Broad Infinity and Generation Principles (Part 1: classical set theory)

Paul Blain Levy, University of Birmingham

August 2, 2023

Abstract

We introduce Broad Infinity, a new set-theoretic axiom scheme that may be considered plausible. It states that three-dimensional trees whose growth is controlled by a specified class function form a set; these trees are called "broad numbers".

Assuming the axiom of choice, or at least the weaker principle known as WISC, Broad Infinity is equivalent to Mahlo's principle: the class of all regular limits is stationary. It also leads to a convenient principle for generating a subset of a class using a "rubric" (family of rules). This directly gives the existence of Grothendieck universes, without requiring a detour via ordinals.

In the absence of choice, Broad Infinity implies that the derivations of elements from a rubric form a set. This yields the existence of Tarski-style universes.

Additionally, the paper reveals a pattern of resemblance between "Broad" principles, that go beyond ZFC, and "Wide" principles , that are provable in ZFC.

Contents

1 Introduction: Broad Infinity and its motivation **3**

1.1 Broad Infinity in a nutshell 3
1.2 Regular limits and stationary classes 4
1.3 Two principles from the literature 4
1.4 Limitations of Mahlo's principle 4
1.5 Motivating Broad Infinity 5

2 Goals and structure of the paper **5**

2.1 Plausible vs useful 5
2.2 Urelements and non-well-founded membership 6
2.3 Weak choice principle 6
2.4 Wide vs Broad 6
2.5 Summary of goals 6

2.6 Related work . . . 8
2.7 Structure of paper . . . 8

3 Basic theory of sets and well-foundedness **9**
3.1 Our base theory . . . 9
3.2 Classes, functions and partial functions . . . 10
3.3 Additional axioms . . . 11
3.4 Fixpoints . . . 12
3.5 Initial algebras . . . 12
3.6 Schematicity and predicativity . . . 13
3.7 Natural numbers . . . 13
3.8 Set-based relations . . . 14
3.9 Set-based well-founded relations . . . 15

4 Scaffolds **17**
4.1 Generating a subclass . . . 17
4.2 Generating a partial function . . . 19
4.3 Introspective scaffolds . . . 19

5 Wide and Broad Infinity **20**
5.1 Wide Infinity . . . 20
5.2 Broad Infinity . . . 23

6 Rubrics **26**
6.1 Generating a subset . . . 26
6.2 Application: Grothendieck universes . . . 29
6.3 Derivations . . . 29
6.4 Application: Tarski-style universes . . . 31
6.5 Proving the Derivation Set principles . . . 33
6.6 Injective rubrics . . . 34
6.7 Quasiwide rubrics . . . 36
6.8 Powerset and rubric functions . . . 38

7 The use of choice **39**
7.1 Consequences of AC . . . 39
7.2 WISC principles . . . 40
7.3 Sufficiency of WISC . . . 42

8 Basic theory of ordinals **44**
8.1 Set-based well-orderings . . . 44
8.2 Ordinals . . . 48

9 From rubrics to sup-generation **49**
9.1 Inductive chains . . . 49
9.2 Sup-generation . . . 49

10 Lindenbaum numbers **52**

11 From sup-generation to Mahlo's principle **54**
11.1 Unbounded and stationary classes 54
11.2 Cofinality 55
11.3 Regular limits 56
11.4 Blass's axiom and Mahlo's principle 56

12 The power of stationarity **57**
12.1 Club classes and continuous ordinal functions 57
12.2 Application: iterated inaccessibility 62

13 Conclusions **63**
13.1 Summary of achievements 63
13.2 Further work 64

Bibliography **64**

1 Introduction: Broad Infinity and its motivation

1.1 Broad Infinity in a nutshell

This paper is about a new axiom scheme of set theory, which is very easy to state.

First, some preliminaries. For the sake of this introduction, assume either ZF or ZFA—the latter is a variant that allows urelements (atoms). The axiom of choice (AC) is not assumed. We write $\mathfrak{T}$ for the class of all things (universal class), and $\mathfrak{S}$ for the class of all sets. In ZF, they are the same.

We define the ordered pair $\langle x, y\rangle \stackrel{\text{def}}{=} \{\{x\},\{x,y\}\}$, and note that $\langle x,y\rangle = \langle x',y'\rangle$ implies $x = x'$ and $y = y'$. We likewise define $\mathsf{Nothing} \stackrel{\text{def}}{=} \emptyset$ and $\mathsf{Just}(x) \stackrel{\text{def}}{=} \{x\}$, and note that $\mathsf{Just}(x) \neq \mathsf{Nothing}$, and that $\mathsf{Just}(x) = \mathsf{Just}(x')$ implies $x = x'$.

For a set K, a *K-tuple* is a function on K. It is written $[x_k]_{k\in K}$ and imagined as a column with K entries. It is *within* a given class C when $x_k \in C$ for all $k \in K$.

Recall Zermelo's *Axiom of Infinity* [Zer08], which is included in ZF and ZFA. It says that there is a set X that contains $\mathsf{Nothing}$ and, for any $x \in X$, contains $\mathsf{Just}(x)$.

The new axiom scheme of *Simple Broad Infinity* is similar. It says that, for any function $F : \mathfrak{T} \to \mathfrak{S}$, there is a set X that contains $\mathsf{Nothing}$ and, for any $x \in X$ and Fx-tuple y within X, contains $\mathsf{Just}\langle x, y\rangle$.

ZF extended with this scheme is called *Broad ZF*. In the following sections I will motivate Simple Broad Infinity in light of a previously studied principle.

1.2 Regular limits and stationary classes

To begin, we reprise some useful notions concerning ordinals. We write Ord for the class of all ordinals and Lim for the class of all *limits*—ordinals that are neither 0 nor a successor. An *initial* ordinal is one that is not the range of a function from a smaller ordinal; examples are the finite ordinals, ω and ω_1.

A limit κ is *regular* when, for all $\alpha < \kappa$, the supremum function $\mathsf{Ord}^\alpha \to \mathsf{Ord}$ restricts to a function $\kappa^\alpha \to \kappa$. (See Section 11.3 for an alternative definition.) It follows that κ is initial, so ω is the only regular limit that is countable. We write Reg for the class of all regular limits.

For a function $F : \mathsf{Ord} \to \mathsf{Ord}$, we say that an limit λ is *F-closed* when F restricts to a function $\beta \to \beta$. Here are some examples:

- Let S be the successor function. Every limit is S-closed.
- For an ordinal α, let Const_α be the constant function $\gamma \mapsto \alpha$. A limit is Const_α-closed iff greater than α.
- For functions $F, G : \mathsf{Ord} \to \mathsf{Ord}$, let $F \vee G$ be the pointwise maximum $\gamma \mapsto F(\gamma) \vee G(\gamma)$. A limit is $(F \vee G)$-closed iff both F-closed and G-closed.

A class of limits D is *stationary* when, for every function $F : \mathsf{Ord} \to \mathsf{Ord}$, it contains an F-closed ordinal. (See Sections 11.1 and 12 for alternative definitions.) It follows that D is unbounded and, for every function $F : \mathsf{Ord} \to \mathsf{Ord}$, contains stationarily many F-closed limits.

1.3 Two principles from the literature

Next we look at two principles that use the above notions.

- *Mahlo's principle*, also known as "Ord is Mahlo", says that Reg is stationary [Ham03, Jor70, Le60, May00, Wan77]. To illustrate its power, note that ZFC + Mahlo's principle proves that there are stationarily many inaccessible cardinals. That is because, in ZFC, an inaccessible cardinal is precisely an uncountable F-closed limit, where F sends α to 2^α if α is a cardinal and to 0 otherwise.
- *Blass's axiom* [Bla83] says merely that Reg is unbounded. It follows from AC, but is it provable in ZF alone? To answer this question, Gitik [Git80] showed that, if ZFC with arbitrarily large strongly compact cardinals is consistent, then ZF cannot even prove the existence of an uncountable regular limit, let alone Blass's axiom.

1.4 Limitations of Mahlo's principle

Appealing though Mahlo's principle may be, I consider it deficient as an axiom scheme, in two respects.

Firstly, it does not meet the ZF standard of simplicity. Each ZF axiom, other than Foundation, expresses the idea that some easily grasped things form a set: the natural numbers (Infinity), the subsets of a set (Powerset), the elements of a set that satisfy a property (Separation), the images of a set's elements under a function (Replacement), and so forth. This is what makes these axioms so compelling. But Mahlo's principle does not do this.

The second problem is that Mahlo's principle, or indeed any addition to ZF that implies the existence of an uncountable regular limit, seems to be *entangled with choice* in light of Gitik's result. Admittedly this view is contentious, as some people would try to justify Mahlo's principle via the following choiceless argument: "For any $J : \mathsf{Ord} \to \mathsf{Ord}$, the property of being a J-closed regular limit can be reflected down from Ord to an ordinal." But such thinking is avoided in this paper.

1.5 Motivating Broad Infinity

In light of the preceding discussion, my primary goal was to obtain an axiom scheme that

1. is equivalent to Mahlo's principle, assuming AC
2. asserts that some easily grasped things form a set
3. does not imply (without AC) that an uncountable regular limit exists.

To this end, I propose Simple Broad Infinity. Does it meet the requirements?

1. Assuming AC, we shall prove that Simple Broad Infinity is equivalent to Mahlo's Principle. So this requirement is met.
2. Simple Broad Infinity asserts, for each function $F : \mathfrak{T} \to \mathfrak{S}$, that the class of F-broad numbers (explained below) forms a set. Arguably this is "easily grasped", but the question is subjective and must be left to the reader's judgement.
3. I see no way to obtain the existence of an uncountable regular limit in Broad ZF. However, an analogue of Gitik's result is lacking.

2 Goals and structure of the paper

2.1 Plausible vs useful

Simple Broad Infinity has been designed to be as *plausible* as possible. In other words, I aimed to minimize the mental effort needed to believe it. This is surely a desirable feature for an axiom scheme. Furthermore, disentanglement from choice is helpful in achieving it. That is because, even for a person who finds AC intuitively convincing (as I do), it is easier to accept one intuition at a time.

My second goal was different: to find an equivalent scheme that is as *useful* as possible. In other words, I wanted to minimize the effort needed to apply it. In particular, assuming AC, it should *obviously* imply the existence of Grothendieck universes, without requiring a detour via notions of ordinal or cardinal.[1]

To this end, I propose a scheme called *Broad Set Generation*. For people who accept AC, this meets the stated goal. For those who do not, I offer instead a principle called *Broad Derivation Set*. The latter yields the existence of "Tarski-style" universes that are sometimes used in the literature [ML84].

2.2 Urelements and non-well-founded membership

In ZF, everything is a set and the membership relation is well-founded. But our results also hold in variants of these theories that allow urelements and/or non-well-founded membership. Making this clear is the third goal.

2.3 Weak choice principle

Although—as stated above—some of our results depend on AC, the full strength of this axiom is not needed. More precisely, a weak form of choice known as WISC (Weakly Initial Set of Covers) is sufficient to give our results. Explaining this fact is the fourth goal.

Caveat: we shall see different versions of WISC, and care must be taken to use an appropriate one. In ZF, they are all equivalent.

2.4 Wide vs Broad

We give the name "Broad" to the principles studied in this paper that go beyond ZFC. It turns out that each of them has a ZFC-provable counterpart that we call "Wide". For example, Mahlo's principle is Broad, and its Wide counterpart is Blass's axiom.

The fifth goal is to convey this pattern of resemblance, which is depicted in Figure 1, a summary of the results in the paper (using a base theory weaker than ZF). The rows within each block are equivalent, and each arrow represents inclusion of theories, i.e., reverse implication. The Wide principles appear on the left and the corresponding Broad principles on the right.

2.5 Summary of goals

To summarize the previous sections, our goals are as follows.

1. To give a simple and plausible axiom scheme, disentangled from choice, that is equivalent over ZFC to Mahlo's principle. **Solution** Simple Broad Infinity.

[1] In the absence of AC, there is no accepted notion of inaccessible cardinal. See [BDL07] for a comparative analysis.

Without assuming the Axiom of Choice

$$\begin{tikzcd}
\begin{array}{l}\text{Powerset} + \text{Infinity}\\ \text{Wide Infinity*}\\ \text{Wide Derivation Set†}\\ \text{Injective Wide Subset Generation†}\\ \text{Derivation Set on a Set}\end{array} \arrow[r] \arrow[d] & \begin{array}{l}\text{Broad Infinity*}\\ \text{Broad Derivation Set}\\ \text{Injective Broad Subset Generation}\end{array} \arrow[d] \\
\begin{array}{l}\text{Powerset} + \text{Wide Sup-generation*}\\ \text{Powerset} + \text{Blass's Axiom}\\ \text{Wide Subset Generation†}\end{array} \arrow[r] & \begin{array}{l}\text{Powerset} + \text{Broad Sup-generation*}\\ \text{Powerset} + \text{Mahlo's Principle}\\ \text{Broad Subset Generation}\end{array}
\end{tikzcd}$$

Assuming the Axiom of Choice or at least WISC

$$\begin{tikzcd}
\begin{array}{l}\text{Powerset} + \text{Infinity}\\ \text{Wide Infinity*}\\ \text{Wide Derivation Set†}\\ \text{Powerset} + \text{Wide Sup-generation*}\\ \text{Powerset} + \text{Blass's Axiom}\\ \text{Wide Subset Generation†}\\ \text{Derivation Set on a Set}\end{array} \arrow[r] & \begin{array}{l}\text{Broad Infinity*}\\ \text{Broad Derivation Set}\\ \text{Powerset} + \text{Broad Sup-generation*}\\ \text{Powerset} + \text{Mahlo's Principle}\\ \text{Broad Subset Generation}\end{array}
\end{tikzcd}$$

* The Simple and Full versions are equivalent.
† The Wide and Quasiwide versions are equivalent.

Figure 1: Diagram of theories, extending the base theory

2. To give an equivalent principle that is convenient for applications. **Solution** Broad Subset Generation for those who accept AC, and Broad Derivation Set for those who do not.

3. To show that our results hold even when urelements and non-well-founded membership are allowed.

4. To show that, for the results that rely on AC, a weak choice principle suffices.

5. To convey the resemblance between Wide principles (which are provable in ZFC) and Broad principles (which are not, provided ZF is consistent).

2.6 Related work

Many formulations of Mahlo's principle have been studied [Jor70, May00, Le60, Mon62, Dow11], and similar principles have been given for type theory [Rat00, Set00] and Explicit Mathematics [KS10]. Other principles have been shown to be equiconsistent with Ord-is-Mahlo [Ham03, Mat77].

Another related topic is the "induction-recursion" principle [DS06, GH16] used in type theory and the proof assistant Agda. It allows the formation of Tarski-style universes, as in Section 6.4, and is modelled in [DS06] using a Mahlo cardinal. Although it is similar to the Broad Derivation Set principle introduced below, the different setting makes it hard to draw a precise comparision.

2.7 Structure of paper

The paper has two main parts: one that is about sets, and one that is about ordinals. Each begins with an introductory section that reviews standard material, but with an emphasis on classes.

The first part begins in Section 3 with an introduction to classes and well-founded relations. Section 4 gives a way to generate subclasses and partial functions that is used throughout the paper, and especially to formulate the Derivation Set principles. Then Section 5 gives the Wide and Broad Infinity principles, followed in Section 6 by the more practical principles of Subset Generation and Derivation Set. In Section 7 we see how AC, or just a weak version, allows us to show the Subset Generation principles equivalent to the Infinity principles.

The second part begins in Section 8 with an introduction to well-orders and ordinals. Section 9 gives "sup-generation" principles that involve both sets and ordinals. Section 10 explains the concept of Lindenbaum numbers that has often appeared in the literature, and then Section 11 develops Mahlo's principle and establishes all its relationships. Lastly, Section 12 presents the traditional use of Mahlo's principle to prove the existence of various kinds of ordinal.

Lastly we wrap up in Section 13 by summarizing the contributions and suggesting further work.

Some readers may just want to see the ZFC proof that Simple Broad Infinity is equivalent to Mahlo's principle. This is divided into several steps:

- Simple Broad Infinity is equivalent to Full Broad Infinity—Proposition 5.4(1).
- Full Broad Infinity implies Broad Derivation Set—Proposition 6.17(2).
- Broad Derivation Set implies Broad Subset Generation—Proposition 7.1(2). Only this step uses AC.
- Broad Subset Generation implies Full Broad Infinity—Proposition 6.7(2).
- Broad Subset Generation is equivalent to Broad Sup-generation—Proposition 9.7(2).
- Broad Sup-generation is equivalent to Mahlo's principle—Proposition 11.11(2).
- Various definitions of stationarity, each giving a different formulation of Mahlo's principle, are equivalent—Proposition 12.9.

3 Basic theory of sets and well-foundedness

3.1 Our base theory

For this paper, I have chosen a base theory that differs from ZF in several ways:

- It allows urelements and non-well-founded membership.
- It excludes Powerset and Infinity, so that we can examine how these axioms relate to other principles.
- It allows undefined (unary) predicate symbols, also known as class variables.

For a given set Pred of predicate symbols, the syntax is as follows:

$$\begin{array}{rcl}\phi, \psi & ::= & P(x) \mid \mathsf{IsSet}(x) \mid x \in y \mid x = y \mid \mathsf{True} \mid \mathsf{False} \\ & & \mid \neg\phi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \phi \Rightarrow \psi \mid \forall x.\, \phi \mid \exists x.\, \phi\end{array}$$

where $P \in \mathsf{Pred}$. The formula $\mathsf{IsSet}(x)$ asserts that x is a set.

We define the *base theory* over Pred to be the classical first-order theory with equality, axiomatized as follows.

- Axiom of *Extensionality*: Any two sets with the same elements are equal.
- Axiom of *Inhabitation*: Anything that has an element is a set.
- Axiom scheme of *Replacement*: For any set A and binary predicate F such that each $x \in A$ has a unique F-image $F(x)$, there is a set $\{F(x) \mid x \in A\}$ of all F-images of elements of A.
- Axiom of *Empty Set*: There is a set $\emptyset$ with no elements.

- Axiom of *Singleton*: There is a set $\{\emptyset\}$ whose sole element is $\emptyset$.
- Axiom of *Doubleton*: There is a set $\{\emptyset, \{\emptyset\}\}$ whose sole elements are $\emptyset$ and $\{\emptyset\}$.
- Axiom of *Union Set*: For any set of sets $\mathcal{A}$, there s is a set $\bigcup \mathcal{A}$ of all elements of elements of $\mathcal{A}$.

Pairing and Separation follow via

$$\begin{array}{rcl} \{x_0, x_1\} & \stackrel{\text{def}}{=} & \{x_i \mid i \in \{0,1\}\} \\ \{x \in A \mid P(x)\} & \stackrel{\text{def}}{=} & \bigcup_{x \in A} \begin{cases} \{x\} & \text{if } P(x) \\ \emptyset & \text{otherwise} \end{cases} \end{array}$$

Henceforth we assume the base theory.

3.2 Classes, functions and partial functions

Since classes are so important in our story, much of this paper is devoted to studying them. As usual in set theory, a class is represented as a predicate formula with parameters. We write $\mathfrak{S}$ for the class of all sets, and Ur for that of all urelements (things that are not sets).

Given classes A and B, we write

$$\begin{array}{rcl} A \times B & \stackrel{\text{def}}{=} & \{\langle x, y \rangle \mid x \in A, y \in B\} \\ A + B & \stackrel{\text{def}}{=} & \{\mathsf{inl}\ x \mid x \in A\} \cup \{\mathsf{inr}\ y \mid y \in B\} \end{array}$$

where $\langle x, y \rangle \stackrel{\text{def}}{=} \{\{x\}, \{x, y\}\}$ and $\mathsf{inl}\ x \stackrel{\text{def}}{=} \langle 0, x \rangle$ and $\mathsf{inr}\ y \stackrel{\text{def}}{=} \langle 1, y \rangle$. The notation $A \subseteq B$ means that A is included in (i.e., a subclass of) B. The notation $F : A \to B$ means that F is a function on A whose range is included in B.

For a function F on a class A, the restriction of F to a subclass C of A is written $F \upharpoonright_C$.

A *family* consists of a set I and a function $x : I \to \mathfrak{T}$. More generally, a *class-family* consists of a class I and function $x : I \to \mathfrak{T}$. It may be written as (I, x) or as $(x_i)_{i \in I}$. It is *within* a class C when, for all $i \in I$, we have $x_i \in C$.

Given class-families $(x_i)_{i \in I}$ and $(y_j)_{j \in J}$, we say that the former is *included* in the latter when $I \subseteq J$ and $x = y \upharpoonright_I$. A *map* $(x_i)_{i \in I} \to (y_j)_{j \in J}$ is a function $f : I \to J$ such that, for all $i \in I$, we have $x_i = y_{f(i)}$.

For a class C, we write $\mathcal{P}C$ for the class of all subsets of C, and $\mathsf{Fam}(C)$ for the class of all families within C. For any set I, we write C^I for the class of all I-tuples within C.

Given a classes A and B, a *partial function* $G : A \rightharpoonup B$ consists of a subclass $\mathsf{Dom}(G)$ of A and a function $\overline{G} : \mathsf{Dom}(G) \to B$. Put differently, it is a class-family (M, F) such that for all $x \in M$ we have $x \in A$ and $F(x) \in B$.

We also speak about "collections", although such talk is informal.

We write Class for the collection of all classes. Given a class C, we write $\mathcal{P}^{\blacksquare}(C)$ for the collection of all subclasses of C, and $\mathsf{Fam}^{\blacksquare}(C)$ for the collection of all class-families within C.

Given a class A, we may speak of an "indexed class" $(B_x)_{x\in A}$, also written $B : A \to \mathsf{Class}$. It is represented as a binary predicate formula $\phi(x,y)$ with parameters, so that, for $x \in A$, we have $B_x = \{y \mid \phi(x,y)\}$. We can then form the class

$$\sum_{x\in A} B_x \quad \stackrel{\text{def}}{=} \quad \{\langle x,y\rangle \mid x\in A, y\in B_x\}$$

The notation $F \in \prod_{x\in A} B_x$ means that F is a function on A that sends each $x\in A$ to an element of B_x. Likewise, a *partial function* $G \in \overset{\rightharpoonup}{\prod}_{x\in A} B_x$ consists of a subclass $\mathsf{Dom}(G)$ of A and function $\overline{G} \in \prod_{x\in\mathsf{Dom}(G)} B_x$. Put differently, it is a class-family (M,F) such that for all $x\in M$ we have $x \in A$ and $F(x) \in B_x$.

Two other kinds of function occur in the paper.

- Given a class A, we speak of a function $F : A \to \mathcal{B}$, where $\mathcal{B}$ is a collection, or of a function $F \in \prod_{x\in A} \mathcal{B}_x$, where $(\mathcal{B}_x)_{x\in A}$ is an indexed collection. In each instance, it is obvious how F can be represented. For example, we can represent $F : A \to \mathsf{Class}^2$ as a pair of functions $A \to \mathsf{Class}$.
- Given collections $\mathcal{A}$ and $\mathcal{B}$, we speak of a function $\mathcal{A} \to \mathcal{B}$. For example, $\mathcal{P}$ is an endofunction on Class.

We often consider *ordered* collections, such as $\mathcal{P}^{\blacksquare}(C)$ or $\overset{\rightharpoonup}{\prod}_{x\in A} B_x$, ordered by inclusion. Given ordered collections $(\mathcal{A}, \leqslant)$ and $(\mathcal{B}, \leqslant')$, a function $f : \mathcal{A} \to \mathcal{B}$ is *monotone* when $x \leqslant y$ implies $f(x) \leqslant' f(y)$.

3.3 Additional axioms

Here are some more axioms. We always state explicitly when they are assumed.

Proposition 3.1. *The following are equivalent.*

- Powerset*: For any set A, the class $\mathcal{P}A$ is a set.*
- Exponentiation*: For any sets A and B, the class B^A is a set.*
- Dependent Exponentiation*: For any family of sets $(B_i)_{i\in I}$, the class $\prod_{i\in I} B_i$ is a set.*

Proof. Via the following constructions.

$$\begin{aligned}
\mathcal{P}A &\stackrel{\text{def}}{=} \{\{x\in A \mid f(x) = 1\} \mid f \in \{0,1\}^A\} \\
B^A &\stackrel{\text{def}}{=} \prod_{x\in A} B \\
\prod_{i\in I} B_i &\stackrel{\text{def}}{=} \{f \in \mathcal{P}\sum_{i\in I} B_i \mid \forall i\in I.\, \exists! b\in B_i.\, \langle i,b\rangle \in f\}
\end{aligned}$$

□

Lastly, the axiom of *Choice* says that, for any family of inhabited sets $(A_i)_{i\in I}$, the class $\prod_{i\in I} A_i$ is inhabited.

3.4 Fixpoints

Since fixpoints appear repeatedly throughout the paper, we present the basic notions here.

Let $\mathcal{E}$ be a collection equipped with an endofunction Γ. Then an element $x\in\mathcal{E}$ is Γ-*fixed* or a Γ-*fixpoint* iff $\Gamma(x) = x$. The prefix Γ can be omitted when clear from the context.

Let $(\mathcal{E}, \leqslant)$ be an ordered collection equipped with a monotone endofunction Γ. An element $x \in \mathcal{E}$ is Γ-*prefixed* or a Γ-*prefixpoint* when $\Gamma(x) \leqslant x$, and Γ-*postfixed* or a Γ-*postfixpoint* when $x \leqslant \Gamma(x)$.[2] Clearly, x is fixed iff both prefixed and postfixed. An infimum of prefixpoints is prefixed, and a supremum of postifxpoints is postfixed.

Γ is *inflationary* when every $x\in\mathcal{E}$ is a postfixpoint, and *deflationary* when every $x\in\mathcal{E}$ is a prefixpoint.

The least prefixpoint of Γ, if it exists, is written $\mu\Gamma$. It is necessarily fixed; this fact is called *inductive inversion*. Dually the greatest postfixpoint of Γ, if it exists, is written $\nu\Gamma$ and is necessarily fixed.

A prefixpoint x is *minimal* when the only prefixpoint $y \leqslant x$ is x itself. A least prefixpoint is minimal, and conversely if $(\mathcal{E}, \leqslant)$ has binary meets, which is always the case in our examples.

Here is an application (not used in the sequel). Note first that a function $f : \prod_{x\in A} B_x \to \prod_{x\in C} D_x$ is monotone iff the following condition holds: for any $(M, F) \in \prod_{x\in A} B_x$ sent by f to (N, G), and any subclass X of M, we have

$$\Delta(X, F\restriction_X) \quad = \quad (Y, G\restriction_Y)$$

for a subclass Y of N, which we call $\Delta^{(M,F)}X$. Now let Δ be a monotone endofunction on $\prod_{x\in A} B_x$. Then a Δ-prefixpoint (M, F) is least iff the only $\Delta^{(M,F)}$-prefixed subclass of M is M itself. This follows from the equivalence of least and minimal prefixpoints.

3.5 Initial algebras

A small amount of category theory is used in this paper, specifically the fact that "initial algebras of isomorphic endofunctors are isomorphic". We briefly review this.

Let $\mathcal{C}$ be a category equipped with an endofunctor F. An F-*algebra* (x, θ) consists of an object x (the *carrier*) and morphism $\theta : Fx \to x$ (the *structure*). Given F-algebras (x, θ) and (y, ϕ), a *map* $(x, \theta) \to (y, \phi)$ is a $\mathcal{C}$-morphism f :

[2]This terminology follows [SP82]. Some authors use the opposite terminology, following [MS78].

$x \to y$ such that the square

$$\begin{tikzcd} Fx \arrow[r, "Ff"] \arrow[d, "\theta"'] & Fy \arrow[d, "\phi"] \\ x \arrow[r, "f"'] & y \end{tikzcd}$$

commutes. An F-algebra (x, θ) is *initial* when it has a unique map to every F-algebra. It follows that θ is an isomorphism; this is *Lambek's lemma*, and inductive inversion is a special case. We come to the desired result:

Proposition 3.2. *Let F and G be endofunctors on a category $\mathcal{C}$, with initial algebras (x, θ) and (y, ϕ) respectively. Given a natural isomorphism $\alpha : F \to G$, there is a unique morphism $f : x \to y$ such that*

$$\begin{tikzcd} Fx \arrow[r, "\rho_f"] \arrow[d, "\theta"'] & Gy \arrow[d, "\phi"] \\ x \arrow[r, "f"'] & y \end{tikzcd}$$

commutes, and it is an isomorphism.

Proof. Straightforward. □

3.6 Schematicity and predicativity

When speaking about classes, we must observe two disciplines.

Firstly, in order to claim that all classes have a given property, we must prove this *schematically*, i.e., for an arbitrary class, and not just for classes that are definable from first-order parameters. This is because our base theory's syntax includes class variables.

Secondly, in order to claim the existence of a class with a given property, we must prove this *predicatively*, i.e., without quantification over class variables. This is because our base theory's syntax does not provide such quantification.

Both requirements are illustrated in the next section.

3.7 Natural numbers

Bearing in mind that we do not assume Infinity, we must carefully define the class $\mathbb{N}$ of all natural numbers. Firstly, we define the monotone endofunction Maybe on Class that sends

$$X \quad \mapsto \quad \{\mathsf{Nothing}\} \cup \{\mathsf{Just}(x) \mid x \in X\}$$

This terminology comes from the functional programming literature [Ha06]. Note that a class X is Maybe-prefixed iff it contains $\mathsf{Nothing}$ and, for any $x \in X$, contains $\mathsf{Just}(x)$.

We want $\mathbb{N}$ to be the least Maybe-prefixed class. (To emphasize, "least" means that $\mathbb{N}$ is included in an arbitrary Maybe-prefixed class; this is the schematicity requirement.) We cannot simply define $\mathbb{N}$ to be the intersection of all Maybe-prefixed classes, as that would be impredicative. Instead we proceed as follows.

Proposition 3.3. *The class* $\mathbb{N} \stackrel{\text{def}}{=} \mu\mathsf{Maybe}$ *exists.*

Proof. First note that a class X is Maybe-postfixed iff every $x \in X$ is either $\mathsf{Nothing}$ or $\mathsf{Just}(x)$ for some $x \in X$. When this is so, say that a subclass U of X is *inductive* when it contains $\mathsf{Nothing}$ if X does, and, for all $x \in U$ contains $\mathsf{Just}(x)$ if X does.

A thing x that is contained in a Maybe-postfixed set is contained in a least such set, written $\downarrow x$. Take the class of all x such that x is contained in a Maybe-postfixed set and the only inductively subset of $\downarrow x$ is $\downarrow x$ itself. This has the required properties. ☐

We often write $0 \stackrel{\text{def}}{=} \mathsf{Nothing}$, and $n+1 \stackrel{\text{def}}{=} \mathsf{Just}(n)$ for $n \in \mathbb{N}$. The standard properties of $\mathbb{N}$ hold, including the following.

Proposition 3.4. (Recursion over $\mathbb{N}$.) *For any class sequence* $(B_n)_{n\in\mathbb{N}}$ *and any* $p \in B_0$ *and* $L \in \prod_{n\in\mathbb{N}}(B_n \to B_{n+1})$*, there is a unique sequence* $b \in \prod_{n\in\mathbb{N}} B_n$ *such that* $b_0 = p$ *and, for all* $n \in \mathbb{N}$*, we have* $b_{n+1} = L_n(b_n)$.

Note that the axiom of Infinity says that a Maybe-prefixed set exists; this is equivalent to $\mathbb{N}$ being a set.

3.8 Set-based relations

Before defining well-foundedness, we consider some preliminary notions.

Definition 3.5. Let $(C, <)$ be a class equipped with a relation.

1. A subclass X of C is *hereditary* when every child of an element of X is in X.

2. Let $x \in C$. An element $y \in x$ is

   - a *child* of x when $y < x$.
   - a *descendant* of x, written $y <^* x$, when there is a sequence

   $$y = z_0 < \cdots < z_n = x$$

   - a *strict descendant* of x, written $y <^+ x$, when there is such a sequence with $n > 0$.

   We write $J_<(x)$ for the class of all children of x, and $J^*_<(x)$ for the class of all descendants, and $J^+_<(x)$ for the class of all strict descendants. Thus $J^*_<(x)$ is the least hereditary subclass of C that contains x. The subscript $<$ is often omitted.

Example 3.6. Consider the membership relation on $\mathfrak{T}$. A class X is *membership-hereditary* or *transitive* when every element of an element of X is in X. For a thing x, we write $\mathcal{E}(x)$ for its element set, which is x or $\emptyset$ according as x is a set or an urelement. We write $\mathcal{E}^*(x)$ for the class of all membership-descendants of x, and $\mathcal{E}^+(x)$ for the class of all strict ones. Thus $\mathcal{E}^*(x)$ is the least transitive class containing x.

Definition 3.7. A relation $<$ on a class C is

- *set-based* when $J(x)$ is a set for all $x \in C$.
- *iteratively set-based* when each $x \in C$ is contained in a hereditary subset of C; this is equivalent to $J^*(x)$ being a set.

Thus $<$ is iteratively set-based iff $<$ is set-based.

Example 3.8. The membership relation on $\mathfrak{T}$ is set-based. It is iteratively set-based iff the *Transitive Containment* axiom holds: Every thing is contained in a transitive set.

Proposition 3.9. *Infinity is equivalent to the statement: "Every set-based relation on a class is iteratively set-based."*

Proof. For ($\Rightarrow$), let $<$ be a set-based relation on a class C, and $x \in C$. By induction on $n \in \mathbb{N}$, the class $J^n(x)$ of descendants of x at depth n is a set, since

$$\begin{array}{rcl} J^0(x) & = & \{x\} \\ J^{n+1}(x) & = & \bigcup_{y \in J^n(x)} J(y) \end{array}$$

So the class $J^*(x) = \bigcup_{n \in \mathbb{N}} J^n(x)$ is a set.

For ($\Leftarrow$), the relation on $\mathfrak{T}$ given by $\{\langle x, y\rangle \mid y = \mathsf{Just}(x)\}$ is set-based, and the descendant class of $\mathsf{Nothing}$ is $\mathbb{N}$. So if set-based implies iteratively set-based, then $\mathbb{N}$ is a set. $\square$

3.9 Set-based well-founded relations

The notion of well-foundedness relies on the following concepts.

Definition 3.10. Let $(C, <)$ be a class equipped with a relation. Let X be a subclass.

1. X is *inductive* when every element of C whose children are all in X is in X.
2. An element of X is *minimal* when it has no child in X.

Now we are ready to formulate well-foundedness.

Proposition 3.11. *Let $(C, <)$ be a class equipped with a set-based relation. The following are equivalent:*

1. *The only inductive subclass of C is C itself.*
2. *Every inhabited subclass of C has a minimal element.*
3. *The relation $<$ is iteratively set-based, and every inhabited subset of C has a minimal element.*

Proof. Conditions (1) and (2) are equivalent because a subclass of C is inductive iff its complement has no minimal element.

To prove (1) implies (3), we need only show that $<$ is iteratively set-based: for all $x \in C$, the class $J^*(x)$ is a set, by induction on x.

To show (3) implies (2), any subclass Y of C inhabited by x gives a subset $Y \cap J^*(x)$ of C inhabited by x. The latter has a minimal element, which is also a minimal element of Y. □

When the above conditions hold, we say that $<$ is a *set-based well-founded relation.*

Example 3.12. Consider again the membership relation. A class X is *membership-inductive* when every thing whose elements are all in X is in X. An element $x \in X$ is *membership-minimal* when it has no element in common with X. We can say that membership is well-founded in each of the following ways.

- The axiom scheme of *Membership Induction*: The only membership-inductive class is $\mathfrak{T}$.
- The axiom scheme of *Class Regularity*: Every inhabited class has a membership-minimal element.
- Transitive Containment + the axiom of *Regularity*: Every inhabited set has a membership-minimal element.

Here are some basic properties of well-founded relations.

Proposition 3.13. *Let $(A, <)$ and $(B, <')$ be classes equipped with a set-based relation, and $f : A \to B$ a function such that $x < y$ implies $f(x) <' f(y)$. If $<'$ is well-founded, then $<$ is too.*

Proof. Let X be an inductive subclass of A. We prove by induction on $y \in B$ that $f^{-1}(y) \subseteq X$. □

Proposition 3.14. *Let $(C, <)$ be a class equipped with a set-based relation.*

1. *The relation $<^+$ is well-founded iff $<$ is.*
2. *If $<$ is well-founded, then there is no infinite sequence $\cdots < x_1 < x_0$. In particular, $<$ is irreflexive.*

Proof.

1. The direction ($\Rightarrow$) is by Proposition 3.13. For ($\Leftarrow$), let X be a $<^+$-inductive subclass of C. For all $x \in C$, we prove $J^*(x) \subseteq C$ by induction on x.

2. Fix such a sequence. We prove that every $x \in C$ fails to appear in it, by induction on x. $\square$

Functions and partial functions can be defined by well-founded recursion:

Proposition 3.15. *Let $(A, <)$ be a class equipped with a set-based well-founded relation, and $(B_x)_{x \in A}$ an indexed class.*

1. *For any $L \in \prod_{x \in A}((\prod_{y \in J(x)} B_y) \to B_x)$, there is a unique function $F \in \prod_{x \in A} B_x$ sending $x \in A$ to $L_x(F \restriction_{J(x)})$. In other words, the endofunction Φ_L on $\prod_{x \in A} B_x$ sending F to $x \mapsto L_x(F \restriction_{J(x)})$ has a unique fixpoint.*

2. *For any $L \in \prod_{x \in A}((\prod_{y \in J(x)} B_y) \rightharpoonup B_x)$, let Ψ_L be the monotone endofunction on $\prod^{\rightharpoonup}_{x \in A} B_x$ sending (M, F) to (N, G), where N is the class of all $x \in A$ such that $J(x) \subseteq M$ and $F \restriction_{J(x)} \in \mathsf{Dom}(L_x)$, and G sends such an x to $\overline{L_x}(F \restriction_{J(x)})$. Then Ψ_L has a least prefixpoint that is also a greatest postfixpoint and therefore a unique fixpoint.*

Proof. We first prove part 2. For a hereditary subclass M of A, an *attempt* on M is a function $F \in \prod_{x \in M} B_x$ such that, for all $x \in M$, we have $F \restriction_{J(x)} \in \mathsf{Dom}(L_x)$ and $F(x) = \overline{L_x}(F_x \restriction_{J(x)})$. Thus a Ψ_L-postfixpoint (M, F) consists of a hereditary subclass M of A, and an attempt F on M. Induction shows that

- any attempt on M and attempt on M' agree on $M \cap M'$
- any postfixpoint is included in any prefixpoint.

Let P be the class of all x such that there is a (necessarily unique) attempt on $J^*(x)$. Let H send each $x \in P$ to its image under the attempt on $J^*(x)$. Then (P, H) is a fixpoint, since any attempt g on $J^+(x)$ such that $g \restriction_{J(x)} \in \mathsf{Dom}(L_x)$ extends to an attempt $g \cup \{\langle x, \overline{L_x}(g \restriction_{J(x)})\rangle\}$ on $J^*(x)$. So part 2 is proved.

If L_x is total for all $x \in A$, then any Ψ_L-prefixpoint is total, and Φ_L is the restriction of Ψ_L to total functions, so part 1 follows. $\square$

4 Scaffolds

4.1 Generating a subclass

It is often useful to know that a monotone endofunction on $\mathcal{P}^{\blacksquare}(C)$ or $\prod^{\rightharpoonup}_{x \in A} B_x$ has a least prefixpoint, and this section provides conditions that guarantee that. First we deal with $\mathcal{P}^{\blacksquare}(C)$, for a given class C.

Definition 4.1.

1. A *scaffold* on C consists of
   - a subclass D
   - a relation $<$ from C to D.

We call $x \in D$ a *parent* and $y < x$ a *child* of x. The scaffold is *set-based* when, for all $x \in D$, the class $J_<(x) \stackrel{\text{def}}{=} \{y \in C \mid y < x\}$ is a set.

2. A scaffold $(D, <)$ on C gives rise to a monotone endofunction $\Gamma_{(D,<)}$ on $\mathcal{P}^{\blacksquare}(C)$, sending X to the class of all $x \in D$ whose children are all in X.

Thus a subclass X of C is

- $\Gamma_{(D,<)}$-prefixed iff every parent whose children are all in X is in X
- $\Gamma_{(D,<)}$-postfixed iff it is a hereditary subclass of D.

Proposition 4.2. *Let $(D, <)$ be a scaffold on C.*

1. (Cogenerated subclass.) *C has a greatest $\Gamma_{(D,<)}$-postfixed subclass.*
2. (Generated subclass.) *If the scaffold is set-based, then C has a least $\Gamma_{(D,<)}$-prefixed subclass, which is also the greatest $\Gamma_{(D,<)}$-postfixed subclass on which $<$ is well-founded.*

Proof.

1. Take the class of all $x \in C$ such that $J^*(x) \subseteq D$.
2. First note that $<$ is a set-based relation on C, since for $x \in C \setminus D$, the class $J(x)$ is $\emptyset$.

   Take the class of all $x \in C$ such that $J^*(x)$ is a subset of D whose only inductive subset is itself. This is clearly the least $\Gamma_{(D,<)}$-prefixed subclass. By inductive inversion, it is $\Gamma_{(D,<)}$-postfixed. The rest is straightforward, using the fact that any set-based well-founded relation is iteratively set-based. □

Note All scaffolds in this paper are on $\mathfrak{T}$, except in the proof of Proposition 5.2, where we use a scaffold on a set.

Example 4.3. The endofunction Maybe arises from the following scaffold on $\mathfrak{T}$: a parent is either $\mathsf{Nothing}$, which has no children, or $\mathsf{Just}(x)$, whose sole child is x. So Proposition 3.3 is an instance of Proposition 4.2(2).

Example 4.4. The endofunction $\Gamma_{(\mathfrak{T},\in)}$ sends a class X to $\mathsf{Ur} \cup \mathcal{P}X$. We define

$$V_{\mathsf{impure}} \stackrel{\text{def}}{=} \mu\Gamma_{(\mathfrak{T},\in)}$$

which is the least $\mathcal{P}$-prefixed class that includes Ur. Proposition 4.2(2) tells us that V_{impure} is the least membership-inductive class, and the greatest transitive class on which membership is well-founded. An element is called a *vonniad*—the name alludes to "von Neumann iteration".

Example 4.5. A thing is *pure* when its membership-descendants are all sets. The class of all pure things is $\nu\mathcal{P}$, which is an instance of Proposition 4.2(1) since $\mathcal{P} = \Gamma_{(\mathfrak{S},\in)}$. Likewise, the class of all pure vonniads is given by

$$V_{\mathsf{pure}} \stackrel{\text{def}}{=} \mu\mathcal{P}$$

which is an instance of Proposition 4.2(2).

Related work In a different context, the question of when a monotone endofunction on Class has a least prefixpoint is considered in [AR10, Theorem 12.1.1].

4.2 Generating a partial function

Next we deal with $\vec{\prod}_{x\in A} B_x$, for a given class A and indexed class $(B_x)_{x\in A}$.

Definition 4.6. Let $(D,<)$ be a set-based scaffold on A.

1. A *functionalization* of $(D,<)$ on $(B_x)_{x\in A}$ is an $L \in \prod_{x\in D}((\prod_{y\in J(x)} B_y) \rightharpoonup B_x)$.

2. Let L be such a functionalization. The monotone endofunction $\Delta^L_{(D,<)}$ on $\vec{\prod}_{x\in A} B_x$ sends (M,F) to (N,G), where

   - N is the class of $x \in D$ such that $J(x) \subseteq M$ and $F\restriction_{J(x)} \in \mathsf{Dom}(L_x)$
   - G sends each such x to $\overline{L_x}(F\restriction_{J(x)})$.

Proposition 4.7. (Generated partial function.) *Let $(D,<)$ be a set-based scaffold on A with functionalization L on $(B_x)_{x\in A}$. Then there is a least $\Delta^L_{(D,<)}$-prefixpoint, which is also the greatest $\Delta^L_{(D,<)}$-postfixpoint (M,F) such that $<$ is well-founded on M.*

Proof. Let E be the subclass of A generated by $(D,<)$. Since E is $\Gamma_{(D,<)}$-prefixed, $\Delta^L_{(D,<)}$ restricts to an endofunction on $\vec{\prod}_{x\in E} B_x$. By Proposition 3.15(2), the latter has a least prefixpoint (M,F) that is also a greatest postfixpoint, since $<$ is well-founded on E. Because (M,F) is a minimal prefixpoint in $\vec{\prod}_{x\in E} B_x$, it is a minimal and therefore least prefixpoint in $\vec{\prod}_{x\in E} B_x$.

For any postfixpoint (N,G), the class N is $\Gamma_{(D,<)}$-postfixed. So if $<$ is well-founded on N, then $N \subseteq E$, giving $(N,G) \in \vec{\prod}_{x\in E} B_x$ and so $(N,G) \leqslant (M,F)$. $\square$

4.3 Introspective scaffolds

This section is not used in the sequel.

If Membership Induction is assumed, then $\mathbb{N}$ is the unique class X such that $x \in X$ iff either $x = \mathsf{Nothing}$ or $x = \mathsf{Just}(y)$ for $y \in X$. Various classes defined later in the paper, such as $\mathsf{Wide}(S)$ and $\mathsf{Broad}(G)$ and $\mathsf{Deriv}_{\mathcal{R}}$ and Ord, have a similar property. That is because they are "introspectively generated", in a sense that I now explain.

Definition 4.8. Let C be a class.

1. A relation $<$ on C is *introspective* when $<$ is included in $\in^+$. In other words: when, for all $x \in C$, we have $J(x) \subseteq \mathcal{E}^+(x)$.

2. A scaffold $(D,<)$ on C is *introspective* when $<$ is included in $\in^+$. In other words: when, for all $x \in D$, we have $J(x) \subseteq \mathcal{E}^+(x)$.

All the scaffolds used in this paper (except in the proof of Proposition 5.2) are introspective, as well as being set-based.

Proposition 4.9.

1. *Transitive Containment is equivalent to the statement: "Every introspective relation on a class is iteratively set-based."*

2. *Membership Induction is equivalent to the statement: "Every introspective relation on a class is well-founded."*

Proof.

1. Transitive Containment is equivalent to membership being iteratively set-based, which is equivalent to $\in^+$ being iteratively set-based, which is equivalent to every introspective relation being iteratively set-based.

2. Similar. □

Now we come to the key result of the section:

Proposition 4.10. *Each of the following is equivalent to Membership Induction.*

1. *Any monotone endofunction on $\mathcal{P}^{\blacksquare}(C)$ arising from an introspective set-based scaffold has a unique fixpoint.*

2. *Any monotone endofunction on $\vec{\prod}_{x \in A} B_x$ arising from an introspective set-based scaffold with functionalization has a unique fixpoint.*

Proof.

1. Membership Induction implies this by Proposition 4.9(2) and Proposition 4.2(2), since a least prefipxoint that is also a greatest postfixpoint is a unique fixpoint. For the converse, since $\mathfrak{T}$ is a fixpoint of $\Gamma_{(\mathfrak{T},\in)}$, it must be the least prefixpoint and so $\in$ is well-founded over it.

2. Membership Induction implies this by Proposition 4.9(2) and Proposition 4.7. For the converse, the scaffold $(\mathfrak{T}, \in)$ on $\mathfrak{T}$ has a functionalization L on $(1)_{x \in \mathfrak{T}}$ that at $*$ takes the empty function to $*$. Since the partial function $(\mathfrak{T}, x \mapsto *)$ is a fixpoint of $\Delta^L_{(\mathfrak{T},\in)}$, it must be the least prefixpoint and so $<$ is well-founded over $\mathfrak{T}$. □

Example 4.11. Let us apply Proposition 4.10(1) to Example 4.3. Assuming Membership Induction, we see that $\mathbb{N}$ is the unique Maybe-fixpoint.

5 Wide and Broad Infinity

5.1 Wide Infinity

Let us review what we have seen previously. The class of natural numbers, written $\mathbb{N}$, is the least Maybe-prefixed class, given by Proposition 3.3. Examples

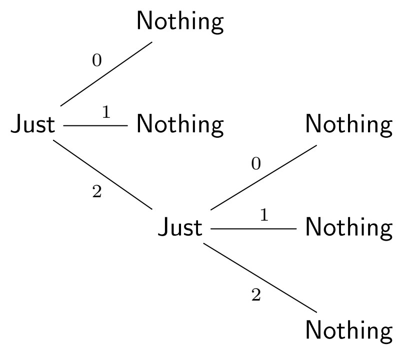

Figure 2: Visualization of a simple wide number

of natural numbers are $\mathsf{Nothing}$ and $\mathsf{Just}(\mathsf{Nothing})$ and $\mathsf{Just}(\mathsf{Just}(\mathsf{Nothing}))$. The axiom of Infinity says that a Maybe-prefixed set exists, which is equivalent to $\mathbb{N}$ being a set.

Now we continue. A set K (which in this context may be called an *arity*) gives rise to a monotone endofunction Maybe_K on Class, sending X to $\mathsf{Maybe}\, X^K$. Thus a class X is Maybe_K-prefixed iff it contains $\mathsf{Nothing}$ and, for any K-tuple y within X, contains $\mathsf{Just}(y)$.

The least such class is called the *class of all simple K-wide numbers*, and written $\mathsf{SimpleWide}(K)$. It exists by Proposition 4.2(2), since Maybe arises from the following scaffold on $\mathfrak{T}$. A parent is either $\mathsf{Nothing}$, which has no children, or $\mathsf{Just}(y)$, for any K-tuple y, in which case the set of children is $\{y_k \mid k \in K\}$.

Example 5.1. Define K to be the arity $\{0, 1, 2\}$. The following are simple K-wide numbers:

- $\mathsf{Nothing}$
- $\mathsf{Just}\begin{bmatrix}\mathsf{Nothing}\\ \mathsf{Nothing}\\ \mathsf{Nothing}\end{bmatrix}$
- $\mathsf{Just}\begin{bmatrix}\mathsf{Nothing}\\ \mathsf{Nothing}\\ \mathsf{Just}\begin{bmatrix}\mathsf{Nothing}\\ \mathsf{Nothing}\\ \mathsf{Nothing}\end{bmatrix}\end{bmatrix}$

We can visualise a wide number as a well-founded tree. For example, the last number in Example 5.1 is visualized in Figure 2, using the vertical dimension for $\begin{bmatrix}\vdots\end{bmatrix}$ and the horizontal dimension for internal structure, with the root appearing at the left and the $\mathsf{Nothing}$-marked leaves at the right. The axiom of *Simple Wide Infinity* says, for every arity K, that a Maybe_K-prefixed set exists; this is equivalent to $\mathsf{SimpleWide}(K)$ being a set.

There is an alternative version, formulated as follows.

A *signature* $S = (K_i)_{i \in I}$ is a family of sets. Explicitly, it consists of a set I of *symbols*, where each symbol $i \in I$ is equipped with a set K_i called its *arity*. It gives rise to a monotone endofunction H_S on Class, sending X to $\sum_{i \in I} X^{K_i}$. Thus a class X is H_S-prefixed iff, for any $i \in I$ and K_i-tuple y within X, it contains $\langle i, y \rangle$. The least such class is called the *class of all S-wide numbers* and written $\mathsf{Wide}(S)$. It exists by Proposition 4.2(2), since H_S arises from a set-based scaffold on $\mathfrak{T}$.

The axiom of *Full Wide Infinity* says, for every signature S, that an H_S-prefixed set exists; this is equivalent to $\mathsf{Wide}(S)$ being a set. Previously this axiom has appeared in [page 15][vdB11] under the name "Smallness of W-types", alluding to the notion of W-type in type theory [ML84, MP00, AAG05].

Categorical remark. Write **Class** for the category of all classes and functions. An endofunctor F on **Class** is *monotone* when it preserves inclusions. It follows that any class X that is an initial F-algebra with identity structure is also a least F-prefixpoint. (Proof: minimality suffices. Let Y be an F-prefixed subset of Y, giving inclusions $i : FY \to Y$ and $j : Y \to X$. Since $j : (Y, i) \to (X, \mathsf{id}_X)$ is an algebra map, it has a section. So it is surjective, giving $Y = X$.)

Here are some examples:

- $\mathcal{P}$ extends to a monotone endofunctor on **Class**, sending $f : A \to B$ to the function

$$\begin{array}{rcl} \mathcal{P}A & \to & \mathcal{P}B \\ X & \mapsto & \{f(x) \mid x \in X\} \end{array}$$

  Moreover, V_{pure} is an initial $\mathcal{P}$-algebra with identity structure.

- Maybe extends to an monotone endofunctor on **Class**. Moreover, $\mathbb{N}$ is an initial Maybe-algebra with identity structure.

- Maybe_K, for a set K, extends to a monotone endofunctor on **Class**. Moreover, $\mathsf{SimpleWide}(K)$ is an initial Maybe_K-algebra with identity structure.

- H_S, for a signature S, extends to a monotone endofunctor on **Class**. Moreover, $\mathsf{Wide}(S)$ is an initial H_S-algebra with identity structure.

Proposition 5.2. *Simple Wide Infinity, Full Wide Infinity and Powerset + Infinity are all equivalent.*

Proof. To show that Full Wide Infinity implies Simple Wide Infinity: given a set K, define S to be the signature $(L_i)_{i \in \{0,1\}}$ where $L_0 \stackrel{\text{def}}{=} \emptyset$ and $L_1 \stackrel{\text{def}}{=} K$. We have a natural isomorphism $\mathsf{Maybe}_K \cong H_S$, so Proposition 3.2 tells us that $\mathsf{SimpleWide}(K)$ is a set iff $\mathsf{Wide}(S)$ is a set.

To show that Simple Wide Infinity implies Infinity, we have a natural isomorphism $\mathsf{Maybe} \cong \mathsf{Maybe}_1$, so $\mathbb{N}$ is a set iff $\mathsf{SimpleWide}(1)$ is a set.

To show that Full Wide Infinity implies Exponentiation, let A and B be sets. Define the signature S with nullary symbols $(\mathsf{Leaf}_b)_{b \in B}$ and an A-ary symbol

Node. Consider the injection $B^A \to \mathsf{Wide}(S)$ sending a function f to the S-wide number

$$\langle \mathsf{Node}, [\langle \mathsf{Leaf}_{f(a)}, [\,]\rangle]_{a \in A}\rangle$$

Since $\mathsf{Wide}(S)$ is a set, B^A is too.

Now we show that Full Wide Infinity follows from Powerset+Infinity and from Simple Wide Infinity. Let $S = (K_i)_{i \in I}$ be a signature.

Observation: if there is an H_S-algebra (X, θ) whose structure θ is injective, then $\mathsf{Wide}(S)$ is a set. To see this, let Y be the subset of X generated by the scaffold where a parent is an element (uniquely) expressible as $\theta\langle i, [x_k]_{k \in K}\rangle$ and its set of children is $\{x_k \mid k \in K\}$. Then θ restricts to a function $H_S Y \to Y$, and the H_S-algebra $(Y, \theta \upharpoonright_{H_S Y})$ is initial. Since Y is a set, $\mathsf{Wide}(S)$ is too (as initial algebras of an endofunctor are isomorphic).

Next, assume Powerset+Infinity. Following [Bar93], we form the ω^{op}-chain

$$1 \xleftarrow{\langle\rangle} H_S 1 \xleftarrow{H_S \langle\rangle} H_S^2 1 \xleftarrow{H_S^2 \langle\rangle} \cdots$$

known as the "final chain". (Intuitively $H_S^n 1$ is the set of S-trees with stumps at level n.) Let M be the limit and $\theta : H_S M \to M$ the canonical map. (Intuitively M is the set of S-trees.) The map θ is bijective, since H_S preserves limits of connected diagrams, so $\mathsf{Wide}(S)$ is a set by our observation.

Lastly, assume Simple Wide Infinity and define the set $\overline{S} \stackrel{\text{def}}{=} I + \sum_{i \in I} K_i$. Then $\mathsf{SimpleWide}(\overline{S})$ is a set. We obtain an injection $H_S \mathsf{SimpleWide}(\overline{S}) \to \mathsf{SimpleWide}(\overline{S})$ sending $\langle i, [a_k]_{k \in K_i}\rangle$ to $\mathsf{Just}([b_p]_{p \in \overline{S}})$, where

$$\begin{array}{rcll} b_{\mathsf{inl}\ i} & \stackrel{\text{def}}{=} & \mathsf{Just}([\mathsf{Nothing}]_{p \in \overline{S}}) & \\ b_{\mathsf{inl}\ j} & \stackrel{\text{def}}{=} & \mathsf{Nothing} & (\text{for } j \in I, j \neq i) \\ b_{\mathsf{inr}\ \langle i,k\rangle} & \stackrel{\text{def}}{=} & a_k & (\text{for } k \in K_i) \\ b_{\mathsf{inr}\ \langle j,k\rangle} & \stackrel{\text{def}}{=} & \mathsf{Nothing} & (\text{for } j \in I, j \neq i, k \in K_j) \end{array}$$

So $\mathsf{Wide}(S)$ is a set by our observation. $\square$

Related work The idea that every signature has an initial algebra (meaning: algebra carried by a set) has often appeared in the literature. Apparently, the earliest ZFC proof was given by Słomiński [Sł58], and the earliest ZF proof by Kerkhoff [Ker65]. Furthermore, as explained in [PS78, Bla83, MP00], the result holds in any topos with a natural number object.

5.2 Broad Infinity

Now we come to the main principle of the paper, which was briefly described in Section 1.1.

A *broad arity* is a function $F : \mathfrak{T} \to \mathbf{Set}$. It gives rise to a monotone endofunction Maybe_F on Class, sending X to $\mathsf{Maybe} \sum_{x \in X} X^{Fx}$. Thus a class X is

Maybe_F-prefixed when it contains $\mathsf{Nothing}$ and, for any $x \in X$ and Fx-tuple y within X, contains $\mathsf{Just}\langle x, y\rangle$.

The least such class is called the *class of all simple F-broad numbers* and written $\mathsf{SimpleBroad}(F)$. It exists by Proposition 4.2(2), since Maybe_F arises from a set-based scaffold on $\mathfrak{T}$.

Although F is defined over $\mathfrak{T}$, only its restriction to $\mathsf{SimpleBroad}(F)$ matters. More precisely, for broad arities F and F' with the same restriction to $\mathsf{SimpleBroad}(F) \cap \mathsf{SimpleBroad}(F')$, we have $\mathsf{SimpleBroad}(F) = \mathsf{SimpleBroad}(F')$. Proof: the class $\mathsf{SimpleBroad}(F) \cap \mathsf{SimpleBroad}(F')$ is both Maybe_F-prefixed and $\mathsf{Maybe}_{F'}$-prefixed.

Let us see some examples of simple broad numbers.

Example 5.3. Define F to be the broad arity that sends $\mathsf{Just}\langle\mathsf{Nothing}, [\,]\rangle$ to $\{0, 1\}$, and everything else to $\emptyset$. The following are simple F-broad numbers:

- $\mathsf{Nothing}$
- $\mathsf{Just}\langle\mathsf{Nothing}, [\,]\rangle$
- $\mathsf{Just}\langle\mathsf{Just}\langle\mathsf{Nothing}, [\,]\rangle, \begin{bmatrix} \mathsf{Nothing} \\ \mathsf{Just}\langle\mathsf{Nothing}, [\,]\rangle \end{bmatrix}\rangle$
- $\mathsf{Just}\langle\mathsf{Just}\langle\mathsf{Just}\langle\mathsf{Nothing}, [\,]\rangle, \begin{bmatrix} \mathsf{Nothing} \\ \mathsf{Just}\langle\mathsf{Nothing}, [\,]\rangle \end{bmatrix}\rangle, [\,]\rangle$

We can visualise a broad number as a well-founded three-dimensional tree, using the vertical dimension for $\begin{bmatrix} \vdots \end{bmatrix}$, the horizontal dimension for $\langle -, - \rangle$ and the depth dimension for internal structure. The root appears at the front, and the $\mathsf{Nothing}$-marked leaves at the rear.

The axiom scheme of *Simple Broad Infinity* says, for every reduced broad signature F, that a Maybe_F-prefixed set exists; this is equivalent to $\mathsf{SimpleBroad}(F)$ being a set.

There is an alternative version, formulated as follows.

Writing Sig for the class of all signatures, a *broad signature* is a function $G : \mathfrak{T} \to \mathsf{Sig}$. It gives rise to a monotone endofunction Maybe_G on Class, sending X to $\mathsf{Maybe} \sum_{x \in X} H_{Gx} X$. Thus a class X is Maybe_G-prefixed iff it contains $\mathsf{Nothing}$ and, for any $x \in X$ with $Gx = (K_i)_{i \in I}$ and any $i \in I$ and K_i-tuple y within X, contains $\mathsf{Just}\langle x, i, y\rangle$.

The least such class is called the *class of all G-broad numbers* and written $\mathsf{Broad}(G)$. It exists by Proposition 4.2(2), since Maybe_G arises from a set-based scaffold on $\mathfrak{T}$. For broad signatures G and G' with the same restriction to $\mathsf{Broad}(G) \cap \mathsf{Broad}(G')$, we have $\mathsf{Broad}(G) = \mathsf{Broad}(G')$.

The axiom scheme of *Full Broad Infinity* says, for every broad signature G, that a Maybe_G-prefixed set exists; this is equivalent to $\mathsf{Broad}(G)$ being a set.

Proposition 5.4.

1. *Simple Broad Infinity and Full Broad Infinity are equivalent.*

2. *Broad Infinity implies Wide Infinity.*

Proof.

1. To show that Full Broad Infinity implies Simple Broad Infinity, let F be a broad arity. We recursively define the injection f on $\mathsf{SimpleBroad}(F)$ that sends

   - $\mathsf{Nothing}$ to $\mathsf{Nothing}$.
   - $\mathsf{Just}\langle w, [a_k]_{k\in Fw}\rangle$ to $\mathsf{Just}\langle f(w), 0, [f(a_k)]_{k\in Fw}\rangle$.

   Define the broad signature G sending

   - $f(x)$, for $x\in\mathsf{SimpleBroad}(F)$, to $(Fx)_{i\in\{0\}}$
   - everything else to the empty signature.

   Observe that f sends each $w\in\mathsf{SimpleBroad}(F)$ to a G-broad number, by induction on w. Since $\mathsf{Broad}(G)$ is a set, $\mathsf{SimpleBroad}(F)$ is too.

   To show the converse, for a signature $S = (K_i)_{i\in I}$, we write $\overline{S} \stackrel{\text{def}}{=} I + \sum_{i\in I} K_i$. Given a broad signature G, we recursively define the injection g on $\mathsf{Broad}(G)$ that sends

   - $\mathsf{Nothing}$ to $\langle\mathsf{Nothing}, [\,]\rangle$.
   - and $\mathsf{Just}\langle w, i, [a_k]_{k\in K_i}\rangle$, where $Gw = (K_i)_{i\in I}$, to $\langle\mathsf{Just}(g(w)), [b_p]_{p\in\overline{Gw}}\rangle$, using the definitions

   $$\begin{array}{rcll} b_{\mathsf{inl}\ i} & \stackrel{\text{def}}{=} & \mathsf{Just}\langle\mathsf{Nothing}, [\,]\rangle & \\ b_{\mathsf{inl}\ j} & \stackrel{\text{def}}{=} & \mathsf{Nothing} & (\text{for } j\in I, j\neq i) \\ b_{\mathsf{inr}\ \langle i,k\rangle} & \stackrel{\text{def}}{=} & \mathsf{Just}(g(a_k)) & (\text{for } k\in K_i) \\ b_{\mathsf{inr}\ \langle j,k\rangle} & \stackrel{\text{def}}{=} & \mathsf{Nothing} & (\text{for } j\in I, j\neq i, k\in K_j). \end{array}$$

   Let F be the broad arity that sends

   - $\mathsf{Just}(g(w))$, for $w\in\mathsf{Broad}(G)$, to $\overline{Gw}$
   - and everything else, including $\mathsf{Nothing}$, to $\emptyset$.

   Say that an F-*pair* is a pair $\langle w, c\rangle$ consisting of $w\in\mathsf{SimpleBroad}(F)$ and $c\in\mathsf{SimpleBroad}(F)^{Fw}$. Clearly, if x is an F-pair, then $\mathsf{Just}(x)$ is an F-broad number. Observe that g sends each $w\in\mathsf{Broad}(G)$ to an F-pair, by induction on w. Since $\mathsf{SimpleBroad}(F)$ is a set, the class of all F-pairs is too, and therefore $\mathsf{Broad}(G)$ is too.

2. We show that Full Broad Infinity implies Full Wide Infinity as follows. Given a signature S, let G be the broad signature sending everything to S. Recursively define the injection $g : \mathsf{Wide}(S) \to \mathsf{Broad}(G)$ sending $\langle i, [y_k]_{k\in K_i}\rangle$ to $\mathsf{Just}\langle\mathsf{Nothing}, i, [g(y_k)]_{k\in Fx}\rangle \in X$. Since $\mathsf{Broad}(G)$ is a set, $\mathsf{Wide}(S)$ is too. $\square$

6 Rubrics

6.1 Generating a subset

Having completed our presentation of the "plausible" principles, we now move on to the "useful" ones. First we consider how to generate a subset of a class using a suitable collection of rules, called a *rubric*.

Definition 6.1. Let C be a class.

1. A *wide rule on* C consists of a set K (the *arity*) and a function $R: C^K \to \mathsf{Fam}(C)$. It is written $\langle K, R\rangle$, and the collection of all such is denoted $\mathsf{WideRule}(C)$.

2. A *wide rubric on* C is a family of wide rules, i.e., a set I and function $r : I \to \mathsf{WideRule}(C)$. It is written $(r_i)_{i\in I}$, and the collection of all such is denoted $\mathsf{WideRub}(C)$.

3. A *broad rule on* C consists of a set L (the *arity*) and a function $T : C^L \to \mathsf{WideRub}(C)$. It is written $\langle L, T\rangle$, and the collection of all such is denoted $\mathsf{BroadRule}(C)$.

4. A *broad rubric on* C is a family of broad rules, i.e., a set J and function $s : J \to \mathsf{BroadRule}(C)$. It is written $(s_j)_{j\in J}$, and the collection of all such is denoted $\mathsf{BroadRub}(C)$.

Example 6.2. Here is a wide rubric on $\mathbb{N}$, consisting of two wide rules.

- Rule 0 is binary and sends $\begin{bmatrix} m_0 \\ m_1 \end{bmatrix} \mapsto (m_0 + m_1 + p)_{p\geqslant 2m_0}$.
- Rule 1 is nullary and sends $[\,] \mapsto (2p)_{p\geqslant 50}$.

Informally, these rules prescribe when an element of $\mathbb{N}$ is acceptable. Rule 0 says that, if m_0 and m_1 are acceptable, then $m_0 + m_1 + p$ is acceptable for all $p \geqslant 2m_0$. Rule 1 says that $2p$ is acceptable for all $p \geqslant 50$. So 100, 102 and 402 are acceptable, and by induction every acceptable number is $\geqslant 100$.

Example 6.3. Here is a broad rubric on $\mathbb{N}$, consisting of two broad rules. Broad rule 0 is nullary and sends $[\,]$ to the wide rubric described in Example 6.2. Broad rule 1 is unary. It sends $[7]$ to the the following wide rubric, consisting of one wide rule.

- Rule 0 is binary and sends $\begin{bmatrix} m_0 \\ m_1 \end{bmatrix} \mapsto (m_0 + m_1 + 500p)_{p\geqslant 9}$.

It sends $[100]$ to the following wide rubric, consisting of three wide rules.

- Rule 0 is ternary and sends $\begin{bmatrix} m_0 \\ m_1 \\ m_2 \end{bmatrix} \mapsto (m_0 + m_1 m_2 + p)_{p\geqslant 17}$.

- Rule 1 is nullary and sends $[\,] \mapsto (p)_{p \geqslant 1000}$.
- Rule 2 is binary and sends $\begin{bmatrix} m_0 \\ m_1 \end{bmatrix} \mapsto (m_1 + p)_{p \geqslant 4}$.

And it sends $[n]$, for $n \in \mathbb{N} \setminus \{7, 100\}$, to the empty wide rubric.

Informally, these rules prescribe when an element of $\mathbb{N}$ is acceptable. For example, if 100 is acceptable and m_0, m_1, m_2 are too, then so is $m_0 + m_1 m_2 + p$ for all $p \geqslant 17$. So 100, 102, 402 and 107 are acceptable, and by induction every acceptable number is $\geqslant 100$.

To make the notion of "acceptable element" precise, we proceed as follows.

Definition 6.4. Let C be a class. A subclass X is

- $\langle K, R \rangle$*-closed*, for a wide rule $\langle K, R \rangle$ on C, when the family $R(y)$ is within X for all $y \in X^K$.
- $\mathcal{R}$*-complete*, for a wide rubric $\mathcal{R} = (r_i)_{i \in I}$ on C, when X is r_i-closed for all $i \in I$.
- $\langle L, T \rangle$*-closed*, for a broad rule $\langle L, T \rangle$ on C, when X is $T(x)$-complete for all $x \in X^L$.
- $\mathcal{S}$*-complete* for a broad rubric $\mathcal{S} = (s_j)_{j \in J}$ on C, when X is s_j-closed for all $j \in J$.

We also give an alternative formulation of completeness.

Definition 6.5. Let C be a class.

1. Let $\mathcal{R} = (\langle K_i, R_i \rangle)_{i \in I}$ be a wide rubric on C. An $\mathcal{R}$*-plate* is a triple $\langle i, y, p \rangle$ consisting of

   - an index $i \in I$ and K_i-tuple y within C, giving
     $$R_i(y) \quad = \quad (z_p)_{p \in P}$$
   - and index $p \in P$

   and the *result* of this plate is z_p. The plate is *within* a subclass X when the tuple y is.

2. Let $\mathcal{R} = (\langle L_j, T_j \rangle)_{j \in J}$ be a broad rubric on C. An $\mathcal{R}$*-plate* is a 5-tuple $\langle j, x, i, y, p \rangle$ consisting of

   - an index $j \in J$ and L_j-tuple x within C, giving
     $$T_j(x) \quad = \quad (\langle K_i, R_i \rangle)_{i \in I}$$
   - an index $i \in I$ and K_i-tuple y within C, giving
     $$R_i(y) \quad = \quad (z_p)_{p \in P}$$

- and an index $p \in P$

and the *result* of this plate is z_p. The plate is *within* a subclass X when the tuples x and y are.

Definition 6.6. For any rubric $\mathcal{R}$ on a class C, we define the monotone endofunction $\Gamma_{\mathcal{R}}$ on $\mathcal{P}^{\blacksquare}(C)$, sending X to the class of all results of $\mathcal{R}$-plates within X.

Clearly a subclass of C is $\mathcal{R}$-complete iff $\Gamma_{\mathcal{R}}$-prefixed.

The least $\mathcal{R}$-complete subclass of C, if it exists, is said to be *generated by* $\mathcal{R}$ and written $\mathsf{Gen}(\mathcal{R})$. I do not know whether this always exists in our base theory. But in any case, a rubric's purpose is to generate a subset, not merely a subclass. So we formulate the following principles.

- The *Wide Subset Generation* scheme says that any wide rubric on a class generates a subset.
- The *Broad Subset Generation* scheme says that any broad rubric on a class generates a subset.

Proposition 6.7.

1. *Wide Subset Generation implies Wide Infinity.*
2. *Broad Subset Generation implies Broad Infinity.*

Proof.

1. Let $S = (K_i)_{i \in I}$ be a signature. A class is H_S-prefixed iff it is $\hat{S}$-complete, writing $\hat{S}$ for the following wide rubric on $\mathfrak{T}$: it consists of I rules, and rule $i \in I$ has arity K_i and sends the tuple x to the singleton $(\langle i, x \rangle)$. So $\mathsf{Wide}(S)$ is generated by $\hat{S}$, and is therefore a set if Wide Subset Generation holds.

2. Let G be a broad signature. A class is Maybe_G-prefixed iff it is $\hat{G}$-complete, writing $\hat{G}$ for the broad rubric on $\mathfrak{T}$ consisting of the following:

   - A nullary broad rule that returns the wide rubric consisting of a nullary broad rule that returns the singleton $(\mathsf{Nothing})$.
   - A unary broad rule that sends $[w]$, where $Gw = (K_i)_{i \in I}$ to the wide rubric consisting of I rules, where rule I has arity K_i and sends x to the singleton $(\mathsf{Just}\langle w, i, x \rangle)$.

   So $\mathsf{Broad}(G)$ is generated by $\hat{G}$, and is therefore a set if Broad Subset Generation holds. □

6.2 Application: Grothendieck universes

For this section, **assume Wide Infinity**.

As promised in Section 2.1, we see the utility of Broad Subset Generation: it gives Grothendieck universes without any detour via notions of cardinal or ordinal.

Definition 6.8. A *Grothendieck universe* is a transitive set $\mathfrak{U}$ with the following properties.

- $\mathbb{N} \in \mathfrak{U}$.
- For every set of sets $\mathcal{A} \in \mathfrak{U}$, we have $\bigcup \mathcal{A} \in \mathfrak{U}$.
- For every set $A \in \mathfrak{U}$, we have $\mathcal{P}A \in \mathfrak{U}$.
- For every set $K \in \mathfrak{U}$ and K-tuple $[a_k]_{k \in K}$ within $\mathfrak{U}$, we have $\{a_k \mid k \in K\} \in \mathfrak{U}$.

Proposition 6.9. *Broad Subset Generation implies the "Axiom of Universes": For every set X, there is a least Grothendieck universe $\mathfrak{U}$ with $X \subseteq \mathfrak{U}$.*

Proof. We define a broad rubric $\mathcal{B}$ on $\mathfrak{T}$, consisting of two broad rules. Broad rule 0 is nullary and sends $[\,]$ to the following wide rubric indexed by $X + 4$.

- To achieve $X \subseteq \mathfrak{U}$, rule $\mathsf{inl}\ x$ (for $x \in X$) is nullary, and sends $[\,]$ to (x).
- To achieve transitivity, rule $\mathsf{inr}\ 0$ is unary, sending $[A]$ to $(b)_{b \in A}$ if A is a set, and the empty family otherwise.
- Rule $\mathsf{inr}\ 1$ is nullary, and sends $[\,]$ to $(\mathbb{N})$.
- Rule $\mathsf{inr}\ 2$ is unary, sending $[\mathcal{A}]$ to $(\bigcup \mathcal{A})$ if $\mathcal{A}$ is a set of sets, and the empty signature otherwise.
- Rule $\mathsf{inr}\ 3$ has arity 1, sending $[A]$ to $(\mathcal{P}A)$ if A is a set, and the empty signature otherwise.

Broad rule 1 is unary. For any set B, it sends $[B]$ to the rubric consisting of one B-ary rule that sends $[a_k]_{k \in B}$ to $(\{a_k \mid k \in B\})$. If b is not a set, then Broad Rule 1 sends $[b]$ to the empty rubric. This completes the definition of $\mathcal{B}$. The set that it generates is the desired Grothendieck universe. $\square$

6.3 Derivations

Intuitively, when we have a rubric on a class C, each acceptable element $x \in C$ has one or more "derivations" that explain why it is acceptable.

Example 6.10. For the wide rubric in Example 6.2:

- $\langle 1, [\,], 50 \rangle$ derives 100.
- $\langle 1, [\,], 51 \rangle$ derives 102.

- $\langle 0, \begin{bmatrix} \langle 1, [\,], 50 \rangle \\ \langle 1, [\,], 50 \rangle \end{bmatrix}, 202 \rangle$ and $\langle 0, \begin{bmatrix} \langle 1, [\,], 50 \rangle \\ \langle 1, [\,], 51 \rangle \end{bmatrix}, 200 \rangle$ derive 402.

Note that each derivation is a triple consisting of an index, a tuple of derivations, and another index.

Example 6.11. For the broad rubric in Example 6.3:

- $\langle 0, [\,], 1, [\,], 50 \rangle$ derives 100.
- $\langle 0, [\,], 1, [\,], 51 \rangle$ derives 102.
- $\langle 0, [\langle 0, [\,], 1, [\,], 50 \rangle], 2, \begin{bmatrix} \langle 0, [\,], 1, [\,], 50 \rangle \\ \langle 0, [\,], 1, [\,], 51 \rangle \end{bmatrix}, 5 \rangle$ derives 107.

Note that each derivation is a 5-tuple consisting of an index, a tuple of derivations, another index, another tuple of derivations, and yet another index.

Given a rubric $\mathcal{R}$ on a class C, we would like to define the class $\mathsf{Deriv}_{\mathcal{R}}$ of all $\mathcal{R}$-*derivations*. Each $\mathcal{R}$-derivation x will have an *overall result* $\mathsf{O}_{\mathcal{R}}(x) \in C$, and we shall call $(\mathsf{Deriv}_{\mathcal{R}}, \mathsf{O}_{\mathcal{R}}(x))$ the $\mathcal{R}$-*derivational* class-family within C.

To do this, we adapt the notion of $\mathcal{R}$-plate (Definition 6.5).

Definition 6.12. Let C be a class.

1. Let $\mathcal{R} = (\langle K_i, R_i \rangle)_{i \in I}$ be a wide rubric on C. For a class-family (M, F) within C, an $(\mathcal{R}, M, F)$-*plate* is a triple $s = \langle i, y, p \rangle$ consisting of
   - an index $i \in I$ and K_i-tuple y within M, giving
   $$R_i[Fx_k]_{k \in K_i} \quad = \quad (b_p)_{p \in P}$$
   - and an index $p \in P$

   and the *result* of this plate is b_p.

2. Let $\mathcal{R} = (\langle L_j, T_j \rangle)_{j \in J}$ be a broad rubric on C. For a class-family (M, F) within C, an $(\mathcal{R}, M, F)$-*plate* is a 5-tuple $\langle j, y, i, x, p \rangle$ consisting of
   - an index $j \in J$ and L_j-tuple y within M, giving
   $$T_j[Fy_l]_{l \in L_j} \quad = \quad (\langle K_i, R_i \rangle)_{i \in I}$$
   - an index $i \in I$ and K_i-tuple x within M, giving
   $$R_i[Fx_k]_{k \in K_i} \quad = \quad (b_p)_{p \in P}$$
   - and an index $p \in P$

   and the *result* of this plate is b_p.

Definition 6.13. Let $\mathcal{R}$ be a rubric on a class C. We define the monotone endofunction $\Delta_{\mathcal{R}}$ on $\mathsf{Fam}^{\blacksquare}(C)$ sending (M, F) to (N, G), where N is the class of all $(\mathcal{R}, M, F)$-plates and G sends each such plate to its result.

Proposition 6.14. *Let $\mathcal{R}$ be a rubric on a class C. Then $\Delta_{\mathcal{R}}$ has a least prefixpoint.*

Proof. A class-family within C is the same thing as a partial function $\mathfrak{T} \rightharpoonup C$. So we use Proposition 4.7, by expressing $\Delta_{\mathcal{R}}$ as $\Delta^{L}_{(D,<)}$, for a set-based scaffold $(D,<)$ with functionalization L on $(C)_{x\in\mathfrak{T}}$.

We give the wide case, as the broad case is similar.

Say that a *wide pre-plate* is a tuple $x = \langle i, a, p\rangle$, where i and p are anything and a is a K-tuple for some set K. Its *component set* is $J(x) \stackrel{\text{def}}{=} \{a_k \mid k \in K\}$. This gives a set-based scaffold $(D,<)$ on $\mathfrak{T}$, where D is the class of all wide pre-plates and $<$ is componenthood. For $x \in D$ and a function f on $J(x)$, we define $\widehat{f}(x) \stackrel{\text{def}}{=} \langle i, [f(a_k)]_{k\in K}, p\rangle$.

Now let $\mathcal{R} = (\langle K_i, R_i\rangle)_{i\in I}$ be a wide rubric on C. Then $(D,<)$ has the following functionalization L to $(C)_{x\in\mathfrak{T}}$. For a wide pre-plate $x = \langle i, [a_k]_{k\in K}, p\rangle$, we define $\mathsf{Dom}(L_x)$ to be the class of all functions $f : J(x) \to C$ such that $\widehat{f}(x)$ is an $\mathcal{R}$-plate, and $\overline{L_x}$ sends such a function f to the result of $\widehat{f}(x)$.

It is then evident that $\Delta_{\mathcal{R}} = \Delta^{L}_{(D,<)}$. □

Thus we define the $\mathcal{R}$-derivational class-family $(\mathsf{Deriv}_{\mathcal{R}}, \mathsf{O}_{\mathcal{R}}) \stackrel{\text{def}}{=} \mu\Delta_{\mathcal{R}}$. We want to know whether this is a family, i.e., whether $\mathsf{Deriv}_{\mathcal{R}}$ is a set. So we formulate the following principles.

- *Wide Derivation Set*: Any wide rubric on a class has a derivation set.
- *Broad Derivation Set*: Any broad rubric on a class has a derivation set.

Categorical remark. For a class C, the category of class-families within C is written $\mathbf{Class}/C$, using "slice category" notation. An endofunctor F on $\mathbf{Class}/C$ is *monotone* when it preserves inclusions. It follows that any class-family that is an initial F-algebra with identity structure is also a least F-prefixpoint.

For example, let $\mathcal{R}$ be a rubric on C. Then $\Delta_{\mathcal{R}}$ extends to a monotone endofunctor on $\mathbf{Class}/C$. Moreover, the $\mathcal{R}$-derivational class-family is an initial $\Delta_{\mathcal{R}}$-algebra with identity structure.,

6.4 Application: Tarski-style universes

For this section, **assume Wide Infinity**.

In the type theory literature [ML84], a "Tarski-style universe" is a family of types that is indexed by a set of "codes" and closed under various constructions, such as $\sum$ and $\prod$ and W-type formation. Furthermore, the existence of such universes follows from various "induction-recursion" principles [DS06, GH16].

In a similar way, we show that Broad Derivation Set yields the existence of that arbitrarily large Tarski-style universes. (It does not, to my knowledge, yield the existence of a Grothendieck universe.)

To begin, we define several injections with disjoint range:

$$\begin{aligned}
\mathrm{embed}(x) &\stackrel{\mathrm{def}}{=} \langle 0, \langle x \rangle \rangle \\
\mathrm{zero} &\stackrel{\mathrm{def}}{=} \langle 1, \langle \rangle \rangle \\
\mathrm{two} &\stackrel{\mathrm{def}}{=} \langle 2, \langle \rangle \rangle \\
\mathrm{eq}(x,y,z) &\stackrel{\mathrm{def}}{=} \langle 3, \langle x,y,z \rangle \rangle \\
\mathrm{sigma}(x,y) &\stackrel{\mathrm{def}}{=} \langle 4, \langle x,y \rangle \rangle \\
\mathrm{wtype}(x,y) &\stackrel{\mathrm{def}}{=} \langle 5, \langle x,y \rangle \rangle
\end{aligned}$$

Definition 6.15. Let $(B_a)_{a \in A}$ be family of sets. A *Tarski-style universe* extending it is a family of sets $(D_m)_{m \in M}$ satisfying the following conditions.

- For all $a \in A$, we have $\mathrm{embed}(a) \in M$ with $D_{\mathrm{embed}(a)} = B_a$.
- We have $\mathrm{zero} \in M$ with $D_{\mathrm{zero}} = \emptyset$.
- We have $\mathrm{two} \in M$ with $D_{\mathrm{two}} = \{0, 1\}$.
- For any $m \in M$ and $a, b \in D_m$, we have $\mathrm{eq}(m, a, b) \in M$ with $D_{\mathrm{eq}(m,a,b)}$ is 1 if $a = b$ and $\emptyset$ otherwise.
- For any $m \in M$ and function $g : D_m \to M$, we have $\mathrm{sigma}(m, g) \in M$ with $D_{\mathrm{sigma}(m,g)} = \sum_{k \in D_m} D_{g(m)}$.
- For any $m \in M$ and function $g : D_m \to M$, we have $\mathrm{wtype}(m, g) \in M$ with $D_{\mathrm{wtype}(m,g)} = \mathsf{Wide}(D_{g(m)})_{k \in D_m}$.

Proposition 6.16. *Broad Derivation Set implies that, for any family of sets, there is a least Tarski-style universe extending it.*

Proof. Let $(B_a)_{a \in A}$ be a family of sets. We define a broad rubric $\mathcal{B}$ on $\mathfrak{S}$, consisting of two broad rules. Broad rule 0 is nullary and sends $[\,]$ to the following rule indexed by $A + 2$.

- Rule $\mathsf{inl}\ a$ (for $a \in A$) has arity 0 and sends $[\,]$ to (B_a).
- Rule $\mathsf{inr}\ 0$ has arity 0 and sends $[\,]$ to $(\emptyset)$.
- Rule $\mathsf{inr}\ 1$ has arity 1 and sends $[\,]$ to $(\{0, 1\})$.

Broad rule 1 is unary, and sends $[D]$ to the following rubric indexed by $D^2 + 2$:

- Rule $\mathsf{inl}\ \langle d, e \rangle$ (for $d, e \in D$) has arity 0 and sends $[\,]$ to $(1_{d=e})$
- Rule $\mathsf{inr}\ 0$ has arity D and sends $[E_k]_{k \in D}$ to $(\sum_{k \in K} E_k)$.
- Rule $\mathsf{inr}\ 1$ has arity D and sends $[E_k]_{k \in D}$ to $(\mathsf{Wide}(E_k)_{k \in D})$.

This completes the definition of $\mathcal{B}$. Let $(E_n)_{n\in N}$ be its derivation-indexed family. We recursively define the injection θ on N that sends

$$\begin{array}{rcl} \langle 0, [\,], \mathsf{inl}\ a, [\,], *\rangle & \mapsto & \text{embed}(a) \\ \langle 0, [\,], \mathsf{inr}\ 0, [\,], *\rangle & \mapsto & \text{zero} \\ \langle 1, [n], \mathsf{inl}\ \langle d, e\rangle, [\,], *\rangle & \mapsto & \text{eq}(\theta n, d, e) \\ \langle 1, [n], \mathsf{inr}\ 0, g, *\rangle & \mapsto & \text{sigma}(\theta n, \theta \circ g) \\ \langle 1, [n], \mathsf{inr}\ 1, g, *\rangle & \mapsto & \text{wtype}(\theta n, \theta \circ g) \end{array}$$

Let M be its range. Then $(E_{\theta^{-1}(m)})_{m\in M}$ is the desired Tarski-style universe. $\square$

6.5 Proving the Derivation Set principles

The following is the central result of the paper (at least for people who do not accept AC), since it says that *plausible principles entail useful ones.* The converse of each statement is also true, but postponed to the next section.

Proposition 6.17.

1. *Wide Infinity implies Wide Derivation Set.*
2. *Broad Infinity implies Broad Derivation Set.*

Proof.

1. Let $\mathcal{R} = (\langle K_i, R_i\rangle)_{i\in I}$ be a wide rubric on a class C. Define the signature S to be $(K_i)_{i\in I}$. We recursively associate to each $t \in \mathsf{Wide}(S)$ a family (M_t, F_t) within C as follows. For $t = \langle i, [t_k]_{k\in K_i}\rangle$, an element of M_t is a triple $\langle i, [m_k]_{k\in K_i}, p\rangle$ where $i \in I$ and $m \in \prod_{k\in K_i} M_{t_k}$ with $R_i[F_{t_k} m_k]_{k\in K_i} = (y_p)_{p\in P}$ and $p \in P$, and F_t sends this element to y_p. For any $t, t' \in \mathsf{Wide}(S)$, if $M_t \cap M'_t$ is inhabited, then $t = t'$, by induction on t.

   We define the family (M, F) within C to be the union of all these. Thus we define $M \stackrel{\text{def}}{=} \bigcup_{t\in\mathsf{Wide}(S)} M_t$, and F sends $m \in M_t$ to $F_t m$. It is then evident that (M, F) is the derivational family of $\mathcal{R}$.

2. To show Broad Infinity implies Broad Derivation Set, let $\mathcal{R} = (\langle L_j, T_j\rangle)_{j\in J}$ be a broad rubric on a class C. We recursively define the function θ on $\mathsf{Deriv}_{\mathcal{R}}$ that sends $\langle j, [x_l]_{l\in L_j}, i, [y_k]_{k\in K_i}, p\rangle$ to

   $$\mathsf{Just}\langle \mathsf{Just}\langle \mathsf{Just}\langle \mathsf{Nothing}, j, [\theta x_i]_{i\in L_j}\rangle, i, [\theta y_k]_{k\in K_i}\rangle, p, [\,]\rangle$$

   By induction, θ is injective. Let G be the broad signature that sends

   - $\mathsf{Nothing}$ to $(L_j)_{j\in J}$
   - $\mathsf{Just}\langle \mathsf{Nothing}, j, [\theta x_l]_{l\in L_j}\rangle$ obtained from

– an index $j \in J$ and $x \in \mathsf{Deriv}_{\mathcal{R}}^{L_j}$, giving

$$T_j[\mathsf{O}_{\mathcal{R}}(x_l)]_{l\in L_j} \quad = \quad (\langle K_i, R_i\rangle)_{i\in I}$$

to $(K_i)_{i\in I}$

- $\mathsf{Just}\langle\mathsf{Just}\langle\mathsf{Nothing}, j, [\theta x_i]_{i\in L_j}\rangle, i, [\theta y_k]_{k\in K_i}\rangle$ obtained from

  – an index $j \in J$ and $x \in \mathsf{Deriv}_{\mathcal{R}}^{L_j}$, giving

$$T_j[\mathsf{O}_{\mathcal{R}}(x_l)]_{l\in L_j} \quad = \quad (\langle K_i, R_i\rangle)_{i\in I}$$

  – and an index $j \in J$ and $x \in \mathsf{Deriv}_{\mathcal{R}}^{L_j}$, giving

$$R_i[\mathsf{O}_{\mathcal{R}}(x_k)]_{k\in K_i} \quad = \quad (b_p)_{p\in P}$$

  to $(\emptyset)_{p\in P}$

- everything else to the empty signature.

For every $\mathcal{R}$-derivation x, we see that θx is a G-broad number, by induction. Since the G-broad numbers form a set, the $\mathcal{R}$-derivations do too. □

6.6 Injective rubrics

Our next task is to prove the converse of each part of Proposition 6.17. In order to achieve this without AC, we have to pay attention to injectivity.

Definition 6.18.

1. An *injective class-family* (M, F) consists of a class M and an injective function F on M.
2. For a class C, the collection of all injective class-families within C is written $\mathsf{InjFam}^{\blacksquare}(C)$.
3. A rubric on a class C is *injective* when any two $\mathcal{R}$-plates with the same result are equal.

Proposition 6.19. *Let $\mathcal{R}$ be an injective rubric on a class C.*

1. *$\Delta_{\mathcal{R}}$ restricts to an endofunction on $\mathsf{InjFam}^{\blacksquare}(C)$, and the square*

$$\begin{tikzcd} \mathsf{InjFam}^{\blacksquare}(C) \arrow[r, "\Delta_{\mathcal{R}}"] \arrow[d, "\mathsf{Range}"'] & \mathsf{InjFam}^{\blacksquare}(C) \arrow[d, "\mathsf{Range}"] \\ \mathcal{P}^{\blacksquare}(C) \arrow[r, "\Gamma_{\mathcal{R}}"'] & \mathcal{P}^{\blacksquare}(C) \end{tikzcd}$$

commutes.

2. *The $\mathcal{R}$-derivational class-family is injective, and its range is generated by $\mathcal{R}$.*

Proof. We prove the wide case, as the broad case is similar.

1. Let (M, F) be an injective class-family within C. We have a result-preserving bijection $\widehat{F}$ from the class of all $(\mathcal{R}, M, F)$-plates to the class of all $\mathcal{R}$-plates within $B \stackrel{\text{def}}{=} \mathsf{Range}(M, F)$, sending $\langle i, [a_k]_{k \in K}, p\rangle$ to $\langle i, [F(a_k)]_{k \in K}, p\rangle$.

   For any $(\mathcal{R}, M, F)$-plates x and x' with the same result c, the $\mathcal{R}$-plates $\widehat{F}(x)$ and $\widehat{F}(x')$ have result c, and are therefore equal by injectivity of $\mathcal{R}$, giving $x = x'$ by injectivity of $\widehat{F}$. So $\Delta_{\mathcal{R}}(M, F)$ is injective. Its range is the class of results of all $(\mathcal{R}, M, F)$-plates, which is the class of results of all $\mathcal{R}$-plates within B, which is $\Gamma_{\mathcal{R}} B$. So the square commutes.

2. The fact that $(\mathsf{Deriv}_{\mathcal{R}}, \mathsf{O}_{\mathcal{R}})$ is a $\Delta_{\mathcal{R}}$-fixpoint means that an $\mathcal{R}$-derivation d with overall result c is the same thing as a $(\mathcal{R}, \mathsf{Deriv}_{\mathcal{R}}, \mathsf{O}_{\mathcal{R}})$-plate with result c.

   Therefore, $\widehat{\mathsf{O}_{\mathcal{R}}}$ is a bijection from $\mathsf{Deriv}_{\mathcal{R}}$ to the class of $\mathcal{R}$-plates within $Q \stackrel{\text{def}}{=} \mathsf{Range}(\mathsf{Deriv}_{\mathcal{R}}, \mathsf{O}_{\mathcal{R}})$. Moreover, the result of $\widehat{\mathsf{O}_{\mathcal{R}}}(d)$ is the overall result of d.

   For any $\mathcal{R}$-derivations d, d' with the same overall result, we show $d = d'$, by induction on d as follows. Write $d = \langle i, [e_k]_{k \in K}, p\rangle$ and $d' = \langle i', [e'_k]_{k \in K'}, p'\rangle$. Their overall result is also the result of the $\mathcal{R}$-plates $\widehat{\mathsf{O}_{\mathcal{R}}}(d)$ and $\widehat{\mathsf{O}_{\mathcal{R}}}(d')$. Injectivity of $\mathcal{R}$ gives

$$\begin{array}{rrcl} & \widehat{\mathsf{O}_{\mathcal{R}}}(d) & = & \widehat{\mathsf{O}_{\mathcal{R}}}(d') \\ \text{i.e.,} & \langle i, [\mathsf{O}_{\mathcal{R}}(e_k)]_{k \in K}, p\rangle & = & \langle i', [\mathsf{O}_{\mathcal{R}}(e'_k)]_{k \in K'}, p'\rangle \end{array}$$

   Thus $i = i'$, and for all $k \in K$ we have $\mathsf{O}_{\mathcal{R}} e_k = \mathsf{O}_{\mathcal{R}} e'_k$ giving $e_k = e'_k$ by the inductive hypothesis, and $p = p'$. So $d = d'$. This means that $(\mathsf{Deriv}_{\mathcal{R}}, \mathsf{O}_{\mathcal{R}})$ is injective.

   We show that Q is $\mathcal{R}$-complete as follows. Let x be an $\mathcal{R}$-plate within Q whose result is c. Then $x = \widehat{\mathsf{O}_{\mathcal{R}}}(d)$ for a unique $\mathcal{R}$-derivation d, and the latter has overall result c. So $c \in Q$.

   Lastly we show that Q is included in any $\mathcal{R}$-complete class X. To do this, we show that every $\mathcal{R}$-derivation d has overall result in X, by induction on d, as follows. For $d = \langle i, [e_k]_{k \in K_i}, p\rangle$, its overall result c is the result of $\widehat{\mathsf{O}_{\mathcal{R}}}(d) = \langle i, [\mathsf{O}_{\mathcal{R}}(e_k)]_{k \in K_i}, p\rangle$. For all $k \in K_i$, we have $\mathsf{O}_{\mathcal{R}}(e_k) \in X$ by the inductive hypothesis. So $c \in X$ by $\mathcal{R}$-completeness of X. □

We define two additional principles:

- The *Injective Wide Subset Generation* scheme says that any injective wide rubric on a class generates a subset.

- The *Injective Broad Subset Generation* scheme says that any injective broad rubric on a class generates a subset.

Proposition 6.20.

1. *Wide Infinity, Wide Derivation Set and Injective Wide Subset Generation are equivalent.*

2. *Full Broad Infinity, Broad Derivation Set and Injective Broad Subset Generation are equivalent.*

Proof. Proposition 6.19(2) gives Wide Derivation Set $\Rightarrow$ Injective Wide Subset Generation and Broad Derivation Set $\Rightarrow$ Injective Broad Subset Generation.

The proof of Proposition 6.7(1) gives Injective Wide Subset Generation $\Rightarrow$ Wide Infinity, since the rubric used is injective. Likewise, the proof of Proposition 6.7(2) gives Injective Broad Subset Generation $\Rightarrow$ Broad Infinity, since the rubric used is injective.

Proposition 6.17 gives the rest. $\square$

6.7 Quasiwide rubrics

This section is not used in the sequel.

So far, we have considered a rubric $\mathcal{R}$ on a class C, and asked whether $\mathcal{R}$ generates a subset and whether $\mathsf{Deriv}_{\mathcal{R}}$ is a set. In the case that C is a set, it is clear that $\mathcal{R}$ generates a subset, but the status of $\mathsf{Deriv}_{\mathcal{R}}$ remains unclear. So let us formulate the *Derivation Set on a Set* scheme: Every rubric on a set has a derivation set. This principle turns out to be equivalent to Wide Infinity, in line with a similar result in [HMG+13].

In this section, we liberalize our principles concerning wide rubrics to include this and other examples. Doing so helps to clarify the difference between what Broad Infinity delivers and what Wide Infinity delivers.

For any rule or rubric $\mathcal{R}$, we write $\mathsf{Arit}(\mathcal{R})$ for the class of all arities that appear in it. The following spells this out.

Definition 6.21. Let C be a class.

1. For a wide rule $r = \langle K, R \rangle$ on C, we write

$$\mathsf{Arit}(r) \stackrel{\text{def}}{=} \{K\}$$

2. For a wide rubric $\mathcal{R} = (r_i)_{i \in I}$ on C, we write

$$\mathsf{Arit}(\mathcal{R}) \stackrel{\text{def}}{=} \bigcup_{i \in I} r_i$$

3. For a broad rule $s = \langle L, S \rangle$ on C, we write

$$\mathsf{Arit}(s) \stackrel{\text{def}}{=} \{L\} \cup \bigcup_{x \in C^L} \mathsf{Arit}(S(x))$$

4. For a broad rubric $\mathcal{S} = (s_j)_{j \in J}$ on C, we write

$$\mathsf{Arit}(\mathcal{S}) \quad \stackrel{\text{def}}{=} \quad \bigcup_{j \in J} s_j$$

For a wide rubric $\mathcal{R}$, it is evident that $\mathsf{Arit}(\mathcal{R})$ is a set. Accordingly, say that a *quasiwide rubric* is a broad rubric $\mathcal{S}$ such that $\mathsf{Arit}(\mathcal{S})$ is a set. For our purposes, wide and quasiwide are essentially the same:

Proposition 6.22. *Let C be a class.*

1. *Any wide rubric $\mathcal{R}$ on C gives*
   - *a quasiwide rubric $\mathcal{S}$ on C that satisfies $\Gamma_{\mathcal{R}} = \Gamma_{\mathcal{S}}$, and is injective iff $\mathcal{R}$ is*
   - *a natural isomorphism $\Delta_{\mathcal{R}} \cong \Delta_{\mathcal{S}}$.*

2. *Conversely, any quasiwide rubric $\mathcal{S}$ on C gives*
   - *a wide rubric $\mathcal{R}$ on C that satisfies $\Gamma_{\mathcal{S}} = \Gamma_{\mathcal{R}}$, and is injective iff $\mathcal{S}$ is*
   - *a natural isomorphism $\Delta_{\mathcal{S}} \cong \Delta_{\mathcal{R}}$.*

Proof.

1. Define $\mathcal{S}$ to consist of single nullary broad rule (indexed by 0) that sends $[\,]$ to $\mathcal{R}$. For a class-family (M, F) within C, an $(\mathcal{R}, M, F)$-plate $\langle i, x, p \rangle$ corresponds to the $(\mathcal{S}, M, F)$-plate $\langle 0, [\,], i, x, p \rangle$.

2. First some notation: given an I-tuple x and a J-tuple y, we write $x \otimes y$ for their *copairing*, i.e., the $(I + J)$-tuple whose kth component is x_i or y_j according as k is $\mathsf{inl}\ i$ or $\mathsf{inr}\ j$.

   Let $\mathcal{S} = (\langle L_j, S_j \rangle)_{j \in J}$. We form the wide rubric

   $$\mathcal{R} \quad \stackrel{\text{def}}{=} \quad (\langle L + K, R_{L,K} \rangle)_{\langle L,K \rangle \in \mathsf{Arit}(\mathcal{R})^2}$$

   where $R_{L,K}$ sends $x \otimes y$, for $x \in C^L$ and $y \in C^K$, to the family $(c_n)_{n \in N}$ defined as follows. An element of N is a triple $\langle j, i, p \rangle$ such that

   - $j \in J$ and $L = L_j$ and $S_j(x) = (\langle K_i, R_i \rangle)_{i \in I}$
   - $i \in I$ and $K = K_i$ and $R_i(y) = (z_p)_{p \in P}$
   - $p \in P$

   with $c_{\langle j,i,p \rangle} \stackrel{\text{def}}{=} z_p$. For a class-family (M, F) within C, an $(\mathcal{S}, M, F)$-plate $\langle j, [x_l]_{l \in L_j}, i, [y_k]_{k \in K_i}, p \rangle$ corresponds to the $(\mathcal{R}, M, F)$-plate $\langle \langle L, K \rangle, x \otimes y, \langle j, i, p \rangle \rangle$. $\square$

Corollary 6.23.

1. *Wide Derivation Set is equivalent to* Quasiwide Derivation Set*: Any quasiwide rubric on a class has a derivation set.*

2. *Wide Subset Generation is equivalent to* Quasiwide Subset Generation*: Any quasiwide rubric on a class generates a subset.*

3. *Injective Wide Subset Generation is equivalent to* Injective Quasiwide Subset Generation*: Any injective quasiwide rubric on a class generates a subset.*

Proof. We apply Proposition 3.2 to get part 1, and the rest is straightforward. □

Now we can easily handle the question of rubrics on a set.

Proposition 6.24.

1. (Assuming Powerset.) *Any broad rubric on a set is quasiwide.*

2. *Derivation Set on a Set is equivalent to Wide Infinity.*

Proof.

1. Let $\mathcal{R} = (\langle L_j, S_j \rangle)_{j \in J}$ be a broad rubric on a set A. Then we have

$$\mathsf{Arit}(\mathcal{R}) \quad = \quad \bigcup_{j \in J} (\{L_j\} \cup \bigcup_{x \in A^{L_j}} \mathsf{Arit}(S_j(x)))$$

whichis a set by Exponentiation.

2. Wide Infinity implies Powerset and Quasiwide Derivation Set. We thus obtain Derivation Set on a Set by part 1.

To show that Derivation Set on a Set implies Wide Infinity: given a signature $S = (K_i)_{i \in I}$, form a wide rubric on 1 via

$$\mathcal{R} \quad \stackrel{\text{def}}{=} \quad (\langle K_i, [*]_{k \in K} \mapsto (*) \rangle)_{i \in I}$$

Recursively define the bijection $\theta : \mathsf{Wide}(S) \cong \mathsf{Deriv}_{\mathcal{R}}$ sending $\langle i, [x_k]_{k \in K_i} \rangle$ to $\langle i, [\theta(x_k)]_{k \in K}, * \rangle$. Thus $\mathsf{Wide}(S)$ is a set iff $\mathsf{Deriv}_{\mathcal{R}}$ is. □

6.8 Powerset and rubric functions

Here is a result that we shall use frequently, e.g., in Proposition 7.1 and the proof of Proposition 9.7.

Proposition 6.25. *Each of the following is equivalent to Powerset.*

1. *For any rubric $\mathcal{R}$ on a class C, the endofunction $\Gamma_{\mathcal{R}}$ restricts to one on $\mathcal{P}C$.*

2. *For any rubric $\mathcal{R}$ on a class C, the endofunction $\Delta_{\mathcal{R}}$ restricts to one on* $\mathsf{Fam}(C)$.

Proof. Recall that Powerset is equivalent to Exponentiation (Proposition 3.1).

Assume Exponentiation. The $\mathcal{R}$-plates within a set form a set, giving (1). For any family (M, F) within C, the $(\mathcal{R}, M, F)$-plates form a set, giving (2).

Conversely, let A and B be sets. Obtain the (injective) wide rubric $\mathcal{R}$ on $\mathfrak{T}$ consisting of a single A-ary rule that sends $x \in \mathfrak{T}^A$ to (x). Then

$$\begin{array}{rcl} \Gamma_{\mathcal{R}}(B) & = & B^A \\ \Delta_{\mathcal{R}}(b)_{b\in B} & = & (x)_{\langle *,x,*\rangle \in 1\times B^A \times 1} \\ & \cong & (x)_{x\in B^A} \quad \text{via } \langle *,x,*\rangle \mapsto x. \end{array}$$

So if either $\Gamma_{\mathcal{R}}(B)$ or $\Delta_{\mathcal{R}}(b)_{b\in B}$ is a set, then so is B^A. □

7 The use of choice

7.1 Consequences of AC

Our next task is to prove the implication Wide Derivation Set $\Rightarrow$ Wide Subset Generation, and likewise Broad, assuming AC. Here is the key result that we use:

Proposition 7.1.

1. (Assuming AC and Wide Infinity.) *Let $\mathcal{R}$ be a wide rubric on a class C. Then the square*

$$\begin{tikzcd} \mathsf{Fam}(C) \arrow[r, "\Delta_{\mathcal{R}}"] \arrow[d, "\mathsf{Range}"'] & \mathsf{Fam}(C) \arrow[d, "\mathsf{Range}"] \\ \mathcal{P}C \arrow[r, "\Gamma_{\mathcal{R}}"'] & \mathcal{P}C \end{tikzcd}$$

commutes, and the range of the $\mathcal{R}$-derivational family is generated by $\mathcal{R}$.

2. (Assuming AC and Broad Infinity.) *Let $\mathcal{R}$ be a broad rubric on a class C. Then the square*

$$\begin{tikzcd} \mathsf{Fam}(C) \arrow[r, "\Delta_{\mathcal{R}}"] \arrow[d, "\mathsf{Range}"'] & \mathsf{Fam}(C) \arrow[d, "\mathsf{Range}"] \\ \mathcal{P}C \arrow[r, "\Gamma_{\mathcal{R}}"'] & \mathcal{P}C \end{tikzcd}$$

commutes, and the range of the $\mathcal{R}$-derivational family is generated by $\mathcal{R}$.

Proof.

1. Let (M, F) be a family within C.

   We have a result-preserving surjection $\widehat{F}$ from the class of all $(\mathcal{R}, M, F)$-plates to the class of all $\mathcal{R}$-plates within $B \stackrel{\text{def}}{=} \mathsf{Range}(M, F)$, sending $\langle A, [x_k]_{k \in K_i}, p\rangle$ to $\langle A, [F(x_k)]_{k \in K_i}, p\rangle$.

   The range of $\Delta_{\mathcal{R}}(M, F)$ is the class of results of all $(\mathcal{R}, M, F)$-plates, which is the class of results of all $\mathcal{R}$-plates in B, which is $\Gamma_{\mathcal{R}} B$. So the square commutes.

   The fact that $(\mathsf{Deriv}_{\mathcal{R}_f}, \mathsf{O}_{\mathcal{R}_f})$ is a $\Delta_{\mathcal{R}_f}$-fixpoint means that an $\mathcal{R}_f$-derivation d with overall result c is the same thing as a $(\mathcal{R}, \mathsf{Deriv}_{\mathcal{R}_f}, \mathsf{O}_{\mathcal{R}_f})$-plate with result c.

   Therefore, $\widehat{\mathsf{O}_{\mathcal{R}_f}}$ is a surjection from $\mathsf{Deriv}_{\mathcal{R}}$ to the class of $\mathcal{R}$-plates within $B \stackrel{\text{def}}{=} \mathsf{Range}(\mathsf{Deriv}_{\mathcal{R}_f}, \mathsf{O}_{\mathcal{R}_f})$. Moreover, the result of $\widehat{\mathsf{O}_{\mathcal{R}_f}}(d)$ is the overall result of d.

   We show that Q is $\mathcal{R}$-complete as follows. Let x be an $\mathcal{R}$-plate within Q whose result is c. Then $x = \widehat{\mathsf{O}_{\mathcal{R}}}(d)$ for some $\mathcal{R}$-derivation d, and the latter has overall result c. So $c \in Q$.

   Lastly we show that B is included in any $\mathcal{R}$-complete class X. To do this, we show that every $\mathcal{R}$-derivation d has overall result in X, by induction on d, as follows. For $d = \langle i, [e_k]_{k \in K_i}, p\rangle$, its overall result c is the result of $\widehat{\mathsf{O}_{\mathcal{R}}}(d) = \langle i, [\mathsf{O}_{\mathcal{R}}(e_k)]_{k \in K_i}, p\rangle$. For cor all $k \in K_i$ we have $\mathsf{O}_{\mathcal{R}}(e_k) \in X$ by the inductive hypothesis. So $c \in X$ by $\mathcal{R}$-completeness of X.

2. Similar. □

Combining this with Propositions 6.7 and 6.17, we obtain the following:

Corollary 7.2. (Assuming AC.)

1. *Wide Subset Generation, Wide Derivation Set and Wide Infinity are equivalent.*

2. *Broad Subset Generation, Broad Derivation Set and Broad Infinity are equivalent.*

7.2 WISC principles

Our goal is to improve Corollary 7.2, by replacing AC with a weak form of choice called WISC, originally studied in [Str05, vdBM14]. We shall formulate three versions of it, using the following notions.

Definition 7.3. Let K be a set.

1. A *K-cover* δ is a K-tuple of inhabited sets.

2. The *unit K-cover* is $1_K \stackrel{\text{def}}{=} [1]_{k \in K}$.

3. Given K-covers A and B, *map* $f : A \to B$ is a K-tuple of functions $[f_k : A_k \to B_k]_{k \in K}$.

4. A set $\mathcal{A}$ of K-covers is *weakly initial* when, for any K-cover B, there is $A \in \mathcal{A}$ and a map $f : A \to B$. We then say that $\mathcal{A}$ is a *WISC* (weakly initial set of covers) for K.

Definition 7.4.

1. A *WISC function* on a class of sets $\mathcal{K}$ sends each set $K \in \mathcal{K}$ to a WISC for it.

2. A *global WISC function* is a WISC function on $\mathfrak{S}$.

Now consider the following principles.

- *Simple WISC*: Every set has a WISC.
- *Local WISC*: Every set of sets has a WISC function.
- *Global WISC*: There is a global WISC function.

These principles are related as follows.

Proposition 7.5.

1. *AC implies Global WISC.*

2. *Global WISC implies Local WISC.*

3. *Simple WISC and Local WISC are equivalent.*

4. **In ZF** *the three WISC principles are equivalent.*

Proof.

1. AC is equivalent to $K \mapsto \{1_K\}$ being a global WISC function.

2. Given a set of sets $\mathcal{K}$, restrict the global WISC function to it.

3. To show Local WISC $\Rightarrow$ Simple WISC: for any set K, obtain a WISC function for the singleton $\{K\}$ and apply it to K.

   To show the converse, let $\mathcal{K}$ be a set of sets, and write $L \stackrel{\text{def}}{=} \sum_{K \in \mathcal{K}} K$. For each $K \in \mathcal{K}$, define $\acute{K}$ to be the functor from the category of L-covers to that of K-covers, sending $[A_l]_{l \in L}$ to $[A_{\langle K,k \rangle}]_{k \in K}$ and likewise for maps.

   Given a WISC $\mathcal{A}$ for L, define f to be the function sending $K \in \mathcal{K}$ to the set of K-covers

   $$\{\acute{K}(A) \mid A \in \mathcal{K}\}$$

   which is weakly initial by the following argument. Any K-cover B is equal to $\acute{K}(C)$, where C is the L-cover whose $\langle M, k \rangle$-component is B_k if $M = K$ and 1 otherwise. Weak initiality of $\mathcal{A}$ gives an L-cover $A \in \mathcal{A}$ and map $g : A \to C$, so we obtain $\acute{K}(A) \in f(K)$ and $\acute{K}(g) : \acute{K}(A) \to \acute{K}(C) = B$. Thus f is a WISC function on $\mathcal{K}$.

4. We write $(V_\alpha)_{\alpha\in\mathsf{Ord}}$ for the cumulative hierarchy in the usual way.

   Suppose Simple WISC. For each set K, there is an ordinal α such that the set $C_\alpha(K)$ of all K-covers in V_α^K is weakly initial, and we define $t(K)$ to be the least such ordinal. This gives a global WISC function $K \mapsto C_{t(K)}(K)$. $\square$

It has been shown that the theory ZF + WISC is strictly between ZF and ZFC, provided ZF is consistent [Kar14, Rob15].

7.3 Sufficiency of WISC

Our task is to weaken the AC assumption of Corollary 7.2. Specifically, we shall prove that Local WISC suffices for part 1, and Global WISC for part 2.

Proposition 7.6. *Let K be a set and A a K-cover. Let C be a class. Then the function $\theta : C^K \to C^{\sum_{k\in K} A_k}$ sending $[x_k]_{k\in K}$ to $[x_k]_{k\in K, a\in A_k}$ is injective. Its range is the class of all $[x_{k,a}]_{k\in K,a\in A_k}$ with the following property: for any $k\in K$ and $a, b\in A_k$, we have $x_{k,a} = x_{k,b}$.*

Proof. Obvious. $\square$

Definition 7.7.

1. For any wide rule $\langle K, R\rangle$ on C and any K-cover A, define the wide rule $\langle K, R\rangle^A$ consisting of the arity $\sum_{k\in K} A_k$ and function $C^{\sum_{k\in K} A_k} \to \mathsf{Fam}(C)$ sending $\theta(x)$ to $R(x)$ and everything else to the empty family.

2. For any wide rubric $\mathcal{R} = (\langle K_i, R_i\rangle)_{i\in I}$ on C, define the set of sets

   $$\overline{\mathcal{R}} \quad \stackrel{\text{def}}{=} \quad \{K_i \mid i \in I\}$$

   Let f be either a WISC function on $\overline{\mathcal{R}}$ or a global WISC function. Then define the wide rubric

   $$\mathcal{R}_f \quad \stackrel{\text{def}}{=} \quad (\langle K_i, R_i\rangle^A)_{i\in I, A\in f(K_i)}$$

3. For any broad rule $\langle L, S\rangle$ on C and any L-cover B and global WISC function f, define the broad rule $\langle L, S\rangle_f^B$ consisting of the arity $\sum_{l\in L} B_l$ and function $C^{\sum_{l\in L} A_k} \to \mathsf{WideRub}(C)$ sending $\theta(x)$ to $S(x)_f$ and everything else to the empty rubric.

4. For any broad rubric $\mathcal{R} = (\langle L_j, S_j\rangle)_{j\in J}$ on C and global WISC function f, define the broad rubric

   $$\mathcal{R}_f \quad \stackrel{\text{def}}{=} \quad (\langle L_j, S_j\rangle_f^B)_{j\in J, B\in f(L_j)}$$

Now we show the following.

Proposition 7.8.

1. (Assuming Wide Infinity.) *For any wide rubric $\mathcal{R}$ on C and any WISC function f on $\overline{\mathcal{R}}$, the square*

$$\begin{tikzcd} \mathsf{Fam}(C) \arrow[r, "\Delta_{\mathcal{R}_f}"] \arrow[d, "\mathsf{Range}"'] & \mathsf{Fam}(C) \arrow[d, "\mathsf{Range}"] \\ \mathcal{P}C \arrow[r, "\Gamma_{\mathcal{R}}"'] & \mathcal{P}C \end{tikzcd}$$

commutes, and the range of the $\mathcal{R}_f$-derivational family is generated by $\mathcal{R}$.

2. (Assuming Broad Infinity.) *For any broad rubric $\mathcal{R}$ on C and any global WISC function f, the square*

$$\begin{tikzcd} \mathsf{Fam}(C) \arrow[r, "\Delta_{\mathcal{R}_f}"] \arrow[d, "\mathsf{Range}"'] & \mathsf{Fam}(C) \arrow[d, "\mathsf{Range}"] \\ \mathcal{P}C \arrow[r, "\Gamma_{\mathcal{R}}"'] & \mathcal{P}C \end{tikzcd}$$

commutes, and the range of the $\mathcal{R}_f$-derivational family is generated by $\mathcal{R}$.

Proof.

1. Let (M, F) be a family within C.

   For any $(\mathcal{R}_f, M, F)$-plate $m = \langle\langle i, A\rangle, [x_{k,a}]_{k\in K_i, a\in A_k}, p\rangle$, we obtain a $\mathcal{R}_f$-plate $\langle\langle i, A\rangle, [F(x_{k,a})]_{k\in K_i, a\in A_k}, p\rangle$ with the same result. It is uniquely expressible as $\langle\langle i, A\rangle, \theta(y), p\rangle$, and we obtain an $\mathcal{R}$-plate $\widehat{F}(m) \stackrel{\text{def}}{=} \langle i, y, p\rangle$.

   We see that $\widehat{F}$ is a result-preserving surjection from the class of $(\mathcal{R}_f, M, F)$-plates to the class of all $\mathcal{R}$-plates within $B \stackrel{\text{def}}{=} \mathsf{Range}(M, F)$. To prove surjectivity, let $n = \langle i, y, p\rangle$ be an $\mathcal{R}$-plate within B. We obtain a K_i-cover $[F^{-1}(y_k)]_{k\in K_i}$, so there is $A \in f(K_i)$ and a K-cover map $g : A \to [F^{-1}(y_k)]_{k\in K_i}$. We then obtain an $(\mathcal{R}_f, M, F)$-plate $m = \langle\langle i, A\rangle, [g_k(a)]_{k\in K_i, a\in A_k}, p\rangle$. Since $F(g_k(a)) = y_k$ for all $k \in K_i$ and $a \in A_k$, we have $\widehat{F}(m) = n$.

   The range of $\Delta_{\mathcal{R}_f}(M, F)$ is the class of results of all $(\mathcal{R}_f, M, F)$-plates, which is the class of results of all $\mathcal{R}$-plates in B, which is $\Gamma_{\mathcal{R}} B$. So the square commutes.

   The fact that $(\mathsf{Deriv}_{\mathcal{R}_f}, \mathsf{O}_{\mathcal{R}_f})$ is a $\Delta_{\mathcal{R}_f}$-fixpoint means that an $\mathcal{R}_f$-derivation d with overall result c is the same thing as a $(\mathcal{R}, \mathsf{Deriv}_{\mathcal{R}_f}, \mathsf{O}_{\mathcal{R}_f})$-plate with result c.

   Therefore, $\widehat{\mathsf{O}_{\mathcal{R}_f}}$ is a surjection from $\mathsf{Deriv}_{\mathcal{R}}$ to the class of $\mathcal{R}$-plates within $Q \stackrel{\text{def}}{=} \mathsf{Range}(\mathsf{Deriv}_{\mathcal{R}_f}, \mathsf{O}_{\mathcal{R}_f})$. Moreover, the result of $\widehat{\mathsf{O}_{\mathcal{R}_f}}(d)$ is the overall result of d.

   We show that Q is $\mathcal{R}$-complete as follows. Let x be an $\mathcal{R}$-plate within Q whose result is c. Then $x = \widehat{\mathsf{O}_{\mathcal{R}_f}}(d)$ for some $\mathcal{R}_f$-derivation d, and the latter has overall result c. So $c \in Q$.

Lastly we show that Q is included in any $\mathcal{R}$-complete class X. To do this, we show that every $\mathcal{R}_f$-derivation d has overall result in X, by induction on d, as follows. For $d = \langle\langle i, A\rangle, [e_{k,a}]_{k\in K_i, a\in A_k}, p\rangle$, its overall result c is the result of $\widehat{\mathsf{O}_{\mathcal{R}_f}}(d) = \langle i, [y_k]_{k\in K_i}, p\rangle$ where, for all $k \in K_i$ and $a \in A_k$, we have $y_k = \mathsf{O}_{\mathcal{R}_f}(e_{k,a})$ and therefore $y_k \in X$ by the inductive hypothesis. Thus, for all $k \in K_i$ we have $y_k \in X$ since A_k is inhabited. So $c \in X$ by $\mathcal{R}$-completeness of X.

2. Similar. □

Combining this with Propositions 6.7 and 6.20, we obtain the following:

Corollary 7.9.

1. (Assuming Local WISC.) *Wide Subset Generation is equivalent to Wide Infinity.*

2. (Assuming Global WISC.) *Broad Subset Generation is equivalent to Broad Infinity.*

8 Basic theory of ordinals

8.1 Set-based well-orderings

Having completed our study of sets, we begin the second part of the paper, which is concerned with ordinals. Since the notion of ordinal is based on that of well-ordering, we consider the latter first. As in Section 3.9, we treat not only relations on a set, but also on a class.

Let C be a class with a relation $<$. Recall the notation $J(x) \stackrel{\text{def}}{=} \{y \in C \mid y < x\}$.. The *extensional preorder* relates $x, y \in C$ when $J(x) \subseteq J(y)$. It is an order iff J is injective, and we then say that $<$ is *extensional*.

A *strict order* on a class C is an irreflexive transitive relation. Strict orders correspond bijectively to orders via

$$\begin{aligned} x \leqslant y \quad &\stackrel{\text{def}}{\Longleftrightarrow} \quad x < y \vee x = y \\ x < y \quad &\stackrel{\text{def}}{\Longleftrightarrow} \quad x \leqslant y \wedge x \neq y \end{aligned}$$

A *linear order* is a strict order $<$ such that, for all $x, y \in C$ we have either $x \leqslant y$ or $y \leqslant x$. It follows that $<$ is extensional, and its extensional order is $\leqslant$.

Now let us formulate the notion of well-ordering.

Proposition 8.1. *Let $(C, <)$ be a class equipped with a set-based transitive relation. The following are equivalent:*

1. *$<$ is well-founded and extensional.*

2. *$<$ is well-founded and a linear order.*

3. *$<$ is a strict order, and any inhabited subclass has a least element.*

4. $<$ *is a strict order, and any inhabited subset has a least element.*

Proof. Firstly, (2) implies (1) since a linear order is extensional.

For (1) $\Rightarrow$ (2) say that $x, y \in C$ are *comparable* when either $x \leqslant y$ or $y \leqslant x$. We claim that, for all $a, b \in C$, if a is comparable with all $y \in J(b)$, and all $x \in J(a)$ with b, then a is with b. It follows by induction that any $a, b \in C$ are comparable.

To prove the claim, it suffices to show that $a \not< b$ and $b \not< a$ implies $a = b$. Any $x \in J(a)$ is comparable with b, and is therefore $< b$ as $b \leqslant x$ would imply $b < a$. Thus $J(a) \subseteq J(b)$, and likewise $J(b) \subseteq J(a)$. Extensionality gives $a = b$.

The rest follows from Proposition 3.11, noting that any set-based transitive relation is iteratively set-based. $\square$

A relation satisfying the above conditions is called a *set-based well-ordering*. Henceforth, we often abbreviate $(C, <)$ as C.

Definition 8.2. Let A and B be set-based well-ordered classes. For a function $f : A \to B$, the following are equivalent.

- f is an isomorphism from A to a hereditary subclass of B.
- The square

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, "J"'] & B \arrow[d, "J"] \\ \mathcal{P}A \arrow[r, "\mathcal{P}f"'] & \mathcal{P}B \end{tikzcd}$$

  commutes. Explicitly: for all $x \in A$, we have $J(f(x)) = \{f(y) \mid y \in J(x)\}$.

Proof. ($\Rightarrow$) is straightforward. For ($\Leftarrow$), we show that $f(x) = f(x')$ implies $x = x'$ by induction on $x \in A$ as follows: $f(x) = f(x')$ implies that for all $y \in J(x)$ there is $y' \in J(x')$ such that $f(y) = f(y')$ and hence $y = y'$ so $J(x) \subseteq J(x')$, and the reverse inclusion likewise, so extensionality of A gives $x = x'$. $\square$

A function satisfying the above conditions is called an *embedding*.

Proposition 8.3.

1. *For set-based well-ordered classes A, B, C, the composite of embeddings $A \to B$ and $B \to C$ is an embedding $A \to C$.*
2. *For set-based well-ordered classes A and B, there is at most one embedding $A \to B$.*

Proof.

1. Follows from the fact that, for any hereditary subclass X of B, an embedding $B \to C$ restricts to an embedding $X \to C$.

2. For embeddings $f, g : A \to B$, we show $f(x) = g(x)$ by induction on $x \in A$. $\square$

Next we consider isomorphisms:

Proposition 8.4. *Let A and B be set-based well-ordered classes.*

1. *For a function $f : A \to B$, the following are equivalent:*
   - *f is an isomorphism.*
   - *f is a surjective embedding.*
   - *f is an embedding, and there is an embedding $B \to A$.*
2. *There is at most one isomorphism $A \cong B$.*

Proof.

1. Clearly, the first two conditions are equivalent and imply the third. Lastly, if f and $g : D \to C$ are embeddings, then Proposition 8.3(1) gives endo-embeddings $g \circ f$ on C and $f \circ g$ on D. By Proposition 8.3(2), both are identity maps.
2. Special case of Proposition 8.3(2). $\square$

Proposition 8.5. *Let C be a set-based well-ordered class.*

1. *The only hereditary subclasses of C are $J(x)$, for $x \in C$, and C itself.*
2. *These are pairwise non-isomorphic. In other words:*
   - *For any $x, y \in C$, if $J(x) \cong J(y)$, then $x = y$.*
   - *For any $x \in C$, we do not have $J(x) \cong C$.*

Proof.

1. Let X be a hereditary subclass. If it has a strict upper bound, it has a strict supremum x and is $J(x)$. Otherwise it is C.
2. For $x \in C$ and a hereditary subclass Y of C, any isomorphism $\theta : J(x) \cong Y$ is an embedding $J(x) \to C$. By Proposition 8.3(2), this must be the inclusion, so $J(x) = Y$. Since $x \notin Y$, we cannot have $Y = C$, nor $X = J(y)$ for $y > x$. Lastly, for $y < x$, we canot have $Y = J(y)$ since $y \in Y$. $\square$

Proposition 8.6.

1. *For set-based well-ordered classes C and D, either C embeds into D or vice versa.*
2. *Let B be a subclass of a set-based well-ordered class $(C, <)$. Then $<$ is a set-based well-ordering on B, and there is a deflationary embedding $B \to C$.*

Proof.

1. Define $R(C, D)$ to be the relation from C to D that relates $x \in C$ to $y \in D$ when $J(x) \cong J(y)$. It is an isomorphism from a hereditary subclass X of C to a hereditary subclass Y of D. If $X = J(x)$ and $Y = J(y)$, then $(x, y) \in R(C, D)$, contradiction. Therefore either $X = C$ or $Y = D$, which gives the two cases.

2. On the subclass B, the relation $<$ is set-based and well-founded by Proposition 3.13, and also linear, hence a well-ordering. Obtain $R(B, C)$ as before. Induction shows that, for each $x \in B$, there is $y \leqslant x$ such that $(x, y) \in R(B, C)$. Thus R is total and deflationary. □

Proposition 8.7. *Let $(C, <)$ be a set-based well-ordered class. The following are equivalent.*

1. *C is a proper class.*
2. *Every subset of C has a strict supremum.*
3. *Every hereditary subset of C has a strict supremum.*
4. *Every set-based well-ordered class embeds into C.*
5. *Every well-ordered set is isomorphic to $J(x)$ for some $x \in C$.*

Proof. Clearly (2) implies (3). For the converse, given a subset A of C, its hereditary closure $\{x \in C \mid \exists y \in A.\, x \leqslant y\}$ has the same strict upper bounds.

(3) implies (1), because C itself does not have a strict upper bound, so is not a set.

(1) implies (4) because, for any set-based well-ordered class B that does not embed into C, Proposition 8.6(1) gives an embedding $f : C \to B$. The range of f is a hereditary subclass of B that is not a set, and therefore is B by Proposition 8.5(1). So f is an isomorphism $C \cong B$, and its inverse is an embedding $B \to C$, contradiction.

Clearly (4) implies (5).

Lastly, (5) implies (3) because, for any hereditary subset A of C, there is an isomorphism $f : A \cong J(x)$ for some $x \in A$, which is the identity by Proposition 8.3(2). So x is a strict supremum of A. □

We say that $(C, <)$ is *complete* when it has the above properties.

Proposition 8.8. *Any two complete set-based well-ordered classes are uniquely isomorphic.*

Proof. By Propositions 8.7(4) and 8.4. □

Related work The equivalence of (1)-(2) in Proposition 8.1 is proved in [nLa11] using a different method.

8.2 Ordinals

Our next task is to define Ord. To do this, we write TrSet for the class of all transitive sets. Thus the function $\Gamma_{(\mathsf{TrSet},\in)}$ sends a class X to the class of all its transitive subsets. By Proposition 4.2(2), we can define

$$\mathsf{Ord} \stackrel{\text{def}}{=} \mu\Gamma_{(\mathsf{TrSet},\in)}$$

Alternatively, we can say that an *ordinal* is a transitive set of transitive pure vonniads. Here is yet another characterization of Ord:

Proposition 8.9. Ord *is the unique class of sets* X *such that* $(X, \in)$ *is a complete set-based well-ordered class.*

Proof. To prove uniqueness, let X, Y be two such classes. Proposition 8.8 gives an isomorphism $f : (X, \in) \Rightarrow (Y, \in)$. For all $x \in X$, we have $f(x) = x$, by induction on x: since $f(x)$ and x have the same elements (by the inductive hypothesis) and are both sets, they are equal. So $X = Y$.

Now we consider Ord. By inductive inversion, every ordinal is a transitive set of ordinals, so membership is extensional and transitive on Ord. Furthermore, membership is (set-based and) well-founded on Ord, so it is a well-ordering. Lastly, $(\mathsf{Ord}, \in)$ is complete since any transitive set of ordinals is an ordinal and its own supremum. □

Say that an *extended ordinal* is a transitive class of ordinals. By Proposition 8.5(1), it is either an ordinal or Ord. By completeness, any set-based well-ordered class C is uniquely isomorphic to an extended ordinal, called the *order-type* of C. The collection of all extended ordinals is written Ord_∞.

The least ordinal is $0 \stackrel{\text{def}}{=} \emptyset$, and the least ordinal greater than $\alpha \in \mathsf{Ord}$ is $\mathsf{S}\alpha \stackrel{\text{def}}{=} \alpha \cup \{\alpha\}$. We recursively define the injection $\iota : \mathbb{N} \to \mathsf{Ord}$ sending $\mathsf{Nothing} \mapsto 0$ and $\mathsf{Just}(n) \mapsto \mathsf{S}(\iota n)$. Its range ω is an extended ordinal; it is is a limit if Infinity holds and Ord otherwise.

Let I be a set. Given $\alpha \in \mathsf{Ord}^I$, we write $\bigvee_{i \in I} \alpha_i$ for the supremum and $\mathsf{ssup}_{i\in I}\alpha_i$ for the strict supremum. Thus we have functions $\bigvee_I, \mathsf{ssup}_I : \mathsf{Ord}^I \to \mathsf{Ord}$ that connected via the equation

$$\mathsf{ssup}_{i\in I}\alpha_i = \bigvee_{i\in I} \mathsf{S}\alpha_i$$

As an application (not used in the sequel), let $(C, <)$ be a class equipped with a set-based well-ordered relation. We recursively define $\mathsf{rank} : C \to \mathsf{Ord}$ by $\mathsf{rank}(x) \stackrel{\text{def}}{=} \mathsf{ssup}_{y\in J(x)}\mathsf{rank}(y)$. Induction on $x \in C$ gives $\mathsf{rank}(x) = \{\mathsf{rank}(y) \mid y \in J^+(x)\}$.

Lastly, given ordinals $\alpha \leqslant \beta$, we define intervals

$$\begin{aligned}(\alpha\,..\,\beta) &\stackrel{\text{def}}{=} \{\gamma \in \mathsf{Ord} \mid \alpha < \gamma < \beta\}\\ [\alpha\,..\,\beta) &\stackrel{\text{def}}{=} \{\gamma \in \mathsf{Ord} \mid \alpha \leqslant \gamma < \beta\}\end{aligned}$$

9 From rubrics to sup-generation

9.1 Inductive chains

Our next task will be to adapt Wide Subset Generation into a similar principle for ordinals, and likewise for Broad. We shall do this using the following notions.

Definition 9.1. Let C be a class, and F a monotone endofunction on $\mathcal{P}C$. The *inductive chain* $(\mu^{\alpha}F)_{\alpha\in\mathsf{Ord}}$ is the increasing sequence of subsets of C, defined recursively by

$$\begin{aligned} \mu^{0}F &\stackrel{\text{def}}{=} \emptyset \\ \mu^{\mathsf{S}\alpha}F &\stackrel{\text{def}}{=} F(\mu^{\alpha}F) \\ \mu^{\lambda}F &\stackrel{\text{def}}{=} \bigcup_{\beta<\lambda}\mu^{\beta}F \quad \text{for a limit } \lambda \end{aligned}$$

Example 9.2. *(Assuming Powerset.)* The von Neumann hierarchy is the inductive chain of $\mathcal{P}$.

In some cases, the inductive chain can be used to obtain a least prefixpoint.

Proposition 9.3. *Let C be a class, and F a monotone endofunction on $\mathcal{P}C$.*

1. *Every prefixpoint is an upper bound of the inductive chain.*
2. *If $\mu^{\alpha}F$ is a prefixpoint, then $(\mu^{\beta}F)_{\beta\geqslant\alpha}$ is constant, and $\mu^{\alpha}F$ is a least prefixpoint. (We say that F* inductively stabilizes *at α.)*
3. *Conversely, if F has a prefixpoint, then it inductively stabilizes at some ordinal.*

Proof.

1. Induction.
2. Follows from part 1.
3. Define the class $B \stackrel{\text{def}}{=} \bigcup_{\alpha\in\mathsf{Ord}}\mu^{\alpha}F$. If F has a prefixpoint X, then $B \subseteq X$, so B is a set. For each $y \in B$, let $\overline{y}$ be the unique ordinal β such that $y \in \mu^{\mathsf{S}\beta}F \setminus \mu^{\beta}F$. Then F inductively stabilizes at $\gamma \stackrel{\text{def}}{=} \mathsf{ssup}_{y\in B}\overline{y}$, since $B \subseteq \mu^{\gamma}F$. □

9.2 Sup-generation

Definition 6.4 gave us the notion of a class being closed or complete. Here are the analogous properties for limits:

Definition 9.4. A limit λ is

- *K-sup-closed* , for a set K, when $\bigvee_K$ (or equivalently ssup_K) restricts to a function $\lambda^K \to \lambda$.

- $\mathcal{K}$*-sup-complete*, for a set of sets $\mathcal{K}$, when it is K-sup-complete for all $K \in \mathcal{K}$.
- F*-sup-closed*, for a function $F : \mathsf{Ord} \to \mathfrak{S}$, when it is $F\alpha$-sup-closed for all $\alpha < \lambda$.
- H*-sup-complete*, for a function $H : \mathsf{Ord} \to \mathcal{P}\mathfrak{S}$, when it is $H\alpha$-sup-complete for all $\alpha < \lambda$.

Here are some examples.

1. For a set of sets $\mathcal{K}$, let $\mathsf{Const}_{\mathcal{K}}$ be the constant function $\gamma \mapsto \mathcal{K}$. A limit is $\mathsf{Const}_{\mathcal{K}}$-sup-complete iff $\mathcal{K}$-sup-complete.
2. For functions $H, H' : \mathsf{Ord} \to \mathcal{P}\mathfrak{S}$, let $H \vee H'$ be the pointwise union $\gamma \mapsto H(\gamma) \cup H'(\gamma)$. A limit is $(H \vee H')$-sup-complete iff both H-sup-complete and H'-sup-complete.

Below (Proposition 11.6) we shall characterize sup-closedness and sup-completeness in an explicit way.

Definition 9.5. A limit is

1. *simply sup-generated* by a set K when it is the least K-sup-closed one.
2. *sup-generated* by a set of sets $\mathcal{K}$ when it is the least $\mathcal{K}$-sup-complete one.
3. *simply sup-generated* by a function $F : \mathsf{Ord} \to \mathfrak{S}$ when it is the least F-sup-closed one.
4. *sup-generated* by a function $H : \mathsf{Ord} \to \mathcal{P}\mathfrak{S}$ when it is the least H-sup-complete one.

This leads to the following principles.

- *Simple Wide Sup-generation*: Any set simply sup-generates a limit.
- *Full Wide Sup-generation*: Any set of sets sup-generates a limit.
- *Simple Broad Sup-generation*: Any function $\mathsf{Ord} \to \mathfrak{S}$ simply sup-generates a limit.
- *Full Broad Sup-generation*: Any function $\mathsf{Ord} \to \mathcal{P}\mathfrak{S}$ sup-generates a limit.

Proposition 9.6.

1. *The two forms of Wide Sup-generation are equivalent.*
2. *The two forms of Broad Sup-generation are equivalent.*

Proof. Full $\Rightarrow$ Simple is obvious. For the converse, we first note that, for a set of sets $\mathcal{K}$, any $\sum_{K\in\mathcal{K}} K$-sup-closed limit λ is $\mathcal{K}$-sup-complete. That is because, for $K\in\mathcal{K}$ and $p\in\mathsf{Ord}^K$, we have

$$\begin{aligned} \bigvee_{k\in K} p_k &= \bigvee_{\langle L,K\rangle\in\sum_{K\in\mathcal{K}} K} \begin{cases} p_k & (L=K) \\ 0 & \text{otherwise} \end{cases} \\ &< \lambda \end{aligned}$$

It follows that, for a function $H:\mathsf{Ord}\to\mathcal{P}\mathfrak{S}$, any $(\beta\mapsto\sum_{K\in H\beta} K)$-sup-closed limit is H-sup-complete. $\square$

Now we give the relationship between sup-generation and subset generation.

Proposition 9.7.

1. *Wide Subset Generation is equivalent to Powerset + Wide Sup-generation.*
2. *Broad Subset Generation is equivalent to Powerset + Broad Sup-generation.*

Proof. Wide Subset Generation implies Powerset, by Propositions 5.2 and 6.7(1). To show that it implies Wide Sup-generation, let $\mathcal{K}$ be a set of sets. A $\mathcal{K}$-sup-complete ordinal is an $\mathcal{R}$-inductive subset of Ord, where the wide rubric $\mathcal{R}$ on Ord consists of the following.

- The unary rule sending $[\alpha]$ to $(\beta)_{\beta<\alpha}$. (Note that a set is closed under this rule iff it is transitive, i.e., an ordinal.)
- For each $K\in\mathcal{K}$, a K-ary rule sending $[\alpha_k]_{k\in K}$ to $(\bigvee_{k\in K}\alpha_k)$.

To show Broad Subset Generation implies Broad Sup-generation, let $H:\mathsf{Ord}\to\mathcal{P}\mathfrak{S}$ be a function. An H-sup-complete ordinal is an $\mathcal{R}$-inductive subset of Ord, where the broad rubric $\mathcal{R}$ on Ord consists of the following.

- The nullary rule returning the wide rubric consisting of just the unary rule sending $[\alpha]$ to $(\beta)_{\beta<\alpha}$. (Note that a set is closed under this rule iff it is transitive, i.e., an ordinal.)
- The unary rule sending $[\alpha]$ to the wide rubric consisting of, for each $K\in H\alpha$, the K-ary rule sending $[\alpha_k]_{k\in K}$ to $(\bigvee_{k\in K}\alpha_k)$.

To show Powerset + Wide Sup-generation implies Wide Set Generation, let $\mathcal{R}=(\langle K_i,R_i\rangle)_{i\in I}$ be a wide rubric on C. We shall show that the endofunction $\Gamma_{\mathcal{R}}$ on $\mathcal{P}C$, obtained from Powerset by Proposition 6.25(1), inductively stabilizes at an ordinal α. Define the set of sets

$$\mathcal{K} \stackrel{\text{def}}{=} \{K_i \mid i\in I\}$$

Let α be a $\mathcal{K}$-sup-complete limit. For any $x\in\mu^{\alpha}\Gamma_{\mathcal{R}}$, put $\overline{x}$ for the unique $\beta<\alpha$ such that $x\in\mu^{\mathsf{S}\beta}\Gamma_{\mathcal{R}}\setminus\mu^{\beta}\Gamma_{\mathcal{R}}$. Given an $\mathcal{R}$-plate

$$w = \langle i,y,p\rangle$$

within $\mu^\alpha\Gamma_{\mathcal{R}}$, put

$$\beta \quad \stackrel{\text{def}}{=} \quad \mathsf{ssup}_{l\in L_j}\overline{x_l}$$

Since α is K_i-sup-closed, and $\overline{y_k} < \alpha$ for all $k \in K_i$, we obtain $\beta < \alpha$. Since w is an $\mathcal{R}$-plate within $\mu^\beta\Gamma_{\mathcal{R}}$, its result is in $\mu^{\mathsf{S}\beta}\Gamma_{\mathcal{R}}$, which is included in $\mu^\alpha\Gamma_{\mathcal{R}}$ since $\mathsf{S}\beta < \alpha$.

To show Powerset + Broad Sup-generation implies Broad Set Generation, let $\mathcal{R} = (\langle L_j, T_j\rangle)_{j\in J}$ be a broad rubric on C. We shall show that the endofunction $\Gamma_{\mathcal{R}}$ on $\mathcal{P}C$, obtained from Powerset by Proposition 6.25(1), inductively stabilizes at an ordinal α. Define the set

$$\mathcal{K} \quad \stackrel{\text{def}}{=} \quad \{L_j \mid j \in J\}$$

and the function $H : \mathsf{Ord} \to \mathcal{P}\mathfrak{S}$ sending β to

$$\bigcup_{j\in J} \quad \bigcup_{\substack{x\in(\mu^\beta\Gamma_{\mathcal{R}})^{L_j} \\ T_j(x)=(\langle K_i,R_i\rangle)_{i\in I}}} \quad \{K_i \mid i \in I\}$$

Let α be a limit that is $(\mathsf{Const}_{\mathcal{K}} \vee H)$-sup-complete, i.e., both $\mathcal{K}$-sup-complete and H-sup-complete. For any $x \in \mu^\alpha\Gamma_{\mathcal{R}}$, put $\overline{x}$ for the unique $\beta < \alpha$ such that $x \in \mu^{\mathsf{S}\beta}\Gamma_{\mathcal{R}} \setminus \mu^\beta\Gamma_{\mathcal{R}}$. Given an $\mathcal{R}$-plate

$$w \quad = \quad \langle j, x, i, y, p\rangle$$

within $\mu^\alpha\Gamma_{\mathcal{R}}$, put

$$\begin{aligned} \beta \quad &\stackrel{\text{def}}{=} \quad \mathsf{ssup}_{l\in L_j}\overline{x_l} \\ \gamma \quad &\stackrel{\text{def}}{=} \quad \mathsf{ssup}_{k\in K_i}\overline{y_k} \end{aligned}$$

Since α is L_j-sup-closed, and $\overline{x_l} < \alpha$ for all $l \in L_j$, we obtain $\beta < \alpha$. We have $x \in (\mu^\beta\Gamma_{\mathcal{R}})^{L_j}$ and $T_j(x) = (\langle K_i, R_i\rangle)_{i\in I}$, so $K_i \in H\beta$, so α is K_i-sup-closed. Since $\overline{y_k} < \alpha$ for all $k \in K$, we obtain $\gamma < \alpha$. Since w is an $\mathcal{R}$-plate within $\mu^{\beta\vee\gamma}\Gamma_{\mathcal{R}}$, its result is in $\mu^{\mathsf{S}(\beta\vee\gamma)}\Gamma_{\mathcal{R}}$, which is included in $\mu^\alpha\Gamma_{\mathcal{R}}$ since $\mathsf{S}(\beta \vee \gamma) < \alpha$. $\square$

10 Lindenbaum numbers

We interrupt our journey towards Mahlo's principle to give some useful and well-known constructions that relate sets to ordinals. First we give some notation:

Definition 10.1. Let A and B be sets.

1. We write $A \preccurlyeq B$ when there is an injection $A \to B$.
2. We write $A \preccurlyeq^* B$ when there is a partial surjection $B \to A$. Equivalently: when either $A = \emptyset$ or there is a surjection $B \to A$.

Thus $A \preccurlyeq B$ implies $A \preccurlyeq^* B$, and conversely if B is well-orderable or AC holds.

Definition 10.2. Let A be a set.

1. We write $\aleph(A)$ for the class of all ordinals α such that $\alpha \preccurlyeq A$.
2. We write $\aleph^*(A)$ for the class of all ordinals α such that $\alpha \preccurlyeq^* A$.

Note that $\aleph(A)$ and $\aleph^*(A)$ are extended ordinals. They are called the *Hartogs number* and the *Lindenbaum number* of A, respectively. We have $\aleph(A) \leqslant \aleph^*(A)$, with equality if A is well-orderable or AC holds. Note that $\aleph(A) \not\preccurlyeq A$ and $\aleph^*(A) \not\preccurlyeq^* A$. So, for any ordinal α, we have $\alpha < \aleph(\alpha) \leqslant \aleph^*(\alpha)$.

We use Lindenbaum numbers extensively in the next section, and introduce some axioms concerning them. The first is *Full Lindenbaum*: For any set K, the extended ordinal $\aleph^*(K)$ is an ordinal.

Proposition 10.3.

1. *Powerset implies Full Lindenbaum.*
2. *Wide Sup-generation implies Full Lindenbaum.*

Proof.

1. For a set K, we can express $\aleph^* K$ as the set of order-types of well-ordered partial partitions of K.
2. For a set K, let λ be a K-closed limit. If $\lambda < \aleph^*(K)$, then there is a surjection $f : K \to \lambda$. So $\lambda = \sup_{k \in K} f(k) < \lambda$, contradiction. □

We divide Full Lindenbaum into two parts:

- *Ordinal Lindenbaum*: For any ordinal α, the extended ordinal $\aleph^*(\alpha)$ is an ordinal.
- *Relative Lindenbaum*: For any set K, there is an ordinal α such that $\aleph^*(K) \subseteq \aleph^*(\alpha)$.

Proposition 10.4. *Full Lindenbaum is equivalent to Ordinal Lindenbaum + Relative Lindenbaum.*

Proof. ($\Leftarrow$) is obvious, and clearly Full Lindenbaum implies Ordinal Lindenbaum. To show that it implies Bounding Lindenbaum, put $\alpha \stackrel{\text{def}}{=} \aleph^* K$ so that $\aleph^*(K) < \aleph^*(\alpha)$. □

Lastly we consider the *Well-orderability* axiom: Every set is well-orderable.

Proposition 10.5.

1. *Powerset + AC implies Well-orderability.*
2. *Well-orderability implies Relative Lindenbaum + AC.*

Proof. Everything is standard, except for Well-orderability $\Rightarrow$ Bounding Lindenbaum. To prove this: given a set K, define α to be the least order-type of a well-ordering of K. Since $\alpha \cong K$, we have $\aleph^*(K) = \aleph^*(\alpha)$. □

11 From sup-generation to Mahlo's principle

11.1 Unbounded and stationary classes

Now at last, it is time to treat Mahlo's principle; but we approach it more slowly than in Section 1. To begin, we revisit the notions of unbounded and stationary class from Section 1.2

Definition 11.1. Let A be a set-based well-ordered class.

1. A subclass B is *cofinal* when, for all $x \in A$, there is $y \in B$ such that $y \geqslant x$. Equivalently: when it has no strict upper bound.
2. A subclass B is *strictly cofinal* when, for all $x \in A$, there is $y \in B$ such that $y > x$. Equivalently: when it has no upper bound.

If A has no greatest element (e.g., when $A = \mathsf{Ord}$), then "cofinal" and "strictly cofinal" are equivalent, and the word "unbounded" is also used.

We turn next to ordinal functions.

Definition 11.2. A limit λ is

- *G-based*, for a function $G : \mathsf{Ord} \to \mathsf{Ord}_\infty$, when, for all $\alpha < \lambda$, we have $G(\alpha) \leqslant \lambda$.
- *F-closed*, for a function $F : \mathsf{Ord} \to \mathsf{Ord}$, when, for all $\alpha < \lambda$, we have $F(\alpha) < \lambda$. In short: when F restricts to an endofunction on α.

Proposition 11.3. *For a class of limits D, the following are equivalent.*

1. *For every function $F : \mathsf{Ord} \to \mathsf{Ord}$, there is a F-based limit in D.*
2. *For every function $F : \mathsf{Ord} \to \mathsf{Ord}$, there is a F-closed limit in D.*

Proof. Since F-closed is the same as $\mathsf{S}F$-based and implies F-based. $\square$

A class of limits with these properties is said to be *stationary.* It is then unbounded, and, for any function $F : \mathsf{Ord} \to \mathsf{Ord}$, contains stationarily many F-closed elements. Here is an application:

Proposition 11.4. *Each of the following is equivalent to Infinity.*

1. Lim *is unbounded.*
2. Lim *is stationary.*

Proof. If Infinity does not hold, then Lim is empty.

Assume Infinity. To show Lim is stationary, let $F : \mathsf{Ord} \Rightarrow \mathsf{Ord}$. Define $G : \mathsf{Ord} \Rightarrow \mathsf{Ord}$ sending α to $\bigvee_{\beta<\alpha}(\mathsf{S}\beta \vee F\beta)$. Put $\lambda \stackrel{\text{def}}{=} \bigvee_{n\in\mathbb{N}} G^n(0)$. If $\beta < \lambda$, then for some $n \in \mathbb{N}$ we have $\beta < G^n(0)$, so $\mathsf{S}\beta, F\beta < G^n(0) \leqslant \lambda$. So α is an F-closed limit. $\square$

11.2 Cofinality

We next treat the (standard) notion of cofinality, using the following result:

Proposition 11.5. *For any extended ordinal α and function $f : \alpha \to \mathsf{Ord}$, the range of f has a cofinal subclass of order-type $\leqslant \alpha$.*

Proof. Let K be the class of all $i < \alpha$ such that $f(i)$ is a strict upper bound of $\{f(j) \mid j < i\}$. We prove by induction on $i < \alpha$ that there is $k \leqslant i$ such that $k \in K$ and $f(k) \geqslant f(i)$, as follows. If $i \in K$, put $k \stackrel{\text{def}}{=} i$, and if not, then there is $j < i$ such that $f(i) \leqslant f(j)$, and we apply the inductive hypothesis to it.

Thus the range of $f \upharpoonright_K$ is cofinal within that of f, and (since $f \upharpoonright_K$ is strictly monotone) has the same order-type as K, which is $\leqslant \alpha$ by Proposition 8.6(2). □

Let λ be a limit. The cofinal and strictly cofinal subclasses are the same (as stated above), and the order-type of each is a limit. The least such order-type is called the *cofinality* of λ, and written $\mathsf{cf}(\lambda)$. Clearly it satisfies $\mathsf{cf}(\lambda) \leqslant \lambda$ and $\mathsf{cf}(\mathsf{cf}(\lambda)) = \mathsf{cf}(\lambda)$.

Now we characterize sup-closedness and sup-completeness using cofinality.

Proposition 11.6. *A limit λ is*

1. *K-sup-closed, for a set K, iff $\aleph^*(K) \leqslant \mathsf{cf}(\lambda)$.*
2. *$\mathcal{K}$-sup-complete, for a set of sets $\mathcal{K}$, iff $\bigvee_{K \in \mathcal{K}} \aleph^*(K) \leqslant \mathsf{cf}(\lambda)$.*
3. *F-sup-closed, for a function $F : \mathsf{Ord} \to \mathfrak{S}$, iff $\mathsf{cf}(\lambda)$ is $(\alpha \mapsto \aleph^*(F(\alpha)))$-based.*
4. *H-sup-complete, for a function $H : \mathsf{Ord} \to \mathcal{P}\mathfrak{S}$, iff $\mathsf{cf}(\lambda)$ is $(\alpha \mapsto \bigvee_{K \in H(\alpha)} \aleph^*(K))$-based.*

Proof. We prove part 1, from which the other parts follow.

For ($\Rightarrow$), take a cofinal subclass B of λ with order-type $\mathsf{cf}(\lambda)$. If $\mathsf{cf}(\lambda) < \aleph^*(K)$, then we have a surjection $K \to \mathsf{cf}(\lambda)$, and hence $K \to B$, so λ is B-sup-closed, so $\bigvee_{\beta < B} \beta < \lambda$, a contradiction.

For ($\Leftarrow$), the range of any $f : K \to \lambda$ has order-type in $\aleph^*(K)$ and hence in $\mathsf{cf}(\lambda)$, so its supremum is $< \lambda$. □

Where the sets in questions are ordinals, we give a simpler characterization:

Proposition 11.7. *A limit λ is*

1. *α-sup-closed, for an ordinal α, iff $\alpha < \mathsf{cf}(\lambda)$.*
2. *ρ-sup-complete, for an ordinal ρ, iff $\rho \leqslant \mathsf{cf}(\lambda)$.*
3. *F-sup-closed, for a function $F : \mathsf{Ord} \to \mathsf{Ord}$, iff $\mathsf{cf}(\lambda)$ is F-closed.*
4. *G-sup-complete, for a function $G : \mathsf{Ord} \to \mathsf{Ord}$, iff $\mathsf{cf}(\lambda)$ is G-based.*

Proof. Part 1 is by Proposition 11.5, and the rest follows. □

11.3 Regular limits

We revisit the notion of regular limit from Section 1.2. First we note that Proposition 11.7(2) gives the following:

Corollary 11.8. *A limit λ is λ-sup-complete iff $\mathsf{cf}(\lambda) = \lambda$.*

A *regular* limit is one that satisfies these conditions. Thus $\mathsf{cf}(\lambda)$ is regular, for any limit λ. Here is another way of obtaining examples:

Proposition 11.9. *A limit that is sup-generated by either*

- *a set K*
- *a set of sets $\mathcal{K}$*
- *a function $F : \mathsf{Ord} \to \mathfrak{S}$*
- *or a function $H : \mathsf{Ord} \to \mathcal{P}\mathfrak{S}$*

is regular.

Proof. We first prove that, for a set of sets $\mathcal{K}$, a minimal $\mathcal{K}$-sup-complete limit λ is regular. Fix $\alpha < \lambda$ and a tuple $[x_i]_{i<\alpha}$ within λ. We shall show that $\bigvee_{i<\alpha} x_i < \lambda$. For $\beta < \lambda$, put $\overline{\beta} \stackrel{\text{def}}{=} \bigvee_{i<\alpha\wedge\beta} x_i$. The class $P \stackrel{\text{def}}{=} \{\beta < \lambda \mid \overline{\beta} < \lambda\}$ is a $\mathcal{K}$-sup-complete limit $\leqslant \lambda$ by the following reasoning.

- For $\gamma \leqslant \beta \in P$ we have $\gamma \in P$, since $\overline{\gamma} \leqslant \overline{\beta}$. So P is an ordinal $\leqslant \lambda$.
- We have $0 \in P$ since $\overline{0} = 0$, and for any $\beta \in P$ we have $\mathsf{S}\beta \in P$, since $\overline{\mathsf{S}\beta}$ is $\overline{\beta} \wedge x_\beta$ if $\beta < \alpha$ and $\overline{\beta}$ otherwise So P is a limit.
- For any $K \in \mathcal{K}$ and tuple $[\beta_k]_{k\in K}$ within P, we have $\bigvee_{k\in K} \beta_k \in P$ since $\overline{\bigvee_{k\in K} \beta_k} = \bigvee_{k\in K} \overline{\beta_k}$. So P is $\mathcal{K}$-sup-complete.

Minimality of λ gives $P = \lambda$, so $\alpha \in P$, meaning that $\bigvee_{i<\alpha} x_i = \overline{\alpha} < \lambda$ as required.

Lastly, for $H : \mathsf{Ord} \to \mathcal{P}^{\blacksquare}(\mathfrak{S})$, any minimal H-sup-complete limit λ is also a minimal $(\bigcup_{\beta<\lambda} H\beta)$-sup-complete limit, and therefore regular. $\square$

11.4 Blass's axiom and Mahlo's principle

Now we can revisit the principles from Section 1.3: *Blass's axiom* says that Reg is unbounded, and *Mahlo's principle* that Reg is stationary. Here is an analogue of Proposition 10.3:

Proposition 11.10. *Blass's axiom implies Ordinal Lindenbaum.*

Proof. Given an ordinal α, take a regular limit $\lambda > \alpha$. It is is α-sup-closed by Proposition 11.7(1). We show $\aleph^*(\alpha) \leqslant \lambda$. For any $\beta < \aleph^*\alpha$, we have a partial surjection $\alpha \to \beta$, so λ is β-sup-closed, so $\beta < \lambda$ by Proposition 11.7(1). $\square$

We arrive at the main result of the section:

Proposition 11.11.

1. *Wide Sup-generation is equivalent to Blass's axiom + Relative Lindenbaum.*

2. *Broad Sup-generation is equivalent to Mahlo's principle + Relative Lindenbaum.*

Proof. We prove part 2, as part 1 is similar.

For ($\Rightarrow$), Broad Sup-generation implies Relative Lindenbaum by Proposition 10.3(2). To show that it implies Mahlo's principle, let $F : \mathsf{Ord} \Rightarrow \mathsf{Ord}$. The limit that F simply sup-generates is F-sup-closed so by Proposition 11.7(3) it is F-closed. It is regular by Proposition 11.9.

We give two proofs of ($\Leftarrow$), as the former may be more appealing to people who accept Well-orderability, and the latter to people who accept Powerset.

- Let $F : \mathsf{Ord} \to \mathfrak{S}$. Define $G : \mathsf{Ord} \to \mathsf{Ord}$ sending α to the least ordinal β such that $\aleph^*(F(\alpha)) \leqslant \aleph^*(\beta)$. Then there is a regular limit that is G-closed. By Proposition 11.7(3) it is G-sup-closed, so by Proposition 11.6(3) it is F-sup-closed.

- Full Lindenbaum holds by Proposition 11.10. For any function $H : \mathsf{Ord} \to \mathcal{P}\mathfrak{S}$, there is a regular limit that is $(\alpha \mapsto \bigvee_{K \in H(\alpha)} \aleph^*(K))$-based. By Proposition 11.6(4) it is H-sup-complete. $\square$

Corollary 11.12. (Assuming Powerset or Well-orderability.)

1. *Wide Sup-generation is equivalent to Blass's axiom.*

2. *Broad Sup-generation is equivalent to Mahlo's principle.*

Proof. Immediate from Proposition 11.11, using the fact that Relative Lindenbaum follows from Powerset and from Well-orderability. $\square$

12 The power of stationarity

12.1 Club classes and continuous ordinal functions

Our final task is to develop the traditional theory of stationarity, in which α-inaccessibles (for $\alpha \in \mathsf{Ord}$) and hyper-inaccessibles are obtained from Mahlo's principle. Throughout this section, *class* will always mean a class of ordinals, and *function* an endofunction on Ord. We use the following constructions:

Definition 12.1.

1. For any monotone function H, we write $\mathsf{Pref}(H)$ for the class of all its prefixpoints.

2. For any family of classes $(C_i)_{i\in I}$, the *intersection* is given by

$$\bigcap_{i\in I} C_i \quad \stackrel{\text{def}}{=} \quad \{\alpha\in\mathsf{Ord} \mid \forall i\in I.\, \alpha \in C_i\}$$

3. For any family of monotone functions $(H_i)_{i\in I}$, the *supremum* is given by

$$\begin{array}{rrcl} & \bigvee_{i\in I} H_i \quad : \quad \alpha & \mapsto & \bigvee_{i\in I} H_i(\alpha) \\ \text{so that} & \mathsf{Pref}(\bigvee_{i\in I} H_i) & = & \bigcap_{i\in I} \mathsf{Pref}(H_i) \end{array}$$

4. For any sequence of classes $(C_\alpha)_{\alpha\in\mathsf{Ord}}$, the *diagonal intersection* is given by

$$\Delta_{\alpha\in\mathsf{Ord}} C_\alpha \quad \stackrel{\text{def}}{=} \quad \{\alpha\in\mathsf{Ord} \mid \forall\beta < \alpha.\, \alpha \in C_\beta\}$$

5. For any sequence of monotone functions $(H_\alpha)_{\alpha\in\mathsf{Ord}}$, the *diagonal supremum* is given by

$$\begin{array}{rrcl} & \nabla_{\alpha\in\mathsf{Ord}} H_\alpha \quad : \quad \alpha & \mapsto & \bigvee_{\beta<\alpha} H(\beta) \\ \text{so that} & \mathsf{Pref}(\nabla_{\alpha\in\mathsf{Ord}} H_\alpha) & = & \Delta_{\alpha\in\mathsf{Ord}} \mathsf{Pref}(H_\alpha) \end{array}$$

Definition 12.2. A function H is *continuous* when it is monotone and sends every limit λ to $\bigvee_{\alpha<\lambda} H(\alpha)$.

Here are some ways to obtain continuous functions:

Proposition 12.3.

1. *For every ordinal α, the function Const_α is continuous.*
2. *For any family of continuous functions $(H_i)_{i\in I}$, the supremum $\bigvee_{i\in I} H_i$ is continuous.*
3. *For any sequence of continuous functions $(H_\alpha)_{\alpha\in\mathsf{Ord}}$, the diagonal supremum $\nabla_{\alpha\in\mathsf{Ord}} H_\alpha$ is continuous.*

Proof.

1. Obvious.
2. Straightforward.

3. Monotonicity is obvious. For continuity, let λ be a limit. Then

$$\begin{aligned}(\nabla_{\alpha \in \mathsf{Ord}} H_\alpha)(\lambda) &= \bigvee_{\alpha<\lambda} H_\alpha(\lambda) \\ &= \bigvee_{\alpha<\lambda} \bigvee_{\beta<\lambda} H_\alpha(\beta) \\ &\leqslant \bigvee_{\alpha<\lambda} \bigvee_{\beta<\lambda} H_\alpha(\beta \vee \mathsf{S}\alpha) \\ &\leqslant \bigvee_{\gamma<\lambda} \bigvee_{\alpha<\gamma} H_\alpha(\gamma) \\ &= \bigvee_{\gamma<\lambda} (\nabla_{\alpha \in \mathsf{Ord}} H_\alpha)(\gamma)\end{aligned}$$

□

Definition 12.4. Let C be a class.

1. A *limit point* of C is a limit λ such that $\lambda \cap C$ is unbounded in λ. The class of all such is written $\mathsf{LimPt}(C)$.

2. A class C is *closed* when it is LimPt-prefixed, i.e., contains every limit λ such that $\lambda \cap C$ is unbounded in λ.

Here are some ways to obtain closed classes:

Proposition 12.5.

1. *For any continuous function H, the class $\mathsf{Pref}(H)$ is closed.*

2. *For any class C, the class $\mathsf{LimPt}(C)$ is closed.*

3. *For a family of closed classes $(C_i)_{i \in I}$, the intersection $\bigcap_{i \in I} C_i$ is closed.*

4. *For a sequence of closed classes $(C_\alpha)_{\alpha \in \mathsf{Ord}}$, the diagonal intersection $\Delta_{\alpha \in \mathsf{Ord}} C_\alpha$ is closed.*

Proof.

1. We must show that a limit point λ of $\mathsf{Pref}(H)$ is in $\mathsf{Pref}(H)$. For any $\gamma < \lambda$, there is $\beta \in [\gamma \,..\, \lambda) \cap \mathsf{Pref}(H)$, giving $H(\gamma) \leqslant H(\beta) \leqslant \beta < \lambda$. So we have

$$\begin{aligned}H(\lambda) &= \bigvee_{\gamma<\lambda} H(\gamma) \\ &\leqslant \lambda\end{aligned}$$

as required.

2. We show that a limit point λ of $\mathsf{LimPt}(C)$ is a limit point of C. For any $\alpha < \lambda$, there is $\beta \in (\alpha \,..\, \lambda) \cap \mathsf{LimPt}(C)$. Since $\beta \in \mathsf{LimPt}(C)$, there is $\gamma \in [\alpha \,..\, \beta) \cap C$. Thus we have $\gamma \in [\alpha \,..\, \lambda) \cap C$ as required.

3. Since an infimum of prefixpoints is a prefixpoint.

4. It suffices to show that $\mathsf{LimPt}(\Delta_{\alpha\in\mathsf{Ord}}C_\alpha) \subseteq \Delta_{\alpha\in\mathsf{Ord}}\mathsf{LimPt}(C_\alpha)$. This means that any limit point λ of $\Delta_{\alpha\in\mathsf{Ord}}C_\alpha$ is, for all $\alpha < \lambda$, a limit point of C_α. For any $\beta \in (\alpha \mathinner{..} \lambda)$, there is $\gamma \in [\beta \mathinner{..} \lambda) \cap \Delta_{\alpha\in\mathsf{Ord}}C_\alpha$. Since $\alpha < \beta \leqslant \gamma$, we have $\gamma \in C_\alpha$ as required. $\square$

Next we consider unbounded classes.

Definition 12.6.

1. A *closure operator* (on Ord) is a function that is inflationary and idempotent.

2. For an unbounded class C, we write H_C for the unique closure operator whose range is C. Explicitly, it sends α to the least element of C that is $\geqslant \alpha$.

Thus we have a bijection between the collection of all unbounded classes and that of all closure operators. Moreover it is *dual*, i.e., for unbounded classes C and D, we have $H_C \leqslant H_D$ iff $D \subseteq C$.

Proposition 12.7. *For any unbounded class C, the following are equivalent:*

- *C is closed.*
- *H_C is continuous.*

Proof. Firstly, H_C is continuous at every $\lambda \not\in \mathsf{LimPt}(C)$, since there is $\alpha < \lambda$ such that $[\alpha \mathinner{..} \lambda)$ has no element in C, so H_C sends every ordinal in this interval to $H_C(\lambda)$. As for $\lambda \in \mathsf{LimPt}(C)$, we have $\bigvee_{\gamma<\lambda} H_C(\gamma) = \lambda$, so H_C is continuous at λ iff $\lambda \in C$. $\square$

Thus we have a dual bijection between the collection of all closed unbounded classes, known as *club classes*, and that of all continuous closure operators. Here are some ways to obtain club classes:

Proposition 12.8. *Each of the following is equivalent to Infinity.*

1. *For any continuous function H, the class $\mathsf{Pref}(H)$ is club.*
2. *For any unbounded class C, the class $\mathsf{LimPt}(C)$ is club.*
3. *For a family of club classes $(C_i)_{i\in I}$, the intersection $\bigcap_{i\in I} C_i$ is club.*
4. *For a sequence of club classes $(C_\alpha)_{\alpha\in\mathsf{Ord}}$, the diagonal intersection $\Delta_{\alpha\in\mathsf{Ord}}C_\alpha$ is club.*

Proof. Firstly, if Infinity does not hold (so that a club class is just an unbounded class of natural numbers), then the statements are all false:

1. $\mathsf{Pref}(\mathsf{S})$ is empty.

2. $\mathsf{LimPt}(\mathsf{Ord})$ is empty.

3. Let C be the class of all even numbers and D that of all odd numbers. They are club, but $C \cap D$ is empty.

4. For each ordinal n, let C_n be the class of ordinals $\geqslant n+2$. These are club, but $\Delta_{n \in \mathsf{Ord}} C_n = \{0\}$.

Now assume Infinity. Because of Proposition 12.5, we need only prove unboundedness.

1. Let α be an ordinal α. Form a strictly increasing sequence of ordinals $(x_n)_{n\in\mathbb{N}}$ via $x_0 \stackrel{\text{def}}{=} \alpha$ and $x_{n+1} \stackrel{\text{def}}{=} \mathsf{S}(x_n) \vee H(x_n)$. Its supremum λ is a limit and satisfies

$$
\begin{aligned}
H(\lambda) &= \bigvee_{\gamma<\lambda} H(\beta) \\
&= \bigvee_{n\in\mathbb{N}} H(x_n) \\
&\leqslant \bigvee_{n\in\mathbb{N}} x_{n+1} \text{ (since } H(x_n) \leqslant x_{n+1}) \\
&= \lambda
\end{aligned}
$$

so λ is an H-prefixpoint $\geqslant \alpha$.

2. Let α be an ordinal. Form a strictly increasing sequence of ordinals $(x_n)_{n\in\mathbb{N}}$ via $x_0 \stackrel{\text{def}}{=} \alpha$ and $x_{n+1} \stackrel{\text{def}}{=}$ the least ordinal in C that is greater than x_n. Then $\bigvee_{n\in\mathbb{N}} x_n$ is in $\mathsf{LimPt}(C)$ and is $> \alpha$.

3. Since

$$
\begin{aligned}
\bigcap_{i\in I} C_i &= \bigcap_{i\in I} \mathsf{Pref}(H_{C_i}) \\
&= \mathsf{Pref}(\bigvee_{i\in I} H_{C_i})
\end{aligned}
$$

which is club by part 1, using Proposition 12.3(2) .

4. Similar. ☐

Proposition 12.9. *For a class of limits D, the following are equivalent.*

1. *D is stationary.*

2. *Every continuous function has a prefixpoint in N.*

3. *Every club class has an element in D.*

Proof. To show that (1) implies (2), let H be a continuous function. Then any H-based limit is H-prefixed. To show the converse, let G be a function. Then the function $H \stackrel{\text{def}}{=} \nabla_{\alpha \in \mathsf{Ord}} \mathsf{const}_{G(\alpha)}$ is continuous by Proposition 12.3, and a limit is G-based iff it is H-prefixed.

To show that (2) implies (1): for any club class C, we have $C = \mathsf{Pref}(H_C)$ and H_C is continuous. To show the converse, we note that the club class Ord has an element in $D \subseteq \mathsf{Lim}$. So Infinity holds and we can apply Proposition 12.8(1). □

Corollary 12.10. *Let D be a stationary class of limits. The intersection of D with any club class is stationary.*

12.2 Application: iterated inaccessibility

In order to formulate iterated inaccessibility, we use the following result.

Proposition 12.11. *Let D be a class of limits.*

1. *There is a sequence of classes $(X_\alpha)_{\alpha \in \mathsf{Ord}}$ and class X_ω uniquely specified by*

$$\begin{aligned} X_\alpha &= D \cap \bigcap_{\beta < \alpha} \mathsf{LimPt}(X_\beta) \\ X_\Omega &= D \cap \Delta_{\beta \in \mathsf{Ord}} \mathsf{LimPt}(X_\beta) \end{aligned}$$

2. *If D is stationary, then so are all these classes.*

Proof.

1. Recall that we cannot (in general) define a sequence of classes recursively, so we proceed as follows. For any ordinal ρ, we recursively define a sequence $(X^\rho_\alpha)_{\beta \in \mathsf{Ord}}$ of subsets of ρ via

$$X^\rho_\alpha = \rho \cap D \cap \bigcap_{\beta < \alpha} \mathsf{LimPt}(X^\rho_\beta)$$

These sequences are compatible in the sense that, for $\rho \leqslant \sigma$, we have $X^\rho_\alpha = \rho \cap X^\sigma_\alpha$. We define

$$X_\alpha \stackrel{\text{def}}{=} \{\rho \in \mathsf{Ord} \mid \rho \in X^{\mathsf{S}\rho}_\alpha\}$$

and obtain the required properties.

2. Firstly, since D is stationary, Infinity holds and we can use Proposition 12.8.

We cannot simply prove stationarity of X_α by induction on α, as this assertion involves second-order quantification. Instead, we prove unboundedness of X_α by induction on α, as follows. For all $\beta < \alpha$, the class X_β

is unbounded, so $\mathsf{LimPt}(X_\beta)$ is club. So $\bigcap_{\beta<\alpha} X_\beta$ is club, making X_α stationary and hence unbounded. This completes the induction.

Next, for any ordinal α, we see (again) that $\bigcap_{\beta<\alpha} \mathsf{LimPt}(X_\beta)$ is club, making X_α stationary. Likewise $\nabla_{\beta\in\mathsf{Ord}}\mathsf{LimPt}(X_\beta)$ is club, making X_Ω stationary. □

For an application **assuming Powerset + Infinity + AC**, let D be the class of all inaccessibles. Then X_α is the class of all *α-inaccessibles*, and X_Ω the class of all *hyper-inaccessibles*. Proposition 12.11(2) tells us that these are stationary if Broad Infinity holds.

This construction can be further iterated, giving hyper-hyper-inaccessibles and more. In Carmody's work [Car17], this is achieved by generalising the subscripts used in Proposition 12.11 to a system of "meta-ordinals".

Related work A standard treatment of stationarity is given in [Jec03]. The focus there is not on classes but on subsets of a given limit. So the predicativity issue does not arise, and additional results are obtained, such as Fodor's pressing-down lemma and Solovay's partitioning theorem. See [GHK21] for an analysis of whether Fodor's lemma applies to classes.

13 Conclusions

13.1 Summary of achievements

We have now established all the relationships in Figure 1. The main technical achievement was proving the equivalence (assuming Powerset + AC) of Simple Broad Infinity and Mahlo's principle. The centrepiece is the implication Full Broad Infinity ⇒ Broad Derivation Set, which relies on Proposition 4.7 to generate the $\mathcal{R}$-derivational class-family.

On the philosophical side, I claim that the notion of F-broad number (for a broad arity F) is easily grasped, making Simple Broad Infinity a plausible axiom scheme. This is for the reader to judge.

On the practical side, we have seen several equivalent principles that are convenient for applications. Specifically:

- Broad Derivation Set yields the existence of Tarski-style universes.
- Broad Subset Generation yields the existence of Grothendieck universes.
- Mahlo's principle in the form "Every club class contains a regular limit" yields the existence of α-inaccessibles (for $\alpha \in \mathsf{Ord}$) and hyper-inaccessibles.

As required in Section 2.5, we have developed our results in a setting that allows urelements and non-well-founded membership, proved the sufficiency of Global WISC for our main AC-reliant results, and seen the pattern of resemblance between Wide and Broad principles throughout the paper.

13.2 Further work

Despite the above, much work remains to be done. Firstly, there are unanswered questions, particularly about the power of Broad ZF.

1. By analogy with Gitik's work [Git80], can we show, under some large cardinal consistency hypothesis, that Broad ZF does not prove the existence of an uncountable regular limit?

2. Jech [Jec82] showed that ZF proves the class of hereditarily countable sets to be a set, and his result has been extended to other cardinalities [Die92, Hol14]. Is there something similar for Broad ZF? In particular, consider the following. By Proposition 4.2(2), for any function $F : \mathfrak{T} \to \mathfrak{S}$, there is a least class X that contains $\mathsf{Nothing}$ and, for any $x \in X$ and $f \in X^{F(x)}$, contains $\mathsf{Just}\langle x, \mathsf{Range}(f)\rangle$. Does Broad ZF prove this to be a set?

3. If Broad ZF does not already prove Mahlo's principle, then what about Broad ZF + Blass's axiom?

Everything in this paper has been done in a base theory that—like ZF—uses classical first-order logic and pays no attention to logical complexity. But some other versions of set theory use intuitionistic logic and/or restrict the use of logically complex sentences [Cro20, Mat01]. The task of adapting our results to such theories (as far as possible) is left to future work.

Lastly, the link between type-theoretic work on induction-recursion [DS06, GH16] and the principles in this paper remains to be developed.

Acknowledgments I thank Asaf Karagila for teaching me about Lindenbaum numbers. I also thank Tom de Jong, for persuading me to write a paper just on classical set theory, and to leave intuitionistic and constructive set theory for Part 2, which hopefully will come to fruition.

Bibliography

[AAG05] Michael Abbott, Thorsten Altenkirch, and Neil Ghani. Containers: Constructing strictly positive types. *Theoretical Computer Science*, 342(1):3–27, 2005. Applied Semantics: Selected Topics.

[AR10] Peter Aczel and Michael Rathjen. Constructive Set Theory. Book draft, 2010.

[Bar93] Michael Barr. Terminal coalgebras in well-founded set theory. *Theoretical Computer Science*, 114(2):299–315, 21 June 1993.

[BDL07] Andreas Blass, Ioanna Dimitriou, and Benedikt Löwe. Inaccessible Cardinals without the Axiom of Choice. *Fundamenta Mathematicae*, 194:179–189, 01 2007.

[Bla83] Andreas Blass. Words, Free Algebras and Coequalizers. *Fundamenta Mathematicae*, 117:117–160, 1983.

[Car17] Erin Kathryn Carmody. Killing them softly: degrees of inaccessible and Mahlo cardinals. *Mathematical Logic Quarterly*, 63(3-4):256–264, 2017.

[Cro20] Laura Crosilla. Set Theory: Constructive and Intuitionistic ZF. In *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2020 edition, 2020.

[Die92] Karl-Heinz Diener. On the transitive hull of a κ-narrow relation. *Mathematical Logic Quarterly*, 38(1):387–398, 1992.

[Dow11] Martin Dowd. Some new axioms for set theory. *International Journal of Pure and Applied Mathematics*, 66(2):121–136, 2011.

[DS06] Peter Dybjer and Anton Setzer. Indexed induction-recursion. *The Journal of Logic and Algebraic Programming*, 66(1):1–49, 2006.

[GH16] Neil Ghani and Peter Hancock. Containers, monads and induction recursion. *Mathematical Structures in Computer Science*, 26(1):89–113, 2016.

[GHK21] Victoria Gitman, Joel Hamkins, and Asaf Karagila. Kelley–Morse set theory does not prove the class Fodor principle. *Fundamenta Mathematicae*, 254, 01 2021.

[Git80] Moti Gitik. All uncountable cardinals can be singular. *Israel Journal of Mathematics*, 35(1–2):61–88, 1980.

[Ham03] Joel David Hamkins. A Simple Maximality Principle. *The Journal of Symbolic Logic*, 68(2):527–550, 2003.

[Ha06] Haskell wiki. Maybe, 2006. `wiki.haskell.org/Maybe`.

[HMG+13] Peter Hancock, Conor McBride, Neil Ghani, Lorenzo Malatesta, and Thorsten Altenkirch. Small induction recursion. In *Proceedings, 11th International Conference, Typed Lambda Calculi and Applications*. Springer, Eindhoven, Netherlands, 2013. Preprint.

[Hol14] Randall Holmes. On hereditarily small sets in ZF. *Mathematical Logic Quarterly*, 60, 05 2014. Preprint.

[Jec82] Thomas Jech. On hereditarily countable sets. *The Journal of Symbolic Logic*, 47(1):43–47, 1982.

[Jec03] Thomas Jech. *Set Theory, the Third Millennium Edition, Revised and Expanded*. Springer, 2003.

[Jor70] Murray Jorgensen. An equivalent form of Lévy's axiom schema. *Proceedings of the American Mathematical Society*, 26:651–654, 04 1970.

[Kar14] Asaf Karagila. Embedding orders into cardinals with DC_κ. *Fundamenta Mathematicae*, 226:143–156, 2014. Preprint.

[Ker65] Robert Kerkhoff. Eine Konstruktion absolut freier Algebren (A construction of absolutely free algebras). *Mathematische Annalen*, 158:109–112, 1965.

[KS10] Reinhard Kahle and Anton Setzer. An extended predicative definition of the Mahlo universe. In Ralf Schindler, editor, *Ways of Proof Theory*, Ontos Series in Mathematical Logic, pages 315–340. Ontos Verlag, Frankfurt (Main), Germany, 2010. Preprint.

[Le60] Levy, Azriel. Axiom schemata of strong infinity in axiomatic set theory. *Pacific Journal of Mathematics*, 10(1):223–238, 1960.

[Mat77] Adrian R.D. Mathias. Happy families. *Annals of Mathematical Logic*, 12(1):59–111, 1977.

[Mat01] Adrian R.D. Mathias. The strength of Mac Lane set theory. *Annals of Pure and Applied Logic*, 110(1-3):107–234, 2001. Preprint.

[May00] John P. Mayberry. *The Foundations of Mathematics in the Theory of Sets*. Cambridge University Press, 2000.

[ML84] Per Martin-Löf. *Intuitionistic type theory*. Bibliopolis, Napoli, 1984.

[Mon62] Richard Montague. Two contributions to the foundation of set theory. In *Proc. International Congress in Logic, Methodology and Philosophy of Science (1960)*, pages 94–110. Stanford University Press, 1962.

[MP00] Ieke Moerdijk and Erik Palmgren. Wellfounded trees in categories. *Annals of Pure and Applied Logic*, 104(1):189–218, 2000.

[MS78] Zohar Manna and Adi Shamir. The convergence of functions to fixedpoints of recursive definitions. *Theoretical Computer Science*, 6(2):109–141, 1978.

[nLa11] nLab. Well-orders are linear, 2011. Explanatory note.

[PS78] Robert Paré and Dietmar Schumacher. Abstract families and the adjoint functor theorems. In *Indexed categories and their applications*, pages 1–125. Springer, 1978.

[Rat00] Michael Rathjen. The superjump in Martin-Löf type theory. In *Proceedings, Logic Colloquium 1998*, volume 13 of *Lecture Notes in Logic*, pages 363–386. Association for Symbolic Logic, 2000. Preprint.

[Rob15] David Michael Roberts. The Weak Choice Principle WISC may Fail in the Category of Sets. *Studia Logica: An International Journal for Symbolic Logic*, 103(5):1005–1017, 2015.

[Set00] Anton Setzer. Extending Martin-Löf Type Theory by one Mahlo-Universe. *Arch. Math. Log.*, 39:155–181, 2000.

[Sł58] Józef Słomiński. Theory of models with infinitary operations and relations. *Bulletin de l'Académie Polonaise des Sciences*, 6:449–456, 1958.

[SP82] Michael Smyth and Gordon D. Plotkin. The category-theoretic solution of recursive domain equations. *SIAM Journal on Computing*, 11:761–783, 1982.

[Str05] Thomas Streicher. Realizability models for CZF and ¬Pow. Unpublished note, 2005.

[vdB11] Benno van den Berg. Categorical semantics of constructive set theory, 2011. Habilitation thesis, Technische Universität Darmstadt.

[vdBM14] Benno van den Berg and Ieke Moerdijk. The axiom of multiple choice and models for constructive set theory. *Journal of Mathematical Logic*, 14(01):1450005, 2014.

[Wan77] Hao Wang. Large sets. In *Logic, Foundations of Mathematics, and Computability Theory*, pages 309–333. Springer, 1977.

[Zer08] Ernst Zermelo. Untersuchungen über die Grundlagen der Mengenlehre (Investigations in the foundations of set theory). *Mathematische Annalen*, 65(2):261–281, 1908.