta\in\alpha} V_\beta\right\} = V_\alpha,
\end{aligned}
$$

and the other inclusion is clear. ∎

Also, the definition of the clauses for $e \Vdash \mathfrak{a} = \mathfrak{b}$ and $e \Vdash \mathfrak{a} \in \mathfrak{b}$ must be justified: Again, we invoke the inductive definition theorem: Let $\langle x, a\rangle$ be a member of Φ iff

$a = \langle \mathfrak{a}, \mathfrak{b}, 0, e\rangle$, where $\mathfrak{a}^\circ = \mathfrak{b}^\circ$ and $\forall \langle \mathfrak{d}, f\rangle \in \mathfrak{a}\, (\langle \mathfrak{d}, \mathfrak{b}, 1, \{(e)_0\}f\rangle \in x) \wedge \forall \langle \mathfrak{d}, f\rangle \in \mathfrak{b}\, (\langle \mathfrak{d}, \mathfrak{a}, 1, \{(e)_1\}f\rangle \in x)$, or
$a = \langle \mathfrak{a}, \mathfrak{b}, 1, e\rangle$, where $\mathfrak{a}^\circ \in \mathfrak{b}^\circ$ and $\exists \mathfrak{d}[\langle \mathfrak{d}, (e)_0\rangle \in \mathfrak{b} \wedge \langle \mathfrak{a}, \mathfrak{d}, 0, (e)_1\rangle \in x]$,

and in both cases $\mathfrak{a}, \mathfrak{b} \in V^*$, $e \in \omega$. The smallest Φ-closed class J defines the relations $e \Vdash \mathfrak{a} = \mathfrak{b}$ iff $\langle \mathfrak{a}, \mathfrak{b}, 0, e\rangle \in J$ and $e \Vdash \mathfrak{a} \in \mathfrak{b}$ iff $\langle \mathfrak{a}, \mathfrak{b}, 1, e\rangle \in J$.

4.3.2 Definition of realizability

We can now give the definition of realizability we will work with. As we want to prove that **CZF** and **IZF** follow the rules discussed in Section 4.2, our kind of realizability needs so refer to truth inside those systems. What we define is hence a variant of "realizability with truth" from [55].

For $\phi = \phi(\mathfrak{a}_1, \dots, \mathfrak{a}_n)$ with all the free variables shown, we define ϕ° by $\phi(\mathring{\mathfrak{a}}_1, \dots, \mathring{\mathfrak{a}}_n)$. The full definition of realizability reads as follows:

Let $\mathfrak{a}, \mathfrak{b}, \mathfrak{d}$ range over V^* and e, f, c over ω. We define recursion

$$\begin{array}{rll}
e \Vdash \mathfrak{a} \in \mathfrak{b} & \text{iff} & \mathfrak{a}^{\circ} \in \mathfrak{b}^{\circ} \wedge \exists \mathfrak{d}[\langle (e)_0, \mathfrak{d} \rangle \in \mathfrak{b}^* \wedge (e)_1 \Vdash \mathfrak{a} = \mathfrak{d}] \\
e \Vdash \mathfrak{a} = \mathfrak{b} & \text{iff} & \mathfrak{a}^{\circ} = \mathfrak{b}^{\circ} \wedge \forall f, \mathfrak{d}[(\langle f, \mathfrak{d} \rangle \in \mathfrak{a}^* \rightarrow \{(e)_0\}f \Vdash \mathfrak{d} \in \mathfrak{b}) \wedge (\langle f, \mathfrak{d} \rangle \in \mathfrak{b}^* \rightarrow \\
 & & \{(e)_1\}f \Vdash \mathfrak{d} \in \mathfrak{a})] \\
e \Vdash \phi \wedge \psi & \text{iff} & (e)_0 \Vdash \phi \wedge (e)_1 \Vdash \psi \\
e \Vdash \phi \vee \psi & \text{iff} & [(e)_0 = \mathbf{0} \wedge (e)_1 \Vdash \phi] \vee [(e)_0 \neq \mathbf{0} \wedge (e)_1 \Vdash \psi] \\
e \Vdash \neg \phi & \text{iff} & (\neg \phi^{\circ}) \wedge \forall f[\neg (f \Vdash \phi)] \\
e \Vdash \phi \rightarrow \psi & \text{iff} & (\phi^{\circ} \rightarrow \psi^{\circ}) \wedge \forall f[(f \Vdash \phi) \rightarrow (\{e\}f \Vdash \psi)] \\
e \Vdash \forall x \in \mathfrak{a}. \phi(x) & \text{iff} & \forall x \in \mathfrak{a}^{\circ}. \phi^{\circ}(x) \wedge \forall \langle f, \mathfrak{c} \rangle \in \mathfrak{a}^* \, [\{e\}f \Vdash \phi(\mathfrak{c})] \\
e \Vdash \exists x \in \mathfrak{a}. \phi(x) & \text{iff} & \exists \mathfrak{c}[\langle (e)_0, \mathfrak{c} \rangle \in \mathfrak{a}^* \wedge (e)_1 \Vdash \phi(\mathfrak{c})] \\
e \Vdash \forall x \, \phi & \text{iff} & \forall \mathfrak{d} \, [e \Vdash \phi(\mathfrak{d})] \\
e \Vdash \exists x \, \phi & \text{iff} & \exists \mathfrak{d} \, [e \Vdash \phi(\mathfrak{d})] \\
e \Vdash \perp & \text{iff} & \perp
\end{array}$$

We write $V^* \vDash \phi$ iff there is some $e \in \omega$ such that $e \Vdash \phi$. Note that unbounded and bounded quantifiers are treated as semantically different types of quantifiers in our definition of realizability. This is not merely for cosmetical reasons: In giving realizers for the axioms of **CZF**, it will be crucial to find a witness of a bounded existential statement $\exists x \in \mathfrak{b}. \phi(x)$ not somewhere in the **class** V^* but rather in the **set** $\mathfrak{b}$.

4.4 A simple Completeness Theorem

Owing to our reference to truth within the system, we can now easily prove the completeness of realizability:

Theorem 4.6 (Completeness): If $V^* \vDash \theta$, then **CZF** $\vdash \theta^{\circ}$.

Before proving the theorem, we need the following notion of standard representatives of sets of **CZF** within V^*:

Definition 4.7: We define for every set x its *standard representative* $\hat{x}$ in V^* by recursion:

$$\hat{x} = \langle x, \{\langle 0, \hat{u} \rangle : u \in x\} \rangle.$$

The property $(\hat{x})^{\circ} = x$ is shown by simple recursion.

We are now ready to give the proof of the completeness-theorem:

Proof of Theorem 4.6: By induction on θ. The base cases are clear by definition and so are the cases of implication, negation, unbounded existential and bounded universal quantification.

Suppose $e \Vdash \forall x\, \phi(x)$. Thus, for any $\mathfrak{a} \in V^*$, $e \Vdash \phi(\mathfrak{a})$. By induction hypothesis, **CZF** $\vdash \left(\phi(\mathfrak{a})\right)^\circ$. In particular, for any set x, we obtain that **CZF** proves $(\phi(\hat{x}))^\circ$ and thus $\phi^\circ(x)$, showing the case for unbounded universal quantification.

For restricted existential quantification, suppose $e \Vdash \exists x \in \mathfrak{a}.\, \phi(x)$. This means, there is $\mathfrak{d} \in V^*$ such that $\langle (e)_0, \mathfrak{d} \rangle \in \mathfrak{a}^* \wedge (e)_1 \Vdash \phi(\mathfrak{d})$. By definition, $\mathfrak{d}^\circ \in \mathfrak{a}^\circ$ and by induction hypothesis, $\phi^\circ(\mathfrak{d}^\circ)$. Altogether we have $\left(\exists x \in \mathfrak{a}.\, \phi(x)\right)^\circ \equiv \exists x \in \mathfrak{a}^\circ.\, \phi^\circ(x)$. ■

4.5 Soundness Theorem

The next sections we will spend with proving the soundness theorem:

Ⓜ**Theorem 4.8 (Soundness)**: For every theorem θ of **CZF**, there exists a closed application term t, such that **CZF** $\vdash (t \Vdash \theta)$.

All we have to do is to give realizing terms for the underlying axioms and inference rules of **HPL** with equality and then do the same for the axioms of **CZF**.

4.5.1 Realizing equality

Lemma 4.9 (closure):

(1) $\mathfrak{b} \in V^*_\alpha \to \exists \beta \in \alpha. \forall \mathfrak{c} \left(V^* \vDash \mathfrak{c} \in \mathfrak{b} \to \mathfrak{c} \in V^*_\beta\right)$
(2) $(\mathfrak{a} \in V^*_\alpha \wedge V^* \vDash \mathfrak{a} = \mathfrak{b}) \to \mathfrak{b} \in V^*_\alpha$

Proof: By simultaneous induction. To prove (1), the inductive hypothesis is that for all $\beta \in \alpha$,

$$\left(\mathfrak{d} \in V^*_\beta \wedge V^* \vDash \mathfrak{c} = \mathfrak{d}\right) \to \mathfrak{c} \in V^*_\beta.$$

For $\mathfrak{b} \in V^*_\alpha$, there is, by definition, some $\beta \in \alpha$, s.t. $\mathfrak{b}^* \subseteq \omega \times V^*_\beta$. Let $e \Vdash \mathfrak{c} \in \mathfrak{b}$. This means,

$$\exists \mathfrak{d} [\langle (e)_0, \mathfrak{d} \rangle \in \mathfrak{b}^* \wedge (e)_1 \Vdash \mathfrak{c} = \mathfrak{d}].$$

By the inductive hypothesis, we conclude $\mathfrak{c} \in V^*_\beta$.

Now for (2), assume (1) for α. Let $\mathfrak{a} \in V^*_\alpha$ and choose $\beta \in \alpha$ such that $\forall \mathfrak{c}\left(V^* \vDash \mathfrak{c} \in \mathfrak{b} \to \mathfrak{c} \in V^*_\beta\right)$. If $e \Vdash \mathfrak{a} = \mathfrak{b}$, then for all $\langle f, \mathfrak{d} \rangle \in \mathfrak{b}^*$, $(ef)_1 \Vdash \mathfrak{d} \in \mathfrak{a}$, which gives $\mathfrak{d} \in V^*_\beta$. Altogether, $\mathfrak{b}^* \subseteq \omega \times V^*_\beta$. Also $e \Vdash \mathfrak{a} = \mathfrak{b}$ entails $\mathfrak{a}^\circ = \mathfrak{b}^\circ$ and hence $\mathfrak{b} \in V^*_\alpha$. ■

Theorem 4.10: There are $\mathbf{i}_r, \mathbf{i}_s, \mathbf{i}_t, \mathbf{i}_0, \mathbf{i}_1 \in \omega$ such that for all $\mathfrak{a}, \mathfrak{b}, \mathfrak{c} \in V^*$,

1) $\mathbf{i}_r \Vdash \mathfrak{a} = \mathfrak{a}$
2) $\mathbf{i}_s \Vdash \mathfrak{a} = \mathfrak{b} \to \mathfrak{b} = \mathfrak{a}$
3) $\mathbf{i}_t \Vdash (\mathfrak{a} = \mathfrak{b} \wedge \mathfrak{b} = \mathfrak{c}) \to \mathfrak{a} = \mathfrak{c}$
4) $\mathbf{i}_0 \Vdash (\mathfrak{a} = \mathfrak{b} \wedge \mathfrak{b} \in \mathfrak{c}) \to \mathfrak{a} \in \mathfrak{c}$
5) $\mathbf{i}_1 \Vdash (\mathfrak{a} = \mathfrak{b} \wedge \mathfrak{c} \in \mathfrak{a}) \to \mathfrak{c} \in \mathfrak{b}$

Moreover, for each formula $\phi(x)$ there exists $\mathbf{i}_\phi \in \omega$ such that for all $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}_1, \dots, \mathfrak{c}_n \in V^*$

$$\mathbf{i}_\phi \Vdash \big(\mathfrak{a} = \mathfrak{b} \wedge \phi(\mathfrak{a}, \mathfrak{c}_1, \dots, \mathfrak{c}_n)\big) \to \phi(\mathfrak{b}, \mathfrak{c}_1, \dots, \mathfrak{c}_n).$$

Proof: It goes without saying that throughout the proof we need not bother about truth in **CZF**, as all the corresponding axioms of identity hold in **CZF**. For 1 we use the Fixed-Point Theorem to define $\mathbf{i}_\phi$ to be such that $\mathbf{i}_\phi \simeq \mathbf{p}(\lambda y.\mathbf{p}y\mathbf{i}_\phi)(\lambda y.\mathbf{p}y\mathbf{i}_\phi)$. Let $\mathfrak{a} \in V^*_\alpha$ and assume $\mathbf{i}_\phi \Vdash \mathfrak{b} = \mathfrak{b}$ for all $\mathfrak{b} \in V^*_\beta$ with $\beta \in \alpha$. If $\langle f, \mathfrak{b}\rangle \in \mathfrak{a}$, then $(\mathbf{i}_\phi)_0 f \simeq \mathbf{p}f\mathbf{i}_\phi$. As $\langle(\mathbf{p}f\mathbf{i}_\phi)_0, \mathfrak{b}\rangle \in \mathfrak{a}$ and $(\mathbf{p}f\mathbf{i}_\phi)_1 \Vdash \mathfrak{b} = \mathfrak{b}$, we have shown $(\mathbf{i}_\phi)_0 f \Vdash \mathfrak{b} \in \mathfrak{a}$ and thus $\mathbf{i}_\phi \Vdash \mathfrak{a} = \mathfrak{a}$.

For 2, let $\mathbf{i}_s \equiv \lambda x.\mathbf{p}(\mathbf{r}x)(\mathbf{l}x)$. This term swaps entries of a pair, which is all we need.

Axioms 3 and 4 are shown by simultaneous induction on the structure of V^*. We first need to define application terms τ_1-τ_6 as follows:

$$\begin{aligned}
\tau_1(x,y) &\equiv \mathbf{r}(\mathbf{l}(\mathbf{l}y)x)\\
\tau_2(y) &\equiv \mathbf{l}(\mathbf{r}y)\\
\tau_3(x,y) &\equiv \mathbf{r}(\mathbf{r}(\mathbf{r}y)x)\\
\tau_4(x,y) &\equiv \mathbf{r}(\mathbf{l}y)\big(\mathbf{l}(\mathbf{r}(\mathbf{r}y)x)\big)\\
\tau_5(x,y) &\equiv \tau_2(y)\big(\mathbf{l}(\mathbf{l}(\mathbf{l}y)x)\big)\\
\tau_6(y) &\equiv \mathbf{p}(\mathbf{l}y)\big(\mathbf{r}(\mathbf{r}y)\big)
\end{aligned}$$

Let σ_1 and σ_2 be defined as follows:

$$\begin{aligned}
&\sigma_1(a,b) \equiv \lambda y.\mathbf{p}\left(\lambda x.a\big(\mathbf{p}\tau_1(x,y)\tau_5(x,y)\big)\right)\left(\lambda x.a\big(\mathbf{p}\tau_4(x,y)\tau_5(x,y)\big)\right),\\
&\sigma_2(a,b) \equiv \lambda y.\mathbf{p}\left(\tau_2(y)\big(b\tau_6(y)\big)\right).
\end{aligned}$$

By Lemma 4.3 **(Double recursion)** on double recursion, there are $\mathbf{i}_t$ and $\mathbf{i}_0$ such that $\mathbf{i}_t \simeq \sigma_1(\mathbf{i}_t, \mathbf{i}_0)$ and $\mathbf{i}_0 \simeq \sigma_2(\mathbf{i}_t, \mathbf{i}_0)$.

For 3 the induction hypothesis is that for all $\mathfrak{c} \in V^*_\alpha$,

$$\mathbf{i}_0 \Vdash (\mathfrak{a} = \mathfrak{b} \wedge \mathfrak{b} \in \mathfrak{c}) \to \mathfrak{a} \in \mathfrak{c}.$$

Let $\mathfrak{e} \in VV^*_\alpha$, $\langle g, \mathfrak{g}\rangle \in \mathfrak{d}$ and h such that

$h \Vdash (\mathfrak{d} = \mathfrak{e} \wedge \mathfrak{e} = \mathfrak{f})$

then $((h)_0)_0 g \Vdash \mathfrak{g} \in \mathfrak{e}$

then $\exists \mathfrak{i}\, \langle (((h)_0)_0 g)_0, \mathfrak{i}\rangle \in \mathfrak{b}$ and $(((h)_0)_0 g)_1 \Vdash \mathfrak{g} = \mathfrak{i}$

then $\underbrace{((h)_0)_1(((h)_0)_0 g)_0}_{i} \Vdash \mathfrak{i} \in \mathfrak{f}$, since $(h)_1 \Vdash \mathfrak{e} = \mathfrak{f}$, where by the closure lemma (2), we may assume $\mathfrak{f} \in V^*_\alpha$.

then $\mathbf{i}_0\mathbf{p}((((h)_0)_0 g)_1 i) \Vdash \mathfrak{g} \in \mathfrak{f}$ by induction hypothesis

As $(((h)_0)_0 g)_1 \equiv \tau_1(g,h)$ and $i \equiv \tau_5(g,h)$ we have shown $\mathbf{i_0}\big(\mathbf{p}\tau_1(g,h)\tau_5(g,h)\big) \Vdash \mathfrak{g} \in \mathfrak{f}$. A similar argument shows $\mathbf{i_0}\big(\mathbf{p}\tau_4(g,h)\tau_5(g,h)\big) \Vdash \mathfrak{g} \in \mathfrak{d}$ for all $\langle g, \mathfrak{g}\rangle \in \mathfrak{f}$. Altogether, we have $\mathbf{i}_t \simeq \sigma_2(\mathbf{i}_t, \mathbf{i_0}) \Vdash \mathfrak{d} = \mathfrak{f}$.

For (4), the inductive hypothesis is that for all $\beta \in \alpha$ and $\mathfrak{e} \in V^*_\beta$,

$$\mathbf{i}_t \Vdash (\mathfrak{d} = \mathfrak{e} \wedge \mathfrak{e} = \mathfrak{f}) \rightarrow \mathfrak{d} = \mathfrak{f}.$$

Let $\mathfrak{c} \in V^*_\alpha$ and h such that

$$\begin{array}{lll}
 & h \Vdash (\mathfrak{a} = \mathfrak{b} \wedge \mathfrak{b} \in \mathfrak{c}) & \\
\text{iff} & (h)_0 \Vdash \mathfrak{a} = \mathfrak{b} & \text{and } (h)_1 \Vdash \mathfrak{b} \in \mathfrak{c} \\
\text{iff} & (h)_0 \Vdash \mathfrak{a} = \mathfrak{b} & \text{and } \exists \mathfrak{d}\, \langle ((h)_1)_0, \mathfrak{d}\rangle \in \mathfrak{c} \wedge ((h)_1)_1 \Vdash \mathfrak{b} = \mathfrak{d}, \text{ where by the closure lemma (1),} \\
 & & \text{we may assume } \mathfrak{b} \in V^*_\beta \text{ for some } \beta \in \alpha \\
\text{then} & \mathbf{i_0}(\mathbf{p}(h)_0((h)_1)_1) \Vdash \mathfrak{a} = \mathfrak{d} & \qquad\qquad \text{by inductive hypothesis} \\
\text{then} & \mathbf{p}(((h)_1)_0)\big(\mathbf{i_0}(\mathbf{p}(h)_0((h)_1)_1)\big) \Vdash \mathfrak{a} \in \mathfrak{c} &
\end{array}$$

We can compute

$$\mathbf{p}(((h)_1)_0)\big(\mathbf{i_0}(\mathbf{p}(h)_0((h)_1)_1)\big) \simeq \mathbf{p}\big(\tau_2(h)\big)\Big(\mathbf{i_0}\big(\tau_6(h)\big)\Big) \simeq \sigma_1(\mathbf{i}_t, \mathbf{i_0})h \simeq \mathbf{i_0}h,$$

thus, we have shown $\mathbf{i_0}h \Vdash \mathfrak{a} \in \mathfrak{c}$. As h was arbitrary, $\mathbf{i_0} \Vdash (\mathfrak{a} = \mathfrak{b} \wedge \mathfrak{b} \in \mathfrak{c}) \rightarrow \mathfrak{a} \in \mathfrak{c}$.

Finally, for 5, the term

$$\mathbf{i_1} \equiv \lambda x. \mathbf{i_0}\Big(\mathbf{p}\big(\mathbf{r}(\mathbf{r}x)\big)\Big(\mathbf{l}(\mathbf{l}x)\big(\mathbf{l}(\mathbf{r}x)\big)\Big)\Big) \simeq \mathbf{i_0}\big(\mathbf{p}((h)_1)_1(((h)_0)_0((h)_1)_0)\big)$$

does the job: Let $h \Vdash (\mathfrak{a} = \mathfrak{b} \wedge \mathfrak{c} \in \mathfrak{a})$, let $\langle ((h)_1)_0, \mathfrak{d}\rangle \in \mathfrak{a}$ such that $((h)_1)_1 \Vdash \mathfrak{d} = \mathfrak{c}$. Then

$$((h)_0)_0((h)_1)_0 \Vdash \mathfrak{d} \in \mathfrak{b}.$$

With the properties of $\mathbf{i_0}$, we have that $\mathbf{i_1}h \Vdash \mathfrak{c} \in \mathfrak{b}$.

The term $\mathbf{i}_\phi$ is constructed by recursion on ϕ. The inductive steps are easy and the base cases are provided by $\mathbf{i}_r, \mathbf{i}_s, \mathbf{i}_t, \mathbf{i_0}$ and $\mathbf{i_1}$. ■

4.5.2 Soundness Theorem for **HPL**

Theorem 4.11: If **HPL** $\vdash \phi$, then $V^* \vDash \overline{\forall}\phi$.

Proof: Let us show soundness for some of the axioms and inference rules. Again, truth in **CZF** is guaranteed automatically.

For HPL8, we claim that $h \equiv \lambda x \lambda y \lambda z. \mathbf{d}\big(y(\mathbf{r}x)\big)\big(z(\mathbf{r}x)\big)(\mathbf{l}x)$ does the job. Assume $e \Vdash \phi \vee \psi$. Then, either $(e)_0 = \mathbf{0} \wedge (e)_1 \Vdash \phi$ or $(e)_0 = \mathbf{1} \wedge (e)_1 \Vdash \psi$. Assume, $f \Vdash \phi \rightarrow \chi$ and $g \Vdash \psi \rightarrow \chi$. Then

$$hefg \simeq \mathbf{d}\big(f(\mathbf{r}e)\big)\big(g(\mathbf{r}e)\big)(\mathbf{l}e)\mathbf{0}.$$

By the properties of $\mathbf{d}$,

$$hefg \simeq \begin{Bmatrix} f(\mathbf{r}e), & \text{, if } (e)_0 = \mathbf{0} \\ g(\mathbf{r}e) & \text{, if } (e)_0 \neq \mathbf{0} \end{Bmatrix} \Vdash \phi.$$

For HPL9, let $e \Vdash \phi \to \psi$, $f \Vdash (\phi \to \neg\psi)$ and assume $g \Vdash \phi$. Thus, $ef \Vdash \psi$, $eg \Vdash \neg\psi$. This shows $eg(ef) \Vdash \perp$, and finally $\neg(g \vdash \phi)$ for all $g \in \omega$.

For HPL13, let $e \Vdash \forall u \in \mathfrak{a}.\, \phi(u)$. This means, for all $\langle f, \mathfrak{c}\rangle \in \mathfrak{a}$, $ef \Vdash \phi(\mathfrak{c})$. If $h \Vdash \mathfrak{u} \in \mathfrak{a}$, then there is some $\langle (h)_0, \mathfrak{c}\rangle \in \mathfrak{a}$ with $(h)_1 \Vdash \mathfrak{c} = \mathfrak{u}$. This implies $\mathbf{i}_{\boldsymbol{\phi}}\big(\mathbf{p}((h)_1)(e(h)_0)\big)e(h)_0 \Vdash \phi(\mathfrak{u})$.

Conversely, if $e \Vdash \forall u[u \in a \to \phi(u)]$ and $\langle f, \mathfrak{c}\rangle \in \mathfrak{a}$, then $\mathbf{p}f\mathbf{i}_{\boldsymbol{r}} \Vdash \mathfrak{c} \in \mathfrak{a}$. Thus, $e(\mathbf{p}f\mathbf{i}_{\boldsymbol{r}}) \Vdash \phi(\mathfrak{c})$.
Altogether, we have shown that $\mathbf{p}\left(\lambda e \lambda h.\, \mathbf{i}_{\boldsymbol{\phi}}\left(\mathbf{p}(\mathbf{r}h)\big(e(\mathbf{l}h)\big)\right) e(\mathbf{l}h)\right)\big(\lambda e \lambda f.\, e(\mathbf{p}f\mathbf{i}_{\boldsymbol{r}})\big)$ realizes this axiom.

For EI, assume $e \Vdash \forall x(\phi(x) \to \psi)$. This shows that for all $\mathfrak{a} \in V^*$, $e \Vdash \phi(\mathfrak{a}) \to \psi$. Assume $g \Vdash \exists x \phi$, i.e. there is $\mathfrak{b} \in V^*$ such that $g \Vdash \phi(\mathfrak{b})$. Altogether, $eg \Vdash \psi$. We have seen that we can take $\lambda x.\, x$ to realize EI.

■

4.5.3 Realizing the set axioms of **CZF**

In the course of realizing the set axioms we will often be in the situation to construct a witness to a formula in V^*. We will therefore apply the following result: Given a subset $A \subseteq V^*$, we define

$$A^c = \{\mathfrak{c}^\circ : \exists k\, \langle k, \mathfrak{c}\rangle \in A\}.$$

Lemma 4.12: Let A be a subset of $\omega \times V^*$. Then $\langle A, A^c\rangle$ and any $\langle A, B\rangle$ with $B \supseteq A^c$ is in V^*.

Proof: For each $\langle k, \mathfrak{c}\rangle \in A$ there is some β such that $\langle k, \mathfrak{c}\rangle \in \omega \times V^*_\alpha$. Using collection, form the set C of all such βs and let $\alpha = \bigcup C$. Then $A \subseteq \omega \times V^*_\alpha$. Also, for each $\langle k, \mathfrak{c}\rangle \in A$, $\mathfrak{c}^\circ \in A^c$ and hence $\langle A, A^c\rangle \in V^*_\alpha$. Increasing the set A^c does not change this. ■

Extensionality

Let $\mathbf{i}_{\boldsymbol{r}}$ be as in the last section and

$$e \equiv \lambda y.\, \mathbf{p}\big(\lambda x.\, \mathbf{l}y(\mathbf{p}x\mathbf{i}_{\boldsymbol{r}})\big)\big(\lambda x.\, \mathbf{r}y(\mathbf{p}x\mathbf{i}_{\boldsymbol{r}})\big).$$

If $h \Vdash \forall z(z \in \mathfrak{a} \leftrightarrow z \in \mathfrak{b})$ and $\langle f, \mathfrak{d}\rangle \in \mathfrak{a}$. As $\mathbf{p}f\mathbf{i}_{\boldsymbol{r}} \Vdash \mathfrak{d} \in \mathfrak{a}$, we have that $(eh)_0 f \simeq (h)_0(\mathbf{p}f\mathbf{i}_{\boldsymbol{r}}) \Vdash \mathfrak{d} \in \mathfrak{b}$. The other direction is symmetric and truth in **CZF** is apparent.

Pairing

We need to find a realizer e such that given $\mathfrak{a}, \mathfrak{b} \in V^*$ there is some $\mathfrak{p}$ with $e \Vdash \mathfrak{a} \in \mathfrak{p} \wedge \mathfrak{b} \in \mathfrak{p}$. We define $\mathfrak{p}$ by $\mathfrak{p}^\circ = \{\mathfrak{a}^\circ, \mathfrak{b}^\circ\}$ and $\mathfrak{p}^* = \{\langle 0, \mathfrak{a}\rangle, \langle 0, \mathfrak{b}\rangle\}$. By Lemma 4.12, $\mathfrak{p} \in V^*$. Obviously,

$$\forall w\big(w \in \mathfrak{p}^\circ \leftrightarrow (w = \mathfrak{a}^\circ \vee w = \mathfrak{b}^\circ)\big),$$

and we easily see that $e \equiv \mathbf{p}(\mathbf{p}\mathbf{0}\mathbf{i}_{\mathbf{r}})(\mathbf{p}\mathbf{0}\mathbf{i}_{\mathbf{r}})$ does the job of realizing the axiom.

Union

Invoking Lemma 4.12, we define for $\mathfrak{a} \in V^*$, the set $\mathrm{Un}(\mathfrak{a}) \in V^*$ by

$$\begin{aligned}\mathrm{Un}(\mathfrak{a})^\circ &= \bigcup \mathfrak{a}^\circ,\\ \mathrm{Un}(\mathfrak{a})^* &= \{\langle h, \mathfrak{y}\rangle : \exists \langle f, \mathfrak{x}\rangle \in \mathfrak{a}^*. \langle h, \mathfrak{y}\rangle \in \mathfrak{x}\}.\end{aligned}$$

Clearly, for all $\mathfrak{x} \in V^*$,

$$\mathfrak{x}^\circ \in \mathrm{Un}(\mathfrak{a})^\circ \leftrightarrow \exists w(\mathfrak{x}^\circ \in w \wedge w \in \mathfrak{a}^\circ).$$

Let $h \Vdash \mathfrak{y} \in \mathfrak{x} \wedge \mathfrak{x} \in \mathfrak{a}$. This means

$$\begin{aligned}&\exists \mathfrak{c}[\langle ((h)_0)_0, \mathfrak{c}\rangle \in \mathfrak{x} \wedge ((h)_0)_1 \Vdash \mathfrak{y} = \mathfrak{c}],\\ &\exists \mathfrak{d}[\langle ((h)_1)_0, \mathfrak{d}\rangle \in \mathfrak{a} \wedge ((h)_1)_1 \Vdash \mathfrak{x} = \mathfrak{d}].\end{aligned}$$

Putting things together, we have that $\langle ((h)_0)_0, \mathfrak{c}\rangle \in \mathrm{Un}(\mathfrak{a})$ and thus

$$\lambda x. \mathbf{p}(\mathbf{ll}x)(\mathbf{rl}x) \Vdash \forall y\big(\exists w(y \in w \wedge w \in \mathfrak{a}) \rightarrow y \in \mathrm{Un}(\mathfrak{a})\big).$$

Empty set

The representative of the empty set is just $\widehat{\emptyset} = \langle \emptyset, \emptyset\rangle \in V_1^*$. Its property is realized by every application term and obviously $\neg(\mathfrak{y}^\circ \in \langle \emptyset, \emptyset\rangle^\circ)$.

Infinity

To represent ω in V^*, we define for $m \in \omega + 1$,

$$\overline{m} = \langle n, \{\langle k, \overline{k}\rangle : k < n\}\rangle,$$

(the base case $\overline{0} = \langle \emptyset, \emptyset\rangle = \widehat{\emptyset}$ is implicit). Then $\overline{\omega} \in V^*_{\omega+1}$. Of course, for all $n \in \omega + 1$, $(\overline{m})^\circ = m$.

Lemma 4.13: For all $n \in \omega + 1$ there are realizers e_1, e_2 and e_3 such that

1) $e_1(k) \Vdash \overline{k} \in \overline{n}$ for all $k < n$.
2) $e_2 \Vdash \overline{m} \subseteq \overline{n}$, for all $m \leq n$.
3) $e_3(n) \Vdash \forall y \in \overline{n+1}. (y \in \overline{n} \vee y = \overline{n})$, for $n \in \omega$.

Proof: Let $\mathbf{i}_r$ be the realizer with $\mathbf{i}_r \Vdash \mathfrak{a} = \mathfrak{a}$ for all $\mathfrak{a} \in V^*$ from Theorem 4.10.

1) $e_1(k) \equiv \mathbf{p}k\mathbf{i}_r \Vdash \overline{k} \in \overline{n}$, since clearly $\left(\overline{k}\right)^\circ = k \in n = (\overline{n})^\circ$ and as $\langle k, \overline{k}\rangle \in \overline{n}^*$.
2) Again, $\left(\overline{k}\right)^\circ = k \subseteq n = (\overline{n})^\circ$. The realizer we need is $e_2 \equiv \lambda x. e_1(x)$ by 1).
3) Truth in **CZF** is again easy. Let $\mathbf{d}$ be the operator with $\mathbf{d}abmn \simeq a$ iff $m = n$ and $\mathbf{d}abmn \simeq b$ iff $m \neq n$. Then, by 1), we have that $e_3(n) \equiv \lambda x. \mathbf{d}\mathbf{i}_r(\mathbf{p}k\mathbf{i}_r)xn$ realizes $\forall y \in \overline{n+1}. (y \in \overline{n} \vee y = \overline{n})$. ■

With this lemma it is now easy to realize $\overline{m+1} = s(\overline{m})$. Note that for n “to be a successor of m” is actually the formula

$$\begin{aligned} & n = s(m) \\ \leftrightarrow\ & n = m \cup \{m\} \\ \leftrightarrow\ & [n \subseteq m \cup \{m\}] \wedge [n \supseteq m \cup \{m\}] \\ \leftrightarrow\ & \forall y \in n\, (y \in m \vee y = m) \wedge \forall y \in m\, (y \in n) \wedge m \in n. \end{aligned}$$

Hence, gathering all the realizers in the proof of the lemma,

$$\mathscr{s}(m) = \mathbf{p}\big[\mathbf{p}\big(e_3(m)\big)e_2\big][e_1(m)] \Vdash \overline{m+1} = s(\overline{m}).$$

Realizing the axiom of infinity is now easy: Truth in **CZF** is trivial and

$$\lambda n.\, \mathscr{s}(n) \Vdash \forall n \in \overline{\omega}.\, s(n) \in \overline{\omega},$$
$$\lambda n.\, \mathbf{di}_r\big(\mathbf{p}(n-1)\mathscr{s}(n-1)\big)n0 \Vdash \forall n \in \overline{\omega}\, \big(n = 0 \vee \exists m \in \omega.\, n = s(m)\big).$$

Bounded Separation

Given $\mathfrak{a} \in V^*$ and any formula ϕ, we define $\mathrm{Sep}_\phi(\mathfrak{a})$ by

$$\mathrm{Sep}_\phi(\mathfrak{a})^* = \{\langle \mathbf{p} f g, \mathfrak{c}\rangle : \langle f, \mathfrak{c}\rangle \in \mathfrak{a}^* \wedge g \Vdash \phi(\mathfrak{c})\},$$
$$\mathrm{Sep}_\phi(\mathfrak{a})^\circ = \{x \in \mathfrak{a}^\circ : \phi^\circ(x)\}.$$

This definition needs some justification: First note that $\mathrm{Sep}_\phi(\mathfrak{a})^*$ is indeed a set, as quantification is bounded. For $\langle \mathbf{p} f g, \mathfrak{c}\rangle \in \mathrm{Sep}_\phi(\mathfrak{a})^*$, we have that $g \Vdash \phi(\mathfrak{c})$ and hence by Completeness, $\phi^\circ(\mathfrak{c}^\circ)$. By definition of V^*, $\mathfrak{c}^\circ \in \mathfrak{a}^\circ$ and hence $\mathfrak{c}^\circ \in \mathrm{Sep}_\phi(\mathfrak{a})^\circ$. By Lemma 4.12 we can conclude that $\mathrm{Sep}_\phi(\mathfrak{a}) \in V^*$.

We need to find e, h with

$$e \Vdash \mathfrak{x} \in \mathrm{Sep}_\phi(\mathfrak{a}) \rightarrow \mathfrak{x} \in \mathfrak{a} \wedge \phi(\mathfrak{x}),$$
$$h \Vdash \mathfrak{x} \in \mathfrak{a} \wedge \phi(\mathfrak{x}) \rightarrow \mathfrak{x} \in \mathrm{Sep}_\phi(\mathfrak{a}).$$

As for truth, $\overset{\circ}{\mathfrak{x}} \in \mathrm{Sep}_\phi(\mathfrak{a})^\circ$ is equivalent to $\mathfrak{x}^\circ \in \mathfrak{a}^\circ \wedge \phi^\circ(\mathfrak{x}^\circ)$. To find e, note that $k \Vdash \mathfrak{x} \in \mathrm{Sep}_\phi(\mathfrak{a})$ means that $(k_0) = \mathbf{p} f g$ and $(k)_1 \Vdash \mathfrak{c} = \mathfrak{x}$ for some $\langle f, \mathfrak{c}\rangle \in \mathfrak{a}^*$ and $g \Vdash \phi(\mathfrak{c})$. Hence, $\mathbf{p} f (k)_1 \Vdash \mathfrak{x} \in \mathfrak{a}$ and with $\mathbf{i}_\phi$ from Theorem 4.10, $\mathbf{i}_\phi(\mathbf{p}(k)_1 g) \Vdash \phi(\mathfrak{x})$.

On the other hand, $\mathbf{p} k m \Vdash \mathfrak{x} \in \mathfrak{a} \wedge \phi(\mathfrak{x})$ means that $\langle (k)_0, \mathfrak{c}\rangle \in \mathfrak{a}^*$ and $(k)_1 \Vdash \mathfrak{x} = \mathfrak{c}$ for some $\mathfrak{c}$. As before, $\mathbf{i}_\phi(\mathbf{p}(k)_1 m) \Vdash \phi(\mathfrak{c})$. All of this shows that we can take e and h to be

$$e \equiv \lambda k.\, \mathbf{p}\big(\mathbf{p}(\mathbf{ll}x)(\mathbf{rl}x)\big)\Big(\mathbf{i}_\phi\big(\mathbf{p}(\mathbf{r}x)(\mathbf{rl}x)\big)\Big),$$
$$h \equiv \lambda x.\, \mathbf{p}\left(\mathbf{p}(\mathbf{ll}x)\left(\mathbf{i}_\phi(\mathbf{p}(\mathbf{rl}x)\mathbf{r}x)\right)\right)(\mathbf{rl}x).$$

Strong collection

Let $\mathfrak{a} \in V^*$ and suppose, $e \Vdash \forall x \in \mathfrak{a}.\, \exists y.\, \phi(x, y)$. In particular, $\forall x \in \mathfrak{a}^\circ.\, \exists y.\, \phi^\circ(x, y)$. Applying strong collection in **CZF** to the latter formula yields a set E such that

$$\forall x \in \mathfrak{a}^\circ.\, \exists y \in E.\, \phi^\circ(x, y) \wedge \forall y \in E.\, \exists x \in \mathfrak{a}^\circ.\, \phi^\circ(x, y).$$

Also, $e \Vdash \forall x \in \mathfrak{a}.\, \exists y.\, \phi(x, y)$ means that

$$\forall \langle f, \mathfrak{x} \rangle \in \mathfrak{a}^*.\, \exists \langle f, \mathfrak{d} \rangle.\, \big(ef \Vdash \phi(\mathfrak{x}, \mathfrak{d})\big).$$

Again, invoking strong collection in **CZF**, there is a set D such that

$$\big[\forall \langle f, \mathfrak{x} \rangle \in \mathfrak{a}^*.\, \exists \langle f, \mathfrak{d} \rangle \in D.\, \big(ef \Vdash \phi(\mathfrak{x}, \mathfrak{d})\big)\big] \land \big[\forall \langle f, \mathfrak{x} \rangle \in D.\, \exists \langle f, \mathfrak{d} \rangle \in \mathfrak{a}^*.\, \big(ef \Vdash \phi(\mathfrak{x}, \mathfrak{d})\big)\big].$$

We claim that if we define $\mathrm{Col}_{\phi(x,y)}$ by

$$\mathrm{Col}^*_{\phi(x,y)} = D,$$
$$\mathrm{Col}^\circ_{\phi(x,y)} = E \cup D^c,$$

this will do the job. Indeed, we have

$$\lambda f.\, \mathbf{p} f(ef) \Vdash \forall x \in \mathfrak{a}.\, \exists y \in \mathrm{Col}_{\phi(x,y)}.\, \phi(x, y),$$
$$\lambda f.\, \mathbf{p} f(ef) \Vdash \forall y \in \mathrm{Col}_{\phi(x,y)}.\, \exists x \in \mathfrak{a}.\, \phi(x, y),$$

as for the first line

$$\forall x \in \mathfrak{a}^\circ.\, \exists y \in E.\, \phi^\circ(x, y)$$

and for the second,

$$\forall y \in E.\, \exists x \in \mathfrak{a}^\circ.\, \phi^\circ(x, y) \quad \text{and} \quad \forall y \in D^c.\, \exists x \in \mathfrak{a}^\circ.\, \phi^\circ(x, y),$$

where all facts follow from the definitions of E and D and completeness.

Finally, the axiom is realized iff

$$\forall x \in \mathfrak{a}^\circ.\, \exists y.\, \phi^\circ(x, y) \to \exists C[\forall x \in \mathfrak{a}^\circ.\, \exists y \in C.\, \phi^\circ(x, y) \land \forall y \in C.\, \exists y \in \mathfrak{a}^\circ.\, \phi^\circ(x, y)],$$

but this is clear from strong collection in **CZF**.

Subset collection

Let $\mathfrak{a}, \mathfrak{b} \in V^*$ ϕ be any formula.

Set

$$B = \{\langle \mathbf{p} ef, \mathfrak{d} \rangle : e, f \in \omega \land ef \downarrow \land \langle (ef)_0, \mathfrak{d} \rangle \in \mathfrak{b}^*\}$$

and let $\psi(e, f, \mathfrak{c}, \mathfrak{u}, z)$ be the formula

$$\mathfrak{u} \in V^* \land e, f \in \omega \land ef \downarrow \land \exists \mathfrak{d}\, \big(\langle \mathbf{p} ef, \mathfrak{d} \rangle = z \land \langle (ef)_0, \mathfrak{d} \rangle \in \mathfrak{b}^* \land (ef)_1 \Vdash \phi(\mathfrak{c}, \mathfrak{d}, \mathfrak{u})\big).$$

By subset collection in **CZF**, there exist a set D such that

$$\forall \mathfrak{u}\, \forall e \left[\forall \langle f, \mathfrak{c}\rangle \in \mathfrak{a}^*. \exists z \in B. \psi(e, f, \mathfrak{c}, \mathfrak{u}, z) \rightarrow \exists w_{\mathfrak{u}} \\ \in D\big(\forall \langle f, \mathfrak{c}\rangle \in \mathfrak{a}^*. \exists z \in w_{\mathfrak{u}}. \psi(e, f, \mathfrak{c}, \mathfrak{u}, z) \wedge \forall z \in w_{\mathfrak{u}}. \exists \langle f, \mathfrak{c}\rangle \in \mathfrak{a}^* \in a. \psi(e, f, \mathfrak{c}, \mathfrak{u}, z)\big)\right].$$

Note that we may assume that $D \subseteq B$. Using subset collection in **CZF**, there is a set C such that

$$\forall \mathfrak{u}\big[\forall x \in \mathfrak{a}^\circ. \exists y \in \mathfrak{b}^\circ. \phi^\circ(x, y, \mathfrak{u}^\circ) \rightarrow \exists v_{\mathfrak{u}} \in C\big(\forall x \in \mathfrak{a}^\circ. \exists y \in v_{\mathfrak{u}}. \phi(x, y, \mathfrak{u}^\circ) \wedge \forall y \in v_{\mathfrak{u}}. \exists x \in \mathfrak{a}^\circ. \phi^\circ(x, y, \mathfrak{u}^\circ)\big)\big].$$

We can now define the witness:

$$\begin{aligned} \mathcal{W} &= \{\langle v \cup w^c, w\rangle : v \in C \wedge w \in D\} \\ E &= C \cup \{\mathfrak{z}^\circ : \mathfrak{z} \in \mathcal{W}\} \\ E^+ &= \{\langle 0, \mathfrak{z}\rangle : \mathfrak{z} \in \mathcal{W}\} \\ \mathfrak{e} &= \langle E, E^+\rangle \end{aligned}$$

First, we need to show that indeed $\mathfrak{e} \in V^*$. First note, that clearly, $\mathcal{W} \in V^*$ by Lemma 4.12. Invoking Lemma 4.12 another time, we can conclude that $\mathfrak{e} \in V^*$, since $\mathfrak{z} \in \omega \times V^*$ and $\mathfrak{z}^\circ \in E$ for all $\langle 0, \mathfrak{z}\rangle \in E^+$.

We are now ready to find a realizer for the axiom. Let $\mathfrak{u} \in V^*$ and suppose $e \Vdash \forall x \in \mathfrak{a}. \exists y \in \mathfrak{b}. \phi(x, y, \mathfrak{u})$. This means

$$\begin{aligned} & e \Vdash \forall x \in \mathfrak{a}. \exists y \in \mathfrak{b}. \phi(x, y, \mathfrak{u}) \\ \Rightarrow\quad & \forall \langle f, \mathfrak{c}\rangle \in \mathfrak{a}^*. \exists \mathfrak{d}. \langle (ef)_0, \mathfrak{d}\rangle \in \mathfrak{b}^* \wedge (ef)_1 \Vdash \phi(\mathfrak{c}, \mathfrak{d}, \mathfrak{u}) \qquad \text{(so } z = \langle \mathbf{p}ef, \mathfrak{d}\rangle \in B\text{)} \\ \Leftrightarrow\quad & \forall \langle f, \mathfrak{c}\rangle \in \mathfrak{a}^*. \exists z \in B. \psi(e, f, \mathfrak{c}, \mathfrak{u}, z) \\ \Leftrightarrow\quad & \forall \langle f, \mathfrak{c}\rangle \in \mathfrak{a}^*. \exists z \in w_u. \psi(e, f, \mathfrak{c}, \mathfrak{u}, z) \text{ and } \forall z \in w_u. \exists \langle f, \mathfrak{c}\rangle \in \mathfrak{a}^*. \psi(e, f, \mathfrak{c}, \mathfrak{u}, z) \\ \Leftrightarrow\quad & \forall \langle f, \mathfrak{c}\rangle \in \mathfrak{a}^*. \exists \langle \mathbf{p}ef, \mathfrak{d}\rangle \in w_u. \big(\langle (ef)_0, \mathfrak{d}\rangle \in \mathfrak{b}^* \wedge (ef)_1 \Vdash \phi(\mathfrak{c}, \mathfrak{d}, \mathfrak{u})\big) \\ & \quad \text{and } \forall \langle \mathbf{p}ef, \mathfrak{d}\rangle \in w_u. \exists \langle f, \mathfrak{c}\rangle \in \mathfrak{a}^*. \big(\langle (ef)_0, \mathfrak{d}\rangle \in \mathfrak{b}^* \wedge (ef)_1 \Vdash \phi(\mathfrak{c}, \mathfrak{d}, \mathfrak{u})\big) \end{aligned}$$

On the other hand, $e \Vdash \forall x \in \mathfrak{a}. \exists y \in \mathfrak{b}. \phi(x, y, \mathfrak{u})$. Implies $\forall x \in \mathfrak{a}^\circ. \exists y \in \mathfrak{b}^\circ. \phi^\circ(x, y, \mathfrak{u}^\circ)$ and hence, $\mathfrak{z}_u = \langle v_u \cup w_u^c, w_u\rangle \in \mathcal{W}$. Continuing the chain of implications,

$$\begin{aligned} \Leftrightarrow\quad & \forall \langle f, \mathfrak{c}\rangle \in \mathfrak{a}^*. \exists \langle \mathbf{p}ef, \mathfrak{d}\rangle \in \mathfrak{z}_u^*. \big(\langle (ef)_0, \mathfrak{d}\rangle \in \mathfrak{b}^* \wedge (ef)_1 \Vdash \phi(\mathfrak{c}, \mathfrak{d}, \mathfrak{u})\big) \\ & \quad \text{and } \forall \langle \mathbf{p}ef, \mathfrak{d}\rangle \in \mathfrak{z}_u^*. \exists \langle f, \mathfrak{c}\rangle \in \mathfrak{a}^*. \big(\langle (ef)_0, \mathfrak{d}\rangle \in \mathfrak{b}^* \wedge (ef)_1 \Vdash \phi(\mathfrak{c}, \mathfrak{d}, \mathfrak{u})\big) \\ \Rightarrow\quad & \forall \langle f, \mathfrak{c}\rangle \in \mathfrak{a}^*. \exists \mathfrak{d}. \big(\langle \mathbf{p}ef, \mathfrak{d}\rangle \in \mathfrak{z}_u^* \wedge (ef)_1 \Vdash \phi(\mathfrak{c}, \mathfrak{d}, \mathfrak{u})\big) \\ & \quad \text{and } \forall \langle g, \mathfrak{d}\rangle \in \mathfrak{z}_u^*. \exists \mathfrak{c}\big(\langle (g)_1, \mathfrak{d}\rangle \in \mathfrak{a}^* \wedge ((g)_0 (g)_1)_1 \Vdash \phi(\mathfrak{c}, \mathfrak{d}, \mathfrak{u})\big) \\ \Rightarrow\quad & m_0(e) \equiv \lambda f. \mathbf{p}\left(\mathbf{p}(ef)\big(\mathbf{r}(ef)\big)\right) \Vdash \forall x \in \mathfrak{a}. \exists y \in \mathfrak{z}_{\mathfrak{u}}. \phi(x, y, \mathfrak{u}) \\ & \quad \text{and } m_1 \equiv \lambda g. \mathbf{p}\left((\mathbf{l}g)\left(\mathbf{l}\big((\mathbf{l}g)(\mathbf{r}g)\big)\right)\right) \Vdash \forall x \in \mathfrak{z}_{\mathfrak{u}}. \exists y \in \mathfrak{a}. \phi(x, y, \mathfrak{u}) \end{aligned}$$

The last implication follows because

$$\forall x \in \mathfrak{a}^\circ. \forall x \in \mathfrak{a}. \exists y \in \mathfrak{z}_{\mathfrak{u}}{}^\circ. \phi^\circ(x, y, \mathfrak{u}^\circ) \quad \text{and} \quad \forall x \in \mathfrak{z}_{\mathfrak{u}}{}^\circ. \forall x \in \mathfrak{a}. \exists y \in \mathfrak{a}^\circ. \phi^\circ(x, y, \mathfrak{u}^\circ).$$

With these ingredients, $\lambda e. \mathbf{p}0(\mathbf{p}m_0(e)m_1)$ realizes

$$\forall x \in \mathfrak{a}. \exists y \in \mathfrak{b}. \phi(x, y, \mathfrak{u}) \rightarrow \exists d \in \mathfrak{e}\big(\forall x \in \mathfrak{a}. \exists y \in d. \phi(x, y, \mathfrak{u}) \wedge \forall y \in d. \exists x \in \mathfrak{a}. \phi(x, y, \mathfrak{u})\big),$$

if only

$$\forall x \in \mathfrak{a}^\circ. \exists y \in \mathfrak{b}^\circ. \phi^\circ(x, y, \mathfrak{u}^\circ) \to \exists d \in \mathfrak{e}^\circ \left(\forall x \in \mathfrak{a}^\circ. \exists y \in d. \phi^\circ(x, y, \mathfrak{u}^\circ) \wedge \forall y \in d. \exists x \in \mathfrak{a}^\circ. \phi^\circ(x, y, \mathfrak{u}^\circ)\right)$$

holds in **CZF**. But we can simply choose $d = v_\mathfrak{u}$, as $v_\mathfrak{u} \in C \subseteq \mathfrak{e}^\circ$.

Set induction

Lemma 4.14: Let ϕ be any formula. If $V^* \vDash \forall \mathfrak{a}\big[\big(\forall x \in \mathfrak{a}. \phi(x)\big) \to \phi(\mathfrak{a})\big]$ for all $\mathfrak{a} \in V^*$, then $\phi^\circ(x)$ holds for all sets x.

Proof: If $V^* \vDash \forall \mathfrak{a}\big[\big(\forall x \in \mathfrak{a}. \phi(x)\big) \to \phi(\mathfrak{a})\big]$, then by the completeness theorem, $\forall x \in \mathfrak{a}^\circ. \phi^\circ(x) \to \phi^\circ(\mathfrak{a}^\circ)$ for all $\mathfrak{a} \in V^*$. This is true in particular for $\mathfrak{a} = \hat{a}$ which shows that $\forall x \in a. \phi^\circ(x) \to \phi^\circ(a)$. We can therefore conclude by set induction in **CZF** that $\forall a. \phi^\circ(a)$. ■

Set induction certainly holds for ϕ°, i.e. the formula

$$\forall a\big[\big(\forall y \in a\ \phi^\circ(y)\big) \to \phi^\circ(a)\big] \to \forall a. \phi^\circ(a)$$

Is valid in **CZF**, hence we do not need to care about truth.

Let e be a fixed point of $\tau(z) \equiv \lambda u. u(\lambda x. zu)$. We show by induction that

$$e \Vdash \forall a\big[\big(\forall y \in a\ \phi(y)\big) \to \phi(a)\big] \to \forall a. \phi(a).$$

Let $g \Vdash\Vdash \forall a\big[\big(\forall y \in a. \phi(y)\big) \to \phi(a)\big]$. The induction hypothesis is that $eg \Vdash \phi(\mathfrak{b})$ for all $\mathfrak{b} \in V^*_\beta$ and $\beta \in \alpha$. If $\mathfrak{a} \in V^*_\alpha$, then $\langle h, \mathfrak{x}\rangle \in \mathfrak{a}^*$, then $\mathfrak{x} \in V^*_\beta$ for some $\beta \in \alpha$. Thus, by the induction hypothesis and the last lemma,

$$\begin{aligned} &\lambda x. eg \Vdash \forall y \in \mathfrak{a}. \phi(y), \\ &g(\lambda x. eg) \Vdash \phi(\mathfrak{a}). \end{aligned}$$

As $eg \equiv \big(\lambda u. u(\lambda x. eu)\big)g \simeq g(\lambda x. eg)$, this completes the proof.

4.6 Proof of metamathematical properties

Having proved completeness and soundness of our realizability structure, we are now ready to prove that **CZF** possesses some of the metamathematical properties discussed in Section 4.2.

4.6.1 Disjunction and Numerical existence property

This section is devoted to proving the numerical existence property for **CZF**:

Ⓜ**Theorem 4.15**: Let $\phi(x)$ be a formula with at most x free. If **CZF** $\vdash \exists n \in \omega\ \phi(n)$, then there is some $n \in \omega$ such that **CZF** $\vdash \phi(n)$.

Proof of Theorem 4.15: Let **CZF** $\vdash \exists n \in \omega . \phi(n)$. By the soundness theorem, $V^* \vDash \exists n \in \omega . \phi(n)$. When realizing the infinity axiom we have shown that $V^* \Vdash \omega = \overline{\omega}$ and hence, by Theorem 4.10, $V^* \vDash \exists n \in \overline{\omega} . \phi(n)$. Let $t \Vdash \exists n \in \overline{\omega} . \phi(n)$, this means that $(t)_0 = n$ for some $n \in \omega$ and $(t)_1 \Vdash \phi(\overline{n})$. By completeness, CZF proves $\phi^\circ(\overline{n}^\circ) \equiv \phi(n)$. ∎

As we have noted in Section 4.2.2, the numerical existence property implies the disjunction property. Hence, we immediately have:

Ⓜ**Corollary 4.16**: If **CZF** $\vdash \phi \vee \psi$, then **CZF** $\vdash \phi$ or **CZF** $\vdash \psi$.

4.6.2 Unzerlegbarkeits- and Uniformity-rule

Lemma 4.17: Let $\phi(x)$ be any formula with all free variables shown. If $t \Vdash \forall a\, \phi(a)$, then **CZF** $\vdash \forall a\, \phi(a)$.

Proof: $t \Vdash \forall a\, \phi(a)$ means that $t \Vdash \phi(\mathfrak{a})$ for all $\mathfrak{a} \in V^*$. This is true in particular for $\mathfrak{a} = \hat{x}$, where x is a set. By the completeness theorem, we obtain **CZF** $\vdash \phi^\circ\big((\hat{x})^\circ\big)$ for all x, i.e. **CZF** $\vdash \forall x\, \phi(x)$. ∎

Ⓜ**Theorem 4.18**: Whenever **CZF** $\vdash \forall x . \exists y \in \omega . \psi(x, y)$, then **CZF** $\vdash \exists y \in \omega . \forall x . \psi(x, y)$.

Proof: Suppose **CZF** $\vdash \forall x . \exists y \in \omega . \psi(x, y)$. By the soundness theorem, there is a closed application term t such that **CZF** $\vdash t \Vdash \forall x . \exists y \in \omega . \psi(x, y)$, i.e. **CZF** $\vdash \forall \mathfrak{a} \in V^* \, [t \Vdash \exists y \in \omega . \psi(\mathfrak{a}, y)]$. As we know that $V^* \vDash \omega = \overline{\omega}$ we even have **CZF** $\vdash \forall \mathfrak{a} \in V^* \, [t \Vdash \exists n \in \overline{\omega} . \psi(\mathfrak{a}, n)]$.

This means in turn that there is some $\mathfrak{y} \in V^*$ such that $\langle (t)_0, \mathfrak{n} \rangle \in \overline{\omega}$ and $(t)_1 \Vdash \psi(\mathfrak{a}, \mathfrak{n})$. But by the definition of $\overline{\omega}$, $\mathfrak{n}$ has to be one of the $\overline{n}$s and hence cannot depend on $\mathfrak{a}$. By Lemma 4.17, we conclude **CZF** $\vdash \forall x . \psi(x, n)$. ∎

As we know, Uniformity-rule is just a special case of Unzerlegbarkeit. We therefore have:

Ⓜ**Corollary 4.19**: Whenever **CZF** $\vdash \forall x . \psi(x) \vee \neg\psi(x)$, then **CZF** $\vdash \forall x\, \psi(x) \vee \forall x\, \neg\psi(x)$.

4.6.3 Church's rule

Ⓜ**Theorem 4.20**: If **CZF** $\vdash \forall x \in \omega . \exists y \in \omega . \phi(x, y)$, then there is some natural number e with **CZF** $\vdash \forall x \in \omega . \phi\big(x, \{e\}(x)\big)$.

Proof: We may assume that **CZF** $\vdash \forall x \in \overline{\omega} \, . \exists y \in \overline{\omega} . \phi(x, y)$. By the soundness theorem, there is some $t \in \omega$ with

	$t \Vdash \forall x \in \overline{\omega} . \exists y \in \overline{\omega} . \phi(x, y)$.
then	If $\langle m, \overline{m} \rangle \in \overline{\omega}^*$, then $tm \Vdash \exists y \in \overline{\omega} . \theta(\overline{m}, y)$.
then	If $m \in \omega$, then $tm \Vdash \exists y \in \omega . \theta(\overline{m}, y)$.
then	If $m \in \omega$ then $\exists n \in \omega$ with $(tm)_0 = n$ and $(tm)_1 \Vdash \theta(\overline{m}, \overline{n})$.

Set $s = \lambda u . (tu)_1$. Then **CZF** $\vdash \lambda u . (tu)_1 \simeq e$ for some natural number e. We can conclude by completeness

$$\mathbf{CZF} \vdash \forall x \in \omega.\, \phi\big(x, \{e\}(x)\big).$$

■

4.6.4 Markov's rule

Ⓜ**Theorem 4.21**: If $\mathbf{CZF} \vdash \big(\forall x \in \omega.\, \phi(n) \vee \neg\phi(n)\big) \wedge \big(\neg\neg\exists n \in \omega.\, \phi(n)\big)$, then $\mathbf{CZF} \vdash \exists n \in \omega.\, \phi(n)$.

Proof: We may assume that there are terms t and u with $t \Vdash \forall x \in \overline{\omega}.\, \phi(n) \vee \neg\phi(n)$ and $u \Vdash \neg\neg\exists n \in \overline{\omega}.\, \phi(n)$. The first means that for each $n \in \omega$, either $(tn)_0 = 0$ and $(tn)_1 \Vdash \phi(\overline{n})$, or $(tn)_0 \neq 0$ and $(tn)_1 \Vdash \neg\phi(\overline{n})$. The second means (by classical reasoning), that there is some term p with $p \Vdash \exists n \in \overline{\omega}.\, \phi(n)$, which means $(p)_1 \Vdash \phi\left(\overline{(p)_0}\right)$.

Hence, $(t(p)_0)_0 = 0$ and $(t(p)_0)_1 \Vdash \phi\left(\overline{(p)_0}\right)$. This shows that with $r \equiv \mu p.\, (t(p)_0)_0$ with μ being the least number operator from Definition 4.4,

$$\mathbf{p}(\mathbf{l}r)(\mathbf{rel}r) \Vdash \exists n \in \overline{\omega}.\, \phi(n).$$

■

4.7 Further results

With the work so far, it is easy to obtain further results while only slightly adjusting the realizability structure V^*.

4.7.1 Metamathematics of IZF

Notice, that our realizability refers to truth (for example in the clause for implication). In the formulation of the soundness theorem is becomes clear, that this truth refers to truth inside **CZF**. If we change this to mean truth in **IZF**, we can easily transfer all proofs to this context.

There is nothing to do to alter the completeness theorem:

Ⓜ**Theorem 4.22**: If $V^* \vDash \theta$, then $\mathbf{IZF} \vdash \theta^\circ$.

The soundness theorem requires more work, as we need to check the additional set axioms:

Ⓜ**Theorem 4.23 (Soundness)**: For every theorem θ of **IZF**, there exists a closed application term t, such that $\mathbf{IZF} \vdash (t \Vdash \theta)$.

Proof: We need to give realizers for the axiom schemas of (unbounded) separation and powerset.

Separation schema

This proof is almost identical to the proof of bounded separation. The only difference in this case is that we do not need to pay attention to the involved formulas to be bounded.

Powerset

For $\mathfrak{a} \in V^*$, we define $\mathrm{Pow}(\mathfrak{a})$ by $\mathrm{Pow}(\mathfrak{a})^* = \{\langle e, \mathfrak{c} \rangle : e \Vdash \mathfrak{c} \subseteq \mathfrak{a}\}$ and $\mathrm{Pow}(\mathfrak{a})^\circ = \mathrm{Pow}(\mathfrak{a}) \cup \mathcal{P}(\mathfrak{a}^\circ)$. Note that since in **IZF**, the $V_\alpha s$ and V_α^*s are sets, so must be $\mathrm{Pow}(\mathfrak{a})^*$ and therefore $\mathrm{Pow}(\mathfrak{a})$ is a well-defined element of V^*. If $g \Vdash \mathfrak{c} \subseteq \mathrm{Pow}(\mathfrak{a})$, then $\langle g, \mathfrak{c} \rangle \in \mathrm{Pow}(\mathfrak{a})^*$ and hence $\mathbf{p}g\mathbf{i_r} \Vdash \mathfrak{c} \in \mathrm{Pow}(\mathfrak{a})$. Of course, $\mathfrak{c}^\circ \subseteq \mathfrak{a}^\circ$ implies $\mathfrak{c}^\circ \in \mathrm{Pow}(\mathfrak{a})^\circ$ and hence $\lambda g.\, \mathbf{p}g\mathbf{i_r}$ realizes the axiom. ■

Ⓜ**Theorem 4.24**: **IZF** has the disjunction and numerical existence property. It follows the rules of Unzerlegbarkeit, Uniformity and Church.

4.7.2 Compatibility with Principles

Changing our realizability structure to not refer to truth inside **CZF** in **IZF**, comes at the cost of giving up the completeness theorem. On the other hand, we gain some leeway in that our realizability structures may possesses some properties, that cannot be proved in the theories. In particular, it is especially easy to transform the proofs we conducted in Section 4.6 into proofs of the corresponding principles (see Section 4.2). We formulate our results in terms of the following theorem:

Ⓜ**Theorem 4.25**: If $V^* \vDash \phi$, then $\mathrm{Cons}(\mathbf{IZF}) \Rightarrow \mathrm{Cons}(\mathbf{IZF} + \phi)$ and similarly for **CZF**.

Proof: Towards a contradiction, suppose that $\mathbf{IZF} + \phi$ is inconsistent. Then $\mathbf{IZF} + \phi \vdash \bot$. By the soundness theorem and by $V^* \vDash \phi$, there is some $t \in \omega$ with $\mathbf{IZF} \vdash t \Vdash \bot$. But by definition of realizability, $\mathbf{IZF} \vdash \forall n \in \omega.\, \neg n \Vdash \bot$, contradiction. ■

We can now show that **IZF** and **CZF** are equiconsistent with the theories augmented by the principles UZ, UzP, CP and MP:

Ⓜ**Theorem 4.26**: $\mathrm{Cons}(\mathbf{IZF}) \Rightarrow \mathrm{Cons}(\mathbf{IZF} + \mathrm{UP} + \mathrm{UzP} + \mathrm{CP} + \mathrm{MP})$ and similarly for **CZF**.

Proof: Let **T** be either **CZF** or **IZF**. Given a proof of the property of the form

$$\text{If } \mathbf{T} \vdash A, \text{ then } \mathbf{T} \vdash B,$$

we can now easily transform it into a proof of the rule

$$\mathbf{T} \vdash A \to B.$$

We realize the rule with a term $\lambda x.\, \tau$ mapping realizers of A to realizers of B. The crucial point is that when realizing $A \to B$ we now need not refer to truth inside **T** of the statement $A^\circ \to B^\circ$ and simply use the same term to realize $A \to B$.

For example, let us show how to show the Uniformity-rule: Suppose $e \Vdash \forall x. \exists y \in \omega.\, \psi(x, y)$. This means, that for all $\mathfrak{a} \in V^*$, $e \Vdash \exists y \in \omega.\, \psi(\mathfrak{a}, y)$. We know that internal ω is represented by $\overline{\omega}$ and hence, $t \Vdash \exists n \in \overline{\omega}.\, \psi(\mathfrak{a}, n)$ for all $\mathfrak{a} \in V^*$.

This means that there is some $\mathfrak{y} \in V^*$ such that $\langle (e)_0, \mathfrak{n} \rangle \in \overline{\omega}$ and $(e)_1 \Vdash \psi(\mathfrak{a}, \mathfrak{n})$. But by the definition of $\overline{\omega}$, $\mathfrak{n}$ has to be one of the $\bar{n}$s and hence both $\mathfrak{n}$ and e cannot depend on $\mathfrak{a}$. By Lemma 4.17, we conclude $e \Vdash \forall x. \psi(x, n)$ and therefore $\lambda e. e \Vdash \big(\forall x. \exists y \in \omega. \psi(x, y)\big) \rightarrow \big(\forall x. \psi(x, n)\big)$.

Comparing these lines to the proof of Theorem 4.18, we see that the proof follows the schema indicated above. ■

We have discussed in section 4.2.6 that **LEM** is incompatible with CP. Therefore, we immediately get the following Corollary:

Ⓜ**Corollary 4.27**: **IZF** $\nvdash$ **LEM**.

In light of this result, all weak counterexamples of sections 2.4.2 and 3.3 like the foundation axiom or the well-ordering of $\mathbb{ON}$ turn out to be unprovable statements in **IZF** and **CZF**.

5 Martin-Löf's set theory

As promised, in chapter 2.3, we will take a closer look on Martin-Löf's set theory **ML**. We will describe all rules and briefly discuss some application. In the final section of this chapter, we will give an interpretation of **CZF** into **ML**. It should be noted that all concepts are closely related to type theory. In fact, it is exactly type theory, if we refer to sets as "types". Martin-Löf changed his notation accordingly in his later work and so will we starting with section 5.2. For now, let us stick with his early notation to emphasize that the presented theory presented defines a set theory on its own right.

5.1 Formulating **ML**

5.1.1 Rules of equality

Reflexivity

$$\frac{a \in A}{a = a \in A} \qquad\qquad \frac{A \text{ set}}{A = A}$$

Symmetry

$$\frac{a = b \in A}{b = a \in A} \qquad\qquad \frac{A = B}{B = A}$$

Transitivity

$$\frac{a = b \in A \qquad\qquad b = c \in A}{a = c \in A} \qquad\qquad \frac{A = B \qquad\qquad B = C}{A = C}$$

Equality of types

$$\frac{a \in A \qquad\qquad A = B}{a \in B} \qquad\qquad \frac{a = b \in A \qquad\qquad A = B}{a = b \in B}$$

5.1.2 Substitution rules

We put down the following rules of substitution. Expressions like $\begin{array}{c}(x \in A)\\ B(x)\ \text{set}\end{array}$ should be read as follows: "$B(x)$ is a well-formed set under the assumption that x is in the set A".

$$\frac{a \in A \qquad\qquad \begin{array}{c}(x \in A)\\ B(x)\ \text{set}\end{array}}{B(a)\ \text{set}} \qquad\qquad \frac{a = c \in A \qquad\qquad \begin{array}{c}(x \in A)\\ B(x)\ \text{set}\end{array}}{B(a) = B(c)}$$

$$\frac{a \in A \qquad\qquad \begin{array}{c}(x \in A)\\ B(x) = C(x)\end{array}}{B(a) = C(a)}$$

$$\frac{a \in A \qquad\qquad \begin{array}{c}(x \in A)\\ b(x) \in B(x)\end{array}}{b(a) \in B(a)} \qquad\qquad \frac{a = c \in A \qquad\qquad \begin{array}{c}(x \in A)\\ b(x) \in B(x)\end{array}}{b(a) = b(c) \in B(a)}$$

$$\frac{a \in A \qquad\qquad \begin{array}{c}(x \in A)\\ b(x) = c(x) \in B(x)\end{array}}{b(a) = c(a) \in B(a)}$$

5.1.3 Π-rules

As a first example of rules allowing for a construction of sets, we discuss Π-rules. They prescribe how to construct, given a set A and for each x in A some set $B(x)$, the set $(\Pi x \in A)B(x)$. One can think of this set as the set of all functions mapping $x \in A$ into $B(x)$:

Π-formation

$$\frac{A \text{ set} \qquad \begin{matrix}(x \in A)\\ B(x) \text{ set}\end{matrix}}{(\Pi x \in A)B(x) \text{ set}} \qquad\qquad \frac{A = C \qquad \begin{matrix}(x \in A)\\ B(x) = D(x)\end{matrix}}{(\Pi x \in A)B(x) = (\Pi x \in C)D(x)}$$

Π-introduction

$$\frac{\begin{matrix}(x \in A)\\ b(x) \in B(x)\end{matrix}}{\lambda x.\, b(x) \in (\Pi x \in A)B(x)} \qquad\qquad \frac{\begin{matrix}(x : A)\\ b(x) = d(x) \in B(x)\end{matrix}}{\lambda x.\, b(x) = \lambda x.\, d(x) \in (\Pi x \in A)B(x)}$$

Π-elimination

$$\frac{c \in (\Pi x \colon A)B(x) \qquad a \in A}{\mathrm{Ap}(c, a) \in B(a)} \qquad\qquad \frac{c = d \in (\Pi x \colon A)B(x) \qquad a = b \in A}{\mathrm{Ap}(c, a) = \mathrm{Ap}(d, b) \in B(a)}$$

Π-equality

$$\frac{a \in A \qquad \begin{matrix}(x \in A)\\ b(x) \in B(x)\end{matrix}}{\mathrm{Ap}(\lambda x.\, b(x), a) = b(a) \in B(a)} \qquad\qquad \frac{c \in (\Pi x \in A)B(x)}{\lambda x.\, \mathrm{Ap}(c, x) = c \in (\Pi x \in A)B(x)}$$

The right column explains equality between the constructed sets and their elements. In the case of Π-introduction, two sets $(\Pi x \in A)B(x)$ and $(\Pi x \in C)D(x)$ are equal if $A = C$ and for all $x \in A$, $B(x) = C(x)$.

The rules of Π-introduction declare how to give canonical elements of $(\Pi x \in A)B(x)$, or in other words, how to give a function mapping $x \in A$ into $B(x)$. This is done in giving a method computing for each $x \in A$ an element $b(x) \in B(x)$. The function $x \mapsto b(x)$ thus defined is denoted by $\lambda x.\, b(x)$. The column on the right explains when two such canonical elements are equal: This is the case if the two associated functions agree on every input.

The symbol Ap can be explained as follows: Given an element $c \in (\Pi x \in A)B(x)$, we know it yields some canonical element $\lambda x.\, b(x)$. $\mathrm{Ap}(c, a)$ computes this element and returns $b(a)$. Thus, the elimination rules describe how elements of the type just defined behave.

Finally, the rules of Π-equality show how Ap operates on canonical elements. Note that they correspond to the rules β-reduction and γ-conversion of λ-calculus.

We will usually write $c(a)$ instead of $\mathrm{Ap}(c, a)$, when we know that c is of a corresponding Π-set. If the family $B(x)$ does not depend on x, we will simply write $A \to B$ instead of $(\Pi x \in A)B$.

In the propositions-as-sets interpretation "→" plays the same role as its logical counterpart. Thus, the Π-rules translate nicely into logical rules about implication. As suggested in the starting example, one role of the Π-type is to give an interpretation to universal quantification: One defines the **set** $\forall x \in A\, B(x)$ to be $(\Pi x \in A)B(x)$. Let us translate the Π-rules into the context of plain implication and universal quantification: We write A prop instead of A set. Capturing the constructive standpoint that truth corresponds directly to provability, we write A true iff there is some $c \in A$, i.e. there is some proof of the proposition A. Now the Π-rules take the following forms familiar from predicate logic:

∀-formation

$$\dfrac{A\ \text{prop} \qquad \begin{matrix}(x \in A)\\ B(x)\ \text{prop}\end{matrix}}{(\forall x \in A)B(x)\ \text{prop}}$$

→-formation

$$\dfrac{A\ \text{prop} \qquad B\ \text{prop}}{A \to B\ \text{prop}}$$

∀-introduction

$$\dfrac{\begin{matrix}(x \in A)\\ B(x)\ \text{true}\end{matrix}}{(\forall x \in A)B(x)\ \text{true}}$$

→-introduction

$$\dfrac{\begin{matrix}(A\ \text{true})\\ B\ \text{true}\end{matrix}}{A \to B\ \text{true}}$$

∀-elimination

$$\dfrac{(\forall x \in A)B(x)\ \text{true} \qquad a \in A}{B(a)\ \text{true}}$$

→-elimination

$$\dfrac{A \to B\ \text{true} \qquad A\ \text{true}}{B\ \text{true}}$$

Example: Let us verify the logical axiom $\big(A \to (B \to C)\big) \to \big((A \to B) \to (A \to C)\big)$. One way to do this is in a derivation tree-like syntax to find an element (=proof) of the set $\big(A \to (B \to C)\big) \to \big((A \to B) \to (A \to C)\big)$ applying diretly the Π-rules from above:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[x \in A \to (B \to C)]^1 \qquad [z \in A]^2}{x(z) \in B \to C}\,(\Pi\text{-elim}) \qquad \dfrac{[y \in A \to B]^3 \qquad [z \in A]^2}{y(z) \in B}\,(\Pi\text{-elim})}{x(z)\big(y(z)\big) \in C}\,(\Pi\text{-elim})}{\lambda z.\, x(z)\big(y(z)\big) \in A \to C}\,(\Pi\text{-intr}), \cancel{[2]}}{\lambda y \lambda z.\, x(z)\big(y(z)\big) \in (A \to B) \to (A \to C)}\,(\Pi\text{-intr}), \cancel{[3]}}{\lambda x \lambda y \lambda z.\, x(z)\big(y(z)\big) \in A \to (B \to C) \to \big((A \to B) \to (A \to C)\big)}\,(\Pi\text{-intr}), \cancel{[1]}$$

If we use →-rules instead we get see how this proof of membership turns into a proof of intuitionistic logic in a Curry-Howard-style fashion:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[A \to (B \to C)\ \text{true}]^1 \qquad [A\ \text{true}]^2}{B \to C\ \text{true}}\,(\to\text{-elim}) \qquad \dfrac{[A \to B\ \text{true}]^3 \qquad [A\ \text{true}]^2}{B\ \text{true}}\,(\to\text{-elim})}{C\ \text{true}}\,(\to\text{-elim})}{A \to C\ \text{true}}\,(\to\text{-intr}), \cancel{[2]}}{(A \to B) \to (A \to C)\ \text{true}}\,(\to\text{-intr}), \cancel{[3]}}{A \to (B \to C) \to \big((A \to B) \to (A \to C)\big)\ \text{true}}\,(\to\text{-intr}), \cancel{[1]}$$

The way we introduced the Π-rules is paradigmatic: All rules in the following will contain formation rules, explaining how to construct the set under discussion from other sets. Introduction rules explain how its canonical elements are formed and elimination rules describe their behavior. Finally, equality rules relate introduction and elimination rules. Note that the rules of the right column explaining equality of the constructed sets and their canonical elements always follow the same pattern. We do therefore not write them down explicitly in the following discussion.

Another remark we need to make is that beginning with the rule of Π-introduction, we skipped some of the assumptions of the rule. In this particular case, they should have included A type and $B(x)$ type $(x \in A)$. For the sake of readability, we will not write down these obvious lines.

5.1.4 Σ-rules

We continue with rules for Σ-sets that one can think of as generalized disjoint union of a family of sets. If we have a set A and for each x of set A some set $B(x)$, we can form the disjoint union of the $B(x)$ indexed by A. We denote this by $(\Sigma x \in A)B(x)$.

Σ-formation

$$\frac{A \text{ set} \qquad \begin{matrix}(x \in A)\\ B(x) \text{ set}\end{matrix}}{(\Sigma x \in A)B(x) \text{ set}}$$

Σ-introduction

$$\frac{a \in A \qquad b \in B(a)}{(a,b) \in (\Sigma x \in A)B(x)}$$

Σ-elimination

$$\frac{c \in (\Sigma x \in A)B(x) \qquad \begin{matrix}\big(x \in A, y \in B(x)\big)\\ d(x,y) \in C(x,y)\end{matrix}}{\mathrm{E}(c,d) \in C(c)}$$

Σ-equality

$$\frac{a \in A \qquad b \in B(a) \qquad \begin{matrix}\big(x \in A, y \in B(x)\big)\\ d(x,y) \in C(x,y)\end{matrix}}{\mathrm{E}\big((a,b),d\big) = d(a,b) \in C((a,b))}$$

The canonical elements of $(\Sigma x \in A)B(x)$ are of the form (a,b), where $a \in A$ and $b \in B(a)$. The symbol E operates as follows: It computes $c \in (\Sigma x \in A)B(x)$ to find its associated canonical element of the form (a,b), where $a \in A$ and $b \in B(a)$. It then substitutes the values into d to obtain $d(a,b)$ of $C\big((a,b)\big)$.

The case where we start this procedure with a canonical element (a,b) instead of a general c is written down in the rule of Σ-equality.

Example: Let us set $d(x,y) \equiv x$ and $C(x,y) \equiv A$ in the rule of Σ-elimination. We then obtain the left projection as $p(c) \equiv \mathrm{E}(c,d) \in A$. By Σ-equality it has the property $p\big((a,b)\big) = E\big((a,b),d\big) = d(a,b) = a$. Similarly, the right projection $q(a,b) = b \in B(a)$ is defined.

If $B(x)$ does not depend on x we write $A \times B$ for $(\Sigma x \in A)B$. In the propositions-as-sets interpretation "×" plays the role of conjunction. Indeed, a proof of $A \wedge B$ must be a pair (a,b) where $a \in A$ and $b \in B$. In the case of existential quantification, we can define $(\exists x \in A)B(x)$ to be $(\Sigma \in A)B$. Here, (a,b) consists of a witness a and a proof b of $B(a)$. Indeed, if we have a $c \in (\exists x \in A)B(x)$, then $p(c) \in A$ and $q(c) \in B(x)$. If we again translate the rules, we have

∃-formation

$$\dfrac{A \text{ prop} \qquad \overset{(x \in A)}{B(x) \text{ prop}}}{(\exists x \in A)B(x) \text{ prop}}$$

∧-formation

$$\dfrac{A \text{ prop} \qquad B \text{ prop}}{A \wedge B \text{ prop}}$$

∃-introduction

$$\dfrac{a \in A \qquad B(a) \text{ true}}{(\exists x \in A)B(x) \text{ true}}$$

∧-introduction

$$\dfrac{A \text{ true} \qquad B \text{ true}}{A \wedge B \text{ true}}$$

∃-elimination

$$\dfrac{(\exists x \in A)B(x) \text{ true} \qquad \overset{(x \in A, B(x) \text{ true})}{C \text{ true}}}{C \text{ true}}$$

∧-elimination

$$\dfrac{A \wedge B \text{ true} \qquad \overset{(A \text{ true}, B \text{ true})}{C \text{ true}}}{C \text{ true}}$$

As an example, let us verify that this interpretation respects the logical rule $A \wedge B \to A$: By the example before, we have that $p(c) \in A$ whenever $c \in A \wedge B$. Thus, $\lambda c. p(c) \in A \wedge B \to A$ by Π-introduction. Alternatively, we can use rules for → and ∧:

$$\dfrac{\dfrac{\overset{\boxed{}}{[A \wedge B \text{ true}]^1} \qquad \overset{(A \text{ true}, B \text{ true})}{A \text{ true}}}{A \text{ true}}\,(\wedge\text{-elim})}{A \wedge B \to A \text{ true}}\,(\to\text{-intr}), \cancel{[1]}$$

5.1.5 +-Rules

Given sets A and B, we can construct their disjoint union $A + B$:

+-formation

$$\dfrac{A \text{ set} \qquad B \text{ set}}{A + B \text{ set}}$$

+-introduction

$$\dfrac{a \in A}{i(a) \in A + B}$$

$$\dfrac{b \in B}{j(b) \in A + B}$$

+-elimination

$$\dfrac{c \in A + B \qquad \overset{(x \in A)}{d(x) \in C\big(i(x)\big)} \qquad \overset{(y \in B)}{e(y) \in C\big(j(y)\big)}}{\mathrm{D}(c, d, e) \in C(c)}$$

+-equality

$$\dfrac{a \in A \qquad \overset{(x \in A)}{d(x) \in C\big(i(x)\big)} \qquad \overset{(y \in B)}{e(y) \in C\big(j(y)\big)}}{\mathrm{D}(a, d, e) = d(a) \in C\big(i(a)\big)}$$

$$\frac{b \in B \qquad \begin{array}{c}(x \in A)\\ d(x) \in C\big(i(x)\big)\end{array} \qquad \begin{array}{c}(y \in B)\\ e(y) \in C\big(j(y)\big)\end{array}}{\mathrm{D}(b,d,e) = e(b) \in C\big(j(b)\big)}$$

The canonical elements of $A + B$ are the "labeled" elements of A and B denoted by $i(a)$ and $j(b)$. The operator D decides whether the canonical element associated to $c \in A + B$ is of the form $i(a)$ or $j(b)$. In the first case it computes $d(a)$, in the second $e(b)$.

One might think that the disjoint union $A + B$ may be eliminated in favor of $(\Sigma x \in \mathbb{N}_2)f(x)$, where $\mathbb{N}_2$ is a type containing two elements $\overline{0}$ and $\overline{1}$ and

$$f(x) = \begin{cases} A, & if\ x = \overline{0}, \\ B, & if\ x = \overline{1}. \end{cases}$$

However, as Beeson remarks in [4], one runs into difficulties in defining in such an f, so Martin-Löf found it more convenient to include these extra rules for the disjoint union of two sets.

In the context of logic, $+$ takes the role of disjunction: If we write $A \vee B$ for $A + B$ we have the rules:

∨-formation

$$\frac{A \text{ prop} \qquad B \text{ prop}}{A \vee B \text{ prop}}$$

∨-introduction

$$\frac{A \text{ true}}{A \vee B \text{ true}}$$

$$\frac{B \text{ true}}{A \vee B \text{ true}}$$

∨-elimination

$$\frac{A \vee B \text{ true} \qquad \begin{array}{c}(A \text{ true})\\ C \text{ true}\end{array} \qquad \begin{array}{c}(B \text{ true})\\ C \text{ true}\end{array}}{C \text{ true}}$$

5.1.6 I-rules

We will not make direct use of I-rules, nevertheless, we will briefly discuss them in this section. Note that although we have judgments asserting the equality of two elements with respect to some type, i.e. judgements of the form $a = b \in A$, we cannot formulate such judgement as propositions. Indeed, it turns out that without I-rules one cannot even formulate arithmetic within **ML**.

Given a type A and two elements a and b of type A, $\mathrm{I}(A, a, b)$ is the set (or in this case, rather the "proposition", set of proofs) asserting that a and b are equal elements of A:

I-formation

$$\frac{A \text{ set} \qquad a \in A \qquad b \in A}{\mathrm{I}(A, a, b) \text{ set}}$$

I-introduction

$$\frac{a = b \in A}{\mathrm{r} \in \mathrm{I}(A, a, b)}$$

I-elimination

$$\frac{c \in \mathrm{I}(A, a, b)}{a = b \ \in A}$$

I-equality

$$\frac{c \in \mathrm{I}(A, a, b)}{c = \mathrm{r} \in \mathrm{I}(A, a, b)}$$

If indeed, $a = b \in A$, then $\mathrm{I}(A, a, b)$ contains a canonical proof r of this fact. On the other hand, if $\mathrm{I}(A, a, b)$ contains an element, then $a = b \in A$. Finally, the equality rule says that r is the only element of $\mathrm{I}(A, a, b)$ up to equality.

As an example, let us verify $(\forall x \in A)\mathrm{I}(A, x, x)$ true saying that all elements of A are equal to themselves. Given $x \in A$, we know that $\mathrm{r} \in \mathrm{I}(A, x, x)$ and thus $\lambda x.\mathrm{r} \in (\forall x \in A)\mathrm{I}(A, x, x)$.

5.1.7 $\mathbb{N}_k$-rules

The $\mathbb{N}_k$-rules are the only rules so far guaranteeing the existence of sets without any assumptions. For each k, we postulate the existence of a k-element set:

$\mathbb{N}_k$-formation

$$\mathbb{N}_k \text{ set}$$

$\mathbb{N}_k$-introduction

$$\overline{m} \in \mathbb{N}_k$$
$$\text{for } m = 0, \dots, k-1.$$

$\mathbb{N}_k$-elimination

$$\frac{c \in \mathbb{N}_k \qquad c_0 \in C(\overline{0}) \ \cdots \ c_{k-1} \in C(\overline{k-1})}{\mathrm{R}_k(c, c_0, \dots, c_{k-1}) \in C(c)}$$

$\mathbb{N}_k$-equality

$$\frac{c_0 \in C(\overline{0}) \ \cdots \ c_{k-1} \in C(\overline{k-1})}{\mathrm{R}_k(\overline{m}, c_0, \dots, c_{k-1}) = c_m \in C(\overline{m})}$$
$$\text{for } m = 0, \dots, k-1.$$

In the elemination rule, R_k operates as follows: It executes c to find out to which element $\overline{m}$ it converges. It then returns the corresponding element c_m.

The type $\mathbb{N}_0$ corresponds to the empty set and takes the role of $\bot$ in the proposition-as-sets interpretation. We set $\bot \equiv \mathbb{N}_0$ and observe the rule

$\bot$-elimination

$$\frac{\bot \text{ true}}{C \text{ true}}$$

This is the usual rule of *ex falso quodlibet*. As usual, we can define negation as $\neg A \equiv A \to\bot$. The rule $A \to \neg\neg A$ is proved as usually:

$$\cfrac{\cfrac{\cfrac{[A\ \text{true}]^1 \qquad [A \to\bot\ \text{true}]^2}{\bot\ \text{true}}\,(\to\text{-elim})}{(A \to\bot) \to\bot\ \ \text{true}}\,(\to\text{-intr}),\ \cancel{[2]}}{A \to \big((A \to\bot) \to\bot\big)\ \text{true}}\,(\to\text{-intr}),\ \cancel{[1]}$$

5.1.8 ℕ-rules

The ℕ-rules postulate the existence of the set of natural numbers:

ℕ-formation

$$\mathbb{N}\ \text{set}$$

ℕ-introduction

$$0 \in \mathbb{N}$$

$$\frac{n \in \mathbb{N}}{s(n) \in \mathbb{N}}$$

ℕ-elimination

$$\frac{c \in \mathbb{N} \qquad c \in C(0) \qquad \begin{matrix}(x \in \mathbb{N}, y \in C\big(s(x)\big) \\ e(x,y) \in C\big(s(x)\big)\end{matrix}}{\mathrm{R}(c,d,e) \in C(c)}$$

ℕ-equality

$$\frac{d \in C(0) \qquad \begin{matrix}(x \in \mathbb{N}, y \in C\big(s(x)\big) \\ e(x,y) \in C\big(s(x)\big)\end{matrix}}{\mathrm{R}(0,d,e) = d \in C(0)}$$

$$\frac{a \in \mathbb{N} \qquad d \in C(0) \qquad \begin{matrix}(x \in \mathbb{N}, y \in C\big(s(x)\big) \\ e(x,y) \in C\big(s(x)\big)\end{matrix}}{\mathrm{R}(s(a),d,e) = e(a, R(a,d,e)) \in C\big(s(a)\big)}$$

successor

$$\frac{x \in \mathbb{N}}{\mathrm{I}(\mathbb{N}, s(x), 0) \in \bot}$$

The R in the above rules gives us the possibility of recursive definitions: If $c \in \mathbb{N}$ it executes c to see wether it yields 0 or not. If c converges to 0, $\mathrm{R}(c,d,e)$ returns $d \in C(0)$. Otherwise, $c = s(\overline{m})$ and $\mathrm{R}(c,d,e)$ will be computed recursively by $e\big(\overline{m}, R(\overline{m}, d, e)\big)$. In this sense, the rules for ℕ-equality define two rewrite rules. The successor rule tells us that there is no proof of the statement that 0 is a successor of $s(x)$.

5.1.9 **ML** *in action – Real numbers and the axiom of choice*

We give two examples to see how to work in **ML**. The first is a construction of the set of real numbers. The rational numbers, as well as basic operations on them are obtained from ℕ in the usual way. We would now like to form ℝ as the set of all Cauchy-sequences ℕ → ℚ. To accomplish this, we will use Σ-rules to separate the Cauchy-sequences from all functions in $\mathbb{N} \to \mathbb{Q}$. Generally speaking, if we want to form the set $\{x \in A: B(x)\}$, i.e. the set of all elements of A satisfying some condition B, this will correspond to a set of pairs (a, p_a), where a is an element of A and p_a is a proof of $B(a)$. In the propositions-as-sets interpretation, this means $a \in A$ and $p_a \in B(a)$. Thus, $(a, p_a) \in (\Sigma x \in A)B(a)$ and we can identify $\{x \in A: B(x)\}$ with this set. In our case we can set

$$\mathbb{R} \equiv (\Sigma x \in \mathbb{N} \to \mathbb{Q})\ \mathrm{Cauchy}(x),$$

where

$$\mathrm{Cauchy}(x) \equiv (\forall e \in \mathbb{Q})\big(e > 0 \to (\exists n \in \mathbb{N})(\forall m \in \mathbb{N})(|x(n) - x(m+n)| < e)\big).$$

Note that $p \in \mathrm{Cauchy}(x)$ is a modulus of convergence, i.e. a method computing to each $e > 0$ a some specific $n \in \mathbb{N}$ such that $\forall m \in \mathbb{N}$ we have $|x(n) - x(m+n)| < e$.

As another example let us consider the axiom of choice. It turns out that this axiom takes very divergent roles and strengths in different versions of set theory, sometimes extending the theory to be highly non-constructive. Interestingly, in **ML**, its strongest formulation

$$\mathbf{AC} \equiv (\forall x \in A)\big(\exists y \in B(x)\big)C(x,y) \to \big(\exists f \in (\Pi x \in A)B(x)\big)(\forall x \in A)C\big(x, f(x)\big),$$

turns out to be derivable: Suppose, $c \in (\forall x \in A)\big(\exists y \in B(x)\big)C(x,y)$. By Π-elimination, $c(a) \in \big(\exists y \in B(x)\big)C(x,y)$ for $a \in A$. Using projection functions, $p\big(c(a)\big) \in B(a)$ and $q\big(c(a)\big) \in C\Big(a, p\big(c(a)\big)\Big)$. Thus, we can define the function f to be $\lambda a.\, p\big(c(a)\big)$. This shows that $\lambda c.\Big(\lambda a.\, p\big(c(a)\big), \lambda a.\, q\big(c(a)\big)\Big) \in \mathbf{AC}$.

Here is how a derivation tree would look like:

$$\dfrac{\dfrac{\dfrac{\dfrac{\big[c \in (\forall x \in A)\big(\exists y \in B(y)\big)C(x,y)\big]^1 \qquad [a \in A]^2}{c(a) \in \big(\exists y \in B(y)\big)C(a,y)}\,(\Pi\text{-elim})}{p\big(c(a)\big) \in B(a)}\,(\Sigma\text{-elim})}{\lambda x.\, p\big(c(x)\big) \in (\Pi x \in A)B(x)}\,(\Pi\text{-intr}), \cancel{[2]} \qquad \dfrac{\dfrac{\dfrac{\big[c \in (\forall x \in A)\big(\exists y \in B(y)\big)C(x,y)\big]^1 \qquad [a \in A]^4}{c(a) \in \big(\exists y \in B(y)\big)C(a,y)}\,(\Pi\text{-elim})}{q\big(c(a)\big) \in C\Big(a.\, p\big(c(a)\big)\Big)}\,(\Sigma\text{-elim})}{\lambda x.\, q\big(c(x)\big) \in (\forall x \in A)C\big(x, p\big(c(x)\big)}\,(\Pi\text{-intr}), \cancel{[4]}}{\dfrac{\Big(\lambda x.\, p\big(c(x)\big), \lambda x.\, q\big(c(x)\big)\Big) \in \big(\exists f \in (\Pi x \in A)B(x)\big)(\forall x \in A)C\big(x, f(x)\big)}{\lambda c.\Big(\lambda x.\, p\big(c(x)\big), \lambda x.\, q\big(c(x)\big)\Big) \in \mathbf{AC}}\,(\Pi\text{-intr}), \cancel{[1]}}\,(\Sigma\text{-intr})$$

5.2 Additional rules

In order to give an interpretation of **CZF** into **ML** we need to discuss some further rules which will increase the expressivity of our system. As announced at the beginning of the chapter, slightly change our terminology. From now on, we will refer to sets as *types* – and we will correspondingly write "A type" instead of "A set"[5] and "$a : A$" instead of "$a \in A$" and say that "a *is of type* A" instead of "a is an element

[5] Martin-Löf himself changed his terminology in this way.

of A" (we could thus add another line to the table we discussed in chapter 2.3. This will be convenient to prevent confusion, as in order to interpret the language of **CZF** we will introduce a special type V in **ML** containing these sets and using the $\in$-sign accordingly.

5.2.1 $\mathcal{W}$-rules

As noted, the rules discussed so far are enough to develop a constructive set theory on its own. It misses, however, the recursive character of **CZF**, as there is no way to conduct induction on sets within the system. In the following we will thus discuss the missing ingredients. We start with $\mathcal{W}$-rules (here, $\mathcal{W}$ stands for "well-order"). Although we will need only a special instance of these rules, understanding the general case will help us in understanding.

Before explicitly stating the rules, let us discuss the following example from algebra: Suppose, we are given constants a and b, the unary operation symbol s and the binary operation symbol $+$. We can then form the term algebra given by $a, b, s, +$: It consist of terms, where "term" is defined as result of applying operation symbols to terms. Clearly, this has recursive character and would yield an empty result if we did not have any functions of arity 0. Thanks to them, examples of terms are a, b, $s(a)$, $s(a) + b$, $s(s(a) + b)$ and so on.

In general, let A be a set of operations and for each operation $*$ of A, let $B(*)$ be a set with cardinality equal to the arity of $*$. In our example $A = \{a, b, s, +\}$ and $B(a) = B(b) = \emptyset$, $B(s) = \{1_s\}$ and $B(+) = \{1_+, 2_+\}$. We can then define the term algebra $\mathcal{T}$ as all functions from $B(*) \to \mathcal{T}$, where $*$ is an operation of A. In our example, some terms are represented in the following way:

term	corresponds to
$\boldsymbol{a}$	empty function $B(a) \to \mathcal{T}$
$\boldsymbol{b}$	empty function $B(b) \to \mathcal{T}$
$\boldsymbol{s(a)}$	$1_s \mapsto a$
$\boldsymbol{s(a) + b}$	$1_+ \mapsto s(a), 2_+ \mapsto b$

Writing these ideas as rules, we have (we write $\sup(a, b)$ to label the function $b: B(a) \to \mathcal{T}$)

$\mathcal{W}$-formation

$$\frac{A \text{ type} \qquad \begin{matrix}(x : A)\\ B(x) \text{ type}\end{matrix}}{(\mathcal{W}x: A)B(x) \text{ type}}$$

$\mathcal{W}$-introduction

$$\frac{a : A \qquad b: B(a) \to (\mathcal{W}x: A)B(x)}{\sup(a, b) : (\mathcal{W}x: A)B(x)}$$

How do we recursively define a function F on the term algebra $\mathcal{T}$? First, we declare its value on all constants. In our example, we need to fix values $F(a)$ and $F(b)$. Next, we define how F behaves in the inductive step, i.e. we need rules of the form $F\big(s(t)\big) = f\big(t, F(t)\big)$ and $F(t_1 + t_2) = g\big(t_1, t_2, F(t_1), F(t_2)\big)$ for appropriate functions f and g.

To fix ideas, let us think of $\mathcal{T}$ as coding terms in the language of arithmetic. Say, we interpret a as 0, b as 1, s as the successor function and $+$ as addition function and we want to denote by F the function $x \mapsto 2x$ of multiplication by 2 Thus, we would define

$$\begin{aligned} F(a) &= a, \\ F(b) &= s(b), \\ F\big(s(t)\big) &= t + s(b), \\ F(a+b) &= F(t_1) + F(t_2). \end{aligned}$$

Note that technically speaking, we could include the base case (the case for operations with arity 0) into the inducive step. If we define d to be the function computing all inductive steps, and T for the recursion operator (in our case $F(\cdot) = T(\cdot, d)$) we can write

$\mathcal{W}$-elimination

$$\frac{c : (\mathcal{W}x \colon A)B(x) \qquad \begin{array}{c} \big(x : A, y : B(x) \to (\mathcal{W}x \colon A)B(x), z : \big(\Pi v \colon B(x)\big)C\big(y(v)\big)\big) \\ d(x,y,z) : C(\sup(x,y)) \end{array}}{\mathrm{T}(c,d) : C(c)}$$

$\mathcal{W}$-equality

$$\frac{c : (\mathcal{W}x \colon A)B(x) \qquad b : B(a) \to (\mathcal{W}x \colon A)B(x) \qquad \begin{array}{c} \big(x : A, y : B(x) \to (\mathcal{W}x \colon A)B(x), z : \big(\Pi v \colon B(x)\big)C\big(y(v)\big)\big) \\ d(x,y,z) : C(\sup(x,y)) \end{array}}{\mathrm{T}(\sup\,(a,b), d) = d\big(a, b, \lambda v.\,\mathrm{T}(b(v), d)\big) : C(\sup\,(a,b))}$$

5.2.2 *U-rules*

Intuitively, U can be thought of as the "type of all types". However, such a definition would clearly result in a paradox like Russel's. What we can do however, is to define U as the type of *small types* – basically all types constructible from $\mathbb{N}$ and the $\mathbb{N}_k$ without using $\mathcal{W}$-rules.

U-formation

$$U \text{ type}$$

U-introduction

$$\frac{A : U \qquad \begin{array}{c} (x : A) \\ B(x) : U \end{array}}{(\Pi x \colon A)B(x) : U}$$

$$\frac{A : U \qquad \begin{array}{c} (x : A) \\ B(x) : U \end{array}}{(\Sigma x \colon A)B(x) : U}$$

$$\frac{A : U \qquad B : U}{A + B : U}$$

$$\mathbb{N} : U$$

$$\mathbb{N}_k : U \text{ for all } k$$

U-elimination

$$\frac{A : U}{A \text{ type}}$$

5.3 Interpreting **CZF** in **ML**

5.3.1 *The universe V*

Finally, we are ready to define the type of *sets*, $V \equiv (\mathcal{W}x\colon U)x$. From the $\mathcal{W}$-rules we can infer the following V-rules (we write $\{f(x)|x : A\}$ instead of $\sup(a, b)$):

V-formation

$$V \text{ type}$$

V-introduction

$$\dfrac{A : U \qquad\qquad f : A \to V}{\{f(x)|x : A\} : V}$$

V-elimination

$$\dfrac{c : V \qquad\qquad \begin{matrix}\big(A : U, f : A \to V, z : (\Pi v\colon A)C\big(f(v)\big)\big) \\ d(A, f, z) : C(\{f(x)|x : A\})\end{matrix}}{\mathrm{T}(c, d) : C(c)}$$

V-equality

$$\dfrac{A : U \qquad\qquad f : A \to V \qquad\qquad \begin{matrix}\big(B : U, h : A \to V, z : (\Pi v\colon B)C\big(h(v)\big)\big) \\ d(B, h, z) : C(\{h(x)|x : B\})\end{matrix}}{\mathrm{T}(\{f(x)|x : A\}, d) = d\big(A, f, \lambda v.\mathrm{T}(f(v), d)\big) : C(\{f(x)|x : A\})}$$

In the following we will often use recursion on V in the following way to define a function $F : (\Pi x\colon V)C(x)$: Suppose, we have defined $F(\beta)$ for all $\beta = f(x)$ for $x : A$. We then set $F(\{f(x)|x : A\}) = d\big(A, f, \lambda v.\mathrm{T}(f(v), d)\big)$, where d is such that $d\big(A, f, \lambda v.\mathrm{T}(f(v), d)\big) : C(\{f(x)|x : A\})$.

5.3.2 *The interpretation*

To interpret propositions in the language of set theory into appropriate types, we define by an easy recursion $\widetilde{\{f(x)|x : A\}} = \lambda x.f(x)$ and $\overline{\{f(x)|x : A\}} = A$ (with $d(x, y, z) = x$ in the first and $d(x, y, z) = \lambda x.y(x)$ in the second case). Now we define

$\lVert\boldsymbol{\alpha} = \boldsymbol{\beta}\rVert$	$(\Pi x\colon \bar{\alpha})\big(\Sigma y\colon \bar{\beta}\big)\big\lVert\tilde{\alpha}(x) = \tilde{\beta}(y)\big\rVert \times \big(\Pi y\colon \bar{\beta}\big)(\Sigma x\colon \bar{\alpha})\big\lVert\tilde{\alpha}(x) = \tilde{\beta}(y)\big\rVert$
$\lVert\boldsymbol{\alpha} \in \boldsymbol{\beta}\rVert$	$(\Sigma x\colon \bar{\beta})\big\lVert\alpha = \tilde{\beta}(x)\big\rVert$
$\lVert\bot\rVert$	$\mathbb{N}_0$
$\lVert\boldsymbol{\phi} \to \boldsymbol{\psi}\rVert$	$\lVert\phi\rVert \to \lVert\psi\rVert$
$\lVert\boldsymbol{\phi} \wedge \boldsymbol{\psi}\rVert$	$\lVert\phi\rVert \times \lVert\psi\rVert$
$\lVert\boldsymbol{\phi} \vee \boldsymbol{\psi}\rVert$	$\lVert\phi\rVert + \lVert\psi\rVert$
$\lVert(\forall \boldsymbol{x} \in \boldsymbol{\alpha})\boldsymbol{\phi}(\boldsymbol{x})\rVert$	$(\Pi x\colon \bar{\alpha})\big\lVert\phi\big(\tilde{\alpha}(x)\big)\big\rVert$
$\lVert(\exists \boldsymbol{x} \in \boldsymbol{\alpha})\boldsymbol{\phi}(\boldsymbol{x})\rVert$	$(\Sigma x\colon \bar{\alpha})\big\lVert\phi\big(\tilde{\alpha}(x)\big)\big\rVert$
$\lVert\forall \boldsymbol{x}\boldsymbol{\phi}(\boldsymbol{x})\rVert$	$(\Pi\alpha\colon V)\lVert\phi(\alpha)\rVert$
$\lVert\exists \boldsymbol{x}\boldsymbol{\phi}(\boldsymbol{x})\rVert$	$(\Sigma\alpha\colon V)\lVert\phi(\alpha)\rVert$

It is clear that this definition needs some justification. Let us, however, start with some simple examples to get used to the notions:

Examples: Let $e : \mathbb{N}_0 \to V$ be the canonical empty function. Intuitively, $\mathrm{emp} = \{e(x) | x : \mathbb{N}_0\}$ should represent the empty set and we will prove this in section 5.3.4. If $e_p^x : \mathbb{N}_0 \to (\Sigma y : \mathbb{N}_0)\|e(x) = e(y)\|$ is the canonical empty function and if we define e_p^y similarly, then for $e_p \equiv \left(\lambda x. e_p^x, \lambda y. e_p^y\right)$ we have that $e_p :$ $\|\mathrm{emp} = \mathrm{emp}\|$. Note that these lines of proof do not work with some other type A instead one of the $\mathbb{N}_0$.

As a slightly more interesting example, let $\alpha : V$ be given by $\bar{\alpha} = \mathbb{N}_1$, where $\tilde{\alpha}\left(\bar{0}\right) = \mathrm{emp}$ and $\beta : V$ by $\bar{\beta} = \mathbb{N}_2$, where $\tilde{\beta}\left(\bar{0}\right) = \mathrm{emp}$ and $\tilde{\beta}\left(\bar{1}\right) = \mathrm{emp}$. Intuitively, α and β represent the sets $\{\emptyset\}$ and $\{\emptyset, \emptyset\}$ and should thus be identified. Indeed, with e_p as above,

$$\left(\lambda x.\left(\left(\bar{0}, e_p\right)\right), \lambda y.\left(\left(\bar{0}, e_p\right)\right)\right) : (\Pi x\colon \mathbb{N}_1)(\Sigma y\colon \mathbb{N}_2)\|\mathrm{emp} = \mathrm{emp}\| \times (\Pi y\colon \mathbb{N}_2)(\Sigma x\colon \mathbb{N}_1)\|\mathrm{emp} = \mathrm{emp}\|,$$

and this type stands for $\|\alpha = \beta\|$.

Let us now justify the definition of $\|\alpha = \beta\|$. This type clearlyy aims to mimic the axiom of extensionality. It is defined using recursion with $C \equiv V \to U$ and

$$d(B, h, z) = \lambda\beta.\left[(\Pi x\colon B)(\Sigma y\colon \bar{\beta})z(x)\left(\tilde{\beta}(y)\right) \times (\Pi y\colon \bar{\beta})(\Sigma x\colon B)z(x)\left(\tilde{\beta}(y)\right)\right].$$

Indeed, $d(B, h, z) : V \to U$ and we have with $F(x) \equiv \mathrm{T}(x, d)$:

$$\begin{aligned} F(\{f(x) | x : A\}) &\equiv d\left(A, f, \lambda v. F\left(f(v)\right)\right) \\ &\equiv \lambda\beta.\left[(\Pi x\colon A)(\Sigma y\colon \bar{\beta})F\left(f(x)\right)\left(\tilde{\beta}(y)\right) \times (\Pi y\colon \bar{\beta})(\Sigma x\colon A)F\left(f(x)\right)\left(\tilde{\beta}(y)\right)\right]. \end{aligned}$$

We then set $\|\alpha \ = \beta\| \equiv F(\alpha)(\beta)$,

$$\begin{aligned} \|\alpha \ = \beta\| \equiv F(\alpha)(\beta) &\equiv (\Pi x\colon \bar{\alpha})(\Sigma y\colon \bar{\beta})F\left(\tilde{\alpha}(x)\right)\left(\tilde{\beta}(y)\right) \times (\Pi y\colon \bar{\beta})(\Sigma x\colon \bar{\alpha})F\left(\tilde{\alpha}(x)\right)\left(\tilde{\beta}(y)\right) \\ &\equiv (\Pi x\colon \bar{\alpha})\left(\Sigma y\colon \bar{\beta}\right)\left\|\tilde{\alpha}(x) = \tilde{\beta}(y)\right\| \times \left(\Pi y\colon \bar{\beta}\right)(\Sigma x\colon \bar{\alpha})\left\|\tilde{\alpha}(x) = \tilde{\beta}(y)\right\|. \end{aligned}$$

By these observations and an easy induction on the structure of formulas we obtain:

Ⓜ**Lemma 5.1**: For each restricted formula ϕ, $\|\phi\|$ has small type.

Definition 5.2: A sentence $\phi(x_1, \dots, x_n)$ is *valid* iff there is an expression $a(\alpha_1, \dots, \alpha_n)$ such that

$$a(\alpha_1, \dots, \alpha_n) : \ \|\phi(\alpha_1, \dots, \alpha_n)\| \quad (\alpha_1 : V, \dots, \alpha_n : V).$$

Ⓜ**Theorem 5.3**: All theorems of **CZF** are valid.

To prove this theorem, we must find validating expressions for all logical axioms as well as inference rules and finally all axioms of **CZF**, which we will do in the following.

5.3.3 *Validating* **HPL**

Note that our interpretation of the logical connectives is the same as in the propositions-as-types interpretation. Regarding logical axioms, we thus content ourselves with the axioms $\|\phi \land \psi \to \phi\|$ and $\|(\phi \to (\psi \to \chi)) \to ((\phi \to \psi) \to (\phi \to \chi))\|$ we have dealt with before.

As for rules of inference, we already have established the Modus Ponens rule. (UG) and (EI) take interesting forms (in (EI), a does not depend on c):

$$\frac{\begin{array}{c}(a : \|\phi\|, c : V)\\ b(a,c) : \|\psi(c)\|\end{array}}{\lambda a \lambda c.\, b(a,c) : \|\phi \to \forall x\, \psi(x)\|}\ \text{(UG)} \qquad \frac{\begin{array}{c}(a : \|\phi(c)\|)\\ b(a) : \|\psi\|\end{array}}{\lambda a.\, b(a) : \|\exists x\, \phi(x) \to \psi\|}\ \text{(EI)}$$

Finally, we need to verify the axioms of identity. Naturally, this is done via recursion and for this reason will be rather technical. We will skip the technical details to make the argumentation a bit more comprehensible:

To give a validating expression $r_0(\alpha) : \|\alpha = \alpha\|$ we may assume $r_0(\tilde{\alpha}(x)) : \|\tilde{\alpha}(x) = \tilde{\alpha}(x)\|$. Thus, we will set $r_0(\alpha) = \lambda x.\left(\left(x, r_0(\tilde{\alpha}(x))\right), \left(x, r_0(\tilde{\alpha}(x))\right)\right)$.

For transitivity, assume, $c_0\left(\tilde{\alpha}(x), \tilde{\beta}(y)\right) : \left\|\tilde{\alpha}(x) = \tilde{\beta}(y) \to \tilde{\beta}(y) = \tilde{\alpha}(x)\right\|$. Let us abbreviate $el(c,y) \equiv p(q(c)(y))$ and $pr(c,y) \equiv p(q(c)(y))$. Then for $c : \|\alpha = \beta\|$ and $y : \bar{\beta}$, $el(c) = x : \bar{\alpha}$ with $pr(c,y) : \tilde{\alpha}(x) = \tilde{\beta}(y)$. Thus, $c_0\left(\tilde{\alpha}(x), \tilde{\beta}(y)\right)(pr(c,y)) : \left\|\tilde{\beta}(y) = \tilde{\alpha}(x)\right\|$. We will therefore define

$$c_0(\alpha, \beta) \equiv \lambda c.\left(\lambda y.\, c_0\left(\tilde{\alpha}(el(c,y)), \tilde{\beta}(y)\right)(pr(c,y)), \dots\right),$$

where "…" is defined symmetrically to the left member of the pair. Note that technically speaking, this construction should be carried out by a kind of double recursion we used to define $\|\alpha = \beta\|$.

As a final example, us give a validating expression for $w_0 : \|\alpha \in \beta \land \beta \in \gamma \to \alpha \in \gamma\|$. Here, we do need to make direct use of recursion: Let $c : \|\alpha \in \beta \land \beta \in \gamma\|$, and set $el_1(c) \equiv p(p(c))$ and $pr_1(c) \equiv q(p(c))$. Then $el_1(c) : \bar{\beta}$ and $pr_1(c) : \left\|\alpha = \beta(el_1(c))\right\|$. If we go on to set $el_2(c,x) \equiv q(q(q(c))(x))$, then $el_2(c,x) = y : \bar{\gamma}$ and $pr_2(c) \equiv p(q(q(c))(x)) : \|\beta(x) = \gamma(y)\|$. Altogether, we will define with t_0 as above,

$$w_0 \equiv \lambda c.\left(el_2(c, el_1(c)), t_0\left(\alpha, \beta(el_1(c)), el_2(c, el_1(c))\right)\left(pr_1(c), pr_2(c, el_1(c))\right)\right).$$

Note that by the above expressions we can, by an easy induction on the structure of the formula ϕ find a validating expression e_ϕ of $x = y \to (\phi(x) \leftrightarrow \phi(y))$.

Ⓜ**Lemma 5.4**: For any formula ϕ there is a validation expression e_ϕ of $\forall x \forall y\, [x = y \to (\phi(x) \leftrightarrow \phi(y))]$.

5.3.4 Validating the axioms of **CZF**

Before starting to validate the axioms, we need a simple observation:

Definition & Lemma 5.5: If we set $\alpha^* \equiv \lambda x.\big(x, r_0(\tilde{\alpha}(x))\big)$ with r_0 as before, i.e. $r_0(\alpha) : \|\alpha = \alpha\|$, then $\alpha^*(x) : \|\tilde{\alpha}(x) \in \alpha\|$ $(\alpha : V, x : \bar{\alpha})$

Extensionality

The validation of extensionality will of course exploit the fact that the definition of $\|\alpha = \beta\|$ is tailored in such a way to satisfy this axiom: Let α and β be sets and $c : \|\forall z\,(z \in \alpha \leftrightarrow z \in \beta)\|$. Then for $x : \bar{\alpha}$ we have $\alpha^*(x) : \|\tilde{\alpha}(x) \in \alpha\|$ and thus $p\left(c\big(\tilde{\alpha}(x)\big)\right)\big(\alpha^*(x)\big) : \|\tilde{\alpha}(x) \in \beta\| \equiv (\Sigma y{:}\,\bar{\beta})\big\|\tilde{\alpha}(x) = \tilde{\beta}(y)\big\|$. The other direction is symmetric.

Pair

Given sets α and β, define $f(c) = R_2(c, \alpha, \beta)$, i.e. $f : \mathbb{N}_2 \to V$ with $f(0) = \alpha$ and $f(1) = \beta$. Let γ be the set $\{f(x)|x : \mathbb{N}_2\}$. Then, $\gamma^*(0) : \|\alpha \in \gamma\|$ and $\gamma^*(1) : \|\beta \in \gamma\|$. Therefore, $\big(\gamma, \big(\gamma^*(0), \gamma^*(1)\big)\big)$ is of the desired type $\|\exists z(\alpha \in z \wedge \beta \in z)\|$.

Union

Let α be a set and define $\gamma \equiv \{g(z)|z : A\}$, where $A \equiv (\Sigma x{:}\,\bar{\alpha})\overline{\tilde{\alpha}(x)}$ and $g\big((x, y)\big) = \widetilde{\tilde{\alpha}(x)}(y)$ (by Σ-rules). Now if we have $c : \|\beta \in \gamma\|$, then $p(c) = (x, y)$ and $\big(y, q(c)\big) : \|\beta \in \tilde{\alpha}(x)\|$ and $\big(x, \alpha^*(x)\big) : \|\tilde{\alpha}(x) \in \alpha\|$.

On the other hand, let $\delta : V$, $a : \|\beta \in \delta\|$ and $b : \|\delta \in \alpha\|$. Then $p(a) = z : \bar{\delta}$ with $q(a) : \big\|\beta = \tilde{\delta}(z)\big\|$ and $p(b) = x : \bar{\alpha}$ with $q(b) : \|\delta = \tilde{\alpha}(x)\|$. The latter yields $p\big(q(b)\big)(z) = y : \overline{\tilde{\alpha}(x)}$ with $q\big(q(b)\big)(z) : \big\|\tilde{\delta}(z) = \widetilde{\tilde{\alpha}(x)}(y)\big\|$. Using a validating expression for transitivity of equality, we can validate $\beta = \widetilde{\tilde{\alpha}(x)}(y)$.

Empty set

Set $emp \equiv \{f(x)|x : \mathbb{N}_0\}$, where f is the canonical empty function $\mathbb{N}_0 \to V$. If $\alpha : V$, then $c : \|\beta \in emp\|$ is impossible, since there is no $x : \mathbb{N}_0$ that could be used to validate $\|\beta \in \widetilde{emp}(x)\|$. This shows that $\lambda x.x$ validates $\|\neg(\exists x\; x \in emp)\|$.

Infinity

Ⓜ**Lemma 5.6**: For each set α we can find an expression $S(\alpha)$ and validate the formula $S(\alpha) = s(\alpha)$ saying that $S(\alpha)$ is the successor of α.

Proof: Let α be a set and define $S(\alpha) \equiv \{h(\alpha)(y)|y \in \bar{\alpha} + \mathbb{N}_1\}$, where $h(\alpha)\big(i(x)\big) = \tilde{\alpha}(x)$ for $x : \bar{\alpha}$ and $h(\alpha)\big(j(1)\big) = \alpha$. Note that $S(\alpha) = s(\alpha)$ in fact stands for $\forall x\big(x \in S(\alpha) \leftrightarrow (x \in \alpha \vee x = \alpha)\big)$.

Now if $c : \|\beta \in S(\alpha)\|$, then

- Either $p(c) = i(x) : \bar{\alpha} + \mathbb{N}_1$, i.e. $x : \bar{\alpha}$. But then $q(c) : \|\beta = \tilde{\alpha}(x)\|$ which shows $\big(l\big(p(c)\big), q(c)\big) : \|\beta \in \alpha\|$ (where l is defined on $A + B$ as $l\big(i(a)\big) = a$ and $l\big(j(b)\big) = b$)
- Or $p(c) = j(1) : \bar{\alpha} + \mathbb{N}_1$. Then $q(c) : \|\beta = \alpha\|$

Altogether, $\lambda c.\,\mathrm{D}\Big(l(p(c), \big(l(p(c)), q(c)\big), q(c)\Big)$ validates the implication from left to right.

The other direction is similar: Let $c : \|\beta \in \alpha \vee \beta = \alpha\|$.

- If $c : i(b)$ with $b : \|\beta \in \alpha\|$, then $p(b) = x : \bar{\alpha}$ with $q(b) : \|\beta = \tilde{\alpha}(x)\|$ and also $q(b) : \big\|\beta = \widetilde{S(\alpha)}(x)\big\|$.
- If $c : i(b)$ with $b : \|\beta = \alpha\|$, then $b : \big\|\beta = \widetilde{S(\alpha)}(1)\big\|$.

Combining these two cases, we find the validating expression $\lambda c.\,\mathrm{D}\Big(l(p(c)), (p(l(c)), q(l(c)), l(c)\Big)$. ■

We can now set $\omega \equiv \{\Delta(n) | n : \mathbb{N}\}$, where Δ is defined by recursion on $\mathbb{N}$ as $\Delta(0) = emp$ and $\Delta(s(n)) = S(\Delta(n))$. The last equality shows that, for any $n : \mathbb{N}$, $\omega^*(s(n)) : \big\|S(\Delta(n)) \in \omega\big\|$. But in the last lemma we have shown how to validate $s(\Delta(n)) = S(\Delta(n))$. We can thus combine this with validating expression for equality to validate $s(\Delta(n)) \in \omega$.

On the other hand, let $n : \mathbb{N}$. We can easily validate $\|\Delta(n) = 0 \vee (\exists m \in \omega)\, \Delta(n) = s(m)\|$ using recursion on $\mathbb{N}$:

- If $n = 0 : \mathbb{N}$, then $\Delta(n) = emp$ and we can use the validating expression for the axiom of Empty set to validate $\Delta(n) = \emptyset$.
- If $n = s(m)$, we can use m, the fact $\Delta(n) = \Delta(s(m)) = S(\Delta(m))$ and a validating expression for $s(\Delta(m)) = S(\Delta(m))$ from the lemma to validate $(\exists m \in \omega)(n = s(m))$.

Bounded separation

Let α be a set and ϕ a bounded formula. We set $\gamma \equiv \{g(u) | u \in A\}$, where $A \equiv \|\exists x \in \alpha.\,\phi(x)\|$ is a small type by Lemma 5.1 and $g : A \to V$ is defined by $g((x, v)) = \tilde{\alpha}(x)$ for $x : \bar{\alpha}$ and $v : \big\|\,\phi(\tilde{\alpha}(x))\big\|$. Then for $c : \|\beta \in \gamma\|$ we have $p(c) = (x, v) : \bar{\gamma}$, where $q(c) : \|\beta = \tilde{\alpha}(x)\|$, where $v : \big\|\phi(\tilde{\alpha}(x))\big\|$. We can use this together with ⓂLemma 5.4 to validate $\beta \in \alpha \wedge \phi(\beta)$.

On the other hand, for $c : \|\beta \in \alpha \wedge \phi(\beta)\|$ we have $p(p(c)) = x : \bar{\alpha}$ with $q(p(c)) : \|\beta = \tilde{\alpha}(x)\|$. The expression $q(c) : \|\phi(\beta)\|$ will yield a validation for $\big\|\phi(\tilde{\alpha}(x))\big\|$ and all in all, we would have expressions validating $\beta \in \gamma$.

Strong collection

Let α be a set, $\phi(x, y)$ a formula and $c : \|(\forall x \in \alpha)\exists y\, \phi(x, y)\|$. Define $\beta \equiv \{b(x) | x : \bar{\alpha}\}$, where we set $b \equiv \lambda x.\, p(c(x))$. Now for all $x : \bar{\alpha}$ $(= \bar{\beta})$, we have $q(c)(x) : \Big\|\phi\left(\tilde{\alpha}(x), \tilde{\beta}(x)\right)\Big\|$. This shows that

$$\Big(\lambda x.\,(x, q(c)(x)), \lambda y.\,(y, q(c)(y))\Big) : \|[(\forall x \in \alpha)(\exists y \in \beta)\, \phi(x, y)] \wedge [(\forall x \in \beta)(\exists y \in \alpha)\, \phi(x, y)]\|.$$

Subset collection

Given sets α and β, define $\gamma \equiv \{g(z) | z : \bar{\alpha} \to \bar{\beta}\}$, where $g \equiv \lambda z.\{\tilde{\beta}(z(x) | x : \bar{\alpha}\}$. Let η be a set and $c : \|\forall x \in a\, \exists y \in b\, \phi(x, y, \eta)\|$. We define $\delta \equiv \gamma\left(\lambda x.p(c(x))\right) \equiv \left\{\tilde{\beta}\left(p(c(x))\right) \middle| x : \bar{\alpha}\right\}$. Then, by Lemma 5.5, $\gamma^*\left(\lambda x.p(c(x))\right)$ validates $\delta \in \gamma$. If $x : \bar{\alpha}$ $(= \tilde{\delta})$, then $q(c(x)) : \|\phi(\tilde{\alpha}(x), \tilde{\delta}(x), \eta)\|$. This shows

$$\left(\lambda x.\left(x, q(c(x))\right), \lambda y.\left(y, q(c(y))\right)\right) : \|\forall x \in \alpha.\exists y \in \delta.\phi(x, y, u) \wedge \forall y \in \delta.\exists x \in \alpha.\phi(x, y, u)\|.$$

Set induction

Naturally, this is validated using recursion: Let us abbreviate $B \equiv \|\forall a(\forall y \in a\, \phi(y) \to \phi(a))\|$, i.e.

$$B \equiv (\Pi\alpha : V)\left((\Pi x : \tilde{\alpha})\|\phi(\tilde{\alpha}(x))\| \to \|\phi(\alpha)\|\right).$$

Suppose, we have defined h such that $h(\alpha) = \lambda b.b(\alpha)\left(\lambda x.h(\tilde{\alpha}(x))(b)\right) : B \to \|\phi(\alpha)\|$. Since for $\alpha : V$,

$$b(\alpha) : \|\phi(\tilde{\alpha}(x))\| \to \|\phi(\alpha)\|,$$

and for $b : B$,

$$\lambda x.h(\tilde{\alpha}(x))(b) : (\Pi x : \bar{\alpha})\|\phi(\tilde{\alpha}(x))\|,$$

we have that

$$b(\alpha)\left(\lambda x.h(\tilde{\alpha}(x))(b)\right) : \|\phi(\alpha)\|.$$

This shows that $\lambda b.\lambda\alpha.h(\alpha)(b)$ does the job.

6 Topological Semantics and Independence of Bar induction

We will show in this section, that the Decidable Bar Theorem, that plays a central role in Brouwer's mathematics, is independent from **IZF**. We will start with a brief discussion of the Bar Theorem in 6.1. Our main tool will be topological semantics, where each logical statement is interpreted as open set in a topological space. We will motivate the usage of topological semantics by generalizing the classical Boolean-valued semantics to Heyting-valued semantics in section 6.2. In section 6.3 we will describe the topological model that we will use in the independence and consistency proofs in 6.4-6.6.

6.1 Bar induction in Brouwer's mathematics

In order to obtain the Uniform Continuity Theorem (every function $f: [0,1] \to \mathbb{R}$ is uniformly continuous), Brouwer showed in a proof of rather metamathematical fashion the Decidable Bar Theorem $\mathrm{BI_D}$ from which he obtained as a corollary the Fan Theorem **F**:

$$(\mathrm{BI_D}) \qquad \begin{bmatrix} \mathrm{bar}(B, \omega^\omega) \wedge \\ \forall u \in B\; Q(u) \wedge \\ \forall u \in \omega^{<\omega}\big(\forall k \in \omega\; Q(u * \langle k \rangle) \to Q(u)\big) \end{bmatrix} \to Q(\langle\, \rangle),$$

$$(\mathbf{F}) \qquad \begin{bmatrix} \mathrm{fan}(T) \wedge \\ \mathrm{bar}(B, T) \end{bmatrix} \to \exists z \in \omega\; \forall \alpha \in T\; \exists x \leq z.\, \alpha|_x \in B.$$

Here, $\mathrm{bar}(B, A)$ means $A \subseteq \omega^\omega$ and B is a decidable bar in B, i.e. $\forall \alpha \in A\; \exists n \in \omega.\, \alpha|_n \in B$ and $\forall u \in A\; (u \in B \vee u \notin B)$. Although constructively not unproblematic, as it requires the acceptance of actual infinity, we will remain agnostic about this proof and follow Heyting, Kleene, Trolestra, Dummet and others in accepting BI_D itself as constructively sensible[6] and adopting it as an axiom. For the original proof and thorough analysis and discussion of BI_D and **F**, see [*70*] or [*50*]. From BI_D, we can derive the Fan theorem and the Uniform Continuity Theorem:

Fan-Theorem 6.1: BI_D implies **F**.

Proof: $Q(u)$ be the formula $\exists z \in \omega. \forall \alpha \supseteq u. \exists x \leq z. \alpha|_x \in B$. It satisfies the two conditions of (BI_D) and we thus obtain $Q(\langle\, \rangle)$, which is exactly the consequence of F. ∎

In chapter 2.2 we defined, in Brouwer's spirit, the real numbers $\mathbb{R}$ as members of the spread S of Cauchy-sequences of rational numbers. Similarly, the closed unit interval $[0,1]$ can be given as binary spread with nodes labelled by $\left[\frac{i}{2^k}, \frac{i+1}{2^k}\right]$ for $i < k$ and $k > 0$. With these definitions we can now prove Brouwer's Uniform Continuity Theorem:

[6] Brouwer himself admitted this in footnote 7 in [*11*]

Theorem 6.2 (**IZF** + **WCN** + **F**): Every function $f: [0,1] \to \mathbb{R}$ is continuous.

Sketch of proof: Given f, we define the functions $f_k: [0,1] \to \mathbb{N}$ as $f_k(\alpha) = f(\alpha)_k$, i.e. f_k assigns to f its k-th value of its expansion as element of the spread S. By **WCN**, for each α, there is some n such that $f_k(\alpha) = f_k(\beta)$, if only $\alpha|_n = \beta|_n$. Effectively, we can write $f_k(\alpha) = f_k(\alpha|_n)$. By **F**, applied to the formula

$$\phi(u) \equiv \forall v\big(u = v \to f_k(u) = f_k(v)\big),$$

we conclude that there is some m such that for all α, β and some $n < m$, we have $\alpha|_n = \beta|_n \to f_k(\alpha|_n) = f_k(\beta|_n)$. Translating this into the language of analysis, this means that for all k, there is some m such that

$$|x - y| < 2^{m+1} \to |f(x) - f(y)| < 2^{k+1},$$

which shows that f is uniformly continuous.[7] ■

The rest of the chapter is devoted to showing that in the context **IZF**, the assumption of BI_D is too strong, i.e. both Fan theorem and Uniformity Theorem can be proved without BI_D.

6.2 Heyting-valued semantics

Heyting-valued semantics (in the literature also referred to as “Heyting-algebra semantics”, [*28*]) are the most natural semantics for intuitionistic logic and can be seen as generalization of Boolean-valued semantics. Recall, that the Boolean algebra $B_{0,1}$ is defined on the set $\{0,1\}$ with the usual operations of $\wedge$, $\vee$, $\to$ and $\neg$. A *Boolean-valued interpretation* of propositional logic is a mapping from the set of propositional variables into $\{0,1\}$. The *truth value* of a formula in the propositional language is then computed inside $B_{0,1}$. As it is known, formula is classically derivable iff its truth value is 1 under every Boolean-valued interpretation. For example, the law of excluded middle **LEM**, holds true in classical logic, as is easily seen by the following truth table:

$\boldsymbol{a}$	$\neg \boldsymbol{a}$	$\boldsymbol{a} \vee \neg \boldsymbol{a}$
0	1	**1**
1	0	**1**

Because of this validity of **LEM** in every Boolean algebra, we will have to pass from Boolean algebras to Heyting-algebras if we want to give a similar semantics for intuitionistic logic. Bluntly speaking, a Heyting-algebra is a Boolean algebra, but without **LEM** required to hold.

Definition 6.3: A *Heyting-algebra* $(H, 0, 1, \wedge, \vee, \leq)$ is a bounded lattice such that the *pseudo-complement of* a *with respect to* b, $a \to b \coloneqq \bigvee\{x: x \wedge a \leq b\}$ always exists. We define the *pseudo-complement* of a as $\neg a \coloneqq a \to 0$. A Heyting-algabra is a *Boolean algebra* iff $a \vee \neg a = 1$ for all a.

[7] Actually, it remains to explain why $|x - y| < 2^{m+1}$ means that we may assume that $x|_m = y|_m$.

A *Heyting-valued interpretation* of a propositional language $\mathcal{L}=$ is a mapping $p \mapsto p^{(H)}$ of propositional variables of $\mathcal{L}$ into a Heyting-algebra H. As usual, this interpretation extends naturally to a mapping $\phi \mapsto [\![\phi]\!]_H$ of all formulas of $\mathcal{L}$ into H:

$$\begin{aligned}
[\![p]\!]_H &= p^{(H)} \\
[\![\bot]\!]_H &= 0 \\
[\![\top]\!]_H &= 1 \\
[\![\phi \wedge \psi]\!]_H &= [\![\phi]\!]_H \wedge [\![\psi]\!]_H \\
[\![\phi \vee \psi]\!]_H &= [\![\phi]\!]_H \vee [\![\psi]\!]_H \\
[\![\phi \to \psi]\!]_H &= [\![\phi]\!]_H \to [\![\psi]\!]_H \\
[\![\neg\phi]\!]_H &= [\![\phi]\!]_H \to 0
\end{aligned}$$

We say that a formula ϕ is valid under a Heyting-valued interpretation iff $[\![\phi]\!]_H = 1$.

With this definition we now have a similar completeness proof as in the classical case (however the special role of the two-valued algebra vanishes):

Ⓜ**Theorem 6.4**: A formula is derivable in intuitionistic propositional logic iff it is valid in every Heyting-valued interpretation.

Sketch of proof: The completeness proof is very similar to the classical case. For soundness, let us show how to verify the axiom $p \to (q \to p)$, i.e. $[\![p \to (q \to p)]\!]_H = p^{(H)} \to \left(q^{(H)} \to p^{(H)}\right) = 1$ for every Heyting-valued interpretation. Indeed, for $a, b \in H$, by the definition of "→", we have that $a \to (b \to c) = 1$ iff $1 \wedge a \leq (b \to a)$. But this is clearly the case, as $b \to a$ is the largest element x with $x \wedge a \leq b$. Our a is such an x and we are done.

Let us show the DET-rule: If $[\![\phi]\!]_H = 1$ and $[\![\phi \to \psi]\!]_H = 1$, then $[\![\psi]\!]_H = 1$. But $[\![\phi \to \psi]\!]_H = [\![\phi]\!]_H \to [\![\psi]\!]_H = 1$ means that $[\![\phi]\!]_H \leq [\![\psi]\!]_H$, and we immediately conclude $q = 1$. ■

Given this completeness result (actually soundness suffices) we are ready to give another independence result of **LEM**:

Theorem 6.5: The law of excluded middle (**LEM**) does not hold in all Heyting-algebras. Hence, it cannot be derived in propositional intuitionistic logic.

Proof: Consider the Heyting-algebra H with the underlying set $\left\{0, \frac{1}{2}, 1\right\}$ and the usual ordering. Then, with $p^{(H)} = \frac{1}{2}$ for a propositional variable p,

$$[\![p \vee \neg p]\!]_H = p^{(H)} \vee \neg p^{(H)} = \frac{1}{2} \vee \neg\frac{1}{2} = \frac{1}{2} \vee 0 = \frac{1}{2} \neq 1.$$

By soundness of Heyting-valued semantics, **LEM** cannot be derivable. ■

Before we pass on to first-order logic, we prove some easy, but useful facts about Heyting-algebras:

Lemma 6.6: In any Heyting-algebra H and for any $a, b, c \in H$ the following rules hold:

1. $a \to a = 1$.
2. $a \to (b \wedge c) = (a \to b) \wedge (a \to c)$
3. If $a \leq b$, then $a \to b = 1$.
4. If $a \leq b$, then $c \to a \leq c \to b$.
5. If $a \leq b$, then $a \to c \geq b \to c$.

Proof:

1. Obvious from the definition.
2. $[a \to (b \wedge c)] \wedge a \leq b \wedge c$, hence $a \to (b \wedge c) \leq (a \to b) \wedge (a \to c)$. The other direction follows from $[(a \to b) \wedge (a \to c)] \wedge a = [(a \to b) \wedge (a \to c)] \wedge (a \wedge a) = [(a \to b) \wedge a] \wedge [(a \to c) \wedge a] = b \wedge c$.
3. Using 2 in the third step, $1 = a \to a = a \to (a \wedge b) = (a \to a) \wedge (a \to b) = 1 \wedge (a \to b) = a \to b$.
4. Since $(c \to a) \wedge c \leq a \leq b$, we also have $c \to a \leq c \to b$.
5. $a \wedge (b \to c) \leq b \wedge (b \to c) = b \wedge c \leq c$ and hence, $b \to c \leq a \to c$. ■

6.2.1 Heyting-semantics for first-order logic

We take the mimicking of classical semantics a bit further and consider classical models. In addition to the symbols $\wedge$, $\vee$, $\to$ and $\neg$, one interprets functions, relations and quantifier symbols as functions functions from a given domain $\mathcal{D}$ onto itself, functions from tuples of elements of $\mathcal{D}$ to $\{0,1\}$ and the truth values $\bigvee_{d \in \mathcal{D}} \phi(d)$ for existential and $\bigwedge_{d \in \mathcal{D}} \phi(d)$ for universal quantifiers. While in the Boolean-algebra $B_{0,1}$ these large infima and suprema always exist, we will have to postulate this for Heyting-algebras:

Definition 6.7: A Heyting-algebra is called *complete* if it is a complete lattice and it satisfies the following $\wedge \bigvee$-*distributive law*:

$$p \wedge \bigvee_{i \in I} q_i = \bigvee_{i \in I} (p \wedge q_i),$$

for all index sets I, $p \in H$ and $\{q_i : i \in I\} \subseteq H$.

Definition 6.8: Given a first-order language $\mathcal{L} = (\mathrm{Var}, \mathrm{Con}, \mathrm{Rel}, \mathrm{Fun})$ of variables Var, constants Con, realtions Rel and functions Fun, we define a *Heyting-valued interpretation* of $\mathcal{L}$ to be a mapping $a \mapsto a^{(H)}$ of constant symbols, $R \mapsto R^{(H)}$ of relation symbols, $F \mapsto F^{(H)}$ of function symbols such that $a^{(H)} \in \mathcal{D}$, $R^{(H)}: \mathcal{D}^n \to H$ if the arity of R is n and $F^{(H)}: \mathcal{D}^n \to \mathcal{D}$, if the arity of F is n.

We define an H-*term* by the following rules: All variables of Var are H-terms and so are all $a^{(H)}$ for $a \in \mathrm{Con}$. If $t_1, \dots, t_n$ are H-terms, then so is $F^{(H)}(t_1, \dots, t_n)$, if $F \in \mathrm{Fun}$ and the arity of F. H-*formulas* are defined as usually: $\bot$ and $\top$ are H-formulas and so is $R^{(H)}(t_1, \dots, t_n)$, whenever $t_1, \dots, t_n$ are H-terms, $R \in \mathrm{Rel}$ and the arity of R is n. If ϕ and ψ are H-formulas, then so are $\phi \wedge \psi$, $\psi \vee \psi$, $\neg\psi$, $\forall x\, \phi$ and $\exists x\, \phi$. As usual, we write $\phi(x)$ if the variable x occurs freely in ϕ and $\phi(d)$ for the result of substituting $d \in \mathcal{D}$ for x in ϕ.

Note that H-formulas may or may not be members of H. We can now extend the Definition 6.3 of Heyting-interpretations of propositional logic to first-order logic[8]: For simplicity, we write $x^{(H)}$ for $x \in$ Var.

$$\begin{aligned}
[\![R(a_1, \dots a_n)]\!]_H &= R^{(H)}\left(a_1^{(H)}, \dots, a_n^{(H)}\right), \quad \text{for } a_1, \dots a_n \in \mathsf{Var} \cup \mathsf{Con}\\
[\![\bot]\!]_H &= 0,\\
[\![\top]\!]_H &= 1,\\
[\![\phi \wedge \psi]\!]_H &= [\![\phi]\!]_H \wedge [\![\psi]\!]_H,\\
[\![\phi \vee \psi]\!]_H &= [\![\phi]\!]_H \vee [\![\psi]\!]_H,\\
[\![\phi \to \psi]\!]_H &= [\![\phi]\!]_H \to [\![\psi]\!]_H,\\
[\![\neg\phi]\!]_H &= [\![\phi]\!]_H \to 0,\\
[\![\forall x\, \phi(x)]\!]_H &= \bigwedge_{d \in \mathcal{D}} [\![\phi(d)]\!]_H\,,\\
[\![\exists x\, \phi(x)]\!]_H &= \bigvee_{d \in \mathcal{D}} [\![\phi(d)]\!]_H.
\end{aligned}$$

We say that a formula $\phi(x_1, \dots x_n)$ is *valid* under a Heyting-interpretation iff $[\![\phi(c_1, \dots c_n)]\!]_H = 1$ for all $c_1, \dots, c_n \in \mathcal{D}$. Note that the Tarski-semantics is just a special case of this definition, where $H = B_{0,1}$. It will therefore come as no surprise, that the proof of completeness of this semantics known from classical first-order logic can be carried over to the intuitionistic setting, see [*69*]:

Ⓜ**Theorem 6.9**: A formula is derivable in intuitionistic firsr-order logic iff it is valid in every Heyting-valued interpretation.

Proof: Let us again exemplarily verify some first-order axioms: Let H be a complete Heyting-algebra and $\mathcal{D}$ a domain.

We verify the axiom $\forall x\, \phi(x) \to \phi(c)$, where c is free for x in ϕ and suppose the existence of $[\![\forall x\, \phi(x)]\!]$ is guaranteed. The verification of the axiom can be now done in one line:

$$[\![\forall x\, \phi(x)]\!] = \bigwedge_{d \in \mathcal{D}} [\![\phi(d)]\!] \leq \left[\!\!\left[\phi\left(c^{(H)}\right)\right]\!\!\right] = [\![\phi(c)]\!],$$

and hence $[\![\forall x\, \phi(x) \to \phi(c)]\!] = 1$ by Lemma 6.6.

[8] Whether or not the expressions for universal and existential quantification are well-defined, depends on the specific domain $\mathcal{D}$. Although it is certainly true if $\mathcal{D}$ is a set, our $\mathcal{D}$ will be a proper class.

We show that the UG-rule holds: Suppose, $[\![\phi \to \psi(c)]\!] = 1$ (i.e. $[\![\phi]\!] \leq [\![\psi(c)]\!]$) holds in any complete Heyting-algebra H, where c is free for x in ϕ and occurs free in neither ϕ nor ψ. In particular, for given H and domain $\mathcal{D}$, we may interpret c as we please and hence $[\![\phi]\!] \leq \bigwedge_{d\in\mathcal{D}}[\![\psi(d)]\!]$. We conclude $[\![\phi]\!] \leq [\![\forall x\, \psi(x)]\!]$ and thus, $[\![\phi \to \forall x\, \psi(x)]\!] = 1$. ■

We will need the next lemma, to show that the interpretations of quantifiers exist, even if the underlying domain is a proper (hierarchical) class:

Ⓜ**Lemma 6.10**: Let H be a complete Heyting-algebra and $F\colon \mathbb{ON} \to H$ an increasing or decreasing (class) function. Then there is a least $\alpha \in \mathbb{ON}$ such that F is constant above α, i.e. $F(\beta) = F(\alpha)$ for all $\beta \geq \alpha$.

Proof: Let F be an increasing function, the proof for decreasing functions is similar. If F is never constant, we can define a strictly increasing class function $G\colon \mathbb{ON} \to H$ by recursion:

$$\begin{aligned} G(0) &= 0_H, & & \\ G(\alpha+1) &= F(\gamma), & & \text{where } \gamma \text{ is least with } \gamma > \alpha \text{ and } F(\gamma) > F(\alpha) \\ G(\lambda) &= \bigvee_{\alpha<\lambda} G(\alpha), & & \text{for limit ordinals } \lambda. \end{aligned}$$

Hence, G is a bijection between the proper class $\mathbb{ON}$ and $\mathrm{Im}(G) \subseteq H$. This shows that H must be a proper class as well, contradiction. ■

In the sections 6.4-6.6 we will prove that BI_D is not needed to prove the uniform continuity theorem from **IZF**. These results rely on the following theorem:

Ⓜ**Theorem 6.11:** Let **T** be any first-order theory over the language $\mathcal{L}$ of set theory. Suppose that in $\mathcal{L}$ we can define a set H (as witness of a formula $\exists!\, x\, \theta(x)$) such that

$$\begin{aligned} &\mathbf{ZFC} \vdash H \text{ is a complete Heyting-algebra} & & \text{and} \\ &\mathbf{ZFC} \vdash [\![\tau]\!]_H = 1_H & & \text{for each theorem } \tau \text{ of } \mathbf{T}. \end{aligned}$$

Then $\mathrm{Cons}(\mathbf{ZFC}) \Rightarrow \mathrm{Cons}(\mathbf{T})$.

Proof: If **T** is inconsistent, then, $\mathbf{T} \vdash \bot$ and by hypothesis,

$$\mathbf{ZFC} \vdash [\![\bot]\!]_H = 1_H.$$

But we know $\mathbf{ZFC} \vdash [\![\bot]\!]_H = 0_H$ and hence $\mathbf{ZFC} \vdash 0_H = 1_H$, showing $\neg\mathrm{Cons}(\mathbf{ZFC})$. ■

In section 6.3.2, we will show that Heyting algebras give rise to Heyting-valued models of **IZF**, i.e. $[\![\tau]\!]_H = 1$ for all theorems of **IZF**. In sections 6.4-6.6 we will investigate a special Heyting-algebra, where $[\![\mathrm{BI}_D]\!] = 0$ and $[\![\mathbf{F}]\!] = [\![\mathbf{WCN}]\!] = 1$. Hence, Theorem 6.11 will yield

$$\mathrm{Cons}(\mathbf{ZFC}) \Rightarrow \mathrm{Cons}(\mathbf{IZF} + \neg\mathrm{BI}_D + \text{Uniform continuity theorem}).$$

6.2.2 Topologies as examples of Heyting-algebras

The most important examples of a Heyting-algebra will be topological spaces. Let us recall the following definitions:

Definition 6.12: Let X be a set. A *topology* or *topological space* on X or simply *topology* is a set $\mathcal{T} \subseteq \mathcal{P}(X)$ such that

1) $\emptyset \in \mathcal{T}, X \in \mathcal{T}$.
2) For $\mathcal{T}' \subseteq \mathcal{T}$: $\bigcup \mathcal{T}' \in \mathcal{T}$.
3) For $\mathcal{T}'_{fin} \subseteq \mathcal{T}$ with $\mathcal{T}'_{fin}$ finite: $\bigcap \mathcal{T}'_{fin} \in \mathcal{T}$.

The elements of $\mathcal{T}$ are called *open* sets.

Lemma 6.13: Every topology defines a complete Heyting-algebra, if we interpret the order as set-inclusion and the operations $O_1 \vee O_2 = O_1 \cup O_2$ and $O_1 \wedge O_2 = O_1 \cap O_2$.

Proof: Note that an arbitrary intersection of open sets it is not guaranteed to be open. Hence, we can easily check that for $\mathcal{O} \subseteq \mathcal{T}$, $\bigwedge \mathcal{O} = \operatorname{int}(\bigcap \mathcal{O})$. Pseudo-complements are given like this:

$$O_1 \to O_2 = \bigcup \{O \in \mathcal{T} : O \cap O_1 \subseteq O_2\}.$$

It is easy to check that this forms a complete lattice. Moreover, the $\wedge\bigvee$-distributive-law holds, as

$$\begin{aligned} & x \in O \cap \bigcup_{U \in \mathcal{U}} U \\ \text{iff} \quad & x \in O \text{ and } x \in U_0 \text{ for some } U_0 \in \mathcal{U} \\ \text{iff} \quad & x \in O \cap U_0 \text{ for some } U_0 \in \mathcal{U} \\ \text{iff} \quad & x \in \bigcup_{U \in \mathcal{U}} (O \cap U). \end{aligned}$$

■

Example: For any set X, $\{\emptyset, X\}$ and $\mathcal{P}(X)$ are topologies on X, called the *trivial* and *discrete* topology respectively.

Definition 6.14: Let $\mathcal{T}$ be a topological space on X. Then $\mathcal{B}$ is called a *basis* for $\mathcal{T}$ if for all $U \in \mathcal{T}$ there is $\mathcal{B}' \subseteq \mathcal{B}$ with $U = \bigcup \mathcal{B}'$ or equivalently, if for all $O \in \mathcal{T}$ and $x \in O$ there is $U \in \mathcal{B}$ with $x \in U \subseteq O$.

Proposition 6.15: Let $\mathcal{B} \subseteq \mathcal{P}(X)$. Then $\mathcal{B}$ is the basis of a topology iff

1) $\bigcup \mathcal{B} = X$.
2) For $O, U \in \mathcal{B}$ there is $\mathcal{B}' \subseteq \mathcal{B}$ with $O \cap U = \bigcup \mathcal{B}'$.
   Equivalently: For all $O_1, O_2 \in \mathcal{B}$ and $x \in O_1 \cap O_2$ there is $U \in \mathcal{B}$ with $x \in U \subseteq O_1 \cap O_2$.

Proof: Set $\mathcal{T} = \{\bigcup \mathcal{B}' \mid \mathcal{B}' \subseteq \mathcal{B}\}$. ■

The archetype of a topological space is the Euclidean topology of the real line:

Example 6.16: Let $\mathcal{B} = \{(a,b) | a,b \in \mathbb{R}\}$ then $\mathcal{B}$ is the basis of a topology on $\mathbb{R}$ since

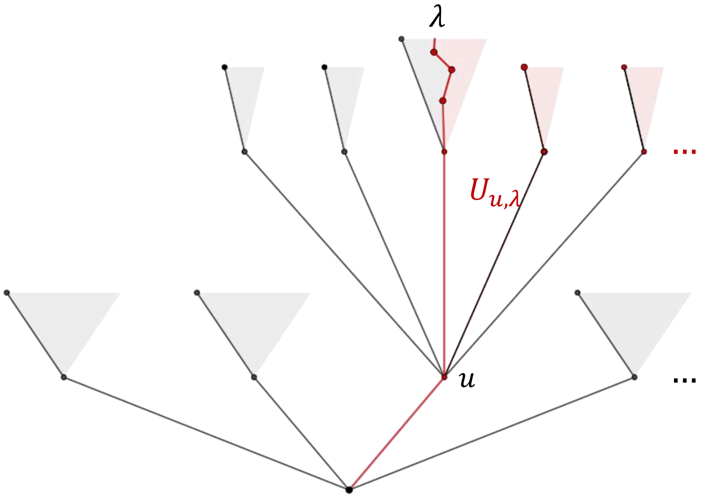
1) $\bigcup \mathcal{B} \supseteq \bigcup\{(n, n+1) | n \in \mathbb{Z}\} = \mathbb{R}$.
2) $(a,b) \cap (c,d) = \begin{cases} \emptyset, & \text{if } d \leq a \text{ or } b \leq a \text{ or } d \leq c \\ (c,d), & \text{if } a \leq c < d \leq b, \\ (c,b), & \text{if } a \leq c < b \leq d, \\ (a,d), & \text{if } c \leq a < d \leq b, \\ (a,b), & \text{if } c \leq a < b \leq d. \end{cases}$

The following example on the space $\mathbb{N}^*$ of finite sequences of natural numbers will play an important role for us:

Example 6.17: We define for $u \in \mathbb{N}^*$ and $\lambda \in \mathbb{N}^{\mathbb{N}}$ with $\lambda|_{\operatorname{dom} u} = u$ (an infinite path through u), define

$$U_{u,\lambda} = \{u\} \cup \left\{v \in \mathbb{N}^* \middle| v \supseteq u \text{ and } v(i_{v,\lambda}) > \lambda(i_{v,\lambda}) \text{ for the least } i_{v,\lambda} \notin \operatorname{dom}(v \cap \lambda)\right\}.$$

$\mathbb{N}^*$ is best thought of as a tree. The $U_{u,\lambda}$ can be pictured as set of all nodes of sequences extending u and lying to the right of λ:



We can also think of $U_{u\lambda}$ as nodes with length at least $\operatorname{dom}(u)$ and coming after λ in the lexicographical ordering. Let us check that this defines a topology $\mathcal{T}$ on $\mathbb{N}^*$:

1) $\bigcup U_{u,\lambda} = \mathbb{N}^*$ is clear.
2) Let $w \in U_{u,\lambda} \cap U_{v,\mu}$, w.lo.g. $v \supseteq u$. Let $i = \max\{i_{v,\lambda}, i_{\mu,\lambda}\}$. Then for $\gamma = \max\{\lambda, \mu\}$ we have
$$w \in U_{w,\gamma} \subseteq U_{u,\lambda} \cap U_{v,\mu}.$$

According to Lemma 6.13, we can define for each topology $(X, \mathcal{T})$ a Heyting-algebra $H^{(X,\mathcal{T})}$. We can also define a mapping in the other direction: Given a Heyting-algebra H, and $h \in H$, we define $O_h =$

$\{w \in H : w \leq h\}$. We may invoke Proposition 6.15 to show that the O_h form a basis of a topology $\mathcal{T}_H$ on H: Indeed, $\bigcup_{h \in H} O_h \supseteq O_1 = H$ and $O_{h_1} \cap O_{h_2} = O_{h_1 \wedge h_2}$. We thus have:

Theorem 6.18: Heyting-algebras and topological spaces are in 1-1 correspondence.

Also, it is clear that both constructions are inverses of one another. To show that both categories are dual, one must pass from general topologies to a certain subcategory of Heyting-spaces, see [*46*].

6.2.3 More facts about topology

In this section we will recall some basic definitions and facts about topological spaces. All of it and more on topology can be found in introductory books on the topic, like [*47*] or [*35*].

Let $(X, \mathcal{T})$ be a topological space. We call complements of open sets *closed*. Sets that are both open and closed are referred to as *clopen*. For any set $M \subseteq X$, we define its *interior* $\operatorname{int}(M)$ to be the largest open set contained in M, i.e. $\operatorname{int}(M) = \bigcup\{O \in \mathcal{T} : O \subseteq M\}$ and its *closure* $\operatorname{cl}(M)$ to be the least closed set containing M, i.e. $\operatorname{cl}(M) = \bigcap\{A : A \text{ closed} \wedge A \supseteq M\}$ (note that arbitrary intersections of closed sets are always closed). A set U is called neighborhood of x iff there is an open set O with $x \in O \subseteq U$. A set M is called *compact* iff each open cover of M has a finite subcover, i.e. $(\mathcal{O} \subseteq \mathcal{T} \wedge \bigcup \mathcal{O} \supseteq M) \rightarrow (\exists \mathcal{O}' \subseteq \mathcal{O}. \mathcal{O}' \text{ finite} \wedge \bigcup \mathcal{O}' \supseteq M)$. The topological space $(X, \mathcal{T})$ is called compact iff X is compact. A set $\mathcal{S} \subseteq \mathcal{T}$ is called *subbasis* iff every open set can be written as arbitrary union of finite intersection of members of $\mathcal{S}$.

Definition 6.19: Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces and $x \in X$. A function $f: X \rightarrow Y$ is called *continuous at* x iff for each neighborhood O of $f(x)$ there is some neighborhood U of x such that $f(y) \in O$ for all $y \in U$. With the definition of the pointwise image $f[A] = \{f(x) : x \in A\}$ we can write this as $f[U] \subseteq O$. A function is called *continuous* if it is continuous at x for each $x \in X$.

We easily see that a function f is continuous iff preimages of open sets are open sets: $f^{-1}[O] \coloneqq \{x \in X : f(x) \in U\} \in \mathcal{T}_X$ for all $O \in \mathcal{T}_Y$. Also, it is clear that our attention in both characterizations may be restricted to basic or subbasic open sets. We have the following simple fact:

Proposition 6.20: Let $f: X \rightarrow Y$ be a continuous function. If X is compact, then so is $f[X]$.

Proof: Let $\mathcal{O}$ be an open cover of Y. Then $\{f^{-1}[O] : O \in \mathcal{O}\}$ is an open cover of X. By compactness, there is a finite subcover $\{f^{-1}[O] : O \in \mathcal{O}'\}$ of X. Then $\mathcal{O}'$ must be an open cover of $f(X)$, since

$$f[X] \subseteq f\left[\bigcup_{O \in \mathcal{O}'} f^{-1}[O]\right] = \bigcup_{O \in \mathcal{O}'} f[f^{-1}[O]] \subseteq \bigcup_{O \in \mathcal{O}'} O.$$

■

Proposition 6.21: A closed subset of a compact topological space is compact itself.

Proof: Let A be a closed subset of the compact space X. Let $\mathcal{O}$ be an open cover of A, then $\mathcal{O} \cup \{X \setminus A\}$ is an open cover of X. By compactness, there is a finite subcover $\mathcal{O}'$ (possibly containing $X \setminus A$). But then $\mathcal{O}' \setminus \{X \setminus A\}$ is a finite cover of A. ■

Lemma 6.22: Let X, Y be topological spaces and g and h continuous functions $X \to Y$. Let O be an open set in X and define

$$f(x) = \begin{cases} g(x), & \text{if } x \in O, \\ h(x), & \text{if } x \notin O, \end{cases}$$

then f is continuous iff it is continuous on $\partial O = \mathrm{cl}(O) \setminus O$.

Proof: For each $x \in O$, continuity of f at x follows from the continuity of g, for $x \in X \setminus \mathrm{cl}(O)$, from h. All that is left to show is continuity at ∂O. ■

Definition 6.23: We say that a sequence $(x_n)_{n\in\omega} \subseteq X$ *converges to* $x \in X$ iff for each neighborhood U of x, there is some m such that for all $n > m$, $x_n \in U$.

6.3 Heyting-valued interpretation of **IZF**

Using the ideas from section 6.2, we will define Heyting-valued models for the set theory **IZF**. Much of the concepts and proofs are slight adaptations of the treatment of **ZF** and Boolean-valued models in [5]. We will define an underlying domain $V^{(H)}$ which will be a reflection of the hierarchical system of all sets V. In interpreting our only relational symbols $\in$ and $=$, and constructing $V^{(H)}$, we are guided by the following idea: Note that for any set A, all information about this set is carried by its characteristic function (classically speaking). We define for all sets x:

$$\mathbb{1}_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A. \end{cases}$$

We may thus as well **identify** A with its characteristic function. What hinders us to set our domain to be the class of all such functions (and generalizing them to take values inside a complete Heyting-algebra rather than $B_{0,1}$) is the fact, that the domain of $\mathbb{1}_A$ does itself not consist of such Heyting-valued valued functions. Sticking to our 2-valued case, we therefore define the class $V^{(0,1)}$ to be the class of functions $V^{(2)} \to B_{0,1}$. Actually, this definition is by recursion on ordinals α:

$$\begin{aligned} V_0^{(2)} &= \emptyset, \\ V_\alpha^{(2)} &= \left\{ \mathfrak{a} \middle| \mathfrak{a} \text{ is a function} \wedge \mathrm{ran}(\mathfrak{a}) \subseteq \{0,1\} \wedge \exists \xi < \alpha \colon \mathrm{dom}(\mathfrak{a}) \subseteq V_\xi^{(2)} \right\}, \\ V^{(2)} &= \bigcup_{\alpha \in \mathbb{ON}} V_\alpha^{(2)}. \end{aligned}$$

Allowing the functions to take values in a complete Heyting-algebra, we define the universe $V^{(H)}$ by

$$
\begin{aligned}
V_\alpha^{(H)} &= \left\{ f \middle| f \text{ is a function} \wedge \operatorname{ran}(f) \subseteq H \wedge \exists \xi < \alpha \colon \operatorname{dom}(f) \subseteq V_\xi^{(H)} \right\},\\
V^{(H)} &= \bigcup_{\alpha \in \mathbb{ON}} V_\alpha^{(H)}.
\end{aligned}
$$

How do we interpret the ∈- and =-relation? To incorporate the axiom of extensionality as well as the logical truth $\mathfrak{u} \in \mathfrak{v} \leftrightarrow \exists y \in \mathfrak{v}.\, \mathfrak{u} = y$ we should have

$$
\begin{aligned}
\llbracket \mathfrak{u} = \mathfrak{v} \rrbracket &= \llbracket \forall x \in \mathfrak{u}\; x \in \mathfrak{v} \wedge \forall y \in \mathfrak{v}\; y \in \mathfrak{u} \rrbracket,\\
\llbracket \mathfrak{u} \in \mathfrak{v} \rrbracket &= \llbracket \exists y \in \mathfrak{v}\; \mathfrak{u} = y \rrbracket.
\end{aligned}
$$

Also, it is reasonable to wish, in the case of bounded quantification, to be able to restrict our attention to elements of the given domain of the set of consideration only, i.e.

$$
\begin{aligned}
\llbracket \forall x \in \mathfrak{v}.\, \phi(x) \rrbracket &= \bigwedge_{\mathfrak{x} \in \operatorname{dom}(\mathfrak{v})} [\mathfrak{u}(\mathfrak{x}) \rightarrow \llbracket \phi(\mathfrak{x}) \rrbracket],\\
\llbracket \exists x \in \mathfrak{u}.\, \phi(x) \rrbracket &= \bigcup_{\mathfrak{x} \in \operatorname{dom}(\mathfrak{u})} [\mathfrak{u}(x) \wedge \llbracket \phi(\mathfrak{x}) \rrbracket].
\end{aligned}
$$

We will later see that this is compatible with the interpretation of unbounded quantification. Combining these two observations, we should set for equality and set-membership:

$$
\begin{aligned}
\llbracket \mathfrak{u} = \mathfrak{v} \rrbracket &= \bigwedge_{\mathfrak{x} \in \operatorname{dom}(\mathfrak{u})} [\mathfrak{u}(\mathfrak{x}) \rightarrow \llbracket \mathfrak{x} \in \mathfrak{v} \rrbracket] \wedge \bigwedge_{\mathfrak{x} \in \operatorname{dom}(\mathfrak{v})} [\mathfrak{v}(\mathfrak{x}) \rightarrow \llbracket \mathfrak{x} \in \mathfrak{u} \rrbracket],\\
\llbracket \mathfrak{u} \in \mathfrak{v} \rrbracket &= \bigvee_{\mathfrak{x} \in \operatorname{dom}(\mathfrak{v})} [\mathfrak{v}(\mathfrak{x}) \wedge \llbracket \mathfrak{u} = \mathfrak{x} \rrbracket].
\end{aligned}
$$

For the sake of completeness, we give a full list of all clauses of the Heyting-valued interpretation of the language of set theory.

List of clauses Heyting-valued interpretation of **IZF**

$$
\begin{aligned}
\llbracket \mathfrak{u} = \mathfrak{v} \rrbracket &= \bigwedge_{\mathfrak{x} \in \operatorname{dom}(\mathfrak{u})} (\mathfrak{u}(\mathfrak{x}) \rightarrow \llbracket \mathfrak{x} \in \mathfrak{v} \rrbracket) \wedge \bigwedge_{\mathfrak{x} \in \operatorname{dom}(\mathfrak{v})} (\mathfrak{v}(\mathfrak{x}) \rightarrow \llbracket \mathfrak{x} \in \mathfrak{u} \rrbracket)\\
\llbracket \mathfrak{u} \in \mathfrak{v} \rrbracket &= \bigvee_{\mathfrak{x} \in \operatorname{dom}(\mathfrak{v})} \mathfrak{v}(\mathfrak{x}) \wedge \llbracket \mathfrak{u} = \mathfrak{x} \rrbracket\\
\llbracket \bot \rrbracket &= 0\\
\llbracket \top \rrbracket &= 1\\
\llbracket \phi \wedge \psi \rrbracket &= \llbracket \phi \rrbracket \wedge \llbracket \psi \rrbracket\\
\llbracket \phi \vee \psi \rrbracket &= \llbracket \phi \rrbracket \vee \llbracket \psi \rrbracket\\
\llbracket \phi \rightarrow \psi \rrbracket &= \llbracket \phi \rrbracket \rightarrow \llbracket \psi \rrbracket\\
\llbracket \neg \phi \rrbracket &= \llbracket \phi \rrbracket \rightarrow 0\\
\llbracket \forall x\, \phi(x) \rrbracket &= \bigwedge_{\mathfrak{d} \in V^{(H)}} \llbracket \phi(\mathfrak{d}) \rrbracket
\end{aligned}
$$

$$[\![\exists x\ \phi(x)]\!] = \bigvee_{\mathfrak{d}\in V^{(H)}} [\![\phi(\mathfrak{d})]\!]$$

$$[\![\forall x \in \mathfrak{u}.\, \phi(x)]\!] = \bigwedge_{\mathfrak{x}\in \mathrm{dom}(\mathfrak{u})} (\mathfrak{u}(\mathfrak{x}) \to [\![\phi(\mathfrak{x})]\!])$$

$$[\![\exists x \in \mathfrak{u}.\, \phi(x)]\!] = \bigvee_{\mathfrak{x}\in \mathrm{dom}(\mathfrak{u})} (\mathfrak{u}(\mathfrak{x}) \wedge [\![\phi(\mathfrak{x})]\!])$$

6.3.1 Some useful facts about $V^{(H)}$

In this section we will gather some useful (but rather technical) results about our interpretation of **IZF**. All of the results are slight adaptations of [5]. Some of them will be the justification of the informal discussion (and the resulting definitions) of the last section. For example, it is a-priori not clear that the expressions for unbounded quantification are justified, since we are taking suprema and infima over proper classes. The following results show that this definition is reasonable:

Ⓜ**Lemma 6.24**: For each formula $\phi(x)$, there is a least α such that $\bigvee_{\mathfrak{d}\in V^{(H)}}[\![\phi(\mathfrak{d})]\!] = \bigvee_{\mathfrak{d}\in V_{\alpha}^{(H)}}[\![\phi(\mathfrak{d})]\!]$ and similarly for infima.

Proof: The function $F\colon \mathbb{ON} \to H$ defined by $F(\alpha) = \bigvee_{\mathfrak{d}\in V_{\alpha}^{(H)}}[\![\phi(\mathfrak{d})]\!]$ is increasing. By Lemma 6.10, it must eventually be constant. The case for infima is similar. ■

Ⓜ**Lemma 6.25**: For $\mathfrak{a}, \mathfrak{b}, \mathfrak{x}, \mathfrak{y}, \mathfrak{u}, \mathfrak{v}, \mathfrak{w} \in V^{(H)}$ and formulas ϕ, we have

1. $[\![\mathfrak{b} = \mathfrak{b}]\!] = 1$
2. $[\![\mathfrak{a} \in \mathfrak{b}]\!] \geq \mathfrak{b}(\mathfrak{a})$ for $\mathfrak{a} \in \mathrm{dom}(\mathfrak{b})$
3. $[\![\mathfrak{a} = \mathfrak{b}]\!] = [\![\mathfrak{b} = \mathfrak{a}]\!]$
4. $[\![\mathfrak{x} = \mathfrak{y}]\!] \wedge [\![\mathfrak{y} \in \mathfrak{w}]\!] \leq [\![\mathfrak{x} \in \mathfrak{w}]\!]$
5. $[\![\mathfrak{x} \in \mathfrak{v}]\!] \wedge [\![\mathfrak{v} = \mathfrak{w}]\!] \leq [\![\mathfrak{x} \in \mathfrak{w}]\!]$
6. $[\![\mathfrak{u} = \mathfrak{v}]\!] \wedge [\![\mathfrak{v} = \mathfrak{w}]\!] \leq [\![\mathfrak{u} = \mathfrak{w}]\!]$
7. $[\![\mathfrak{u} = \mathfrak{v}]\!] \wedge [\![\phi(\mathfrak{u})]\!] \leq [\![\phi(\mathfrak{v})]\!]$
8. $[\![\exists y. y = \mathfrak{x} \wedge \phi(y)]\!] = [\![\phi(\mathfrak{x})]\!]$

Proof: 1 and 2 are shown by simultaneous induction: Suppose 2. Holds for $\mathfrak{b} \in V_{\alpha}^{(H)}$, then,

$$[\![\mathfrak{b} = \mathfrak{b}]\!] = \bigwedge_{\mathfrak{a}\in \mathrm{dom}(\mathfrak{b})} \mathfrak{b}(\mathfrak{a}) \to [\![\mathfrak{a} \in \mathfrak{b}]\!] = 1.$$

Now suppose, 1. holds for all $\mathfrak{b} \in V_{\beta}$ with $\beta < \alpha$ and let $\mathfrak{a} \in \mathrm{dom}(\mathfrak{b})$.

$$[\![\mathfrak{a} \in \mathfrak{b}]\!] = \bigvee_{y\in \mathrm{dom}(\mathfrak{b})} (\mathfrak{b}(y) \wedge [\![\mathfrak{a} = y]\!]) \geq \mathfrak{b}(\mathfrak{a}) \wedge [\![\mathfrak{a} = \mathfrak{a}]\!] = \mathfrak{b}(\mathfrak{a}).$$

3 Holds by symmetry of the definition.

The numbers 4-6 are shown by induction: Assume that for all $\mathfrak{u}, \mathfrak{v}, \mathfrak{w} \in V_\alpha^{(H)}$ and all $\beta < \alpha$ and for $\mathfrak{x}, \mathfrak{y}, \mathfrak{z} \in V_\beta^{(H)}$,

$$\begin{array}{ll} (IH_1) & \llbracket \mathfrak{x} = \mathfrak{y} \rrbracket \wedge \llbracket \mathfrak{y} \in \mathfrak{w} \rrbracket \leq \llbracket \mathfrak{x} \in \mathfrak{w} \rrbracket, \\ (IH_2) & \llbracket \mathfrak{x} = \mathfrak{y} \rrbracket \wedge \llbracket \mathfrak{y} = \mathfrak{z} \rrbracket \leq \llbracket \mathfrak{x} = \mathfrak{z} \rrbracket, \\ (IH_3) & \llbracket \mathfrak{x} \in \mathfrak{u} \rrbracket \wedge \llbracket \mathfrak{u} = \mathfrak{v} \rrbracket \leq \llbracket \mathfrak{x} \in \mathfrak{v} \rrbracket. \end{array}$$

Then, we can infer 4,

$$\begin{aligned} \llbracket \mathfrak{x} = \mathfrak{y} \rrbracket \wedge \llbracket \mathfrak{y} \in \mathfrak{w} \rrbracket &= \llbracket \mathfrak{x} = \mathfrak{y} \rrbracket \wedge \bigvee_{\mathfrak{z} \in \mathrm{dom}(\mathfrak{w})} \big(\llbracket \mathfrak{z} = \mathfrak{y} \rrbracket \wedge \mathfrak{w}(\mathfrak{z}) \big) \\ &= \bigvee_{\mathfrak{z} \in \mathrm{dom}(\mathfrak{w})} \big(\llbracket \mathfrak{x} = \mathfrak{y} \rrbracket \wedge \llbracket \mathfrak{y} = \mathfrak{z} \rrbracket \wedge \mathfrak{w}(\mathfrak{z}) \big) \leq^{IH_2} \bigvee_{\mathfrak{z} \in \mathrm{dom}(\mathfrak{w})} \big(\llbracket \mathfrak{x} = \mathfrak{z} \rrbracket \wedge \mathfrak{w}(\mathfrak{z}) \big) = \llbracket \mathfrak{x} \in \mathfrak{w} \rrbracket. \end{aligned}$$

And 5,

$$\begin{aligned} \llbracket \mathfrak{x} \in \mathfrak{v} \rrbracket \wedge \llbracket \mathfrak{v} = \mathfrak{w} \rrbracket = \llbracket \mathfrak{v} = \mathfrak{w} \rrbracket \wedge \bigvee_{\mathfrak{y} \in \mathrm{dom}(\mathfrak{v})} (\mathfrak{v}(\mathfrak{y}) \wedge \llbracket \mathfrak{x} = \mathfrak{y} \rrbracket) &= \bigvee_{\mathfrak{y} \in \mathrm{dom}(\mathfrak{v})} (\llbracket \mathfrak{v} = \mathfrak{w} \rrbracket \wedge \mathfrak{v}(\mathfrak{y}) \wedge \llbracket \mathfrak{x} = \mathfrak{y} \rrbracket) \\ &\leq \bigvee_{\mathfrak{y} \in \mathrm{dom}(\mathfrak{v})} \left(\left(\bigwedge_{\mathfrak{a} \in \mathrm{dom}(\mathfrak{v})} (\mathfrak{v}(\mathfrak{a}) \to \llbracket \mathfrak{a} \in \mathfrak{w} \rrbracket) \right) \wedge \mathfrak{v}(\mathfrak{y}) \wedge \llbracket \mathfrak{x} = \mathfrak{y} \rrbracket \right) \\ &\leq \bigvee_{\mathfrak{y} \in \mathrm{dom}(\mathfrak{v})} \big((\mathfrak{v}(\mathfrak{y}) \to \llbracket \mathfrak{y} \in \mathfrak{w} \rrbracket) \wedge \mathfrak{v}(\mathfrak{y}) \wedge \llbracket \mathfrak{x} = \mathfrak{y} \rrbracket \big) \\ &\leq \bigvee_{\mathfrak{y} \in \mathrm{dom}(\mathfrak{v})} (\llbracket \mathfrak{y} \in \mathfrak{w} \rrbracket \wedge \llbracket \mathfrak{x} = \mathfrak{y} \rrbracket) \leq^{IH_1} \bigvee_{\mathfrak{y} \in \mathrm{dom}(\mathfrak{v})} (\llbracket \mathfrak{y} \in \mathfrak{w} \rrbracket \wedge \llbracket \mathfrak{x} = \mathfrak{y} \rrbracket) = \bigvee_{\mathfrak{y} \in \mathrm{dom}(\mathfrak{v})} (\llbracket \mathfrak{x} \in \mathfrak{w} \rrbracket) \\ &= \llbracket \mathfrak{x} \in \mathfrak{w} \rrbracket. \end{aligned}$$

And finally 6,

$$\llbracket \mathfrak{u} = \mathfrak{v} \rrbracket \wedge \mathfrak{u}(\mathfrak{x}) \wedge \llbracket \mathfrak{v} = \mathfrak{w} \rrbracket \leq \llbracket \mathfrak{u} = \mathfrak{v} \rrbracket \wedge \llbracket \mathfrak{x} \in \mathfrak{u} \rrbracket \wedge \llbracket \mathfrak{v} = \mathfrak{w} \rrbracket \leq^{IH_3} \llbracket \mathfrak{x} \in \mathfrak{v} \rrbracket \wedge \llbracket \mathfrak{v} = \mathfrak{w} \rrbracket \leq^{IH_3} \llbracket \mathfrak{x} \in \mathfrak{w} \rrbracket$$

Hence,

$$\llbracket \mathfrak{u} = \mathfrak{v} \rrbracket \wedge \llbracket \mathfrak{v} = \mathfrak{w} \rrbracket \leq \mathfrak{u}(\mathfrak{x}) \to \llbracket \mathfrak{x} \in \mathfrak{w} \rrbracket$$

and similarly,

$$\llbracket \mathfrak{u} = \mathfrak{v} \rrbracket \wedge \llbracket \mathfrak{v} = \mathfrak{w} \rrbracket \leq \mathfrak{w}(\mathfrak{z}) \to \llbracket \mathfrak{z} \in \mathfrak{v} \rrbracket$$

And altogether,

$$\llbracket \mathfrak{u} = \mathfrak{v} \rrbracket \wedge \llbracket \mathfrak{v} = \mathfrak{w} \rrbracket \leq \bigwedge_{\mathfrak{x} \in \mathrm{dom}(\mathfrak{u})} (\mathfrak{u}(\mathfrak{x}) \to \llbracket \mathfrak{x} \in \mathfrak{w} \rrbracket) \wedge \bigwedge_{\mathfrak{y} \in \mathrm{dom}(\mathfrak{z})} (\mathfrak{w}(\mathfrak{z}) \to \llbracket \mathfrak{z} \in \mathfrak{v} \rrbracket) = \llbracket \mathfrak{u} = \mathfrak{w} \rrbracket.$$

7 is shown by induction on ϕ. The base cases are 4.-6., all other cases except for "→" are easy. Implication is shown like this:

$$\begin{aligned}&[\![\mathfrak{u} = \mathfrak{v}]\!] \wedge [\![\phi(\mathfrak{u}) \to \psi(\mathfrak{u})]\!] \wedge [\![\phi(\mathfrak{v})]\!] \\ &\qquad = [\![\mathfrak{u} = \mathfrak{v}]\!] \wedge [\![\phi(\mathfrak{u}) \to \psi(\mathfrak{u})]\!] \wedge [\![\phi(\mathfrak{v})]\!] \wedge [\![\mathfrak{u} = \mathfrak{v}]\!] \leq^{IH} [\![\mathfrak{u} = \mathfrak{v}]\!] \wedge ([\![\phi(\mathfrak{u})]\!] \to [\![\psi(\mathfrak{u})]\!]) \wedge [\![\phi(\mathfrak{u})]\!] \\ &\qquad = [\![\mathfrak{u} = \mathfrak{v}]\!] \wedge [\![\psi(\mathfrak{u})]\!] \leq^{IH} [\![\psi(\mathfrak{v})]\!],\end{aligned}$$

and hence,

$$[\![\mathfrak{u} = \mathfrak{v}]\!] \wedge [\![\phi(\mathfrak{u}) \to \psi(\mathfrak{u})]\!] \leq [\![\phi(\mathfrak{v})]\!] \to [\![\psi(\mathfrak{v})]\!].$$

Finally, for 8, we use 7:

$$[\![\exists y.\, y = \mathfrak{x} \wedge \phi(y)]\!] = \bigvee_{\mathfrak{d} \in V^{(H)}} [\![\mathfrak{d} = \mathfrak{x} \wedge \phi(\mathfrak{d})]\!] \leq \bigvee_{\mathfrak{d} \in V^{(H)}} [\![\phi(\mathfrak{x})]\!] = [\![\phi(\mathfrak{x})]\!],$$

on the other hand, clearly $[\![\phi(\mathfrak{x})]\!] = [\![\mathfrak{x} = \mathfrak{x}]\!] \wedge [\![\phi(\mathfrak{x})]\!]$ is bounded by the above supremum. ■

Ⓜ**Lemma 6.26**: Let $f\colon V^{(H)} \to H$ be a (set) function. Then $f \in V^{(H)}$.

Proof: For each $x \in \mathrm{dom}(f)$ there is some (least) α_x such that $x \in V_{\alpha_x}^{(H)}$. By replacement, we can form $\gamma = \bigcup\{\alpha_x\colon x \in \mathrm{dom}(f)\}$. Hence, $\mathrm{dom}(f) \subseteq V_\gamma^{(H)}$. ■

The next lemma shows that our definition of the interpretation of bounded quantification is compatible with the unbounded case:

Ⓜ**Lemma 6.27**:

1. $[\![\exists x \in \mathfrak{u}.\, \phi(x)]\!] = [\![\exists x.\, x \in \mathfrak{u} \wedge \phi(x)]\!]$
2. $[\![\forall x \in \mathfrak{u}.\, \phi(x)]\!] = [\![\forall x.\, x \in \mathfrak{u} \to \phi(x)]\!]$

Proof: For 1., we compute

$$\begin{aligned}[\![\exists x.\, x \in \mathfrak{u} \wedge \phi(x)]\!] &= \bigvee_{\mathfrak{d} \in V^{(H)}} [\![\mathfrak{d} \in \mathfrak{u} \wedge \phi(\mathfrak{d})]\!] = \bigvee_{\mathfrak{d} \in V^{(H)}} ([\![\mathfrak{d} \in \mathfrak{u}]\!] \wedge [\![\phi(\mathfrak{d})]\!]) \\ &= \bigvee_{\mathfrak{d} \in V^{(H)}} \left(\left(\bigvee_{\mathfrak{x} \in \mathrm{dom}(\mathfrak{u})} \mathfrak{u}(\mathfrak{x}) \wedge [\![\mathfrak{d} = \mathfrak{x}]\!]\right) \wedge [\![\phi(\mathfrak{d})]\!]\right) = \bigvee_{\mathfrak{d} \in V^{(H)}} \left(\bigvee_{\mathfrak{x} \in \mathrm{dom}(\mathfrak{u})} \mathfrak{u}(\mathfrak{x}) \wedge [\![\mathfrak{d} = \mathfrak{x}]\!] \wedge [\![\phi(\mathfrak{d})]\!]\right) \\ &= \bigvee_{\mathfrak{x} \in \mathrm{dom}(\mathfrak{u})} \left(\bigvee_{\mathfrak{d} \in V^{(H)}} \mathfrak{u}(\mathfrak{x}) \wedge [\![\mathfrak{d} = \mathfrak{x}]\!] \wedge [\![\phi(\mathfrak{d})]\!]\right) = \bigvee_{\mathfrak{x} \in \mathrm{dom}(\mathfrak{u})} \mathfrak{u}(\mathfrak{x}) \wedge \left(\bigvee_{\mathfrak{d} \in V^{(H)}} [\![\mathfrak{d} = \mathfrak{x} \wedge \phi(\mathfrak{d})]\!]\right) \\ &= \bigvee_{\mathfrak{x} \in \mathrm{dom}(\mathfrak{u})} \mathfrak{u}(\mathfrak{x}) \wedge [\![\exists y.\, y = \mathfrak{x} \wedge \phi(y)]\!] = \bigvee_{\mathfrak{x} \in \mathrm{dom}(\mathfrak{u})} \mathfrak{u}(\mathfrak{x}) \wedge [\![\phi(\mathfrak{x})]\!] = [\![\exists x \in \mathfrak{u}.\, \phi(x)]\!],\end{aligned}$$

where the last step is by Lemma 6.25. For 2., we apply this lemma again together with Lemma 6.6:

$$\llbracket \forall x \in \mathfrak{u}\, \phi(x) \rrbracket = \bigwedge_{\mathfrak{x}\in \mathrm{dom}(\mathfrak{u})} (\mathfrak{u}(\mathfrak{x}) \to \llbracket \phi(\mathfrak{x}) \rrbracket) \geq \bigwedge_{\mathfrak{x}\in \mathrm{dom}(\mathfrak{u})} (\llbracket \mathfrak{x} \in \mathfrak{u} \rrbracket \to \llbracket \phi(\mathfrak{x}) \rrbracket) \geq \bigwedge_{\mathfrak{x}\in V^{(H)}} (\llbracket \mathfrak{x} \in \mathfrak{u} \rrbracket \to \llbracket \phi(\mathfrak{x}) \rrbracket)$$
$$= \llbracket \forall x. x \in \mathfrak{u} \to \phi(x) \rrbracket.$$

On the other hand, for any $\mathfrak{y} \in V^{(H)}$,

$$\llbracket \mathfrak{y} \in \mathfrak{u} \rrbracket \wedge \bigwedge_{\mathfrak{x}\in \mathrm{dom}(\mathfrak{u})} (\mathfrak{u}(\mathfrak{x}) \to \llbracket \phi(\mathfrak{x}) \rrbracket) = \left(\bigvee_{\mathfrak{z}\in \mathrm{dom}(\mathfrak{u})} \mathfrak{u}(\mathfrak{z}) \wedge \llbracket \mathfrak{y} = \mathfrak{z} \rrbracket \right) \wedge \left(\bigwedge_{\mathfrak{x}\in \mathrm{dom}(\mathfrak{u})} (\mathfrak{u}(\mathfrak{x}) \to \llbracket \phi(\mathfrak{x}) \rrbracket) \right)$$
$$= \bigvee_{\mathfrak{z}\in \mathrm{dom}(\mathfrak{u})} \left(\mathfrak{u}(\mathfrak{z}) \wedge \llbracket \mathfrak{y} = \mathfrak{z} \rrbracket \wedge \left(\bigwedge_{\mathfrak{x}\in \mathrm{dom}(\mathfrak{u})} (\mathfrak{u}(\mathfrak{x}) \to \llbracket \phi(\mathfrak{x}) \rrbracket) \right) \right)$$
$$\leq \bigvee_{\mathfrak{z}\in \mathrm{dom}(\mathfrak{u})} (\mathfrak{u}(\mathfrak{z}) \wedge \llbracket \mathfrak{y} = \mathfrak{z} \rrbracket \wedge (\mathfrak{u}(\mathfrak{z}) \to \llbracket \phi(\mathfrak{z}) \rrbracket)) \leq \bigvee_{\mathfrak{z}\in \mathrm{dom}(\mathfrak{u})} (\llbracket \phi(\mathfrak{z}) \rrbracket \wedge \llbracket \mathfrak{y} = \mathfrak{z} \rrbracket) \leq \bigvee_{\mathfrak{z}\in \mathrm{dom}(\mathfrak{u})} \llbracket \phi(\mathfrak{y}) \rrbracket$$
$$= \llbracket \phi(\mathfrak{y}) \rrbracket.$$

And hence,

$$\llbracket \forall x \in \mathfrak{u}\, \phi(x) \rrbracket = \bigwedge_{\mathfrak{x}\in \mathrm{dom}(\mathfrak{u})} (\mathfrak{u}(\mathfrak{x}) \to \llbracket \phi(\mathfrak{x}) \rrbracket) \leq \llbracket \mathfrak{y} \in \mathfrak{u} \rrbracket \to \llbracket \phi(\mathfrak{y}) \rrbracket \leq \bigwedge_{\mathfrak{y}\in V^{(H)}} (\llbracket \mathfrak{y} \in \mathfrak{u} \rrbracket \to \llbracket \phi(\mathfrak{y}) \rrbracket)$$
$$= \llbracket \forall x. x \in \mathfrak{u} \to \phi(x) \rrbracket.$$

■

6.3.2 *Soundness theorem for* **IZF**

Ⓜ**Theorem 6.28**: Every theorem τ of **IZF** are valid, i.e. $\llbracket \tau \rrbracket = 1$.

Proof: We have given examples of validity of propositional axioms as well as examples of proofs of validity of first-order axioms and rules in 6.2. Validity of axioms of equality has been shown in Lemma 6.25. It remains to show validity of set axioms, which we will do in the following. ■

Extensionality

$$\llbracket \forall z. z \in \mathfrak{u} \leftrightarrow z \in \mathfrak{v} \rrbracket = \llbracket (\forall z \in \mathfrak{u}. z \in \mathfrak{v} \wedge \forall z \in \mathfrak{v}. z \in \mathfrak{u}) \rrbracket$$
$$= \left(\bigwedge_{\mathfrak{z}\in \mathrm{dom}(\mathfrak{u})} (\mathfrak{u}(\mathfrak{z}) \to \llbracket \mathfrak{z} \in \mathfrak{v} \rrbracket) \wedge \bigwedge_{\mathfrak{z}\in \mathrm{dom}(\mathfrak{v})} (\mathfrak{v}(\mathfrak{z}) \to \llbracket \mathfrak{z} \in \mathfrak{u} \rrbracket) \right) = \llbracket \mathfrak{u} = \mathfrak{v} \rrbracket,$$

and hence,

$$\llbracket \forall z (z \in \mathfrak{u} \leftrightarrow z \in \mathfrak{v}) \rrbracket \to \llbracket \mathfrak{u} = \mathfrak{v} \rrbracket = 1.$$

Pair

Let $\mathfrak{u} \in V_\alpha^{(H)}, \mathfrak{v} \in V_\beta^{(H)}$ and let $\{\mathfrak{u}, \mathfrak{v}\}^{(H)}$ be defined on $\{\mathfrak{u}, \mathfrak{v}\}$ as $\mathfrak{u} \mapsto 1$ and $\mathfrak{v} \mapsto 1$. By Lemma 6.26, $\{\mathfrak{u}, \mathfrak{v}\}^{(H)} \in V^{(H)}$.

We have

$$[\![\mathfrak{w} \in \{\mathfrak{u},\mathfrak{v}\}^{(H)}]\!] = \bigvee_{\mathfrak{x}\in \mathrm{dom}(\{\mathfrak{u},\mathfrak{v}\}^{(H)})} \{\mathfrak{u},\mathfrak{v}\}^{(H)}(\mathfrak{x}) \wedge [\![\mathfrak{w}=\mathfrak{x}]\!] = (1\wedge[\![\mathfrak{w}=\mathfrak{u}]\!]) \vee (1\wedge [\![\mathfrak{w}=\mathfrak{v}]\!]),$$

Hence $[\![\mathfrak{w} \in \{\mathfrak{u},\mathfrak{v}\}^{(H)} \leftrightarrow ([\![\mathfrak{w}=\mathfrak{u}]\!] \vee [\![\mathfrak{w}=\mathfrak{v}]\!])]\!] = 1$.

Union

Let $\mathfrak{A} \in V_\alpha^{(H)}$. We give $\mathrm{Un}(\mathfrak{A})$ as $\mathrm{dom}(\mathrm{Un}(\mathfrak{A})) = \bigcup\{\mathrm{dom}(\mathfrak{w}) : \mathfrak{w} \in \mathrm{dom}(\mathfrak{A})\}$ and $\mathrm{Un}(\mathfrak{A})(\mathfrak{x}) = [\![\exists w \in \mathfrak{A}\ (\mathfrak{x} \in w)]\!]$. This is a well-defined element of $V^{(H)}$ by Lemma 6.26. Also, we have

$$\begin{aligned}[\![\forall x \in \mathrm{Un}(\mathfrak{A})\ \exists w \in A.\, x \in w]\!] &= \bigwedge_{\mathfrak{x}\in\mathrm{dom}(\mathrm{Un}(\mathfrak{A}))} (\mathrm{Un}(\mathfrak{A})(\mathfrak{x}) \to [\![\exists w \in \mathfrak{A}.\, \mathfrak{x} \in w]\!]) \\ &= \bigwedge_{\mathfrak{x}\in\mathrm{dom}(\mathrm{Un}(\mathfrak{A}))} ([\![\exists w \in \mathfrak{A}.\, \mathfrak{x} \in w]\!] \to [\![\exists w \in \mathfrak{A}.\, \mathfrak{x} \in w]\!]) = 1.\end{aligned}$$

We have that for $\mathfrak{x} \in \mathrm{dom}(\mathfrak{w})$, $\mathfrak{w}(\mathfrak{x}) \leq [\![\mathfrak{x} \in \mathfrak{w}]\!]$ and hence

$$\mathfrak{A}(\mathfrak{w}) \wedge \mathfrak{w}(\mathfrak{x}) \leq \mathfrak{A}(\mathfrak{w}) \wedge [\![\mathfrak{x} \in \mathfrak{w}]\!] \leq \bigvee_{\mathfrak{w}\in\mathrm{dom}(\mathfrak{A})} (\mathfrak{A}(\mathfrak{w}) \wedge [\![\mathfrak{x} \in \mathfrak{w}]\!]) = [\![\exists w \in \mathfrak{A}.\, \mathfrak{x} \in w]\!] = \mathrm{Un}(\mathfrak{A})(\mathfrak{x}).$$

This shows

$$\begin{aligned}[\![\exists w \in \mathfrak{A}.\, \mathfrak{x} \in w]\!] &= \bigvee_{\mathfrak{w}\in\mathrm{dom}(\mathfrak{A})} \mathfrak{A}(\mathfrak{w}) \wedge [\![\mathfrak{x} \in \mathfrak{w}]\!] = \bigvee_{\mathfrak{w}\in\mathrm{dom}(\mathfrak{A})} \mathfrak{A}(\mathfrak{w}) \wedge \left(\bigvee_{\mathfrak{y}\in\mathrm{dom}(\mathfrak{w})} \mathfrak{w}(\mathfrak{x}) \wedge [\![\mathfrak{y} = \mathfrak{x}]\!]\right) \\ &= \bigvee_{\mathfrak{w}\in\mathrm{dom}(\mathfrak{A})} \left(\bigvee_{\mathfrak{y}\in\mathrm{dom}(\mathfrak{w})} \mathfrak{A}(\mathfrak{w}) \wedge \mathfrak{w}(\mathfrak{x}) \wedge [\![\mathfrak{y} = \mathfrak{x}]\!]\right) \leq \bigvee_{\mathfrak{w}\in\mathrm{dom}(\mathfrak{A})} \left(\bigvee_{\mathfrak{y}\in\mathrm{dom}(\mathfrak{w})} \mathrm{Un}(\mathfrak{A})(\mathfrak{x}) \wedge [\![\mathfrak{y} = \mathfrak{x}]\!]\right) \\ &= \bigvee_{\mathfrak{y}\in\bigcup\{\mathrm{dom}(\mathfrak{w}):\mathfrak{w}\in\mathrm{dom}(\mathfrak{A})\}} (\mathrm{Un}(\mathfrak{A})(\mathfrak{x}) \wedge [\![\mathfrak{y} = \mathfrak{x}]\!]) = [\![\mathfrak{x} \in \mathrm{Un}(\mathfrak{A})]\!],\end{aligned}$$

and hence $[\![\exists w \in \mathfrak{A}\ (\mathfrak{x} \in w) \to \mathfrak{x} \in \mathrm{Un}(\mathfrak{A})]\!] = 1$.

Empty set

For a set x, define by recursion $\hat{x} = \{\langle \hat{y}, 1\rangle : y \in x\}$. It is shown by induction on the rank of x that this is an element of $V^{(H)}$. We show that $\widehat{\emptyset}$ is a suitable witness for the empty set axiom: Since $\mathrm{dom}(\widehat{\emptyset}) = \emptyset$ and $\bigvee \emptyset = 0$, we have

$$[\![\forall y\, \neg\big(y \in \widehat{\emptyset}\big)]\!] = \bigwedge_{y \in V^{(H)}} \neg[\![y \in \widehat{\emptyset}]\!] = \bigwedge_{y \in V^{(H)}} \big(\neg[\![y \in \widehat{\emptyset}]\!] \to 0\big) = \bigwedge_{y \in V^{(H)}} \left(\neg \bigvee_{\mathfrak{x} \in \mathrm{dom}(\widehat{\emptyset})} \widehat{\emptyset}(\mathfrak{x}) \wedge [\![y = \mathfrak{x}]\!]\right)$$

$$= \bigwedge_{y \in V^{(H)}} \neg 0 = 1.$$

Infinity

We show that $\widehat{\omega}$ does the job:

$$V^{(H)} \vDash (\forall n \in \widehat{\omega}.\, s(n) \in \widehat{\omega}) \wedge \big(\forall n \in \omega.\, n = \emptyset \vee \exists m \in \widehat{\omega}.\, n = s(m)\big).$$

Ⓜ**Lemma 6.29**: Define for $\mathfrak{u} \in V^{(H)}$, $\mathfrak{s}_{\mathfrak{u}} = \mathfrak{u} \cup \{\langle \mathfrak{u}, 1\rangle\}$. Then $[\![\mathfrak{s}_{\mathfrak{u}} = s(\mathfrak{u})]\!] = 1$, or in more detail,

$$[\![\forall x.\, x \in \mathfrak{s}_{\mathfrak{u}} \leftrightarrow (x \in \mathfrak{u} \vee x = \mathfrak{u})]\!] = 1.$$

Proof: Indeed,

$$[\![\mathfrak{x} \in \mathfrak{s}_{\mathfrak{u}}]\!] = \bigvee_{\mathfrak{y} \in \mathrm{dom}(\mathfrak{s}_{\mathfrak{u}})} (\mathfrak{s}_{\mathfrak{u}}(\mathfrak{y}) \wedge [\![\mathfrak{x} = \mathfrak{y}]\!]) = (\mathfrak{s}_{\mathfrak{u}}(\mathfrak{u}) \wedge [\![\mathfrak{x} = \mathfrak{u}]\!]) \vee \bigvee_{\mathfrak{y} \in \mathrm{dom}(\mathfrak{u})} (\mathfrak{u}(\mathfrak{y}) \wedge [\![\mathfrak{x} = \mathfrak{y}]\!]) = [\![\mathfrak{x} = \mathfrak{u}]\!] \vee [\![\mathfrak{x} \in \mathfrak{u}]\!].$$

■

Ⓜ**Lemma 6.30**: For all $n \in \omega$, $\left[\!\!\left[\widehat{n+1} = s(\hat{n})\right]\!\!\right] = 1$. ■

Proof: It is clear that $\left[\!\!\left[\widehat{n+1} = \mathfrak{s}_{\hat{n}}\right]\!\!\right] = 1$, so the result follows by transitivity. ■

We can now verify the infinity axiom:

$$[\![\forall x \in \widehat{\omega}.\, s(x) \in \widehat{\omega}]\!] = \bigwedge_{n \in \omega} (\widehat{\omega}(\hat{n}) \to [\![s(\hat{n}) \in \widehat{\omega}]\!]) = \bigwedge_{n \in \omega} [\![s(\hat{n}) \in \widehat{\omega}]\!]$$

For any $n \in \omega$,

$$[\![s(\hat{n}) \in \widehat{\omega}]\!] = \bigvee_{m \in \omega} [\![s(\hat{n}) = \hat{m}]\!] \geq \left[\!\!\left[s(\hat{n}) = \widehat{n+1}\right]\!\!\right] = 1.$$

On the other hand,

$$[\![\forall n \in \omega.\, n = \emptyset \vee \exists m \in \omega.\, n = s(m)]\!] = \bigwedge_{n \in \omega} ([\![n = \emptyset]\!] \vee [\![\exists m \in \omega.\, n = s(m)]\!]).$$

Now for $n = 0$, we already know that $\left[\!\!\left[\hat{0} = \emptyset\right]\!\!\right] = 1$. For $n = m + 1$,

$$\left[\!\!\left[\exists m \in \omega\, \big(\hat{n} = s(m)\big)\right]\!\!\right] = \bigvee_{m \in \omega} [\![\hat{n} = s(\hat{m})]\!] \geq \left[\!\!\left[\widehat{m+1} = s(\hat{m})\right]\!\!\right] = 1.$$

Separation

For $\mathfrak{a} \in V^{(H)}$ and a formula $\phi(x)$, set $\mathrm{Sep}_{\phi(x)}(\mathfrak{a}) = \{\langle \mathfrak{x}, \llbracket \phi(\mathfrak{x}) \rrbracket \wedge \mathfrak{a}(\mathfrak{x}) \rangle : \mathfrak{x} \in \mathrm{dom}(\mathfrak{a})\}$. Then for $\mathfrak{y} \in V^{(H)}$,

$$\llbracket \mathfrak{y} \in \mathrm{Sep}_{\phi(x)}(\mathfrak{a}) \rrbracket = \bigvee_{\mathfrak{y} \in \mathrm{dom}\left(\mathrm{Sep}_{\phi(x)}(\mathfrak{a})\right)} \left(\mathrm{Sep}_{\phi(x)}(\mathfrak{a})(\mathfrak{x}) \wedge \llbracket \mathfrak{y} = \mathfrak{x} \rrbracket\right) = \bigvee_{\mathfrak{y} \in \mathrm{dom}(\mathfrak{a})} \left(\mathfrak{a}(\mathfrak{x}) \wedge \llbracket \phi(\mathfrak{x}) \rrbracket \wedge \llbracket \mathfrak{y} = \mathfrak{x} \rrbracket\right)$$
$$= \llbracket \exists x \in \mathfrak{a}\, (\phi(x) \wedge \mathfrak{y} = x) \rrbracket = \left\llbracket \exists x \left(\left(x \in \mathfrak{a} \wedge \phi(\mathfrak{x})\right) \wedge \mathfrak{y} = x\right) \right\rrbracket = \llbracket \mathfrak{y} \in \mathfrak{a} \wedge \phi(\mathfrak{y}) \rrbracket.$$

Powerset

Let $\mathfrak{a} \in V^{(H)}$ and define $\mathrm{Pow}(\mathfrak{a})$ by $\mathrm{dom}\big(\mathrm{Pow}(\mathfrak{a})\big) = H^{\mathrm{dom}(\mathfrak{a})}$ and $\mathrm{Pow}(\mathfrak{a})(\mathfrak{x}) = \llbracket \mathfrak{x} \subseteq \mathfrak{a} \rrbracket$ for $\mathfrak{x} \in \mathrm{dom}\big(\mathrm{Pow}(\mathfrak{a})\big)$. By Lemma 6.26, $\mathrm{Pow}(\mathfrak{a}) \in V^{(H)}$.

$$\llbracket \forall x \in \mathrm{Pow}(\mathfrak{a}).\, x \subseteq \mathfrak{a} \rrbracket = \bigwedge_{\mathfrak{x} \in \mathrm{dom}(\mathrm{Pow}(\mathfrak{a}))} \left(\mathrm{Pow}(\mathfrak{a})(\mathfrak{x}) \to \llbracket \mathfrak{x} \subseteq \mathfrak{a} \rrbracket\right) = \bigwedge_{\mathfrak{x} \in \mathrm{dom}(\mathrm{Pow}(\mathfrak{a}))} \left(\llbracket \mathfrak{x} \subseteq \mathfrak{a} \rrbracket \to \llbracket \mathfrak{x} \subseteq \mathfrak{a} \rrbracket\right) = 1.$$

On the other hand, define for $\mathfrak{x} \in V^{(H)}$, $\mathfrak{x}' = \{\langle \mathfrak{y}, \llbracket \mathfrak{y} \in \mathfrak{x} \rrbracket \rangle : \mathfrak{y} \in \mathrm{dom}(\mathfrak{a})\}$. We need the following two claims:

Claim 1: $\llbracket \mathfrak{x} \subseteq \mathfrak{a} \to \mathfrak{x} = \mathfrak{x}' \rrbracket = 1$.

Proof: For any $\mathfrak{y} \in V^{(H)}$, we have

$$\llbracket \mathfrak{y} \in \mathfrak{x}' \rrbracket = \bigvee_{\mathfrak{z} \in \mathrm{dom}(\mathfrak{x}')} \left(\mathfrak{x}'(\mathfrak{z}) \wedge \llbracket \mathfrak{y} = \mathfrak{z} \rrbracket\right) = \bigvee_{\mathfrak{z} \in \mathrm{dom}(\mathfrak{a})} \left(\llbracket \mathfrak{z} \in \mathfrak{x} \rrbracket \wedge \llbracket \mathfrak{y} = \mathfrak{z} \rrbracket\right) \leq \llbracket \mathfrak{y} \in \mathfrak{x} \rrbracket,$$

hence $\llbracket \mathfrak{x}' \subseteq \mathfrak{x} \rrbracket = \llbracket \forall y\, (y \in \mathfrak{x}' \to y \in \mathfrak{x}) \rrbracket = 1$. Furthermore,

$$\llbracket \mathfrak{y} \in \mathfrak{a} \wedge \mathfrak{y} \in \mathfrak{x} \rrbracket = \llbracket \mathfrak{y} \in \mathfrak{x} \rrbracket \wedge \bigvee_{\mathfrak{z} \in \mathrm{dom}(\mathfrak{a})} \left(\mathfrak{a}(\mathfrak{z}) \wedge \llbracket \mathfrak{y} = \mathfrak{z} \rrbracket\right) = \bigvee_{\mathfrak{z} \in \mathrm{dom}(\mathfrak{a})} \left(\mathfrak{a}(\mathfrak{z}) \wedge \llbracket \mathfrak{y} = \mathfrak{z} \rrbracket \wedge \llbracket \mathfrak{y} \in \mathfrak{x} \rrbracket\right)$$
$$\leq \bigvee_{\mathfrak{z} \in \mathrm{dom}(\mathfrak{a})} \left(\llbracket \mathfrak{y} = \mathfrak{z} \rrbracket \wedge \llbracket \mathfrak{y} = \mathfrak{z} \rrbracket \wedge \llbracket \mathfrak{y} \in \mathfrak{x} \rrbracket\right) \leq \bigvee_{\mathfrak{z} \in \mathrm{dom}(\mathfrak{a})} \left(\llbracket \mathfrak{y} = \mathfrak{z} \rrbracket \wedge \llbracket \mathfrak{z} \in \mathfrak{x} \rrbracket\right)$$
$$= \bigvee_{\mathfrak{z} \in \mathrm{dom}(\mathfrak{x}')} \left(\llbracket \mathfrak{y} = \mathfrak{z} \rrbracket \wedge \mathfrak{x}'(\mathfrak{z})\right) = \llbracket \mathfrak{y} \in \mathfrak{x}' \rrbracket.$$

and therefore, $\llbracket \mathfrak{a} \cap \mathfrak{x} \subseteq \mathfrak{x}' \rrbracket = 1$. Putting these things together,

$$\llbracket \mathfrak{x} \subseteq \mathfrak{a} \rrbracket = 1 \wedge \llbracket \mathfrak{x} \subseteq \mathfrak{a} \rrbracket \wedge 1 = \llbracket \mathfrak{x}' \subseteq \mathfrak{x} \rrbracket \wedge \llbracket \mathfrak{x} \subseteq \mathfrak{a} \rrbracket \wedge \llbracket \mathfrak{a} \cap \mathfrak{x} \subseteq \mathfrak{x}' \rrbracket = \llbracket \mathfrak{x}' \subseteq \mathfrak{x} \rrbracket \wedge \llbracket \mathfrak{x} \subseteq \mathfrak{a} \wedge \mathfrak{a} \cap \mathfrak{x} \subseteq \mathfrak{x}' \rrbracket$$
$$\leq \llbracket \mathfrak{x}' \subseteq \mathfrak{x} \rrbracket \wedge \llbracket \mathfrak{x} \subseteq \mathfrak{x}' \rrbracket \leq \llbracket \mathfrak{x} = \mathfrak{x}' \rrbracket,$$

which shows the claim. ■

Claim 2: $\llbracket \mathfrak{x} \subseteq \mathfrak{a} \to \mathfrak{x}' \in \mathrm{Pow}(\mathfrak{a}) \rrbracket = 1$.

Proof:

$$\begin{aligned}[\![\mathfrak{x} \subseteq \mathfrak{a}]\!] = [\![\forall y. y \in \mathfrak{x} \to y \in \mathfrak{a}]\!] = \bigwedge_{\mathfrak{y}\in V^{(H)}} ([\![\mathfrak{y}\in\mathfrak{x}]\!] \to [\![\mathfrak{y}\in\mathfrak{a}]\!]) \le \bigwedge_{\mathfrak{y}\in\mathrm{dom}(\mathfrak{x}')} (\mathfrak{x}'(\mathfrak{y}) \to [\![\mathfrak{y}\in\mathfrak{a}]\!]) = [\![\forall y \in \mathfrak{x}'\ (y\in\mathfrak{a})]\!] \\ = [\![\mathfrak{x}'\subseteq\mathfrak{a}]\!] = \mathrm{Pow}(\mathfrak{a})(\mathfrak{x}') \le [\![\mathfrak{x}'\in\mathrm{Pow}(\mathfrak{a})]\!].\end{aligned}$$

■

Using both claims, we can finally validate the other direction of the powerset axiom:

$$\begin{aligned}[\![\forall x. x \subseteq \mathfrak{a} \to x \in \mathrm{Pow}(\mathfrak{a})]\!] &= \bigwedge_{\mathfrak{x}\in V^{(H)}} ([\![\mathfrak{x}\subseteq\mathfrak{a} \to \mathfrak{x}\in\mathrm{Pow}(\mathfrak{a})]\!]) \\ &\ge \bigwedge_{\mathfrak{x}\in V^{(H)}} ([\![\mathfrak{x}\subseteq\mathfrak{a}\to\mathfrak{x}=\mathfrak{x}']\!] \wedge [\![\mathfrak{x}\subseteq\mathfrak{a}\to\mathfrak{x}'\in\mathrm{Pow}(\mathfrak{a})]\!]) = 1.\end{aligned}$$

Collection schema

Let $\mathfrak{u}\in V^{(H)}$ and $\phi(x,y)$ be any formula. For each $\mathfrak{x}\in V^{(H)}$, there is a least $\alpha_{\mathfrak{x}}$ such that

$$\bigvee_{\mathfrak{y}\in V^{(H)}} [\![\phi(\mathfrak{x},\mathfrak{y})]\!] = \bigvee_{\mathfrak{y}\in V^{(H)}_{\alpha_{\mathfrak{x}}}} [\![\phi(\mathfrak{x},\mathfrak{y})]\!].$$

By collection in V, we can set $\alpha = \bigcup\{\alpha_{\mathfrak{x}} : \mathfrak{x}\in\mathrm{dom}(\mathfrak{u})\}$ and $\mathrm{Col}_{\phi(x,y)}(\mathfrak{u}) = \left\{\langle\mathfrak{y},1\rangle : \mathfrak{y}\in V^{(H)}_{\alpha}\right\}$. As in the validation of the previous axioms, this set is a well-defined member of $V^{(H)}$ by Lemma 6.26. Furthermore,

$$\begin{aligned}[\![\forall x\in\mathfrak{u}. \exists y. \phi(x,y)]\!] &= \bigwedge_{\mathfrak{x}\in\mathrm{dom}(\mathfrak{u})} (\mathfrak{u}(\mathfrak{x}) \to [\![\exists y. \phi(\mathfrak{x},y)]\!]) = \bigwedge_{\mathfrak{x}\in\mathrm{dom}(\mathfrak{u})} \left(\mathfrak{u}(\mathfrak{x}) \to \bigvee_{\mathfrak{y}\in V^{(H)}} [\![\phi(\mathfrak{x},\mathfrak{y})]\!]\right) \\ &= \bigwedge_{\mathfrak{x}\in\mathrm{dom}(\mathfrak{u})} \left(\mathfrak{u}(\mathfrak{x}) \to \bigvee_{\mathfrak{y}\in V^{(H)}_{\alpha_{\mathfrak{x}}}} [\![\phi(\mathfrak{x},\mathfrak{y})]\!]\right) \le \bigwedge_{\mathfrak{x}\in\mathrm{dom}(\mathfrak{u})} \left(\mathfrak{u}(\mathfrak{x}) \to \bigvee_{\mathfrak{y}\in V^{(H)}_{\alpha}} [\![\phi(\mathfrak{x},\mathfrak{y})]\!]\right) \\ &= \bigwedge_{\mathfrak{x}\in\mathrm{dom}(\mathfrak{u})} \left(\mathfrak{u}(\mathfrak{x}) \to \bigvee_{\mathfrak{y}\in V^{(H)}_{\alpha}} (\mathrm{Col}_{\phi(x,y)}(\mathfrak{u})(\mathfrak{y}) \wedge [\![\phi(\mathfrak{x},\mathfrak{y})]\!])\right) \\ &= \bigwedge_{\mathfrak{x}\in\mathrm{dom}(\mathfrak{u})} \left(\mathfrak{u}(\mathfrak{x}) \to \left[\!\!\left[\exists y\in\mathrm{Col}_{\phi(x,y)}(\mathfrak{u}). \phi(\mathfrak{x},y)\right]\!\!\right]\right) = \left[\!\!\left[\forall x\in\mathfrak{u}. \exists y\in\mathrm{Col}_{\phi(x,y)}(\mathfrak{u}). \phi(x,y)\right]\!\!\right].\end{aligned}$$

The validity of the collection schema follows.

Set induction

Obviously, this is shown by induction: Let $\mathfrak{a}\in V^{(H)}$ and suppose that for all $\beta<\alpha$ and $\mathfrak{y}\in V^{(H)}_{\beta}$,

$$\left[\!\!\left[\forall x\left(\forall y\in x\, \phi(y) \to \phi(x)\right)\right]\!\!\right] \le [\![\phi(\mathfrak{y})]\!].$$

But then

$$\big\llbracket \forall x\, \big(\forall y \in x\, \phi(y) \to \phi(x)\big)\big\rrbracket \leq \bigwedge_{\mathfrak{y}\in\mathrm{dom}(\mathfrak{a})} \llbracket \phi(\mathfrak{y})\rrbracket \leq \bigwedge_{\mathfrak{y}\in\mathrm{dom}(\mathfrak{a})} (\mathfrak{a}(\mathfrak{y}) \to \llbracket \phi(\mathfrak{y})\rrbracket) = \llbracket \forall y \in \mathfrak{a}.\, \phi(y)\rrbracket,$$

hence,

$$\begin{aligned}\llbracket \forall x\, (\forall y \in x\, \phi(y) \to \phi(x)\rrbracket &\leq \llbracket \forall x\, (\forall y \in x\, \phi(y) \to \phi(x)\rrbracket \wedge \llbracket \forall y \in \mathfrak{a}\, \phi(y)\rrbracket \\ &= \bigwedge_{\mathfrak{x}\in V^{(H)}} \llbracket \forall y \in \mathfrak{x}\, \phi(y) \to \phi(\mathfrak{x})\rrbracket \wedge \llbracket \forall y \in \mathfrak{a}\, \phi(y)\rrbracket \leq \llbracket \forall y \in \mathfrak{a}\, \phi(y) \to \phi(\mathfrak{a})\rrbracket \wedge \llbracket \forall y \in \mathfrak{a}\, \phi(y)\rrbracket \\ &\leq \llbracket \phi(\mathfrak{a})\rrbracket.\end{aligned}$$

This shows

$$\llbracket \forall x\, (\forall y \in x\, \phi(y) \to \phi(x)\rrbracket \leq \bigwedge_{\mathfrak{a}\in V^{(H)}} \llbracket \phi(\mathfrak{a})\rrbracket = \llbracket \forall x\, \phi(x)\rrbracket.$$

6.3.3 *Internal set of natural numbers*

As we have seen, when validating the axiom of infinity, $\widehat{\omega}$ plays the role of the set of natural numbers in the model $V^{(H)}$ (we also say that internally, $\widehat{\omega}$ **is** the set of natural numbers). We will need one simple lemma saying that the order on the natural numbers is definite inside the model mirroring the fact that the order is decidable in **IZF**:

Ⓜ**Lemma 6.31**: For all $n, m \in \omega$, we have

1. $\llbracket \hat{n} = \widehat{m}\rrbracket = \begin{cases} 1, & \text{if } n = m, \\ 0, & \text{if } n \neq m. \end{cases}$
2. $\llbracket \hat{n} \in \widehat{m}\rrbracket = \begin{cases} 1, & \text{if } n < m, \\ 0, & \text{else.} \end{cases}$

Proof: 1. and 2. Are shown by simultaneous induction: Suppose that 1. holds for all $k < n$, then

$$\llbracket \widehat{m} \in \hat{n}\rrbracket = \bigvee_{\mathfrak{x}\in\mathrm{dom}(\hat{n})} \llbracket \hat{n}(\mathfrak{x}) \wedge \llbracket \widehat{m} = \mathfrak{x}\rrbracket\rrbracket = \bigvee_{k\in n} \llbracket \widehat{m} = \hat{k}\rrbracket = \begin{cases} 1, & \text{if } m \in n, \\ 0, & \text{if } m \notin n. \end{cases}$$

On the other hand, suppose, 2. holds for all $k < m$ and $l < n$. Then

$$\llbracket \widehat{m} = \hat{n}\rrbracket = \bigwedge_{\mathfrak{x}\in\mathrm{dom}(\hat{n})} \llbracket \hat{n}(\mathfrak{x}) \to \llbracket \mathfrak{x} \in \widehat{m}\rrbracket\rrbracket \wedge \bigwedge_{\mathfrak{x}\in\mathrm{dom}(\widehat{m})} \llbracket \widehat{m}(\mathfrak{x}) \to \llbracket \mathfrak{x} \in \hat{n}\rrbracket\rrbracket = \bigwedge_{k\in n} \llbracket \hat{k} \in \widehat{m}\rrbracket \wedge \bigwedge_{l\in m} \llbracket \hat{l} \in \hat{n}\rrbracket = \begin{cases} 1, & \text{if } m = n, \\ 0, & \text{if } m \neq n. \end{cases}$$

■

6.3.4 *Internal Cross product*

When validating the axiom of pair, we have verified that the internal pair of $\mathfrak{u}, \mathfrak{v} \in V^{(H)}$ is given by $\{\mathfrak{u}, \mathfrak{v}\}^{(H)}$, where $\mathrm{dom}\big(\{\mathfrak{u}, \mathfrak{v}\}^{(H)}\big) = \{\mathfrak{u}, \mathfrak{v}\}$ and $\{\mathfrak{u}, \mathfrak{v}\}^{(H)}(\mathfrak{u}) = \{\mathfrak{u}, \mathfrak{v}\}^{(H)}(\mathfrak{v}) = 1$. It is then clear that the internal ordered pair of $\mathfrak{u}, \mathfrak{v}$ is given as $(\mathfrak{u}, \mathfrak{v})^{(H)} = \big\{\{\mathfrak{u}, \mathfrak{u}\}^{(H)}, \{\mathfrak{u}, \mathfrak{v}\}^{(H)}\big\}^{(H)}$. In this section, we will show that the cross product of $\mathfrak{a}, \mathfrak{b} \in V^{(H)}$ can be given as

$$\mathrm{cp}(\mathfrak{a},\mathfrak{b}) = \left\{\langle(\mathfrak{u},\mathfrak{v})^{(H)}, \mathfrak{a}(\mathfrak{u}) \wedge \mathfrak{b}(\mathfrak{v})\rangle : \mathfrak{u} \in \mathrm{dom}(\mathfrak{a}), \mathfrak{v} \in \mathrm{dom}(\mathfrak{b})\right\},$$

hence avoiding taking internal collection and separation.

Ⓜ**Lemma 6.32**: For $\mathfrak{a},\mathfrak{b} \in V^{(H)}$, the just defined $\mathrm{cp}(\mathfrak{a},\mathfrak{b})$ gives us the internal cross-product.

Proof: We know that all theorems of **IZF** hold inside $V^{(H)}$, hence is suffices to verify that

$$V^{(H)} \vDash \forall w \left(w \in \mathrm{cp}(\mathfrak{a},\mathfrak{b}) \leftrightarrow \exists x \in \mathfrak{a}\, \exists y \in \mathfrak{b}\; w = (x,y)\right).$$

So, let $\mathfrak{w} \in V^{(H)}$, then

$$[\![\mathfrak{w} \in \mathrm{cp}(\mathfrak{a},\mathfrak{b})]\!] = \bigvee_{\mathfrak{y} \in \mathrm{dom}(\mathrm{cp}(\mathfrak{a},\mathfrak{b}))} [\mathrm{cp}(\mathfrak{a},\mathfrak{b})(\mathfrak{y}) \wedge [\![\mathfrak{y} = w]\!]] = \bigvee_{\substack{\mathfrak{u} \in \mathrm{dom}(\mathfrak{a}) \\ \mathfrak{v} \in \mathrm{dom}(\mathfrak{b})}} \left[\mathfrak{a}(\mathfrak{u}) \wedge \mathfrak{b}(\mathfrak{v}) \wedge \left[\!\left[(\mathfrak{u},\mathfrak{v})^{(H)} = \mathfrak{w}\right]\!\right]\right]$$

$$= \bigvee_{\substack{\mathfrak{u} \in \mathrm{dom}(\mathfrak{a}) \\ \mathfrak{v} \in \mathrm{dom}(\mathfrak{b})}} [\mathfrak{a}(\mathfrak{u}) \wedge \mathfrak{b}(\mathfrak{v}) \wedge [\![(u,\mathfrak{v}) = \mathfrak{w}]\!]] = \left[\!\left[\left(\exists x \in \mathfrak{a}\, \exists y \in \mathfrak{b}\; \mathfrak{w} = (x,y)\right)\right]\!\right]$$

As $\mathfrak{w}$ was arbitrary, this shows the lemma. ■

The special case that is important for us is the internal cross product $\mathrm{cp}(\omega,\omega)$. Since $\mathrm{cp}(\omega,\omega) = \widehat{\omega \times \omega}$, the lemma implies that $V^{(H)} \vDash \omega \times \omega = \widehat{\omega \times \omega}$.

6.3.5 Internal set of finite sequences of natural numbers

It is easy to see that internally, $\widehat{\omega^n} = \omega^n$. To define the set of finite sequences of natural numbers, we may mimic the definition of ω^ωin **IZF**, using the axioms of separation, union and collection. We thus immediately get that $V^{(H)} \vDash \omega^{<\omega} = \mathrm{Un}\left(\mathrm{Sep}_{\exists n\; y=\omega^n}\left(\mathrm{Col}_{y=\omega^n}(\omega)\right)\right)$. In this section, we want to prove that $\omega^{<\omega}$ has a simpler internal representative, namely $\widehat{\omega^{<\omega}}$.

Ⓜ**Lemma 6.33**: $V^{(H)} \vDash \{\widehat{\omega^n : n \in \omega}\} = \mathrm{Sep}_{\exists n\; y=\omega^n}\left(\mathrm{Col}_{y=\omega^n}(\omega)\right)$

Proof:

$$\left[\!\left[\widehat{\omega^n} \in \mathrm{Sep}_{\exists n\; y=\omega^n}\left(\mathrm{Col}_{y=\omega^n}(\omega)\right)\right]\!\right] = \bigvee_{\mathfrak{y} \in V_\omega^{(H)}} \left[\mathrm{Sep}_{\exists n\; y=\omega^n}\left(\mathrm{Col}_{y=\omega^n}(\omega)\right)(\mathfrak{y}) \wedge \left[\!\left[\widehat{\omega^n} = \mathfrak{y}\right]\!\right]\right] =$$

$$= \bigvee_{\mathfrak{y} \in V_\omega^{(H)}} \left[[\![\exists n\; \mathfrak{y} = \omega^n]\!] \wedge \left[\!\left[\widehat{\omega^n} = \mathfrak{y}\right]\!\right]\right] \geq \left[\!\left[\exists n\; \widehat{\omega^n} = \omega^n\right]\!\right] \wedge \left[\!\left[\widehat{\omega^n} = \widehat{\omega^n}\right]\!\right] = 1.$$

On the other hand, for $\mathfrak{y} \in V_\omega^H$

$$[\![\exists n\; \mathfrak{y} = \omega^n]\!] \to \left[\!\left[\mathfrak{y} \in \{\widehat{\omega^n : n \in \omega}\}\right]\!\right] = [\![\exists n\; \mathfrak{y} = \omega^n]\!] \to \bigvee_{n \in \omega} \left[\!\left[\widehat{\omega^n} = \mathfrak{y}\right]\!\right] =$$

$$= [\![\exists n\ \mathfrak{y} = \omega^n]\!] \to \left[\!\!\left[\exists n\ \mathfrak{y} = \widehat{\omega^n}\right]\!\!\right] = 1.$$

■

Ⓜ**Proposition 6.34**: $V^{(H)} \vDash \widehat{\omega^{<\omega}} = \mathrm{Un}\left(\mathrm{Sep}_{\exists n\ y=\omega^n}\left(\mathrm{Col}_{y=\omega^n}(\omega)\right)\right)$ and hence $V^{(H)} \vDash \omega^{<\omega} = \widehat{\omega^{<\omega}}$

Proof: In the light of the last lemma, it suffices to show that

$$V^{(H)} \vDash \widehat{\omega^{<\omega}} = \mathrm{Un}\big(\widehat{\{\omega^n : n \in \omega\}}\big).$$

Recall, that the domain of $U = \mathrm{Un}\big(\widehat{\{\omega^n : n \in \omega\}}\big)$ is given by the union of the domains of $L = \widehat{\{\omega^n : n \in \omega\}}$ and $U(\mathfrak{x}) = [\![\exists w \in L.\, \mathfrak{x} \in w]\!]$, for $\mathfrak{x}$ in one of these domains.

So let $\widehat{\big((0,a_0), \dots (n-1, a_{n-1})\big)} \in \mathrm{dom}\big(\widehat{\omega^{<\omega}}\big)$, then

$$\left[\!\!\left[\widehat{\big((0,a_0), \dots (n-1, a_{n-1})\big)} \in U\right]\!\!\right] = \bigvee_{\mathfrak{x} \in \mathrm{dom}(U)} \left([\![\exists w \in L.\, \mathfrak{x} \in w]\!] \wedge \left[\!\!\left[\widehat{\big((0,a_0), \dots (n-1, a_{n-1})\big)} = \mathfrak{x}\right]\!\!\right]\right)$$
$$\geq \left[\!\!\left[\exists w \in L.\, \widehat{\big((0,a_0), \dots (n-1, a_{n-1})\big)} \in w\right]\!\!\right] \geq \left[\!\!\left[\widehat{\big((0,a_0), \dots (n-1, a_{n-1})\big)} \in \widehat{\omega^n}\right]\!\!\right] = 1.$$

On the other hand, let $\mathfrak{x} \in \mathrm{dom}(U)$, then

$$U(x) \to \left[\!\!\left[\mathfrak{x} \in \widehat{\omega^{<\omega}}\right]\!\!\right] = [\![\exists w \in L.\, \mathfrak{x} \in w]\!] \to \bigvee_{\substack{n \in \omega \\ a_0, \dots, a_{n-1} \in \omega}} \left[\!\left[\mathfrak{x} = \big((0,a_0), \dots (n-1, a_{n-1})\big)\right]\!\right]$$
$$= \bigvee_{\substack{n \in \omega \\ a_0, \dots, a_{n-1} \in \omega}} \left[\!\left[\mathfrak{x} = \big((0,a_0), \dots (n-1, a_{n-1})\big)\right]\!\right] \to \bigvee_{\substack{n \in \omega \\ a_0, \dots, a_{n-1} \in \omega}} \left[\!\left[\mathfrak{x} = \big((0,a_0), \dots (n-1, a_{n-1})\big)\right]\!\right] = 1.$$

■

6.3.6 Internal function space

We want to find a representative for the internal function space ω^ω. Externally, this space is defined as subset of $\mathcal{P}(\omega \times \omega)$. Hence, if we repeat this construction inside $V^{(H)}$ i.e. if we set

$$\mathfrak{w}^\omega = \mathrm{Sep}_{\alpha \text{ is total } \wedge \alpha \text{ is single-valued}}\left(\mathrm{Pow}(\widehat{\omega \times \omega})\right),$$

where

α is total $\equiv \forall m \in \omega\ \exists n \in \omega.\, (m,n) \in \alpha$,
α is single-valued $\equiv \forall m \in \omega\ \forall n_1, n_2 \in \omega\ \Big(\big((m,n_1) \in \alpha \wedge (m,n_2) \in \alpha\big) \to n_1 = n_2\Big)$.

we immediately get $V^{(H)} \vDash (\omega^\omega = \mathfrak{w}^\omega)$. However, another representation of ω^ω will be useful in the case where the underlying Heyting-algebra has a clopen basis. Note that $\mathrm{dom}(\mathfrak{w}^\omega)$ is the whole set $H^{\widehat{\omega \times \omega}}$. We want to restrict our attention to those $\beta \in H^{\widehat{\omega \times \omega}}$, where $[\![\beta \text{ is total}]\!] \wedge [\![\beta \text{ is single-valued}]\!] = 1$:

Ⓜ**Theorem 6.35**: Let H be a complete Heyting algebra with clopen basis. Then $\mathfrak{F}$, defined as $\mathrm{dom}(\mathfrak{F}) = \{\beta \in H^{\widehat{\omega\times\omega}} : \llbracket\beta \text{ is total}\rrbracket \wedge \llbracket\beta \text{ is single-valued}\rrbracket = 1\}$ and $\mathfrak{F}(\beta) = 1$, for $\beta \in \mathrm{dom}(\mathfrak{F})$, is another representative of internal ω^ω, i.e. $V^{(H)} \vDash (\mathfrak{F} = \omega^\omega)$.

We will show the rather technical proof of $\llbracket\mathfrak{F} = \mathfrak{w}^\omega\rrbracket = 1$ in a sequence of several lemmas. We start with calculating the following values:

Ⓜ**Lemma 6.36**: For all $\alpha \in H^{\widehat{\omega\times\omega}}$, the following hold

1. $\llbracket(m,n) \in \alpha\rrbracket = \alpha\big(\widehat{(m,n)}\big)$, for all $m,n \in \omega$,
2. $\llbracket\alpha \text{ is total}\rrbracket = \bigwedge_m \bigvee_n \alpha\big(\widehat{(m,n)}\big)$,
3. $\llbracket\alpha \text{ is single-valued}\rrbracket = \bigwedge_{\substack{m \\ n_1\neq n_2}} \neg\Big(\alpha\big(\widehat{(m,n_1)}\big) \wedge \alpha\big(\widehat{(m,n_2)}\big)\Big)$

Proof: Using Lemma 6.31, we have for 1.,

$\llbracket(m,n) \in \alpha\rrbracket = \bigvee_{\mathfrak{x}\in\mathrm{dom}(\alpha)}\big[\alpha(\mathfrak{x}) \wedge \llbracket\widehat{(m,n)} = \mathfrak{x}\rrbracket\big] = \bigvee_{l,k\in\omega}\big[\alpha\big(\widehat{(l,k)}\big) \wedge \llbracket\widehat{(m,n)} = \widehat{(l,k)}\rrbracket\big] = \alpha\big(\widehat{(m,n)}\big) \wedge \llbracket\widehat{(m,n)} = \widehat{(m,n)}\rrbracket = \alpha\big(\widehat{(m,n)}\big)$.

Using this,

$$\llbracket\alpha \text{ is total}\rrbracket = \llbracket\forall m \in \omega\, \exists n \in \omega.\, (m,n) \in \alpha\rrbracket = \llbracket\forall m \in \hat{\omega}\, \exists n \in \hat{\omega}.\, \widehat{(m,n)} \in \alpha\rrbracket = \bigwedge_m \bigvee_n \llbracket\widehat{(m,n)} \in \alpha\rrbracket$$
$$= \bigwedge_m \bigvee_n \alpha\big(\widehat{(m,n)}\big).$$

And finally,

$$\begin{aligned}
\llbracket\alpha \text{ is single-valued}\rrbracket &= \Big\llbracket\forall m \in \omega\, \forall n_1, n_2 \in \omega\, \Big(\big((m,n_1) \in \alpha \wedge (m,n_2) \in \alpha\big) \to n_1 \neq n_2\Big)\Big\rrbracket \\
&= \Big\llbracket\forall m \in \hat{\omega}\, \forall n_1, n_2 \in \hat{\omega}\, \Big(\big(\widehat{(m,n_1)} \in \alpha \wedge \widehat{(m,n_2)} \in \alpha\big) \to n_1 \neq n_2\Big)\Big\rrbracket \\
&= \bigwedge_{m,n_1,n_2} \big[\big(\llbracket\widehat{(m,n_1)} \in \alpha\rrbracket \wedge \llbracket\widehat{(m,n_2)} \in \alpha\rrbracket\big) \to \llbracket n_1 \neq n_2\rrbracket\big] \\
&= \bigwedge_{\substack{m \\ n_1,n_2}} \Big[\Big(\alpha\big(\widehat{(m,n_1)}\big) \wedge \alpha\big(\widehat{(m,n_2)}\big)\Big) \to 0\Big] = \bigwedge_{\substack{m \\ n_1,n_2}} \neg\Big(\alpha\big(\widehat{(m,n_1)}\big) \wedge \alpha\big(\widehat{(m,n_2)}\big)\Big).
\end{aligned}$$

■

Ⓜ**Lemma 6.37**: Let $h \in H$ be clopen and $\alpha \in H^{\widehat{\omega\times\omega}}$ such that $h \leq \llbracket\alpha \text{ is total}\rrbracket \wedge \llbracket\alpha \text{ is single-valued}\rrbracket$. Then there is some $\beta_{h,\alpha} \in \mathfrak{F}$ such that $h \leq \llbracket\alpha = \beta_{h,\alpha}\rrbracket$.

Proof: For the sake of readability of this proof, let us write $\alpha(m,n)$ instead of $\alpha\big(\widehat{(m,n)}\big)$. In the light of Lemma 6.36, the hypothesis $h \leq \llbracket\alpha \text{ is total}\rrbracket \wedge \llbracket\alpha \text{ is single-valued}\rrbracket$ reads

$$h \leq \bigwedge_m \bigvee_n \alpha(m,n) \wedge \bigwedge_{\substack{m \\ n_1 \neq n_2}} \neg\big(\alpha(m,n_1) \wedge \alpha(m,n_2)\big).$$

We define the element $\beta_{h,\alpha} = \beta$ as

$$\begin{aligned} &\beta(m,0) = (\alpha(m,0) \wedge h) \vee \neg h, \\ &\beta(m,n) = \alpha(m,n) \wedge h \qquad\qquad \text{for } n > 0. \end{aligned}$$

Let us first show that $\beta \in \mathrm{dom}(\mathfrak{F})$, i.e. $[\![\beta \text{ is total}]\!] \wedge [\![\beta \text{ is single-valued}]\!] = 1$.

For totality, we have for each m,

$$\begin{aligned} \bigvee_n \beta(m,n) = \beta(m,0) \vee \bigvee_{n>0} \beta(m,n) = (\alpha(m,0) \wedge h) \vee \neg h \vee \bigvee_{n>0} (\alpha(m,n) \wedge h) = \neg h \vee \left(h \wedge \bigvee_n \alpha(m,n)\right) \\ = \neg h \vee h = 1. \end{aligned}$$

Note that $[\![\beta \text{ is single-valued}]\!] = 1$ means that $\beta(m,n_1) \wedge \beta(m,n_2) = 0$ for $n_1 \neq n_2$. We will use the fact that $h \leq \neg\big(\alpha(m,n_1) \wedge \alpha(m,n_2)\big)$, i.e. $h \wedge \big(\alpha(m,n_1) \wedge \alpha(m,n_2)\big) = 0$ for $n_1 \neq n_2$. Indeed, for $n > 0$,

$$\begin{aligned} \beta(m,0) \wedge \beta(m,n) &= [(\alpha(m,0) \wedge h) \vee \neg h] \wedge (\alpha(m,n) \wedge h) \\ &= [(\alpha(m,0) \wedge h) \wedge (\alpha(m,n) \wedge h)] \vee [\neg h \wedge (\alpha(m,n) \wedge h)] \\ &\leq \big[\big(\alpha(m,0) \wedge \alpha(m,n)\big) \wedge h\big] \vee [\neg h \wedge h] = 0 \vee 0 = 0. \end{aligned}$$

For $n_1, n_2 \neq 0$ and $n_1 \neq n_2$,

$$\beta(m,n_1) \wedge \beta(m,n_2) = (\alpha(m,n_1) \wedge h) \wedge (\alpha(m,n_2) \wedge h) = \big(\alpha(m,n_1) \wedge \alpha(m,n_2)\big) \wedge h = 0,$$

hence $\beta \in \mathrm{dom}(\mathfrak{F})$. Finally, again invoking Lemma 6.36,

$$\begin{aligned} [\![\alpha = \beta]\!] &= \bigwedge_{(m,n)} [\alpha(m,n) \to [\![(m,n) \in \beta]\!]] \wedge \bigwedge_{(m,n)} [\beta(m,n) \to [\![(m,n) \in \alpha]\!]] \\ &= \bigwedge_{(m,n)} [\alpha(m,n) \to \beta(m,n)] \wedge \bigwedge_{(m,n)} [\beta(m,n) \to \alpha(m,n)]. \end{aligned}$$

We easily see that $\alpha(m,n) \wedge h \leq \beta(m,n)$ and $\beta(m,n) \wedge h \leq \alpha(m,n)$ for all m,n. Hence, $h \leq \alpha(m,n) \to \beta(m,n)$ and $h \leq \beta(m,n) \to \alpha(m,n)$ which shows that $h \leq [\![\alpha = \beta]\!]$. ■

Having shown this lemma, the proof of Theorem 6.35 follows easily:

Proof: Clearly, for $\beta \in \mathfrak{F}$, we have

$$\begin{aligned} [\![\beta \in \mathfrak{w}^\omega]\!] = \bigvee_{\alpha \in \mathrm{dom}(\mathfrak{w}^\omega)} [\mathfrak{w}^\omega(\alpha) \wedge [\![\beta = \alpha]\!]] = \bigvee_{\alpha \in H^{\widehat{\omega\times\omega}}} [[\![\alpha \text{ is total}]\!] \wedge [\![\alpha \text{ is single-valued}]\!] \wedge [\![\beta = \alpha]\!]] \\ \geq [\![\beta \text{ is total}]\!] \wedge [\![\beta \text{ is single-valued}]\!] \wedge [\![\beta = \beta]\!] = 1. \end{aligned}$$

Using Lemma 6.37, and writing for $\alpha \in H^{\widehat{\omega\times\omega}}$, the element $[\![\alpha \text{ is total}]\!] \wedge [\![\alpha \text{ is single-valued}]\!]$ as join of clopen elements, $\bigvee B'$, we have

$$[\![\alpha \in \mathfrak{F}]\!] = \bigvee_{\beta\in\mathrm{dom}(\mathfrak{F})} [\mathfrak{F}(\beta) \wedge [\![\alpha = \beta]\!]] \geq \bigvee_{h\in B'} [\![\alpha = \beta_{h,\alpha}]\!] \geq \bigvee_{h\in B'} h = [\![\alpha \text{ is total}]\!] \wedge [\![\alpha \text{ is single-valued}]\!].$$

Putting these things together,

$$\begin{aligned}[\![\mathfrak{F} = \mathfrak{w}^\omega]\!] &= \bigwedge_{\alpha\in\mathrm{dom}(\mathfrak{F})} [\mathfrak{w}^\omega(\alpha) \to [\![\alpha \in \mathfrak{F}]\!]] \wedge \bigwedge_{\beta\in\mathrm{dom}(\mathfrak{F})} [\mathfrak{F}(\beta) \to [\![\beta \in \mathfrak{w}^\omega]\!]] \\ &= \bigwedge_{\alpha\in\mathrm{dom}(\mathfrak{F})} [[\![\alpha \text{ is total}]\!] \wedge [\![\alpha \text{ is single-valued}]\!] \to [\![\alpha \in \mathfrak{F}]\!]] \wedge \bigwedge_{\beta\in\mathrm{dom}(\mathfrak{F})} [\![\beta \in \mathfrak{w}^\omega]\!] = 1 \wedge 1 = 1.\end{aligned}$$

■

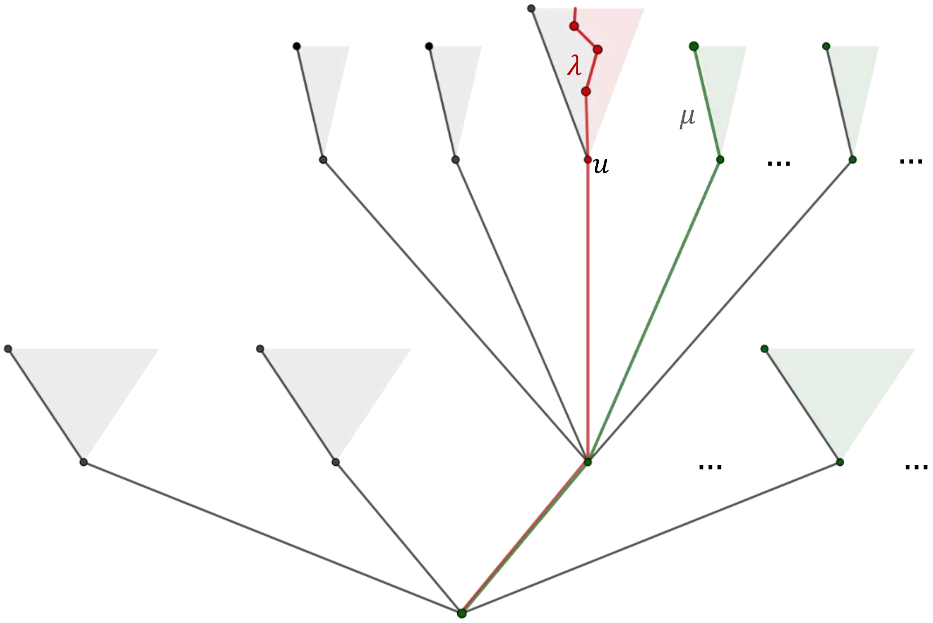
6.4 Independence of Bar induction

In this chapter, we will show that

Using this theorem, we will show using the topology $\mathcal{T}$ on $\mathbb{N}^*$ from Example 6.17 that

$$\mathrm{Cons}(\mathbf{ZF}) \Rightarrow \mathrm{Cons}(\mathbf{IZF} + \neg\mathrm{BI_D}).$$

The proofs in this and the next two chapters is taken from [59]. In order to apply Theorem 6.35, we need an observation about the basic open sets $U_{u,\lambda}$: For simplicity, consider the finite path $u = \langle 2,2\rangle$. In the image we mark $U_{u,\lambda}$ in red and $U_{\langle\,\rangle}, \mu$ in green, where $\langle\,\rangle$ is the empty sequence and $\mu = \langle 2,3,0,0,0,\dots\rangle$.



Note that $U_{u,\lambda}$, $U_{\langle\,\rangle,\mu}$ and the grey area are disjoint and their union is $\mathbb{N}^*$. The grey area may be written as $\bigcup_{x \notin U_{u,\lambda} \cup U_{\langle\,\rangle,\mu}} \{v : v \sqsupseteq x\}$. Since $\{v : v \sqsupseteq x\} = U_{x, x*\bar{0}}$, where $\bar{0} = \langle 0,0,0,\dots\rangle$, this shows that $U_{u,\lambda}$ is clopen. Let us formalize this argument:

Proposition 6.38: The basic sets $U_{u,\lambda}$ are clopen and so are the sets $O_u = \mathbb{N}^* \setminus \{v : v \sqsupseteq u\}$ for $u \in \mathbb{N}^*$.

Proof: Let u be of length n, i.e. $u = \langle u_0, \dots, u_{n-1}\rangle$. We define the infinite path $\mu = \langle u_0, \dots, u_{n-1} + 1, 0, 0, \dots\rangle$. Notice that $U_{u,\lambda}$ and $U_{\langle\,\rangle,\mu}$ are disjoint and all points not in $U_{u,\lambda} \cup U_{\langle\,\rangle,\mu}$ lie to the left of it.

Hence, we may write $\mathbb{N}^*$ as the following disjoint union of open sets:

$$\mathbb{N}^* = U_{u,\lambda} \cup U_{\langle\,\rangle,\mu} \cup \bigcup_{x \notin U_{u,\lambda} \cup U_{\langle\,\rangle,\mu}} \{v : v \sqsupseteq x\},$$

Each O_u is clopen: This works almost the same as before: Set $\lambda = u * \bar{0}$, then $\{v : v \sqsupseteq u\} = U_{u,\lambda}$ and let μ be as before. Again, we can write:

$$\mathbb{N}^* = U_{u,\lambda} \cup \underbrace{U_{\langle\,\rangle,\mu} \cup \bigcup_{x \notin U_{u,\lambda} \cup U_{\langle\,\rangle,\mu}} \{v : v \sqsupseteq x\}}_{O_u},$$

where all three sets are clopen and disjoint. ■

Ⓜ**Lemma 6.39**: Let $t \in \mathbb{N}^*$, $\alpha \in \mathfrak{F}$ and $m \in \omega$. Then there is some $u_t \in \omega^{<\omega}$ such that $t \in \llbracket \alpha|_m = \widehat{u_t} \rrbracket$.

Proof: $\alpha \in \mathfrak{F}$ means that for each n,

$$\dot{\bigcup_k} \alpha\big(\widehat{(n,k)}\big) = \mathbb{N}^*.$$

Since this union is disjoint, for each n, there is a unique k_n such that $t \in \alpha\big(\widehat{(n,k_n)}\big)$. In particular, the set $U = \bigcap_{n<m} \alpha\big(\widehat{(n,k_n)}\big)$ is an open neighborhood of t. Let u_t be the sequence $(k_n)_{n<m}$, then

$$\begin{aligned}
\llbracket \alpha|_m = \widehat{u_t} \rrbracket &= \bigwedge_{\widehat{(n,k)} \in \mathrm{dom}(\alpha|_m)} \left[\alpha\big(\widehat{(n,k)}\big) \to \llbracket \widehat{(n,k)} \in \widehat{u_t} \rrbracket\right] \wedge \bigwedge_{\widehat{(n,k)} \in \mathrm{dom}(u_t)} \left[u_t\big(\widehat{(n,k)}\big) \to \llbracket \widehat{(n,k)} \in \alpha \rrbracket\right] \\
&= \bigwedge_{\substack{\widehat{(n,k)} \in \mathrm{dom}(\alpha) \\ u_t(n) \neq k \\ n<m}} \left[\neg\alpha\big(\widehat{(n,k)}\big)\right] \wedge \bigcap_{n<m} \alpha\big(\widehat{(n,k_n)}\big).
\end{aligned}$$

We already know that t is contained in the second conjunct. For each $n < m$, $\alpha\big(\widehat{(n,k_n)}\big) \subseteq \neg\alpha\big(\widehat{(n,k)}\big)$ for $k \neq k_n$, by disjointness. Hence,

$$t \in U \subseteq \alpha\left(\widehat{(n, k_n)}\right) \subseteq \bigwedge_{\substack{\widehat{(n,k)} \in \mathrm{dom}(\alpha) \\ u_t(n) \neq k}} \left[\neg\alpha\left(\widehat{(n,k)}\right)\right],$$

for each $n < m$, which shows that the first conjunct is an open neighborhood of t as well. ∎

We can now show that the special case of $\mathrm{BI_D}$

$$\begin{bmatrix} \forall \alpha \in \omega^{\omega}. \exists n \in \omega. \alpha|_n \in \mathfrak{B} \wedge \\ \forall u \in \omega^{<\omega} \; (u \in \mathfrak{B} \vee u \notin \mathfrak{B}) \wedge \\ \forall u \in \omega^{<\omega} (\forall k \in \omega \; u * \langle k \rangle \in \mathfrak{B} \to u \in \mathfrak{B}) \end{bmatrix} \to \langle\,\rangle \in \mathfrak{B},$$

is not valid in $V^{(\mathcal{T})}$. Here $\mathfrak{B}$ is given by $\mathrm{dom}(\mathfrak{B}) = \mathrm{dom}(\mathfrak{w}^{<\omega})$ and $\mathfrak{B}(\hat{u}) = O_u$, where $O_u = \mathbb{N}^* \setminus \{v \in \mathbb{N}^* : v \supseteq u\}$. Note that for $u \in \omega^{<\omega}$, we have that

$$[\![\hat{u} \in \mathfrak{B}]\!] = \bigcup_{v \in \omega^{<\omega}} [\mathfrak{B}(\hat{u}) \cap [\![\hat{v} = \hat{u}]\!]] = \mathfrak{B}(\hat{u}) = O_u.$$

We will show that all the antecedents evaluate to $\mathbb{N}^*$ while the consequent has truth value $\emptyset$. Indeed, we immediately get that $[\![\langle\,\rangle \in \mathfrak{B}]\!] = O_{\langle\,\rangle} = \mathbb{N}^* \setminus \{v \in \mathbb{N}^* : v \supseteq \langle\,\rangle\} = \mathbb{N}^* \setminus \mathbb{N}^* = \emptyset$.

For the first antecedent, let $\alpha \in \mathfrak{F}$. For each $t \in \mathbb{N}^*$ let $m = \mathrm{length}(t) + 1$. Then clearly, $t \in O_u$ for each $u \in \mathbb{N}^*$. As in Lemma 6.39, let $u_t \in \omega^{<\omega}$ be such that $t \in [\![\alpha|_m = \widehat{u_t}]\!]$. Then

$$\begin{aligned} t \in [\![\alpha|_m = \widehat{u_t}]\!] \cap O_{u_t} = [\![\alpha|_m = \widehat{u_t}]\!] \cap [\![\widehat{u_t} \in \mathfrak{B}]\!] \subseteq \bigcup_{u \in \omega^{<\omega}} [[\![\alpha|_m = u]\!] \cap [\![u \in \mathfrak{B}]\!]] \\ \subseteq \bigcup_{n \in \omega} \bigcup_{u \in \omega^{<\omega}} [[\![\alpha|_n = u]\!] \cap [\![u \in \mathfrak{B}]\!]] = [\![\exists n \in \omega. \exists u \in \omega^{<\omega} (\alpha|_n = u \wedge u \in \mathfrak{B})]\!] \\ = [\![\exists k \in \omega. \alpha|_k \in \mathfrak{B}]\!]. \end{aligned}$$

The second antecedent is basically the fact that the O_us are clopen (

Note that $U_{u,\lambda}$, $U_{\langle\,\rangle}, \mu$ and the grey area are disjoint and their union is $\mathbb{N}^*$. The grey area may be written as $\bigcup_{x \notin U_{u,\lambda} \cup U_{\langle\,\rangle,\mu}} \{v : v \supseteq x\}$. Since $\{v : v \supseteq x\} = U_{x, x*\bar{0}}$, where $\bar{0} = \langle 0,0,0,\ldots\rangle$, this shows that $U_{u,\lambda}$ is clopen. Let us formalize this argument:

Proposition 6.38): Let $u \in \mathbb{N}^*$, then

$$[\![\hat{u} \in \mathfrak{B} \vee \hat{u} \notin \mathfrak{B}]\!] = \mathfrak{B}(\hat{u}) \cup \neg\mathfrak{B}(\hat{u}) = \mathfrak{B}(\hat{u}) \cup \mathrm{Int}\left(\mathbb{N}^* \setminus \mathfrak{B}(\hat{u})\right) = O_u \cup \mathrm{Int}(\mathbb{N}^* \setminus O_u) = O_u \cup (\mathbb{N}^* \setminus O_u) = \mathbb{N}^*.$$

Finally, the third antecedent: For $u \in \mathbb{N}^*$, we have

$$\llbracket \forall k \in \omega . \widehat{u * \langle k \rangle} \in \mathfrak{B} \rrbracket = \bigwedge_{k \in \omega} \llbracket \widehat{u * \langle k \rangle} \in \mathfrak{B} \rrbracket = \bigwedge_{k \in \omega} \llbracket \widehat{u * \langle k \rangle} \in \mathfrak{B} \rrbracket = \operatorname{Int}\left(\bigcap_{k \in \mathbb{N}} O_{u * \langle k \rangle} \right)$$
$$= \operatorname{Int}\left(\bigcap_{k \in \mathbb{N}} \mathbb{N}^* \setminus \{ v \in \mathbb{N}^* | v \supseteq u * \langle k \rangle \} \right) = \operatorname{Int}\left(\mathbb{N}^* \setminus \bigcup_{k \in \mathbb{N}} \{ v \in \mathbb{N}^* | v \supseteq u * \langle k \rangle \} \right)$$
$$= \operatorname{Int}\big(\{u\} \cup (\mathbb{N}^* \setminus \{ v \in \mathbb{N}^* | v \supseteq u \}) \big) = \mathbb{N}^* \setminus \{ v \in \mathbb{N}^* | v \supseteq u \} = \mathfrak{B}(u) = \llbracket \hat{u} \in \mathfrak{B} \rrbracket.$$

where the equality before the last holds, since any open set containing u must also contain some $u * \langle k \rangle$. We thus have

$$\llbracket \forall k \in \omega . \widehat{u * \langle k \rangle} \in \mathfrak{B} \rightarrow \hat{u} \in \mathfrak{B} \rrbracket = \llbracket \forall k \in \omega . \widehat{u * \langle k \rangle} \in \mathfrak{B} \rrbracket \rightarrow \llbracket \hat{u} \in \mathfrak{B} \rrbracket = \llbracket \hat{u} \in \mathfrak{B} \rrbracket \rightarrow \llbracket \hat{u} \in \mathfrak{B} \rrbracket = \mathbb{N}^*.$$

This concludes the proof of independence of $\mathrm{BI_D}$.

6.5 Compatibility with the Fan theorem

In this section we will show that **IZF** is compatible with the fan theorem. To be more precise, we will use Theorem 6.11 and the topology $\mathcal{T}$ on $\mathbb{N}^*$ from Example 6.17 to show:

Ⓜ**Theorem 6.40**: $\mathrm{Cons}(\mathbf{ZF}) \Rightarrow \mathrm{Cons}(\mathbf{IZF} + \mathbf{F})$.

We will start with discussing some equivalent formulations of the fan theorem.

6.5.1 Fan theorem and equivalent formulations

First, we want to topologize any fan in the following way: Let T be a fan (thought of as infinite paths of the underlying tree). We define the following topology on T: Let $T^{<\omega}$ denote the set of finite paths through T, starting at the root. For each $u \in T^{<\omega}$, let $U_u = \{\alpha \in T : \alpha \supseteq u\}$. Then $\mathcal{B} = \{U_u : u \in T^{<\omega}\}$ forms a basis of a topology on T. By T_2 we denote the important case of the binary tree. We consider the following versions of the fan theorem:

$$\big(\forall \alpha \in T\, \exists n \in \omega\, \phi(\alpha|_n)\big) \rightarrow \big(\exists m \in \omega\, \forall \alpha \in T\, \exists n < m\, \phi(\alpha|_n)\big) \qquad (\mathbf{F}_T)$$

The space T is compact. $(\mathbf{F}_T^{\mathbf{c}})$

$$\big(\forall \alpha \in 2^\omega\, \exists n \in \omega\, \phi(\alpha|_n)\big) \rightarrow \big(\exists m \in \omega\, \forall \alpha \in 2^\omega\, \exists n < m . \phi(\alpha|_n)\big) \qquad (\mathbf{F}_2)$$

The space T_2 is compact. $(\mathbf{F}_2^{\mathbf{c}})$

We will show in the following that all these formulations come down to the same thing. We start with the following two results:

Lemma 6.41: Each fan T is homeomorphic to a subfan of T_2.

Proof: We can think of the underlying tree of T as subtree of ω^ω, where the nodes are labelled by finite sequences of natural numbers in such a way that $u * \langle k \rangle \in T^{<\omega} \rightarrow u * \langle i \rangle \in T^{<\omega}$ for all $i < k$. Given $u \in T^{<\omega}$ of length n, we define $\Phi(u)$ as the finite sequence

$$\underbrace{0,\dots,0}_{\substack{u(0)+1\\ \text{many}}},\underbrace{1,\dots,1}_{\substack{u(1)+1\\ \text{many}}},\underbrace{0,\dots,0}_{\substack{u(2)+1\\ \text{many}}},\underbrace{1,\dots,1}_{\substack{u(3)+1\\ \text{many}}},\dots,\underbrace{*,\dots,*}_{\substack{u(n-1)+1\\ \text{many}}}.$$

Then $\Phi: T \to T_2$ is defined by $\Phi(\alpha) = \bigcup_{n\in\omega} \Phi(\alpha|_n)$. By our assumption on the structure of T, Φ is a mapping of fans, i.e. $\Phi[T]$ is a fan as well. Clearly, Φ is injective. Given $\alpha \in T$ and an open neighborhood U of $\Phi(\alpha)$, we may assume that U is of the form $U_{\Phi(\alpha|_n)}$. Then for all $\beta \in U_{\alpha|_n}$, we have that $\Phi(\beta) \in U$.

For the continuity of Φ^{-1} the fact that T is finitely branching is crucial: For each n, let $\beta_i = \max\{\alpha(i): \alpha \in T\}$. Now, for any open neighborhood U of α (again, assume $U = U_{\alpha|_n}$), we have to find a neighborhood V of $\Phi(\alpha)$ such that for all $\beta \in V$, $\Phi^{-1}(\beta) \in U$. Such a V can be given as $U_{\Phi(\alpha|_m)}$, where $m = \sum_{i=0}^{n-1} \beta_i + n$. ■

Lemma 6.42: Every subfan T' of a fan T is closed.

Proof: For each path $\alpha \in T$ not in T', there is some initial segment $\alpha|_n \notin T^n$. Hence $U_{\alpha|_n}$ is an open neighborhood of α disjoint from T. We can write $T \setminus T'$ as the union of all such neighborhoods. ■

We can now establish equivalence of the different versions of the fan theorem:

Theorem 6.43: For each fan T, $\mathbf{F}_T$ and $\mathbf{F}_T^{\mathbf{c}}$ are equivalent over **IZF**.

Proof: $\mathbf{F}_T^{\mathbf{c}} \to \mathbf{F}_T$: Suppose, $\forall\alpha \in T.\exists n \in \omega.\phi(\alpha|_n)$. We construct an open cover $\mathcal{O}$ for T: Add for each α and n with $\phi(\alpha|_n)$ the set $U_{\alpha|_n}$ to $\mathcal{O}$. By the premise and the fact $\alpha \in U_{\alpha|_n}$, $\mathcal{O}$ is indeed an open cover. By compactness, there are $\alpha_0, \dots, \alpha_{k-1} \in T$ and $n_1, \dots, n_{k-1} \in \omega$ such that $\left(U_{\alpha_i|_{n_i}}\right)_{i<k}$ covers T as well. Let $m = \max\{n_i: i < k\}$. For each $\alpha \in T$, there is some $i < k$ such that $\alpha \in U_{\alpha_i|_{n_i}}$. Since $\phi\left(\alpha_i|_{n_i}\right)$ holds, and $\alpha_i|_{n_i} = \alpha|_{n_i}$, we conclude $\phi\left(\alpha|_{n_i}\right)$.

$\mathbf{F}_T \to \mathbf{F}_T^{\mathbf{c}}$: We show that every open cover of basic sets has a finite subcover: Let $\mathcal{O} \subseteq \mathcal{B}$ with $\bigcup \mathcal{O} = T$. We define the formula ϕ on finite sequences as $\phi(u) \equiv U_u \in \mathcal{O}$. Since $\mathcal{O}$ covers T,

$$\begin{aligned} &\forall\alpha \in T.\exists U_u \in \mathcal{O}.\alpha \in U_u \\ \Leftrightarrow\ &\forall\alpha \in T.\exists u \in T^{<\omega}.\left(\alpha|_{\mathrm{dom}(u)} = u \wedge \phi(u)\right) \\ \Leftrightarrow\ &\forall\alpha \in T.\exists n \in \omega.\alpha \in \phi(\alpha|_n). \end{aligned}$$

Hence, there is some $m \in \omega$ such that for all $\alpha \in T$, there is some $n \in \omega$ with $\phi(\alpha|_n)$. This means that for each $\alpha \in T$, $U_{\alpha|_n} \in \mathcal{O}$ for some $n < m$. As $\alpha \in U_{\alpha|_n}$, we have shown that the finite set $\{U_u \in \mathcal{O}: \mathrm{length}(u) < m\}$ covers T. ■

Proposition 6.44: Let T' be a subfan of T. Then $\mathbf{IZF} \vdash \mathbf{F}_T \to \mathbf{F}_{T'}$.

Proof: $\mathbf{F}_T \to \mathbf{F}_T^{\mathbf{c}}$. T' is a closed subset of a compact space, hence compact itself. This shows $\mathbf{F}_{T'}^{\mathbf{c}}$ and therefore $\mathbf{F}_{T'}$. ■

Theorem 6.45: **IZF** ⊢ $\mathbf{F}_T^{\mathbf{c}}$ for each fan $T \leftrightarrow \mathbf{F}_2^{\mathbf{c}}$.

Proof: The direction from left to right is trivial. Let T_2 be compact and T any fan. We show that T is compact as well. By Lemma 6.41, T is homeomorphic to a subfan T' of T_2. By Lemma 6.42, T' is closed and by compactness of T_2 and Proposition 6.21 it is compact itself. We conclude that T is compact by Proposition 6.20. ■

Corollary 6.46: Over **IZF**, all the formulations of the fan theorem are equivalent.

6.5.2 Proof of compatibility

We are now ready to our proof of compatibility:

Ⓜ**Proposition 6.47**: $V^{(\mathcal{T})} \vDash \widehat{T_2}$ is compact.

Proof: Let $\mathcal{O} \subseteq \mathcal{B}$ and let $t \in \mathbb{N}^*$ and $t \in [\![\forall \hat{\alpha} \in \widehat{2^\omega}\ \exists \hat{O} \in \hat{\mathcal{O}}\ \hat{\alpha} \in \hat{O}]\!]$, i.e. for each $\alpha \in T_2$,

$$t \in \bigcup_{O \in \mathcal{O}} [\![\hat{\alpha} \in \hat{O} \wedge \hat{O} \in \hat{\mathcal{O}}]\!]$$

$$\Leftrightarrow \alpha \in \bigcup_{t \in [\![\hat{O} \in \hat{\mathcal{O}}]\!]} O.$$

This means that $\{O \in \mathcal{O} : t \in [\![\hat{O} \in \hat{\mathcal{O}}]\!]\}$ is an open cover of 2^ω. Classically, König's lemma 2.1 holds and is equivalent to the fan theorem (Theorem 2.4). This implies that T_2 is externally compact (i.e. provably compact in **ZF**). Hence, there is some finite subcover $O_1, \dots O_n$ of 2^ω. This shows

$$t \in [\![O_1 \in \mathcal{O} \wedge \dots \wedge O_n \in \mathcal{O}]\!] \wedge [\![\forall \alpha \in 2^\omega. (\alpha \in O_1 \vee \dots \vee. \alpha \in O_n)]\!].$$

As t was arbitrary, this concludes the proof. ■

We would want to apply Corollary 6.46 to conclude that $\mathbf{F}_2$ (and hence $\mathbf{F}_T$ for all trees T) holds in $V^{(\mathcal{T})}$. However, at this point, all we can say is the following weaker version:

Ⓜ**Corollary 6.48**: $V^{(\mathcal{T})} \vDash \left(\forall \alpha \in \widehat{T_2}\ \exists n \in \omega\ \phi(\alpha|_n)\right) \to \left(\exists m \in \omega\ \forall \alpha \in \widehat{T_2}\ \exists n < m\ \phi(\alpha|_n)\right)$.

To infer the general form of $\mathbf{F}_2$, let us argue inside $V^{(\mathcal{T})}$: Suppose, the weaker form of $\mathbf{F}_2{}'$ of the previous Corollary holds. Then, if $\forall \alpha \in T_2\ \exists n \in \omega\ \phi(\alpha|_n)$, then in particular, the premise of the weak $\mathbf{F}_2{}'$ is satisfied, Hence,

$$\exists m \in \omega\ \forall \alpha \in \widehat{T_2}\ \exists n < m\ \phi(\alpha|_n).$$

We will show that this m works for T_2 as well: Let $\alpha \in T_2$. We know that for each $n < m$, there is some some $u_n \in T_2^{<\omega}$ such that $\alpha|_n = u_n$. But $T_2^{<\omega} = \widehat{T_2^{<\omega}}$ and hence ϕ must hold true for one of the u_ns, say $u_{\tilde{n}}$. We can therefore conclude $\phi(\alpha|_n)$. Formally, the proof looks like this:

$$\begin{aligned}\llbracket \forall \alpha \in T_2\, \exists n \in \omega\, \phi(\alpha|_n) \rrbracket \leq \left\llbracket \forall \alpha \in \widehat{T_2}\, \exists n \in \omega\, \phi(\alpha|_n) \right\rrbracket \leq \left\llbracket \exists m \in \omega\, \forall \alpha \in \widehat{T_2}\, \exists n < m.\, \phi(\alpha|_n) \right\rrbracket \\ \leq \left\llbracket \exists m \in \omega\, \forall \alpha \in \widehat{T_2}\, \exists n < m\, \phi(\alpha|_n) \right\rrbracket \\ \leq \left\llbracket \exists m \in \omega\, \forall \alpha \in T_2\, \exists u \in \widehat{T_2^m}\, \big(u = \alpha|_m \wedge (\exists n < m.\, \phi(u|_n))\big) \right\rrbracket \\ \leq \left\llbracket \exists m \in \omega\, \forall \alpha \in T_2\, \exists n < m\, \exists u \in \widehat{T_2^n}\, \big(u = \alpha|_n \wedge \phi(u|_n)\big) \right\rrbracket \\ \leq \llbracket \exists m \in \omega\, \forall \alpha \in T_2\, \exists n < m\, \phi(\alpha|_n) \rrbracket.\end{aligned}$$

This shows that $V^{(\mathcal{T})} \vDash \mathbf{F}_2{}'$ (and thus $V^{(\mathcal{T})} \vDash \mathbf{F}_T$ for any tree T) and concludes the proof of Theorem 6.40.

6.6 Compatibility with the Weak Continuity Principle

In this section, we will show that the Weak Continuity Principle **WCN** is compatible with **IZF**:

Ⓜ**Theorem 6.49**: $\mathrm{Cons}(\mathbf{ZF}) \Rightarrow \mathrm{Cons}(\mathbf{IZF} + \mathbf{WCN})$

Lemma 6.50: There are continuous functions

1. $f\colon \mathbb{N}^* \to \{v : v \supseteq u\}$ with $f|_{\{v:v\supseteq u\}} = \mathrm{id}_{\{v:v\supseteq u\}}$.
2. $g_1\colon \mathbb{N}^* \to \{v : v \supseteq u\}$ homeomorphism with $g_1(\langle\,\rangle) = u$,
3. $g_2\colon \mathbb{N}^* \to U_{\lambda,w}$ homeomorphism with $g_2(\langle\,\rangle) = w$,
4. $h\colon \mathbb{N}^* \to U_{\lambda,w}$ with $h|_{\{v:v\supseteq u\}}\colon \{v : v \supseteq u\} \to U_{\lambda,w}$ homeomorphism and $h(u) = w$.

Proof: Let $\mathrm{length}(u) = n$. We define f as follows: For any $v \in \mathbb{N}^*$, we set

$$\begin{aligned} f(v) &= u, && \text{if } \mathrm{length}(v) < n \\ f(v) &= u_0, \dots, u_{n-1}, v_n, \dots, v_{\mathrm{length}(v)-1}, && \text{if } \mathrm{length}(v) \geq n. \end{aligned}$$

This f is continuous: Let $f(v) \in U_{\lambda,w}$. Without loss of generality, $w \supseteq u$. Let $\tilde{\lambda} = u * \langle \lambda_n, \lambda_{n+1}, \dots \rangle$. Then for all $x \in U_{\tilde{\lambda},v}$, we have that $f(x) \in U_{\lambda,w}$.

g_1 is defined as $g_1(v) = u * v$ and is clearly continuous and g_2 is defined similarly. Finally, for h, we can set $h = g_2 \circ g_1^{-1} \circ f$. ■

Ⓜ**Proposition 6.51**: Let $h\colon H \to G$ be a homomorphism of complete Heyting-algebras. Then for each formula $\phi(x_1, \dots, x_n)$ with all free variables shown and $\mathfrak{a}_1, \dots, \mathfrak{a}_n \in V^{(H)}$, we have that

$$h(\llbracket \phi(\mathfrak{a}_1, \dots, \mathfrak{a}_n) \rrbracket) = \left\llbracket \phi\big(\mathfrak{a}_1^h, \dots, \mathfrak{a}_n^h\big) \right\rrbracket,$$

where $\mathfrak{a}^h$ is defined recursively by $\mathfrak{a}^h = \big\{\langle \mathfrak{b}^h, h\big(\mathfrak{a}(\mathfrak{b})\big)\rangle : \mathfrak{b} \in \mathrm{dom}(\mathfrak{a})\big\}$.

Proof: This is shown by a straightforward induction on the structure of ϕ, the cases for atomic ϕ are shown by simultaneous induction. ■

Remember, that in our topology on $\mathbb{N}^*$, the function space ω^ω, as given by $\mathfrak{F}$ from Theorem 6.35, is the set of all $\alpha \in \widehat{\omega \times \omega}$ such that for all n,

$$\dot{\bigcup_m} \alpha\left(\widehat{(n,m)}\right) = \mathbb{N}^*.$$

Hence, for each $t \in \mathbb{N}^*$ and each n, there is a unique m such that $t \in \alpha\left(\widehat{(n,m)}\right)$. Hence, externally, α defines a map $\tilde{\alpha}: \mathbb{N}^* \to \omega^\omega$. If we equip ω^ω with the usual tree topology from the last section, we see that $\tilde{\alpha}$ is continuous: Note that the sets $V_{n,m} = \{\alpha \in \omega^\omega : \alpha(n) = m\}$ form a subbasis, so we need to check continuity only for these sets. We have

$$\begin{aligned}\tilde{\alpha}^{-1}\left[V_{n,m}\right] &= \left\{t \in T : \tilde{\alpha}(t) \in V_{n,m}\right\} \\ &= \{t \in T : \tilde{\alpha}(t)(n) = m\} \\ &= \left\{t \in T : t \in \alpha\left(\widehat{(n,m)}\right)\right\} \\ &= \alpha\left(\widehat{(n,m)}\right).\end{aligned}$$

On the other hand, let $\eta: \mathbb{N}^* \to \omega^\omega$ be continuous, then $\tilde{\eta}$ defined as $\tilde{\eta}(n,m) = \eta^{-1}\left[V_{m,n}\right]$ defines a function $\omega \times \omega \to \mathbb{N}^*$. Clearly both constructions are inverses of one another and we have proved

Proposition 6.52: The assignment $\alpha \mapsto \tilde{\alpha}$ is a 1-1 correspondence between continuous functions $\mathbb{N}^* \to \omega^\omega$ and functions $\omega \times \omega \to \mathbb{N}^*$.

In regard with this proposition and Proposition 6.51, we may ask how $\tilde{\alpha}$ behaves under continuous functions: As any continuous function $\mathbb{N}^* \to \mathbb{N}^*$ gives rise to the homomorphism of complete Heyting-algebras of the topological spaces f^{-1}. We can thus apply Proposition 6.51 to get that $f^{-1}(\llbracket \phi(\mathfrak{a}_1, \dots, \mathfrak{a}_n) \rrbracket) = \left\llbracket \phi\left(\mathfrak{a}_1^{f^{-1}}, \dots, \mathfrak{a}_n^{f^{-1}}\right)\right\rrbracket$. We are interested how the elements of $\mathfrak{F}$ look like under this mapping:

Ⓜ**Lemma 6.53**: Let $\alpha \in \mathfrak{F}$ and $f: \mathbb{N}^* \to \mathbb{N}^*$ continuous. Then $\tilde{\alpha}^{f^{-1}} = \tilde{\alpha} \circ f$.

Proof: We have the chain of equivalences

$$\begin{aligned}&\tilde{\alpha}^{f^{-1}}(t)(m) = n \\ \Leftrightarrow\ & t \in \alpha^{f^{-1}}\left(\widehat{(m,n)}\right) \\ \Leftrightarrow\ & t \in f^{-1}\left[\alpha\left(\widehat{(m,n)}\right)\right] \\ \Leftrightarrow\ & f(t) \in \alpha\left(\widehat{(m,n)}\right) \\ \Leftrightarrow\ & \tilde{\alpha}\big(f(t)\big)(m) = n,\end{aligned}$$

That show the desired equality of functions. ■

With these preparations, we are now ready to prove the main result of this section:

Proof of Theorem 6.49: We will show that $V^{(\mathcal{T})} \vDash \mathbf{WCN}$, where $\mathcal{T}$ is again the topology from Example 6.17. Suppose, towards a contradiction, that for some $U_{u,\kappa}$ we have $U_{u,\kappa} \subseteq \llbracket \forall\alpha \in \omega^\omega\ \exists x \in \omega\ \phi(\alpha, x) \rrbracket$, but

$$\begin{aligned} U_{u,\kappa} \nsubseteq & \left\llbracket \forall\alpha \in \omega^\omega \exists y, b \in \omega\ \forall\delta \in \omega^\omega \left[\forall\delta \in \omega^\omega\ \xi|_m = \bar{\delta}(m) \to \phi(\delta, n)\right]\right\rrbracket \\ & = \operatorname{int}\left(\bigcap_{\xi \in \omega^\omega} \llbracket \exists y, b \in \omega\ [\forall\delta \in \omega^\omega . \xi|_m = \delta|_m \to \phi(\delta, n)] \rrbracket \right). \end{aligned}$$

As the interior of a set is the largest open set contained in it, this means that there is some $t \in U_{u,\kappa}$ and $\xi \in \omega^\omega$ such that $t \notin \llbracket \exists y, b \in \omega\ [\forall\delta \in \omega^\omega . \xi|_m = \delta|_m \to \phi(\delta, n)] \rrbracket$.

Hence, for all $m, n \in \omega$, $t \notin \llbracket \forall\delta \in \omega^\omega . \xi|_m = \delta|_m \to \phi(\delta, n) \rrbracket$. As in Proposition 6.52, we may interpret ξ as continuous function $\tilde{\xi}\colon \mathbb{N}^* \to \omega^\omega$. By this continuity, for every $m \in \mathbb{N}$ there exists a basic open $U_{w_m,\lambda_m} \subseteq U_{u,\kappa}$ with $t \in U_{w_m,\lambda_m}$ and $t * \langle m \rangle \notin U_{w_m,\lambda_m}$ such that for every $x \in U_{w_m,\lambda_m}$ we have $\xi(x)(m) = \xi(t)(m)$.

Claim: There are $t_m^n \in U_{w_m,\lambda_m}$ and $t_m^n \in \omega$ such that

$$\llbracket \xi|_m = \eta_m^n|_m \rrbracket = \mathbb{N}^* \text{ and } t_m^n \notin \llbracket \phi(\eta_m^n, n) \rrbracket.$$

Proof: As $t \notin \operatorname{int}\left(\bigcap_{\delta \in \mathfrak{F}} \llbracket \xi|_m = \delta|_m \to \phi(\delta, n) \rrbracket\right)$, for each m, n we can use a form of the axiom of choice to obtain some $t_m^n \in U_{u,\kappa}$ but not in the intersection. Hence there are $\eta_m^n \in \mathfrak{F}$ such that

$$t_m^n \notin \llbracket \xi|_m = \eta_m^n|_m \rrbracket \to \llbracket \phi(\eta_m^n, n) \rrbracket = (\mathbb{N}^* \setminus \llbracket \xi|_m = \eta_m^n|_m \rrbracket) \cup \llbracket \phi(\eta_m^n, n) \rrbracket$$

Possibly redefining η_m^n by β_m^n, we can assume that $\llbracket \xi|_m = \eta_m^n|_m \rrbracket = \mathbb{N}^*$: We simply define $\widetilde{\beta_m^n}$ as

$$\widetilde{\beta_m^n}(t)(k) = \begin{cases} \xi(t)(k), & \text{for } k < m, \\ \eta_m^n(t)(k), & \text{for } k \geq m. \end{cases}$$

Then clearly $\llbracket \xi|_m = \beta_m^n|_m \rrbracket = \mathbb{N}^*$. If t_m^n were in $\llbracket \phi(\beta_m^n, n) \rrbracket$, then

$$\begin{aligned} t_m^n \in \llbracket \phi(\beta_m^n, n) \rrbracket = \llbracket \phi(\beta_m^n, n) \rrbracket \cap \mathbb{N}^* = \llbracket \phi(\beta_m^n, n) \rrbracket \cap \llbracket \xi|_m = \beta_m^n|_m \rrbracket = \llbracket \phi(\beta_m^n, n) \rrbracket \cap \llbracket \eta_m^n = \beta_m^n \rrbracket \\ \subseteq \llbracket \phi(\eta_m^n, n) \rrbracket, \end{aligned}$$

a contradiction. ■

Let $(n_i)_{i \in \omega}$ be a sequence of natural numbers such that each $n \in \omega$ appears in it infinitely many times. Set $s_i = t * \langle i \rangle$ and $V_i = \{v \in \mathbb{N}^* : v \supseteq s_i\}$. By Lemma 6.50, there are continuous functions $\rho_i\colon \mathbb{N}^* \to U_{w_i,\lambda_i}$ such that $\rho_i(s_i) = t_i^{n_i}$ and $\rho|_{V_i}\colon V_i \to U_{w_i,\lambda_i}$ are homeomorphisms.

We define $\eta\colon \mathbb{N}^* \to \omega^\omega$ by

$$\eta(s) = \begin{cases} \eta_i^{n_i}\big(\rho_i(s)\big), & \text{if } s \in V_i, \\ \xi(s), & \text{if } s \notin \bigcup\limits_{i \in \omega} V_i\,. \end{cases}$$

η is well defined, as the V_i are mutually disjoint. To show that it is continuous it suffices to show that it is continuous at $\partial(\bigcup_{i\in\omega} V_i) = \mathrm{cl}(\bigcup_{i\in\omega} V_i) \setminus \bigcup_{i\in\omega} V_i = \{t\}$ (Lemma 6.22): Given k, we need to find an open neighborhood U of t such that $\eta(x)|_k = \eta(t)|_k = \xi(t)|_k$ for all $x \in U$. Let λ be the path $t * \langle k, 0,0,0, \dots \rangle$. We check that $U = U_{\lambda,t}$ does the job: Indeed for $x \neq t$, $x \in U_{\lambda,t}$ means that $x \in V_i$ for some $i \geq k$. Thus $\eta(x) = \eta_i^{n_i}\big(\rho_i(x)\big)$ and $\rho_i(x) \in U_{w_i,\lambda_i}$. Hence,

$$\eta(x)|_i = \eta_i^{n_i}\big(\rho_i(x)\big)\big|_i = \xi\big(\rho_i(x)\big)\big|_i = \xi(t)|_i.$$

Since $t \in U_{u,\kappa} \subseteq [\![\forall \alpha \in \omega^\omega. \exists x \in \omega. \phi(\alpha, x)]\!]$ for $\alpha = \tilde{\eta}$, there is an $n \in \omega$ with $t \in [\![\phi(\tilde{\eta}, n)]\!]$. The $s_i = t * \langle i \rangle$ converge to t and since n appears infinitely many times in $(n_i)_{i\in\omega}$, there is some $i \in \omega$ with $s_i \in [\![\phi(\tilde{\eta}, \eta_i^{n_i})]\!]$, where $n_i = n$.

As $t_i^{n_i} = \rho(s_i) \in [\![\phi(\tilde{\eta} \circ \rho_i^{-1}, n_i)]\!]$ (Lemma 6.53) and $t_i^{n_i} \in [\![\tilde{\eta} \circ \rho_i^{-1} = \eta_i^{n_i}]\!]$, we conclude, that $t_i^{n_i} \in [\![\phi(\eta_i^{n_i}, n_i)]\!]$, a contradiction to how the t_m^n were constructed. ■

Conclusion

In this thesis we have analyzed and compared three different kinds of set theories: Brouwerian set theory, Martin-Löf's set theory **ML** and the constructive axiomatic set theories **IZF** and **CZF**.

Being explicitly based on paradigms from programming, **ML** seems to be the best suited for applications in programming languages and automated proof assistants. Various implementations like Agda, Idris, NuPRL and Coq have been given in the past (see, for example, [2], [33], [39], [BO]).

There are two main ideas behind the axiomatic set theories **IZF** and **CZF**: First, they are based on the most common and widely accepted formalization of mathematics, namely **ZFC**. Hence, their usage is oriented towards existing everyday mathematical practice. Second, the two theories are tailored to meet different levels of constructive demand. As discussed in section 2.4, the rather minimalistic requirement for **IZF** is that it does not go beyond the realm of intuitionistic logic, i.e. **IZF** $\nvdash$ **LEM**. The ultimate confirmation that this is indeed so, is given only in chapter 4 via the metamathematical tool of realizability.

The motivation behind **CZF** to further restrict **IZF** is to end up with a theory that is predicative. Giving a justification that **CZF** is indeed fully predicative, however, is more intricate. Schütte and Feferman have developed a proof-theoretical analysis of the notion of predicativity (see [23]), but presenting it would have been a thesis on its own. Instead we have justified **CZF** in chapter 5, by giving a meaning-preserving interpretation of **CZF** into **ML** – a theory that is considered to give a constructively clear and well-justified notion of sets.

Another result of chapter 4 was that both **CZF** and **IZF** possess the disjunction- and numerical existence property – a feature that it certainly expected from constructive theories. Anyways, it has been shown by more elaborate metamathematical tools, that the stronger existence property does not hold for **IZF** and **CZF** (see [*25*], [*63*]). This may raise doubts for the aptness of **IZF** and **CZF** as constructive set theories. At least in the case of **CZF** this doubt is cleared up by the aforementioned interpretation of **CZF** in **ML**, where witnesses can be constructed.

A further interesting question is which mathematical loss we have when passing form **ZFC** to the weaker theories **IZF** and **CZF**. Since **LEM** is not derivable in **IZF**, it follows that all weak counterexamples of section 2.4.2 are actually underivable statements in **IZF**. For example, **IZF** $\nvdash$ Foundation. As a more interesting, not purely set theoretical result, we have shown in chapter 6, that **IZF** $\nvdash \mathrm{BI}_D$. As for **CZF**, we have shown, relying on the result that **CZF** is compatible with **GUP**, that the powerset axiom does not hold true in **CZF** in section 2.4.3. The implications are far-reaching: The V_αs in the hierarchical structure of the universe of all sets V are not sets any more. The mathematical discipline of topology has to be dealt with in a different way in **CZF**. Instead of topological spaces one passes to point-free topologies, thus preserving many results from classical topology relying on the set-character of topological spaces and especially the axiom of choice, like Tychonoff's theorem. It turns out however, that existence of the Stone-Cech compactification is not always guaranteed in **CZF** (see, for example [*19*], [*20*]).

In section 2.2, we have discussed some basic concepts of Brouwerian set theory, enough to understand Brouwer's justification of the continuum and his fundamentally different approach to continuous functions. Brouwer's set theory goes far beyond what we have sketched in this thesis: In [*7*], he develops concepts like ordinals and cardinals. It would be interesting to see how his notions differ from their vis-à-vis in **IZF** and **ZFC**.

We have shown in chapter 6, that Brouwerian analysis is, in principle, compatible to **IZF**. The principle of bar induction BI_D turns out to be to strong an assumption for Brouwer's proof of the fan theorem. This result, showing that **IZF** may be equiconsistently extended with the fan theorem, Brouwer's continuity principle and $\neg\mathrm{BI}_D$ has been proved by Ščedrov in [*59*]. Another result of this paper that we have not dealt with in this thesis is, that the same holds true for BI_D in place of $\neg\mathrm{BI}_D$ making Bar induction independent from Brouwerian analysis based on **IZF**.

Another topic that we touched only marginally is that of choice principles in constructive set theories. The classical axiom of choice **AC** plays very different roles in the set theories at hand. This is due to the different ways to interpret its meaning. While **AC** is a weak counterexample and hence extends **IZF** to full **ZFC** (section 2.4.2), it is an easy theorem in **ML** (section 5.1.9). Further choice principles are the axioms of countable choice, dependent choice **DC**, the representation axiom and the regular extension axiom **REA** (for an overview, see [*53*]). Actually, interpretation of **CZF** into **ML** works as well for **CZF** + **DC**, and our realizability structure for **IZF** and **CZF** from section 3.3 works for **IZF** + **REA** and **CZF** + **REA**, too (see

[*1*], [*55*]). Hence both extensions also enjoy the metamathematical properties and compatibilities with principles we discussed in chapter 4.

References

[1] P. Aczel, "The type theoretic interpretation of contructive set theory," in *Logic Colloquim 77*, A. Macintyre, L. Pacholski, and J.B. Paris, Eds. Amsterdam: North Holland Publishing Company, 1978, pp. 55-65.

[2] AgdaWiki. Agda. [Online]. Retrieved: 02.12.2019. https://wiki.portal.chalmers.se/agda/pmwiki.php?n=Main.HomePage

[3] I.H. Anellis, "The first Russel paradox," in *Perspectives on the History of Mathematical Logic*, T. Drucker, Ed. Cambridge: Birkäuser, 1993.

[4] M.J. Beeson, *Foundations of Constructive Mathematics: Metamathematical Studies*. Berlin: Springer, 1985.

[5] J.L. Bell, *Boolean-valued Models and Independence Proofs*. Oxford: Clarendon, 2005.

[6] E. Bishop, *Foundations of Constructive Analysis*. New York: Mcgraw-Hill, 1967.

[7] L.E.J. Brouwer, "Begründung der Mengenlehre unabhängig von Satz vom ausgeschlossenen Dritten. Erster Teil: Allgemeine Mengenlehre," in *L.E.J. Brouwer - Collected Works*, A. Heyting, Ed. Amsterdam: North-Holland, 1918.

[8] L.E.J. Brouwer, *Brouwer's Cambridge Lectures on Intuitionism*, D. van Dalen, Ed. Cambridge: Cambridge University Press, 1981.

[9] L.E.J. Brouwer, "Intuitionism and Formalism," *Bulletin of the American Mathematical Society*, no. 20, pp. 81-96, 1913.

[10] L.E.J. Brouwer, "On the foundations of mathematics (PhD thesis)," in *L.E.J. Brouwer - Collected Works*, A. Heyting, Ed. Amsterdam, 1907.

[11] L.E.J. Brouwer, "Über Definitionsbereiche von Funktionen," in *Brouwer: Collected Works*, A. Heyting, Ed. Amsterdam: North-Holland, 1927.

[12] C. Burali-Forti, "Una questione sui numeri transfiniti," pp. 154-164, 1897.

[13] G. Cantor, "Beiträge zur Begründung der transfiniten Mengenlehre," *Mathematische Annalen*, vol. 46, no. 4, pp. 481-512, 1895.

[14] G. Cantor, *Contributions to the founding of the theory of transfinite numbers*. New York: Dover Publications, 1897.

[15] G. Cantor, "letter to Hilbert," in *Georg Cantor: Briefe*, Meschkowski H. and Nilson W., Eds. Berlin: Springer, 1991, p. 388.

[16] P. Cohen, "The Independence of the Continuum Hypothesis," *Proceedings of the National Acadamy of Sciences of the United States*, vol. 50, no. 6, pp. 1143-1148, 1963.

[17] I. Copi, "The Burali-Forti Paradox," *Philosophy of Science*, vol. 4, pp. 281-286, 1958.

[18] Coq Reference Manual. Credits. [Online]. Retrieved: 02.12.2019. https://coq.inria.fr/distrib/current/refman/credits.html?highlight=l%C3%B6f

[19] T. Coquand, "Compact spaces and distributive lattices," *Journal of Pure and Applied Algebra*, no. 184, pp. 86-100, 2003.

[20] G. Curi, "On the existence of Stone-Čech compactification," *Journal of Symbolic Logic*, vol. 75, no. No. 4, pp. 1137-1146, 2010.

[21] R. Diaconescu, "Axiom of choice and complementation," *Proceedings of the American Mathematical Society*, 1975.

[22] M. Džamonja, "Set Theory and its Place in the Foundations of Mathematics: A New Look at and Old Question," *Journal of Indian Council of Philosophical Research*, vol. 34, pp. 415-424, 2017.

[23] S. Feferman, "Predicativity," in *The Oxford Handbook of Philosophy of Mathematics and Logic*.: Oxford University Press, 2005.

[24] S. Feferman, "Systems of predicative analysis," *Journal of Symbolic Logic*, no. 29, pp. 1-30, 1964.

[25] H. Friedman and A. Ščedrov, "Set existence property for intuitionistic theories with dependent choice," *Annals of Pure and Applied Logic*, vol. 25, no. 2, pp. 129-140, 1983.

[26] A. O. Gelfond, "Sur le septième problème de Hilbert," *Bulletin de l'Académie des Sciences de l'URSS. Classe des sciences mathématiques et naturelles*, no. 4, pp. 623-634, 1934.

[27] G. Gonthier, "Formal proof - the four-color theorem," *Notices of the AMS*, vol. 55, no. 11, pp. 1382-1393, 2008.

[28] R.J. Grayson, "Heyting-valued Semantics," in *Logic Colloquium '82 : Proceedings of the colloquium held in Florence, 23 - 28 August, 1982*. 1984, Amsterdam: North-Holland, 1984, pp. 181-208.

[29] A. Heyting, "Die intuitionistische Grundlegung der Mathematik," in *Erkenntnis*., 1931, vol. 2, pp. 106-115.

[30] A. Heyting, *Mathematische Grundlagenforschung. Intuitionismus. Beweistheorie*. Berlin: Springer, 1943.

[31] A. Heyting, "Sur la logique intuitionniste," *Academie Royale de Belgique. Bulletin de la Classe des Sciences*, vol. 16, pp. 957-963, 1930.

[32] J.R. & Seldin, J.P. Hindley, *Lambda-calculus and combinators: An introduction*. Cambridge: Cambridge University Press, 2008.

[33] Idris Wiki. Home. [Online]. Retrieved: 02.12.2019. https://github.com/idris-lang/Idris-dev/wiki

[34] R. Iemhoff, "Intuitionism in the Philosophy of Mathematics," *The Stanford Encyclopedia of Philosophy (Winter 2019 Edition)*, 2008. [Online]. Retrieved: 02.12.2019. https://plato.stanford.edu/archives/win2019/entries/intuitionism/

[35] K. Jänich, *Topologie*, 8th ed. Berlin: Springer, 2005.

[36] S.C. Kleene, "λ-Definability and Recursiveness," *Duke Mathematical Journal*, pp. 340-353, 1936.

[37] A.N. Kolmogorov, "Zur Deutung der intuitionistischen Logik," *Mathematische Zeitschrift*, pp. 58-65, 1932.

[38] J. König, "Über die Grundlagen der Mengenlehre und das Kontinuumsproblem," *Mathematische Annalen*, pp. 156–160, 1905.

[39] Christoph Kreitz. The Nuprl Proof Development System, Version 5 - Reference Manual and User's Guide. [Online]. Retrieved: 02.12.2019. http://www.nuprl.org/html/02cucs-NuprlManual.pdf

[40] K. Kunen, *Set Theory*. London: College Publications, 2011.

[41] B.A. Kushner, "The Constructive Mathematics of A. Markov," *American Mathematical Monthly*, vol. 1113, no. 6, pp. 559-566, 2006.

[42] P. Martin-Löf, "Constructive mathematics and computer programming," in *Logic, methodology and philosophy of science VI, Proceedings of the 1979 international congress at Hannover, Germany*, L.J. Cohen et al., Eds. Amsterdam: North-Holland Publishing Company, 1982, pp. 153-175.

[43] P. Martin-Löf, "Intuitionistic Type Theory : Notes by Giovanni Sambin of a Series of Lectures given in Padua, June 1980," *Studies in Proof Theory*, 1984.

[44] D.C. McCarty, "Realizability and Recursive Mathematics (PhD thesis)," 1984.

[45] D.C. McCarty, "Realizability and recursive set theory," *Annals of Pure and Applied Logic*, pp. 153-183, 1986.

[46] P.J. Morandi. (2005) Dualities in Lattice Theory. [Online]. Retrieved: 02.12.2019. http://sierra.nmsu.edu/morandi/notes/Duality.pdf

[47] J. Munkres, *Topology: A first course*. Englewood Cliffs, NJ: Prentice-Hall, 1975.

[48] J. Myhill, "Constructive Set Theory," *The Journal of Symbolic Logic*, vol. 3, pp. 347-382, 1975.

[49] J. Myhill, "The undefinability of the set of natural numbers in ramified Principia," in *Bertrand Russel's Philosophy*, G. Nakhnikian, Ed. London: Duckworth, 1974, pp. 19-27.

[50] I. Petrakis, "Brouwer's Fan Theorem (Diploma Thesis)," 2010.

[51] H. Poincaré, "Les mathématiques et la logique," *Revue de métaphysique et de morale*, pp. 294-317, 1906.

[52] H. Poincaré, B. Russel, E. Zermelo, and G. Peano, "Les mathématiques et la logique," in *Textes de la discussion (1906-1912) sur les fondements des math´ematiques: des antinomies à la prédicativité*, G. Heinzmann, Ed. Paris: Albert Blanchard, 1986.

[53] M. Rahtjen, "Choice principles in constructive and classical set theories," in *Logic Colloquium '02*, Z. Chatzidakis, P. Koepke, and W. Pohlers, Eds. Cambridge: Cambridge University Press, 2017, pp. 299-326.

[54] M. Rathjen, "Realizability for Constructive Zermelo-Fraenkel set theory," *Lecture Notes in Logic*, 2004.

[55] M. Rathjen, "The Disjunction and Related Properties for Constructive Zermelo-Fraenkel Set Theory," *The Journal of Symbolic Logic*, vol. 4, pp. 233-254, 2005.

[56] J. Richard, "Les Principes des Mathématiques et le Problème des Ensembles," *Revue Générale des Sciences Pures et Appliquées*, 1905.

[57] B. Russel, "Mathematical logic as based on the theory of types," *American Journal of Mathematics*, pp. 222-262, 1908.

[58] B. Russel, *The Principles of Mathematics*. Cambridge: University Press, 1903.

[59] A. Ščedrov, "Consistency and independence results in intuitionistic set theory," in *Constructive mathematics : proceedings of the New Mexico State University conference, held at Las Cruces, New Mexico, August 11 - 15, 1980*. Berlin: Springer, 1981, pp. 54-86.

[60] K. Schütte, "Eine Grenze für die Beweisbarkeit der Transfiniten Induktion in der verzweigten Typenlogik," *Archiv für Logik und Mathematische Grundlagenforschung*, no. 7, pp. 45-60, 1965.

[61] K. Schütte, "Predicative well-orderings," in *Formal Systems and Recursive Functions*, J. & Dummet, M. Crossley, Ed. Amsterdam: North-Holland, 1965, pp. 279-302.

[62] van W.P. Stigt, *Brouwer's Intuitionism*. Amsterdam: North-Holland, 1990.

[63] A.W. Swan, "CZF does not have the existence property," *Annals of Pure and Applied Logic*, vol. 165, no. 5, pp. 1115-1147, 2014.

[64] A.S. & van Dalen, D. Trolestra, *Constructivism in Mathematics*. Amsterdam: North-Holland, 1988, vol. 1.

[65] A.S. Trolestra, "History of constructivism in the 20th century," in *Set Theory, Arithmetic, and Foundations of Mathematics: Theorems, Philosophies*. Cambridge: Cambridge University Press, 2011, pp. 150-179.

[66] A.S Trolestra and D. van Dalen, *Constructivism in Mathematics: An Introduction*. Amsterdam: North-Holland, 1988, vol. 1.

[67] A.M. Turing, "Computability and λ-Definability," *The journal of symbolic logic*, vol. 2, no. 4, pp. 153-163, 1937.

[68] A.M. Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem," *Proceedings of the London Mathematical Society*, vol. 1, pp. 230-265, 1937.

[69] S. Valentini, "A simple proof of the Completeness Theorem of the Intuitionistic Predicate Calculus with Respect to the Topological Semantics," 2004.

[70] M. van Atten, *On Brouwer*. Singapore: Thomson Wadsworth, 2004.

[71] B. Van Den Berg and I. Moerdijk, "Aspects of predicative algebraic set theory, II: Realizability," *Theoretical Computer Science*, vol. 412, no. 20, pp. 1916-1940, 2011.

[72] H. Weyl, *Das Kontinuum : Kritische Untersuchungen über die Grundlagen der Analysis*. Leipzig: Veit, 1918.

Index

- $\{\mathfrak{u}, \mathfrak{v}\}^{(H)}$ 102
- +-Rules 74
- basis (of a topology) 93
- BHK-interpretation 14
- $\mathrm{BI_D}$ (decidable Bar induction) 87
- Boolean algebra 89
- Bounded separation schema 34
- Brouwer 15
- choice sequence 17
- class 35
- clopen sets 95
- closed sets 95
- $\mathtt{Col}_{\phi(x,y)}(\mathfrak{u})$ 105
- Collection schema 34
- Constructivism 7
- CP (Church's principle) 53
- CR (Church's rule) 52
- DET 33
- Disjunction property 50
- EI 33
- Empty set, axiom of 24, 34
- existence property 51
- Extensionality axiom 23, 34
- Foundation axiom 25, 34
- Heyting-algebra 89
- Infinity, axiom of 24, 34
- Intuitionism 15
- I-rules 75
- MP (Markov's principle) 53
- MR (Markov's rule) 52
- $\mathbb{N}_k$-rules 76
- $\mathbb{N}$-rules 77
- numerical existence property 51
- open sets 93
- ordinal 36
- Pair, axiom of 23, 34
- $\mathtt{Pow}(\mathfrak{a})$ 104
- Powerset axiom 25, 34
- Replacement schema 25, 34
- $\mathtt{Sep}_{\phi(x)}(\mathfrak{a})$ 104
- Separation schema 24, 34
- Set induction 34
- species 16
- spread 16
- Strong collection schema 34
- Subset collection 34
- topology 93
- UG 33
- $\mathtt{Un}(\mathfrak{A})$ 102
- Union, axiom of 24, 34
- UP (Uniformity principle) 53
- UR (Uniformity rule) 52
- UzP (Unzerlegbarkeits principle) 53
- UzR (Unzerlegbarkeits-rule) 51
- Π-rules 71
- Σ-rules 73
- **AC** (axiom of choice) 26
- **BST** 35
- **CZF** 35
- **F** (fan theorem) 87
- **HPL** (Heyting's predicate logic) 33
- **IZF** 35
- **LEM** (law of excluded middle) 7
- **ML** (Martin-Löf's set theory 19, 70
- **ZF** 35
- Kl 49
- $U_{u,\lambda}$ 94
- U-rules 80
- $V^{(H)}$ 97
- V^* 54
- $\hat{\mathfrak{x}}$ 103
- $\mathcal{W}$-rules 79
- $\mathfrak{F}$ 109
- $\mathbb{ON}$ (class of ordinals) 36