Mathematics from First Principles

Randall Holmes

April 2, 2025

Contents

1 Introduction

As with many of my writings, this is an experimental math textbook, which may never be used as a textbook. The format is good for foundational exposition.

Of course, there is malice here. The theory we discuss is an unsorted first order theory with equality and membership, but it is actually the simple typed theory of sets presented as an unsorted theory, following a suggestion of Quine. In many ways this is like set theory as usually presented, and in many ways it is different.

2 A first pass at logic

Statements are either true or false.

If A is a statement, $\neg A$ (it is not the case that A) is false if A is true and true if A is false.

If A and B are statements, $A \wedge B$ (A and B) is a statement which is true if A is true and B is true, and otherwise false.

We define $A \vee B$ (A or B) as $\neg(\neg A \wedge \neg B)$. This is false if A and B are both false, and otherwise true.

We define $A \to B$ (if A then B) as $\neg(A \wedge \neg B)$. This is false exactly if A is true and B is false.

We define $A \leftrightarrow B$ (A if and only if B) as $(A \to B) \wedge (B \to A)$.

Basic statements of our mathematical language are $x = y$ (x is the same object as y) and $x \in y$ (x is an element of y). We write $x \neq y$ for $\neq x = y$ and $x \notin y$ for $\neg x \in y$.

If $P(x)$ is a statement, x being a variable, $(\forall x : P(x))$ is a statement, which is true iff $P(a)$ is true for every choice of object a, and $(\exists x : P(x))$ is a statement, which is true precisely if there is an object a such that $P(a)$ is true.

If $P(x)$ is a statement *in which no variable appears which is bound in a larger context* (a qualification which may surprise the knowledgeable reader), we use the notation $(\epsilon x : P(x))$ to represent an object a such that $P(a)$. Such an expression is called a Hilbert symbol. Note that

$$(\exists x : P(x)) \leftrightarrow P((\epsilon x : P(x))),$$

when both sides are defined. This does not mean that the existential quantifier is eliminable in favor of the Hilbert symbol: the fact that variable binding into Hilbert symbols is not allowed obstructs this. It is true that all quantifiers (existential and universal) in a statement *which are not inside Hilbert symbols* can be systematically eliminated by a top down procedure in favor of Hilbert symbols. However, the quantifiers are still needed as primitive because they have syntactical privileges that the Hilbert symbol does not, and in fact quantifiers may still appear in the Hilbert symbols after the top down procedure. They are also needed in practice because the expansion to Hilbert symbols causes an explosion in length. Allowing variable binding into Hilbert symbols would amount to assuming the axiom of choice.

3 Nonempty sets

We give our first

Definition: An object x is a *nonempty set* if and only if $(\exists y : y \in x)$. We write $\mathtt{Set}(x)$ for "x is a nonempty set". For the moment, when we say "set", we mean nonempty set.

Definition: An object is *empty* iff it has no elements. We may choose to call certain empty objects *empty sets* (and if we do, $\mathtt{set}(x)$ will mean that x is either a nonempty set or an empty set); empty objects which are not empty sets we call *atoms*.

Sets are abstract objects, not part of the ordinary furniture of the world (though in the approach to mathematics we take here, we preserve the possibility that all objects in the ordinary furniture of the world are in the domain of discourse), and we must be very clear about their characteristics (as with any mathematical objects).

Clarity about a class of objects is enhanced if we can tell when objects of this class are the same and when they are different. This is the subject of our first

Axiom (extensionality): If A and B are nonempty sets, $A = B$ iff A and B have the same elements. In symbols,

$$(\mathtt{Set}(A) \wedge \mathtt{Set}(B)) \to (A = B \leftrightarrow (\forall x : x \in A \leftrightarrow x \in B)).$$

This is enough about nonempty sets for a first section.

4 Sets, kinds, and properties

We have said when two sets are the same. But we have not said what sets there are. Our basic intuitive idea is that if we have a kind K of objects we can for every property $P(x)$ of a variable x ranging over kind K define a set containing exactly the objects of that kind with that property.

We will express everything in this intuition in a series of definitions and axioms. First, we give a definition of the notion that two objects are of a kind.

Definition (of a kind): We say that objects x and y are "of a kind" (written $x \sim y$) just in case $(\exists z : x \in z \land y \in z)$.

From our intuition we can see that if a and b are of the same kind K, they belong to a common set: for example, they both belong to the set of objects x of kind K such that $x = a \lor x = b$, for which we will later introduce the notation $\{a, b\}$. The idea that if two objects belong to the same set they are of the same kind is not as evident, but can be viewed as part of the firming up of our understanding of the abstraction we are developing[1].

We then introduce an axiom providing for kinds.

Axiom of Kinds: For each object x there is an object K such that $x \in K$ and $(\forall y : y \in K \leftrightarrow y \sim x)$.

Observation and Definition: Notice that for any x, an object K provided by the axiom is a nonempty set (since it contains x) and there is only one such object by extensionality, since its members are exactly determined. This justifies the definition of $\kappa(x)$, the kind of x, as the unique object K such that $x \in K$ and $(\forall y : y \in K \leftrightarrow y \sim x)$.

Theorem ($\sim$ is an equivalence relation): For any x, y, z, (1) $x \sim x$, (2) $x \sim y \rightarrow y \sim x$, and (3) $(x \sim y) \land (y \sim z) \rightarrow x \sim z$.

Proof: $x \in \kappa(x)$ and $x \in \kappa(x)$, so $x \sim x$.

If $x \sim y$, choose z such that $x \in z$ and $y \in z$: it follows that $y \in z$ and $x \in z$ so $y \sim z$.

If $x \sim y$ and $y \sim z$, then $x \sim y$ and $z \sim y$, so $x \in \kappa(y)$ and $z \in \kappa(y)$, so $x \sim z$.

---

[1] In fact, this is the profound difference between the set theory presented here and more usual treatments.

This justifies reading $x \sim y$ as "x and y are of the same kind".
Reference to kinds of kinds is not unusual and motivates a

Definition: For any object x, we define $\kappa^1(x)$ as $\kappa(x)$, and for any positive integer constant n define $\kappa^{n+1}(x)$ as $\kappa(\kappa^n(x))$. The superscript n cannot be quantified over[2]: this is an independent definition for each n.

Now we introduce the formalization of our intuition as to what sets there are.

Axiom (separation): For each sentence $P(x)$ about an object x, we assert

$$(\forall u : (\exists A \in \kappa^2(u) : (\forall x : x \in A \leftrightarrow x \in \kappa(u) \wedge P(x)).$$

Definition: If $(\exists x \in \kappa(u) : P(x))$, the witness A to

$$(\forall u : (\exists A \in \kappa^2(u) : (\forall x : x \in A \leftrightarrow x \in \kappa(u) \wedge P(x))$$

is uniquely determined by extensionality and we call it

$$\{x \in \kappa(u) : P(x)\}.$$

Note that Separation also implies that there are objects without elements in each kind $\kappa^2(u)$, but it doesn't allow us to pick out a unique one. It is important to notice that to use a notation $\{x \in \kappa(u) : P(x)\}$ requires a demonstration that the intended extension is nonempty [unless we provide for empty sets, as discussed under the next heading].

Empty sets, an option: We reserve the right to stipulate that there is a distinguished empty object $\emptyset_{\kappa(x)} \in \kappa^2(x)$ for each kind $\kappa(x)$. These objects we call empty sets and we define $\texttt{set}(x)$ as above, and note that we get the following generalization of extensionality:

$$(\forall xy : x \sim y \wedge \texttt{set}(x) \wedge \texttt{set}(y) \wedge (\forall z : z \in x \leftrightarrow z \in y) \rightarrow x = y),$$

and if $P(x)$ is uniformly false for x in $\kappa(u)$, we can define $\{x \in \kappa(u) : P(x)\}$ as $\emptyset_{\kappa(u)}$.

This is an option. We do not officially choose this option; we do note its consequences where relevant. The reason that we are reluctant to

---

[2]We explore the possibility of doing this in a much later section.

officially adopt it is that the ability to choose a specific empty set from each kind looks like an external process of making infinitely many arbitrary choices; we will show below that our basic theory without empty sets interprets the theory with empty sets without difficulty.

We are also interested in what foundations look like without an empty set. Our impression is that the advantages of having empty sets are clear, but the overhead of not allowing them is not very great.

One could introduce an empty set in a given type $\kappa^2(a)$ as

$$(\epsilon x : x \in \kappa^2(a) \land (\forall y : y \notin x)),$$

using the Hilbert symbol, but there is less to this than meets the eye: the variable a here could not be quantified over: one could not use this definition to make general statements about empty sets. Similar problems would arise if one tried to define set builder notation using the Hilbert symbol.

We introduce some specific very simple sets.

Definition: We define $\{x\}$ or $\iota(x)$, the singleton set of x, as

$$\{y \in \kappa(x) : y = x\}.$$

We define $\iota^0(x)$ as x and $\iota^{n+1}(x)$ for each nonnegative integer x as $\iota(\iota^n(x))$. This is a separate definition for each natural number n.

If $x \sim y$ we define $\{x, y\}$, the unordered pair of x and y, as $\{u \in \kappa(x) : u = x \lor u = y\}$. Notice that $\{x, y\} \in \kappa^2(x) = \kappa^2(y)$.

If $x \sim y$, we define (x, y), the ordered pair of x and y, as $\{\{x\}, \{x, y\}\}$. This is well defined because $\kappa(\{x\}) = \kappa(\{x, y\}) = \kappa^2(x) = \kappa^2(y)$. Notice that $(x, y) \in \kappa^3(x) = \kappa^3(y)$.

Observations: It should be clear from repeated applications of the second clause of the axiom of separation that $\kappa(\iota^n(x)) = \kappa^{n+1}(x)$.

It should be noted that $x \sim y$ is a prerequisite for $\{x, y\}$ or (x, y) to be defined.[3]

---

[3]This is a distinct difference from set theory as you might have learned it on another occasion.

Note that x is the only object which belongs to every element of $\{\{x\}, \{x, y\}\}$, and y is the only object which belongs to exactly one element of $\{\{x\}, \{x, y\}\}$. From this it follows readily that

$$(x, y) = (z, w) \to x = z \land y = w.$$

We introduce a technical axiom, for the moment having no new content.

Axiom of Ordered Pairs: We postulate a concrete natural number constant Δ and for each $x \sim y$ an object $(x, y) \in \kappa^{1+\Delta}(x) = \kappa^{1+\Delta}(y)$. We assert that for all x, y, z, w, $(x, y) = (z, w) \to x = z \land y = w$. For the moment, this axiom is satisfied without taking any additional measures, setting $\Delta = 2$ and defining (x, y) as above. However, we will consider redefining the ordered pair and want to present everything about it at a suitable level of abstraction so we can shift definitions.

To make this work, it is necessary to note that we do nothing with ordered pairs except what is justified by this axiom, in what follows.

Example: The first definition of the ordered pair as a set that was given was $(x, y) = \{\{\{x\}, \emptyset_{\kappa(x)}\}, \{\{y\}\}\}$, for which $\Delta = 3$.

5   Hierarchies

We introduce the final axiom of our preamble and an extension of the definition above. The motivation for this axiom will become clear.

Axiom (diversity): For any x, y, if $\kappa(x) \sim \kappa(y)$ then $\kappa(x) = \kappa(y)$. Kinds of the same kind are equal.

Definition: For any integer n for which $\kappa^n(u)$ is defined, we define $\kappa^{n-1}(u)$ as the unique kind which belongs to $\kappa^n(u)$, if there is one. The axiom of diversity tells us there can be at most one such kind. Clearly this definition will not conflict with the definition already given for n nonnegative.

We establish that there are many kinds.

Theorem: For any x, $\kappa^2(x) \neq \kappa(x)$.

Proof: Suppose that we have an x such that $\kappa(x) = \kappa^2(x)$.

Define R as $\{y \in \kappa(x) : y \notin y\}$. $R \sim \kappa(x)$ so $R \in \kappa^2(x) = \kappa(x)$.

This requires one further remark: we need to show that R exists (that its intended extension is not empty). If $\kappa^2(x) = \kappa(x)$, we know that $\kappa^2(x)$ has at least two elements ($\kappa(x)$ and at least one atom) so $\kappa(x)$ has at least two elements, one of which would be $\kappa(x)$ if the kinds were equal, and we would have $\{\kappa(x)\}$ an element of $\kappa^2(x)$ (and so of $\kappa(x)$) and not an element of itself (because $\kappa(x)$ has at least two distinct elements), so R would be nonempty. This paragraph is not needed if we provide empty sets.

So $R \in R$ iff $R \in \kappa(x)$ [just shown to be true] and $R \notin R$. This is a contradiction.

Theorem: For each concrete $n > 1$, $\kappa^n(x) \neq \kappa(x)$.

Proof: The argument is very similar to the argument above but with some devious use of iteration of the singleton operation. Suppose $\kappa^n(x) = \kappa(x)$.

Let $R_{n,x} = \{\iota^{n-2}(y) \in \kappa^{n-1}(x) : \iota^{n-2}(y) \notin y\}$. To make it entirely clear that the existence of this set follows from Separation, rewrite it as $\{u \in \kappa^{n-1}(x) : (\exists y : u = \iota^{n-2}(y) \wedge \iota^{n-2}(y) \notin y)\}$

We also need to show that $R_{n,x}$ is nonempty: let y be any set all of whose elements are not singletons and it will belong to $R_{n,x}$; the singleton of the suitably indexed kind will work: $y = \{\kappa^{n-1}(x)\}$ works under the tortured counterfactual hypotheses of the argument. Again, this paragraph is not needed if we have empty sets.

Notice that for any $y \in \kappa(x)$, $\iota^{n-2}(y) \in \kappa^{n-1}(x)$. Notice that $R_{n,x} \in \kappa^n(x) = \kappa(x)$. It follows that $\iota^{n-2}(R_{n,x}) \in \kappa^{n-1}(R_{n,x})$. It then follows that

$$\iota^{n-2}(R_{n,x}) \in R_{n,x} \leftrightarrow \iota^{n-2}(R_{n,x}) \notin R_{n,x},$$

which is impossible.

Observation: This tells us that there are as many distinct kinds as we like (up to any concrete finite number). We will see that we cannot talk about an infinite sequence of types (all terms of a sequence are of the same kind, and the axiom of diversity tells us that there cannot be two distinct kinds of the same kind), and we cannot even express the idea that all kinds belong to one sequence of iterated kinds, though this might be a natural supposition.

A situation which we might contemplate though we cannot even express it formally is the possible existence of a kind $\kappa(x)$ such that $\kappa^i(x)$ exists for each integer i, positive and negative.

We prove some useful lemmas about kinds.

Lemmas: If $x \in y$ and $x \in \kappa^n(u)$ then $y \in \kappa^{n+1}(u)$. If $x \in y$ and $y \in \kappa^n(u)$ then $x \in \kappa^{n-1}(u)$.

Proof: If $x \in y$ then any $u \in y$ is an element of $\kappa(x) = \kappa^n(u)$, so

$$y = \{u \in \kappa^n(u) : u \in y\}$$

because both are nonempty sets and they have the same extension, and by separation,

$$y = \{u \in \kappa^n(u) : u \in y\} \in \kappa^{n+1}(u).$$

If $x \in y$ and $y \in \kappa^n(u)$, by the first part of this argument $y \in \kappa^2(x)$, so $\kappa^n(u) = \kappa^2(x)$, so by diversity $\kappa^{n-1}(u)$ exists and is equal to $\kappa(x)$.

6 Sets, atoms, and individuals: some taxonomy

Sets have already been defined as objects with elements (with the additional option of empty sets if we choose).

Objects can fail to be sets in two different ways.

Definition (individual): An object x is an *individual* if it is not of the same kind as a kind. This is equivalent to saying that $\kappa^0(x)$ does not exist, which precludes the object being a set (empty or inhabited). Anything of the same kind as an individual is an individual. There is nothing in our theory which requires two distinct individuals to be of the same kind.

Definition (atom): An object x is an *atom* if it is of the same kind as a kind and is not a set (empty sets are not atoms if we provide them). There is nothing in our theory which requires empty objects of the same kind to be equal.

Observation: Our theory does not prove the existence of either individuals or atoms, but they are important formal possibilities.

7 Introducing empty sets if desired

We could have introduced empty sets by a stipulation that each type $\kappa^2(x)$ contains a specific object $\emptyset_{\kappa(x)}$ with no elements, then defining a set as an object which belongs to a kind $\kappa^2(x)$ and either is a nonempty set or is equal to $\emptyset_{\kappa(x)}$. It is then possible to define $\{x \in \kappa(u) : P(x)\}$ as $\emptyset_{\kappa(x)}$ when $P(x)$ is uniformly false, and to note that *sets* of the same kind with the same (possibly empty) extension are equal.

We remarked on this above, and we will continue to comment on its consequences as we continue with our development.

We prefer not to arrange this by stipulation: the reader might feel that we are appealing to something like choice in an external sense by selecting one atom from each kind to be the empty set of that kind.

Instead, we demonstrate that the theory modified to allow empty sets can be interpreted in the basic theory (and we will continue to work in the basic theory, commenting when the presence of empty sets would make a difference).

We redefine membership and equality deviously.

Definition: We define $x =_{new} y$ as holding if x and y are individuals and $x = y$, and otherwise iff $x \sim y$ and $(\forall z : z \in x \leftrightarrow z \in y)$. The atoms in each type $\kappa^2(u)$ are identified, and otherwise the new equality relation coincides with the old one.

We define $\mathtt{set}(x)$ as false if x is an individual, and otherwise as

$$(\forall zw \in \kappa^0(x) : z =_{new} w \to (z \in x \leftrightarrow w \in x)).$$

Something is a set if membership in it respects the new equality relation. Notice that the atoms are sets in this new sense (and the atoms in any kind are all the same empty set in the new sense of equality). A nonempty set in the old sense is a set in the new sense iff it either contains all of the old sense atoms or excludes all of the old sense atoms; the atoms in the new sense (empty but not the empty set) are the nonempty sets which "cut" the set of atoms in the original sense.

We define $x \in_{new} y$ as $x \in y \land \mathtt{set}(y)$.

Clearly any new sense sets which have the same elements in the new sense are equal.

13

Verifying that the axioms of our original theory (as modified when empty sets are provided) hold in this theory is an extensive exercise.

We do not dispute the proposition that mathematics with the empty set is cleaner. But we think it is interesting to see that even in practice, the empty set is not indispensible, and there is a strong difference between the treatment of empty and nonempty extensions in this theory which suggests that strange though it may be, the view that all sets are nonempty may be more appropriate here.

8 Basic constructions of set theory

We begin by describing the operations of Boolean algebra of sets, the basis for the parlor game of Venn diagrams.

Definition: If A is a set, we define A^c, the complement[4] of A, as

$$\{a \in \kappa^0(A) : a \notin A\}.$$

If A and B are sets and $A \sim B$, we define $A \cap B$, the intersection of A and B, as $\{c \in \kappa^0(A) : c \in A \wedge c \in B\}$, $A \cup B$, the union of A and B, as $\{c \in \kappa^0(A) : c \in A \vee c \in B\}$, and $A - B$ (often written $A \setminus B$), the set difference of A and B, as $\{c \in \kappa^0(A) : c \in A \wedge c \notin B\} = A \cap B^c$.

Note that $\kappa^0(A)^c$ is undefined, $A - B$ is undefined if $A \subseteq B$, and $A \cap B$ is undefined if A and B have no common elements. We say that A and B are *disjoint*, written $A\#B$, iff $A \sim B$ and there is no x belonging to both A and B. If we provide empty sets, these operations become total and $A\#B$ becomes equivalent to $A \cap B = \emptyset_{\kappa^0(A)} = \emptyset_{\kappa^0(B)}$.

Notice that any element of A or of $B \sim A$ belongs to $\kappa^0(A)$, so the bounds of the set abstracts here do nothing to restrict the extensions of these sets.

Observation: Note that each of these sets is bounded in $\kappa^0(A)$, whose uniqueness is a consequence of the axiom of diversity. It is worth noting that an axiom asserting the existence of $A \cup B$ [defined simply as a set containing exactly the elements of A and the elements of B] for sets $A \sim B$, in the presence of the other axioms, implies the axiom of diversity: if $\kappa(x) \sim \kappa(y)$, then $\kappa(x) \cup \kappa(y)$ would exist and contain both x and y as elements, whence $x \sim y$, whence $\kappa(x) = \kappa(y)$.

The axiom of diversity is precisely equivalent to the axiom of binary union in the presence of the other axioms: both are basic to our understanding of sets, but diversity is more closely tied to the exact concepts of this theory, and the order of exposition needs to be slightly tortured if binary union is used as the axiom.

---

[4]More usual treatments of set theory actually do not support complement as a construction, though texts like to talk about complements anyway.

Definition: If A is a set of sets (whence $\kappa^{-1}(A)$ exists) we define $\bigcup A$ as

$$\{x \in \kappa^{-1}(A) : (\exists a \in \kappa^0(A) : x \in a \wedge a \in A)\}.$$

Since every element of A is in $\kappa^0(A)$ and every element of an element of A is in $\kappa^{-1}(A)$, this is simply the collection of all elements of elements of A. If a set does not have a union, it is a set of individuals.

Similarly, we define $\bigcap A$ as

$$\{x \in \kappa^{-1}(A) : (\forall a \in \kappa^0(A) : a \in A \to x \in a)\}.$$

Notice that $\bigcup A$ and $\bigcap A$ both belong to $\kappa^0(A)$.

We define an important relation on sets which should already be familiar. The definition is a little more baroque than the version you might have encountered before.

Definition: We define $A \subseteq B$ (A is a subset of B) as holding iff

$$A \sim B \wedge \texttt{Set}(A) \wedge \texttt{Set}(B) \wedge (\forall x \in \kappa^0(A) : x \in A \to x \in B).$$

If we provide empty sets, we modify this to

$$A \sim B \wedge \texttt{set}(A) \wedge \texttt{set}(B) \wedge (\forall x \in \kappa^0(A) : x \in A \to x \in B).$$

Definition: We define $\mathcal{P}(A)$ as $\{B \in \kappa(A) : B \subseteq A\}$. This is called the power set of A and is the set of all subsets of A. Notice that $\mathcal{P}(A) \in \kappa^2(A)$.

If we provide empty sets, this definition is materially changed, because $\mathcal{P}(A)$ acquires an additional element $\emptyset_{\kappa^0(A)}$. This is a real mathematical difference: the power set of a set with three elements has 7 elements in our base theory and 8 in the theory with empty sets.

We extend our set builder notation.

Definition: We define $\{x \in A : P(x)\}$ as $\{x \in \kappa^0(A) : x \in A \wedge P(x)\}$.

Where $F(t_1, \ldots, t_n)$ is any complicated term representing an object, we define $\{F(t_1, \ldots, t_n) \in \kappa(x) : P(t_1, \ldots, t_n)\}$ as

$$\{u \in \kappa(x) : (\exists t_1, \ldots, t_n : u = F(t_1, \ldots, t_n) \wedge P(t_1, \ldots, t_n))\}.$$

Definition: If A and B are sets with $A \sim B$, we define $A \times B$ as

$$\{(a, b) \in \kappa^\Delta(A) : a \in A \wedge b \in B\}.$$

Note that $A \times B \in \kappa^{1+\Delta}(A) = \kappa^{1+\Delta}(B)$.

9 Relations, functions, and cardinality

In this section we talk about the theory of binary relations, functions, and basics of cardinal arithmetic over not necessarily finite sets.

Definition: A relation is a set of ordered pairs. If R is a relation, we write $x\,R\,y$ to mean $(x, y) \in R$. Notice that if $x\,R\,y$ we must have $x \sim y$ and $R \in \kappa^{2+\Delta}(x) = \kappa^{2+\Delta}(y)$.

Observation: We use binary relation symbols $\in$, $=$, $\sim$, $\subseteq$ in ways which look similar to this, but these are not relations. We might refer to these as "logical relations" as opposed to the "set relations" just introduced. The problem is that a set relation must relate items both of the same kind.

Definitions: If R is a relation, we define $\text{dom}(R)$ (the domain of R) as $\{x \in \kappa^{-\Delta}(R) : (\exists y : x\,R\,y)\}$. We define R^{-1} (the converse of R) as $\{(y, x) \in \kappa^0(R) : x\,R\,y\}$. We define $\text{rng}(R)$ (the range of R) as $\text{dom}(R^{-1})$. For any set $A \in \kappa^{1-\Delta}(R)$, we define $R{\restriction}A$ (R restricted to A) as $R \cap (A \times \kappa^0(A))$ and $R``A$ (the image of A under R) as $\text{rng}(R{\restriction}A)$.

Definition: If F is a relation, we say that F is a *function* iff $F``\{x\}$ has exactly one element for each $x \in \text{dom}(F)$. If $F$ is a function, we define $F(x)$ implicitly by $F``\{x\} = \{F(x)\}$. If F is a function and A and B are sets, we say $F : A \to B$ (F is a function from A into B) iff $\text{dom}(F) = A$ and $\text{rng}(F) \subseteq B$. We say that F is a function from A onto B iff $\text{dom}(F) = A$ and $\text{rng}(F) = B$. We say that F is *injective* or one-to-one iff F^{-1} is a function, and in this case we call F the *inverse* of F. A function from A onto B which is an injection is called a bijection from A to B.

Observation: We make the choice here of defining relations and functions as sets of ordered pairs. This means that there is no notion of surjection or onto function without specific reference to the intended codomain. The alternative would be to add the domain and codomain as features of a relation. This has its advantages, but as objects of set theory, sets of ordered pairs are much simpler than sets of ordered pairs adorned with intended domains and codomains.

Definition: If $F : A \to B$ and $G : B \to C$, then $G \circ F : A \to C$ is defined by $(G \circ F)(x) = G(F(x))$. This is called the composition of G and F. Note that compositions of injections are injections.

Definition: We define $|A|$, the cardinality of A, as the set of all sets B such that there is an injection from A to B and there is an injection from B to A. Note that $|A| \in \kappa^2(A)$. $|A| = |B|$ is clearly equivalent to $B \in |A|$. We further define $|A| \le |B|$ as holding iff there is an injection from A to B (noting that this does not depend on the choice of representatives of the cardinals), and define $|A| < |B|$ as $|A| \le |B| \wedge |A| \ne |B|$.

It should also be noted that a change in the definition of the ordered pair (for which we have provided the abstract framework) will have no effect on cardinals.

Of course the existence of a bijection from A to B implies that $|A| = |B|$, and we have the following

Theorem: It is a well-known theorem (expressing a more usual definition of cardinality of sets) that $|A| = |B|$ (as we have defined it) implies that there is a bijection from A to B, and so existence of a bijection from A to B is equivalent to $|A| = |B|$. The expository need to prove this theorem is less with our definition, but we will still do it presently.

Definition: A relation R is said to be an *equivalence relation* iff it is

reflexive: $x \, R \, x$ for every $x \in \mathtt{dom}(R) = \mathtt{rng}(R)$

symmetric: $x \, R \, y \to y \, R \, x$ for every x, y.

transitive: $(x \, R \, y \wedge y \, R \, z) \to x \, R \, z$ for every x, y, z.

We say that R is an equivalence relation on A if it is an equivalence relation and has domain A.

For each $x \in \mathtt{dom}(R)$, we define $[x]_R$ as $\{y : y \, R \, x\}$. This is called the equivalence class of x under R.

Definition: A set of sets P is said to be a partition iff each $A \in P$ is a nonempty set and any two distinct A, B in P are disjoint. A partition of A is a partition P such that $\bigcup P = A$.

Theorem: The equivalence classes under an equivalence relation R on A make up a partition of A. Each partition P of A determines an equivalence relation on A holding between x, y iff they belong to the same element of P. These two concepts are thus exactly correlated.

Axiom (choice): For each partition P there is a set C such that the intersection of C with each element of P is a singleton set. This is called a *choice set* for P. This axiom expresses the idea that we can choose one element of each compartment in P, which is a theorem if P is finite, but much less obvious if P is infinite.

10   Introduction to cardinal arithmetic

We define some familiar ideas in this new context. (We intend to add selected proofs later, but all are suitable as exercises).

Definition: We define $1_{\kappa^2(x)}$ as $|\{x\}|$. If we provide empty sets, we define $0_{\kappa^2(x)} = |\emptyset_{\kappa(x)}| = \{\emptyset_{\kappa(x)}\}$.

Definition: We define $\iota``A$ as $\{\{x\} : x \in A\}$ and more generally $\iota^n``A$ as $\{\iota^n(x) : x \in A\}$. We define $T(|A|)$ as $|\iota``A|$ and $T^n(|A|)$ as $|\iota^n``A|$ (it is straightforward to show that these definitions do not depend on the choice of the representative A of the cardinal). These are cardinals which externally seem to be the same as $|A|$. T and T^n are injective operations, so they have partial inverse operations T^{-1} and T^{-n}.

Definition: We define $|A| + |B|$ as $|A \cup B|$ when A and B are disjoint. Note that this does not depend on the choice of representatives of the cardinals. If $|A|$ and $|B|$ do not have disjoint representatives, this is undefined. We define $|A| \cdot |B|$ as $T^{-\Delta}(A \times B)$. We define B^A as the set of functions from A to B, and $|B|^{|A|}$ as $T^{-1-\Delta}(|B^A|)$. $2^{|A|}$ can also be defined as $T^{-1}(|\mathcal{P}(A)|)$, noting that characteristic functions on a given domain are two iterated kinds higher than subsets of the given domain.

Addition and multiplication will be total if a suitable axiom of infinity is assumed (or if $\Delta = 0$). We will see below that cardinal exponentiation is provably not total.

Alternative Definitions: Each of addition and multiplication has a definition more in the style of the definition we have chosen for the other.

$|A| + |B|$ can be defined as

$$T^{-\Delta}(|(A \times \{x\}) \cup (B \times \{y\})|),$$

where $x \neq y$.

$|A| \cdot |B|$ can be defined as the cardinality of the union of a partition $P \in T(|A|)$ each of whose elements is in $|B|$.

Proofs that these definitions are equivalent to the ones given above would be nice exercises.

Definition: For any cardinal $|A|$ we define $\sigma(|A|)$, the successor of $|A|$, as $|A| + 1_{\kappa^2(A)}$. This is the same as $\{B \cup \{x\} : B \in |A| \wedge x \notin B\}$.

Theorem: For all sets A, $\sigma(|A|) = \sigma(|B|)$ implies $|A| = |B|$.

Proof: If $\sigma(|A|) = \sigma(|B|)$ then for any $x \notin A$ and any $y \notin B$, there is an injection f from $A \cup \{x\}$ to $B \cup \{y\}$. Define $f^* : A \to B$ as

$$\{(u, v) : ((u, v) \in f \wedge v \neq y) \vee ((u, y) \in f \wedge (x, v) \in f)\}.$$

This is clearly an injection from A to B, and an injection from B to A can be defined in the same way.

Definition: We say that a set I is inductive if it contains $1_{\kappa^2(x)}$ (for some x) [$0_{\kappa^2(x)}$ (for some x) instead, if we provide empty sets and so 0] and $(\forall y \in I : \sigma(y) \in I)$. We define $\mathbb{N}_{\kappa^3(x)}$, the set of natural numbers which is of a kind with $\kappa^3(x)$, as the intersection of all inductive sets belonging to $\kappa^4(x)$.

To be a finite set is to belong to a natural number.

Axiom of Infinity (1): The Axiom of Infinity may be phrased as

$$(\forall x : |\kappa(x)| \notin \mathbb{N}_{\kappa^2 x)}).$$

This is not our final axiom of infinity: we will adopt a slightly stronger statement as the axiom of infinity below.

Definition: $(\aleph_0)_{\kappa^4(x)}$ is defined as $|\mathbb{N}_{\kappa^3(x)}|$, if Infinity is assumed.

Enough examples have been given to motivate a

Convention: When we define a notion depending on the kind of a single variable x, we have been writing it with a subscript $\kappa^n(x)$, which is of the same kind as the defined notion. We introduce an abbreviation: when we define a notion with a single parameter x, the notion being of the same kind as $\kappa^n(x)$, we give the parameter the shape $\mathbf{x}$ (which we will use only for this purpose) and we abbreviate the subscript $\kappa^n(x)$ as $\mathbf{n}$. So $1_{\mathbf{5}}$ is the natural number 2 of the version of the same kind as $\kappa^5(\mathbf{x})$, belonging to $\kappa^6(\mathbf{x})$. These subscripts might be omitted if they can be deduced from context: defining $\sigma(|x|) = |x| + 1$, we deduce that since 1 is of kind $\kappa^2(\mathbf{x})$ for some $\mathbf{x}$. $|x|$ and $\sigma(|x|)$ must be of the same kind $\kappa^2(\mathbf{x})$ and also of type $\kappa(x)$, so we deduce that $\kappa(\mathbf{x})$ is in this context $\kappa^0(x)$.

Thoerems of cardinal arithmetic :

1. $T^n(|A|) + T^n(|B|) = T^n(|A| + |B|)$; $T^n(|A|) \cdot T^n(|B|) = T^n(|A| + |B|)$

2. $|A| + |B| = |B| + |A|$; $|A| \cdot |B| = |B| \cdot |A|$.

3. $(|A|+|B|)+|C| = |A|+(|B|+|C|)$; $(|A| \cdot |B|) \cdot |C| = |A| \cdot (|B| \cdot |C|)$.

4. $|A| \cdot (|B| + |C|) = |A| \cdot |B| + |A| \cdot |C|$

5. $|A| \cdot 1 = |A|$; $[|A| + 0 = |A|; |A| \cdot 0 = 0]$

6. $|A|^{|B|+|C|} = |A|^{|B|} \cdot |A|^{|C|}$; $(|A|^{|B|})^{|C|} = |A|^{|B| \cdot |C|}$; $(|A|^{|C|}) \cdot (|B|^{|C|}) = (|A| \cdot |B|)^{|C|}$. Exponentiation is partial: if either side of one of these equations is defined, so is the other.

7. $1^{|A|} = 1$; $|A|^1 = |A|$; $[0^{|A|} = 0; |A|^0 = 1]$

These all look like familiar algebraic principles (and they specialize to familiar algebraic principles on the natural numbers). They can be proved by explicit constructions of bijections between appropriate sets. Some statements in brackets are provided which hold if we provide empty sets and so the cardinal 0. The principles of additive and multiplicative cancellation do not hold for general cardinal numbers. We do have

Theorem: $|A| + n = |B| + n \rightarrow |A| = |B|$ if $n \in \mathbb{N}$.

There is a certain scandal about infinite cardinals. These facts will witness the failure of general cancellation laws.

Theorem: $|\mathbb{N}| + n = |\mathbb{N}|$ (for $n \in \mathbb{N}$); $|\mathbb{N}| + |\mathbb{N}| = |\mathbb{N}|$; $|\mathbb{N}| \cdot |\mathbb{N}| = \mathbb{N}$

11 An alternative approach to Infinity and an improvement to the ordered pair

We adopt a slightly stronger

Axiom of Infinity (2): For every cardinal $|A|$, $\sigma(|A|) = |A| + 1$ exists.

This is not quite equivalent to the axiom of infinity stated above[5] but certainly implies it. For any kind $\kappa(x)$, $|\kappa(x)| + 1$, if it exists, is less than or equal to $|\kappa(x)|$, and it is easy to prove by mathematical induction that $n + 1$ is not less than or equal to n for any natural number n.

It is a theorem noted above that $\sigma(|A|) = \sigma(|B|) \to |A| = |B|$.

For any set A, define $\sigma_1(A)$ as $\sigma``(A)$ and $\sigma_2(A) = \sigma``(A) \cup \{1\}$. These operations are defined for sets A which are of the same kind as a set of cardinals. These operations are injective and have disjoint ranges.

This allows the definition of $\langle A, B \rangle$ for any sets A, B of the same kind as a set of sets of cardinals as $\sigma_1``A \cup \sigma_2``B$. This is an ordered pair, since we can recover A as $\sigma_1^{-1}``\langle A, B \rangle$ and $B$ as $\sigma_2^{-1}``\langle A, B \rangle$.

The pair $\langle A, B \rangle$ is of the same kind as A and B, which is a serious technical advantage[6]. But it is only defined for sets of sets of elements of kinds $\kappa^2(x)$.

But we can arrange our universe so that this type level pair exists in all types, by apparently cutting it down (though we will argue that the change can also be viewed as an expansion).

Define `subworld(x)` (read "x is in the subworld") as $(\exists u : x \in \mathcal{P}^2(\kappa^2(u)))$. Notice that if x and y are in the subworld and $x \sim y$, $\langle x, y \rangle$ exists.

Define $x \varepsilon y$ as `subworld(x)` $\land (\forall z \in y : $ `subworld(z)` $) \land x \in y$.

We claim that all of our axioms are true in the subworld with ε replacing $\in$.

Extensionality is true because nonempty sets with the same ε-extension are actually also nonempty sets with the same $\in$-extension.

Kinds is true because for each x in the subworld, the intersection $\kappa_\varepsilon(x)$ of $\kappa(x)$ with the subworld has the same ε-extension as its $\in$-extension, x belongs to it, and any y which belongs to it belongs to a common set with x in the sense of ε, namely $\kappa_\varepsilon(x)$, and any y which belongs in the sense of ε to a common set A with x also cohabits in the $\in$ sense with x, so belongs to $\kappa(x)$, and so also to $\kappa_\varepsilon(x)$ because it is in the subworld.

The axiom of separation is true because $\{x \in \kappa_\varepsilon(u) : P(x)\}$ exists for any formula $P(x)$, so certainly one in ε-language, belongs to the subworld because all of its elements do, and has the same ε extension as its $\in$ extension because all its elements are in the subworld. Moreover, it is of the same kind

---

[5]For the knowledegeable, it rules out the possibility that a kind is Dedekind-finite but infinite.

[6]For the knowledgeable, this pair was originally defined by Quine, using natural numbers instead of cardinals. The author introduced this way of defining the Quine pair, preferring it because it is predicative.

as its bounding set in the subworld, because it is of the same kind as its bounding set in the actual world.

The axiom of diversity again clearly holds in the subworld.

Now the subworld satisfies the additional feature that there is an ordered pair $\langle x, y \rangle$ (defined in the language of the entire world as above, but a primitive operation in the subworld; notice that a pair of two nonsets in the sense of the subworld will be a nonset) which is of the same kind as x and y. Note that any kind in the subworld has at least two elements.

What we have shown is that it would be safe to adopt as an additional axiom of our theory

Axiom of Ordered Pairs: For each $x \sim y$ we postulate an object $(x, y) \sim x$. For any x, y, z, w, if $(x, y) = (z, w)$ then $x = z$ and $y = w$. Further, each kind has at least two elements. This is the same as the axiom of ordered pairs above, but with $\Delta = 0$; the proviso that each kind has at least two elements ensures that kinds of individuals are infinite.

We adopt this axiom, and redefine all notions involving relations and functions in terms of this pair, with resulting changes in kinds of relevant objects (which are effected by setting $\Delta = 0$ instead of 2). We have shown that this axiom is no more dangerous than the axiom of infinity in the form given in this section, by showing that we can cut down the world of our theory to a subworld in which all axioms of our theory hold and the axiom of ordered pairs holds.

We note that in the subworld as described, every object is a pair (because every set of sets of objects of a kind including natural numbers is a pair $\langle x, y \rangle$) but the claim we make in our axiom is more modest. We are no more inclined to assert that everything is a pair than to assert that everything is a set.

We have already noted that if $\Delta = 0$, the operations of cardinal addition and multiplication become total (but exponentiation does not). In particular, this implies that the Axiom of Infinity (in either version) is a theorem, so we do not need to assert either form of Infinity as an axiom.

We have a further remark about the subworld construction. It can be viewed as removing all kinds which are not of the form $\kappa^4(x)$ and cutting $\kappa^4(x)$ down to $\mathcal{P}^2(\kappa^2(x))$. But it can also be viewed as preserving all kinds and fattening up $\kappa(x)$ to $\mathcal{P}^2(\kappa^2(x))$, noting that this includes $\iota^4 \text{``} \kappa(x)$ as a subset, which is an isomorphic copy of $\kappa(x)$ and can be viewed as covertly the original $\kappa(x)$. We think that it is important to notice that we are not

eliminating any combinatorial possibilities in passing to the subworld: every set in the larger world is externally the same size as a set in the subworld.

Notice that if we assume infinity and assume that all objects of all kinds are sets (there are no atoms and no individuals: of course this requires us to provide empty sets) then $\langle x, y \rangle$ is always defined to begin with, because all objects are then sets of sets (and certainly there are at least two objects of each kind). This assumption has some charm but also some odd consequences, as we will see.

A final remark is that it should be clear that this modification of our world can be carried out if empty sets are provided. The passage from the basic theory to the theory with empty sets must be executed before the passage to the subworld with the pair with $\Delta = 0$, because the maneuver providing empty sets would not preserve the properties of the pair.

12 The natural numbers; iteration of functions; cardinality definable in terms of bijections

The following five principles have been shown or follow readily from things that have been shown. We give alternative versions in brackets: we start the natural numbers with 0 if we provide empty sets.

1. $1 \in \mathbb{N}$ $[0 \in \mathbb{N}]$

2. if $n \in \mathbb{N}$, then $\sigma(n) \in \mathbb{N}$

3. if $n \in \mathbb{N}$, $\sigma(n) \neq 1$ [if $n \in \mathbb{N}$, $\sigma(n) \neq 0$]

4. if $m, n \in \mathbb{N}$, $\sigma(m) = \sigma(n) \rightarrow m = n$

5. if $A \subseteq \mathbb{N}$ and $1 \in A$ $[0 \in A]$and $\sigma``A = \{\sigma(n) : n \in A\} \subseteq A$, then $A = \mathbb{N}$ (this is the familiar principle of mathematical induction, and is built into the definition of $\mathbb{N}$).

The following additional principles, now usually also presented as axioms of Peano arithmetic, are readily provable.

6: if $m, n \in \mathbb{N}$, then $m + n$ and $m \cdot n$ belong to $\mathbb{N}$.

7: for any $m \in \mathbb{N}$, $m + 1 = \sigma(m)$ [for any $m \in \mathbb{N}$, $m + 0 = m$)]

8: for any $m, n \in \mathbb{N}$, $m + \sigma(n) = \sigma(m + n)$

9: for any $m \in \mathbb{N}$, $m \cdot 1 = m$ [for any $m \in \mathbb{N}$, $m \cdot 0 = 0$]

10: for any $m, n \in \mathbb{N}$, $m \cdot \sigma(n) = m \cdot n + m$

Principles 7-10 follow directly from cardinal arithmetic theorems in the previous section, and they provide the basis for a straightforward proof by induction of principle 6.

An important tool for our nefarious purposes is the

Iteration Theorem: Let A be a set, let $a \in A$ and let $f : A \rightarrow A$ be a function. Then there is a unique function $g : \mathbb{N} \rightarrow A$ such that $g(1) = f(a)$ $[g(0) = a$ if we have empty sets and 0] and $g(\sigma(n)) = f(g(n))$ for every $n \in \mathbb{N}$. We define $f^n(a) = g(n)$.

Indication of Proof: g is the smallest set of ordered pairs which contains $(1, f(a))$ $[(0, a)]$ and for each n, x, contains $(\sigma(n), f(x))$ if it contains (n, x). There is actual work to do to show that this is a function and the unique function with the properties stated.

Of course the kind of the natural numbers used and every other object here depends on the kind of a. Further the exact definition here would not work if a were an individual or set of individuals: a has to be of the same type as a natural number.

We outline a method of adapting this to any type. For any relation R, define R^{ι^n} as $\{(\iota^n(x), \iota^n(y)) : x\,R\,y\}$. It is then always possible to define $(f^{\iota^2})^n(\iota^2(a))$, as above, and define $f_*^n(a)$ as the sole element of an element of $(f^{\iota^2})^n(\iota^2(a))$. Note that in this definition, $n \in \kappa^3(a)$, and $f_*^n(a) = f^{T^{-2}(n)}(a)$ when the latter is defined.

This is a good point to remind the gentle reader that iterated applications of operations taking objects of one kind to objects of another, such as κ^n or ι^n, cannot be defined via the definition in the Iteration Theorem, because κ and ι are not functions.

We do not have a single system of natural numbers, but a different system for counting elements of each kind. There is an intimate connection between the natural numbers of kinds which are connected.

Theorem: For every $n \in \mathbb{N}_{i+1}$, $T(n) \in \mathbb{N}_{i+2}$.

Similarly, $T^m(n) \in \mathbb{N}_{i+m+1}$.

Given the axiom of infinity and $i > 2$, $T^{-1}(n)_{i-1} \in \mathbb{N}_i$, and if $i > m+1$, $T^{-m}(n) \in \mathbb{N}_{i-m+1}$

We have already noted that all these operations are injective: if $T^i(m) = T^i(n)$ for i an integer, then $m = n$.

Counting principles involving sets need to be approached with care. For example, the cardinality of $\{1, \ldots, n\}$ is not n, but $T^2(n)$, because the numerals used to count objects of a kind are not of the same kind as the objects counted, but "two higher".

Another example: it is usual to define $_nC_r$ as the cardinality of the collection of r-element subsets of an n-element set: however, we really would like it to be of the same kind as n and r. A set we might think of as being of this size is $\{|B| : B \subseteq \{1, \ldots, n\} \wedge B \in r\}$ but in fact $_nC_r$ should be

defined as $T^{-2}|\{|B| \in \kappa^3(n) : B \subseteq \{1, \ldots, n\} \wedge B \in T^2(r)\}|$. We leave it as an exercise to track down the details. If $A \in n$ (A is an n element set), then $\{B \in \mathcal{P}(A) : B \in r\}$ is a set we think of as of size ${}_nC_r$, and it is of kind $\kappa^2(r)$. So "${}_{|A|}C_{|B|} = T^{-1}(|\{C \in \mathcal{P}(A) : C \in |B|\}|)$" is the more natural definition. This is nicer than the previous one because we choose a set actually of size n to take our subsets from. The shift of kind seems inevitable because we are counting sets of objects of the kinds counted by n and r (then using T^{-1} to shift numbers back down in kind).

A meditation on number systems: The Iteration Theorem suggests a different approach to implementing the natural numbers. For any function we could define $0(f)(x) = x$ [zero does make sense here, even in the basic theory], and for any function g on functions define $\sigma(g)(f)(x)$ as $f(g(f)(x))$, and then define $\mathbb{N}$ as the minimal set containing 0 and closed under σ. For $n \in \mathbb{N}$ thus defined we would write $f^n(x)$ for $n(f)(x)$. We further remark that if we restrict the natural numbers thus defined to invertible functions, we have also implemented the integers: $(-n)(f)(x) = (f^n)^{-1}(x)$. Addition and multiplication are definable: $f^{m+n}(x) = f^m(f^n(x))$ and $f^{m \cdot n}(x) = (f^m)^n(x)$ give implicit definitions of these operations for natural numbers thus defined or for integers. We further note that for natural numbers $n(m)$ acts as $m^{T(n)}$, defining exponentiation, but extending this to allow integer exponents, thus implementing the rational numbers, would require restricting the functions on which the natural numbers, integers, and implemented rationals are taken to act to a very carefully chosen family (defining such a family might require in effect implementing the rationals already), a family of functions closed under composition and inverse on which each natural number (considered as an iteration operator) is invertible. Once one shows that one can produce such a family of functions, we have an implementation of the rationals with the natural numbers and integers as subsets in the way we expect, which is not the case in usual implementations of the number systems in set theory. The easiest way to produce such a family of functions is to code fractions $\frac{m}{n}$ as pairs (m, n) of natural numbers with $\mathtt{gcd}(m, n) = 1$, for convenience define $\mathtt{simp}(x, y) = (\frac{x}{\mathtt{gcd}(x,y)}, \frac{x}{\mathtt{gcd}(x,y)})$, and define $(m, n) \oplus (p, q) = \mathtt{simp}(mq + np, nq)$. Then the family of functions $(x \mapsto x \oplus r)$ has the desired characteristics: it supports iteration by rational values. Note that the rationals support integer powers,

but not rational powers: in $n(m)$ the two iterators are acting on different families of functions.

There is a perhaps unexpected application of the Iteration Theorem to general cardinal arithmetic.

Theorem: If $|A| = |B|$, there is a bijection from A to B (the converse is obvious from the definition of cardinal number).

Proof: If $|A| = |B|$, then $A \in |B|$, so there is an injection $f : A \to B$ and an injection $g : B \to A$.

For each $a \in A$, define a (possibly partial) integer indexed sequence by $a_0 = a; a_{2i+1} = f(a_{2i}); a_{2i+2} = g(a_{2i+1}); a_{2i-1} = g^{-1}(a_{2i}); a_{2i-2} = f^{-1}(a_{2i})$. This can be firmed up as an elaborate consequence of the Iteration Theorem. These terms may have all integer terms defined or they may have a first term (with a possibly negative index). Note that every element of A will appear as an even indexed term of some sequence and every element of B will appear as an odd-indexed term of some sequence. The same sequence in effect appears with many indexings, but we will see this will not be a problem for us. Our strategy is then to define $h(a_{2i})$ as $f(a_{2i}) = a_{2i+1}$ for all terms of each sequence $\{a_{2i}\}$ in which the lowest indexed term is an element of A, and to define $h(a_{2i})$ as $g^{-1}(a_{2i}) = a_{2i-1}$ in all other cases. This procedure clearly defines an injection, determines a value at every element of A and also clearly produces every value in B.

So there is a bijection h from A to B as desired.

The description above used a rather elaborate recursive definition and talked about integers. It is equivalent to the following definition of h, which is entirely justified in terms we have used.

$h = (f \restriction \bigcup_{i \in \mathbb{N}} (g \circ f)_*^i \, {}^{\shortmid\shortmid}(A \setminus g {}^{\shortmid\shortmid} B)) \cup (g^{-1} \restriction (A - \bigcup_{i \in \mathbb{N}} (g \circ f)_*^i \, {}^{\shortmid\shortmid}(A \setminus g {}^{\shortmid\shortmid} B)))$

13 There are many infinite cardinals: Cantor's theorem, infinite sums and products, König's Lemma

So far we have not established that there is more than one infinite cardinal. We do this rapidly. It is important to note that we are assuming $\Delta = 0$ from here on in the text. In this section in particular, the iteration of kinds would be a bit more complex if we did not assume this, though everything would work.

Theorem: For any set A with $|A| > 1$, $|\iota``A| < |\mathcal{P}(A)|$. An interesting corollary is that $|\iota``(\kappa(x))| < |\mathcal{P}(\kappa(x))| \leq |\kappa^2(x)|$ (The kind of a kind is larger than the kind in an external sense, the local resolution of the Cantor paradox).

Proof: Clearly $|\iota``A| \leq |\mathcal{P}(A)|$: the identity map on $\iota``A$ witnesses this.

Suppose for the sake of a contradiction that $|\iota``A| = |\mathcal{P}(A)|$. Thus there would be an injection $f$ from $\mathcal{P}(A)$ to $\iota``A$. Now consider the set $R = \{x \in \bigcup \mathtt{rng}(f) : x \notin f^{-1}(\{x\})\}$. Let $r \in f(R)$. Then $r \in R$ iff $r \in \bigcup \mathtt{rng}(f)$ [this is true] and $r \notin f^{-1}(\{r\}) = R$. This is a contradiction, as long as R exists (is nonempty). If $|A| = 1$, R does not exist (the intended extension has no elements) and there is no contradiction; if $|A| > 1$, let $f(A) = \{v\}$ and let $f(\{v\}) = \{w\}$: $w \notin f^{-1}(\{w\})$, so $w \in R$ and the argument for a contradiction goes through. If we provide empty sets, the argument works for $|A| \leq 1$ as well.

Corollary: $(\aleph_0)_{\kappa^5(x)} = |\iota``\mathbb{N}_{\kappa^3(x)}| < |\mathcal{P}(\mathbb{N})_{\kappa^3(x)}|$. We define $c_{\kappa^5(x)}$ as $|\mathcal{P}(\mathbb{N}_{\kappa^3(x)})|$. Thus, at least in some kinds, there is more than one infinite cardinal.

Theorem: $|A| < 2^{|A|}$ if the latter is defined. Note that $2^{|A|} = |\mathcal{P}(A)| + 1$ (one of the two constant functions corresponds to the nonexistent empty set). $[2^{|A|} = |\mathcal{P}(A)|$ if empty sets are provided$]$.

Proof: $2^{|A|} = T^{-1}|\{x, y\}^A|$ (where $x \neq y$) [one of the two constant functions corresponds to the nonexistent empty set]. $|A| \leq 2^{|A|}$ because there is an injection from $\iota``A$ taking $\{a\}$ (for $a \in A$) to

$$(b \in A \mapsto x) \cup (b \in A^c \mapsto y).$$

Suppose $|A| = T^{-1}|\{x,y\}^A|$. So there would be an injection f from $|\{x,y\}^A|$ to $\iota``A$. We could then define a map $R : \bigcup \mathtt{rng}(f) \to \{x,y\}$ taking a to $\{x,y\} - f^{-1}(\{a\})(a)$. Now consider r, the sole element of $f(R)$: $R(r) = \{x,y\} - R(r)$ would follow, which is absurd.

This might seem redundant but in fact it has a little more extent: it is not a consequence of $\{x,y\}^A$ being the same size as a set of singletons that A is the same size as a set of singletons.

Corollary: $|A| < |B|^{|A|}$ if the latter cardinal exists and $|B| > 1$.

Observation: the two theorems proved just above are related by the device of *characteristic functions*: define χ_A, for A a set, as the function taking $x \in \kappa^0(A)$ to $\mathbf{t}$ if $x \in A$ and to $\mathbf{f}$ if $x \notin A$, where $\mathbf{t}$ and $\mathbf{f}$ are objects chosen to represent the truth values: it is usual to use 1 and 0, but these do not exist in all kinds. Now the oddity in the basic theory is that $(A \mapsto \chi_A)$ is not onto $\{\mathbf{t},\mathbf{f}\}^A$: the constant function with value $\mathbf{f}$ corresponds to the absent empty set. This is fixed if empty sets are provided.

Power sets and function spaces over a kind can be larger than the kind (in a suitable external sense), but are represented in the kind of the kind. Of course $\mathcal{P}(A)$ and B^A exist for all sets $A \sim B$, but they are in $\kappa^2(A) = \kappa^2(B)$ and may fail to be the same size (via projection using singletons) as any set in $\kappa(A) = \kappa(B)$.

We now consider more complex operations on possibly infinite indexed families of sets and cardinals. Material in this part depends on the axiom of choice in ways that we point out.

Definition: An indexed family of sets is a function A with domain an index set I. We write A_i instead of $A(i)$.

Definition: The infinite cartesian product of a family A with index set $I = \iota``J$, written $\prod_{i \in I} A_i$ or just $\prod A$, is defined as

$$\{f \in (\textstyle\bigcup \mathtt{rng}(A))^J : (\forall j \in J : f(j) \in A_{\{j\}}\}.$$

Note that this is of the same kind as $(\bigcup \mathtt{rng}(A))^J$, which belongs to $\kappa^2(J) = \kappa(I) = \kappa(A)$.

Note also that it is a consequence of the axiom of choice (in fact, an equivalent statement to the axiom of choice) that the product of a family of nonempty sets is in all cases nonempty. Note that the set of sets

$$\{\{(j, x) : x \in A_{\{j\}}\} \in J \times \bigcup \mathtt{rng} A : j \in J\}$$

is a partition: a choice set for this partition is an element of the infinite product.

We note an alternative definition of the infinite cartesian product:

$$\{f \in (\iota``(\bigcup \mathtt{rng}(A)))^I : (\forall i \in I : f(i) \subseteq A_i\}.$$

This set belongs to $\kappa^2(A)$. This definition supports products of larger (but not much larger) families of sets.

Definition: The infinite disjoint union of a family A with domain an index set I, written $\sum_{i \in I} A_i$ or just $\sum A$, is defined as

$$\{(\{x\}, i) \in \iota`` \bigcup (\mathtt{rng} A) \times I : x \in A_i\}.$$

Note that this belongs to $\kappa(I) = \kappa(A)$.

Now we define associated operations on families of cardinals, using the same symbols. This overloading is traditional and in practice should not lead to confusion.

Definition: An indexed family of cardinals will be a function $\kappa : I \to \kappa^2(u)$ where $I = \iota^2``J$ and $\kappa_i = \kappa(i) = \kappa(\{\{j\}\}) = |A_j|$ for some indexed family of sets $A : j \to \kappa(u)$ (the existence of a such a family of sets is a consequence of the axiom of choice).

Definition: The infinite product $\prod \kappa = \prod_{i \in I} \kappa_i = \prod_{j \in J} |A_j|$ (where a family of sets is chosen as in the previous definition) is defined as $T^{-1}(|\prod_{j \in J} A_j|)$.

We discuss the application of T^{-1}. The kind of the product of cardinals should be the same as the kind of the κ_i's. The kind of a κ_i has index two higher than that of the kind of an A_i, so index one higher than that of the kind of A, which is the kind of the product of sets. So the

cardinality of the product of sets has kind with index one higher than that of the cardinality of the κ_i's, so we apply T^{-1} to get the product of cardinals. The same analysis applies to infinite sums of cardinals below.

One can show that the cardinality does not depend on the indexed family of representatives chosen. If another family A' is used, choose a bijection $f_i : A_i \to A'_i$ for each $i \in I$, and these can be used to show that $|\prod_{j \in J} |A_j|| = |\prod_{j \in J} |A'_j||$.

Notice that I here is a set of triple singletons.

If the alternative definition were used, the infinite product $\prod \kappa = \prod_{i \in I} \kappa_i = \prod_{j \in J} |A_j|$ (where a family of sets is chosen as in the previous definition) would be defined as $T^{-2}(|\prod_{j \in J} A_j|)$. The alternative definition might allow products of somewhat larger sets of cardinals, but not much larger. The index set I would in this case be a set of double singletons.

Definition: The infinite sum $\sum \kappa = \sum_{i \in I} \kappa_i = \sum_{j \in J} |A_j|$ (where a family of sets is chosen as in the previous definition) is defined as $T^{-1}(|\sum_{j \in J} A_j|)$.

One can show that the cardinality does not depend on the indexed family of representatives chosen. If another family A' is used, choose a bijection $f_i : A_i \to A'_i$ for each $i \in I$, and these can be used to show that $|\sum_{j \in J} |A_j|| = |\sum_{j \in J} |A'_j||$.

Theorem: If $|A_i| < |B_i|$ for all indices i in a suitable index set I, then $\sum_{i \in I} |A_i| < \prod_{i \in I} |B_i|$. Notice that $I = \iota``J$ is required by the statement to be proved: this is part of suitability.

Proof: $\sum_{i \in I} |A_i| \leq \prod_{i \in I} |B_i|$: Choose injections $g_i : A_i \to B_i$. Choose an element b_j from each $B_{\{j\}}$ which does not belong to $g_i``A_{\{j\}}$. The cardinality conditions (and the axiom of choice) ensure that we can do this. Then map each element $(\{x\}, \{j\})$ of $\sum_{i \in I} A_i$ to the function $f \in \prod_{i \in I} B_i$ sending each k to $g_{\{j\}}(x)$ if $j = k$ and otherwise to b_k. This is clearly an injection: find the projection of a product element in its range which is in an A_i and you can determine the element of the sum from which it came.

Suppose for the sake of a contradiction that there is a bijection f from $\sum_{i \in I} A_i$ to $\prod_{i \in I} B_i$. Construct a function h in the product of the B_i's

such that the image $h(j)$ of each j is chosen to be distinct from all values $f(\{\{x\}, \{j\}\})(j)$ for $x \in A_{\{j\}}$, which can be done because there are fewer possible values for this expression than there are elements of $B_{\{j\}}$. But this means that h differs in its value at some j from every image of an element of the product of the A_i's under f, so it cannot be such an image. $f^{-1}(h)$ is of the form $(\{x\}, \{j\})$ for some $x \in A_{\{j\}}$. Now $h(j)$ was chosen so that for any j (including this one!), for every $x \in A_{\{j\}}$ (including this one!) it is distinct from $f(\{\{x\}, \{j\}\})(j)$, that is in this case $h(j)$ itself, which is absurd.

This proof could be adapted to use the alternative definition of infinite cartesian product with somewhat more elaborate discussion of iterated kinds. The basic argument would be the same.

Observation: Something which may be perceived as a difficulty which does not happen in other presentations of set theory is the need to index correlated families of different kinds with indices of different kinds. However, the singleton operation provides a natural correlation.

14 Well-orderings, ordinals, aleph and beth numbers, the Well-Ordering Theorem

We define some special kinds of relation.

Definition: A *partial order* on a set A is a subset $\leq$ of $A \times A$ which is

> **reflexive:** ($a \leq a$ for every $a \in A$),
>
> **antisymmetric:** (for any a, b, $a \leq b \wedge b \leq a \to a = b$) and
>
> **transitive:** (for any a, b, c, $a \leq b$ and $b \leq c$ implies $a \leq c$).
>
> By convention, when we denote a partial order as $\leq$ (possibly adorned in some way), we use the symbol $<$ (adorned in the same way, if applicable) to denote the correlated strict partial order, defined as $a \leq b \wedge a \neq b$. We will also by convention use $\leq$, $>$, for $\leq^{-1}$, $<^{-1}$.

Definition: A *linear order* on A is a partial order $\leq$ on A with the additional property that it is total: ($\forall a, b \in A : a \leq b \vee b \leq A$)).

Definition: A *well-founded relation* on a set A is a relation W on A with the property that for every nonempty $B \subseteq A$ there is $m \in B$ (referred to as a W-minimal element of B) such that for all $b \in B$, if $b\,W\,m$ then $b = m$. We say that a well-founded relation is strictly well-founded if in addition it is irreflexive (it relates no element of its domain to itself).

Definition: A *well-ordering* on A is a well-founded linear order $\leq$ on A. Notice that a $\leq$-minimal element m of $B \subseteq A$ will satisfy $m \leq b$ for all elements b of B (it will be minimum, not just minimal), because a linear order is total.

Definition: For any relation R, we define $\mathtt{fld}(R)$ as $\mathtt{dom}(R) \cup \mathtt{rng}(R)$. For any relations R, S, we define $R \approx S$ (R is isomorphic to S) as "there exists a bijection f from $\mathtt{fld}(R)$ to $\mathtt{fld}(S)$ such that for all x, y, $x\,R\,y \leftrightarrow f(x)\,S\,f(y)$." Isomorphism is an equivalence relation on relations (on any particular kind).

Definition: For any well-ordering $\leq$, we define the *order type* of $\leq$, written $\mathtt{ot}(\leq)$, as $[\leq]_{\approx}$, the set of all relations isomorphic to $\leq$ (these will all

be well-orderings). A set which is the order type of some well-ordering is called an *ordinal number*.

Note that if empty sets are provided, the empty set is a well-ordering and has an order type 0. Otherwise the first ordinal is the order type of a well-ordering whose domain has one element, which we call 1.[7]

Definition: If $\leq$ is a well-ordering and $x \in \mathtt{dom}(\leq)$, we define $\mathtt{seg}_{\leq}(x)$ as $\{y : y \leq x\}$ and define $[\leq]_x$ as $\leq \cap ((\kappa(x) \times \mathtt{seg}_{\leq}(x)))$ (this is a well-ordering). We define $\leq_{\alpha}$, where α is an ordinal number, as the x (if there is one) such that $\mathtt{ot}([\leq]_x) = \alpha$. Notice that if $\leq \in \kappa^2(x)$, then the ordinal index in $\leq_{\alpha}$ belongs to $\kappa^3(x)$.

Definition and Theorem: The natural order on ordinals is defined thus: $\alpha \leq \beta$ iff $\alpha = \beta \vee (\exists \leq_0 \in \beta : (\exists x : [\leq_0]_x \in \alpha))$. The Theorem is that this is a well-ordering on the ordinal numbers.

Definition: For any set A such that there is a well-ordering on A, we define $\mathtt{init}(|A|)$ as the minimum order type of a well-ordering of A. This clearly does not depend on the choice of representative from the cardinal $|A|$ (if A is finite there is only one such order type, but otherwise there are many). Such objects are referred to as *initial ordinals*: they are in exact correspondence with cardinals; initial ordinals and cardinals are not identified here as they are in other treatments of set theory. We define ω as $\mathtt{init}(\aleph_0)$: this is the smallest infinite ordinal.

Definition: For any relation R, we define R^{ι} as $\{(\{x\}, \{y\}) : (x, y) \in R\}$ and $T([R]_{\approx})$ as $[R^{\iota}]_{\approx}$. We define R^{ι^n} and $T^n([R]_{\approx})$ similarly. Note that the $\cdot^{\iota}$ operation is applicable to well-orderings (producing well-orderings) and the T operation is applicable to ordinal numbers (producing ordinal numbers).

Definition: We define an order $<^*$ on ordinals: $\alpha <^* \beta$ is defined as $\alpha \leq \beta$ if β is finite, and otherwise as $\alpha < \beta$. If empty sets are provided, $<^*$ coincides with $\leq$.

---

[7]There is more to say on this. If we use isomorphism classes of strict well-orderings instead of isomorphism classes of (reflexive) well-orderings as order types, then the order types of 0 and 1 are not distinguishable if we have empty sets (because both are empty) and both fail to exist if we do not provide empty sets, so the first ordinal number turns out to be 2, which did happen in the treatment in *Principia Mathematics*.

Lemma: For any ordinal α, the order type of the natural order on ordinals restricted to the ordinals $<^* \alpha$ is $T^2(\alpha)$. The isomorphism is the bijection taking $\{\{\beta\}\}$ (for $\beta < \alpha$) to $\mathtt{ot}(\leq \lceil \{\gamma \in \kappa(\beta) : \gamma <^* \beta\})$. That this is a welldefined function is clear: that it is a bijection requires the fact that no two distinct initial segments of a well-ordering are isomorphic.

Definition: Let A be an infinite set. We define $\Omega(|A|)$ as the order type of the natural order on the ordinals restricted to order types of well-orderings of subsets of A. A well-ordering of a subset of A is in $\kappa(A)$; its order type is in $\kappa^2(A)$; the natural order on these is in $\kappa^3(A)$; the order type of the natural order on all of these is in $\kappa^4(A)$.

Theorem: For every order type α of a well-ordering on a subset of A, $T^2(\alpha) < \Omega(|A|)$.

A corollary of this is that $T^{-2}(\Omega(|\kappa(x)|))$ cannot exist. If it did, it would be of the same kind as the order type of well-orderings of $\kappa(x)$, but it would then have to be the order type of a well-ordering of $\kappa(x)$ and by the inequality just shown its image under T^2 would be less than itself. This is the local version of the Burali-Forti paradox.

Proof: The order type of the ordinals $<^*$ such an α is $T^2(\alpha)$, and this is a proper initial segment of the order types of well-orderings of subsets of A under the natural order. We do appeal implicitly here to the fact that A is infinite (in a way which is not necessary if empty sets are provided).

Definition: For an ordinal α, define $\mathtt{card}(\alpha)$ as the cardinality of the domain of any well-ordering belonging to A. Note that $\mathtt{card}(\alpha) \in \kappa^0(\alpha)$.

Theorem: $\Omega(|A|)$ is an initial ordinal.

Proof: An ordinal $\alpha < \Omega(|A|)$ is an ordinal $T^2(\beta)$ where β is the order type of a well-ordering of a subset of A. Now any ordinal γ with $\mathtt{card}(\gamma) = \mathtt{card}(T^2(\alpha))$ contains well-orderings W with size $T^2(\mathtt{card}(\alpha))$: W will be isomorphic to an V^{ι^2} with $\mathtt{card}(V) = \mathtt{card}(\alpha)$. But then it is clear that V is isomorphic to a well-ordering of a subset of A with some order type δ and $\gamma = T^2(\delta) < \Omega(|A|)$, which establishes that $\Omega(|A|)$ is an initial ordinal.

Definition: We define $\aleph(|A|)$ as $T^{-2}(\mathtt{init}^{-1}(\Omega(|A|)))$. Note that $\mathtt{init}^{-1}(\Omega(|A|)) \in \kappa^4(A)$: we apply T^{-2} to bring the output to the same kind $\kappa^2(A)$ as $|A|$. The notation $|A|^+$ is also used for $\aleph(|A|)$, usually when A is itself well-orderable.

Theorem: $\aleph(|A|) \not\leq |A|$. A corollary of this is that $\aleph(|\kappa(x)|)$ does not exist.

Proof: If $\aleph(|A|) \leq |A|$, then there is a well-ordering of A with field of size $\aleph(|A|)$, and so there is a well-ordering of A with order type $\mathtt{init}(\aleph(|A|) = T^{-2}(\Omega(|A|))$ and we have already seen that this is impossible.

Discussion: The order types of well-orderings of subsets of A all belong to $\mathcal{P}^2(A \times A)$. So among the well-orderings of subsets of $\mathcal{P}^2(A \times A)$ we find the actual well ordering on the ordinals less than $\Omega(|A|)$, which is of order type $T^2(\Omega(|A|))$, so we have $T^2(\Omega(|A|)) < \Omega(|\mathcal{P}^2(A \times A)|)$. If we use $\mathtt{exp}(|A|)$ to denote $2^{|A|}$, we can conveniently remark that $\Omega(|A|) < \Omega(|\exp^2(|A| \cdot |A|)$ and that $\aleph(|A|) < \aleph(\exp^2(|A| \cdot |A|)$, which will turn out to be a result of interest.

Transfinite Induction and Recursion: Any set S of ordinals which contains 1 and is closed under strict suprema contains all ordinals. If A is a set of ordinals, $\mathtt{sup}^+(A)$ is the smallest ordinal in the set $\{\alpha : (\forall \beta \in S : \beta > \alpha)\}$ of strict upper bounds of A. S is said to be closed under strict suprema iff it contains 1 [this clause not needed if empty sets are provided] and for every $A \subseteq S$, $\mathtt{sup}^+(A) \in S$ if $\mathtt{sup}(A)$ exists. To see that this is true, consider the smallest $\alpha \notin S$. The ordinal α cannot be 1, and if it is not 1, then $\alpha = \mathtt{sup}^+\{\beta : \beta < \alpha\}$, the strict supremum of a subset of S!

This is often broken into subcases: if S contains 1 $[0, (\mathtt{sup}^+(\emptyset))]$ and contains $\alpha + 1$ $(\mathtt{sup}^+(\{\alpha\}))$ for each $\alpha \in S$, and contains each limit ordinal λ $(\lambda = \mathtt{sup}^+(\{\beta : \beta < \lambda\})$ when every ordinal less than λ belongs to S, then S contains all ordinals.

Recursion goes hand in hand with induction as always. For any object x and function G sending functions on initial segments of the ordinals to singletons, there is a unique function F such that $F(1) = x$ and otherwise $F(\alpha)$ is the sole element of $G(F\lceil\{\beta : \beta < \alpha\})$. The special treatment of 1 would not be required if we provided empty sets (and the ordinal 0).

This can be broken into subcases: the function might have separate definitions stated for 1 [0], for successors, and for limit ordinals.

The results above do not involve choice. The following results do.

Theorem (Zorn's Lemma): Any partial order $\leq$ with the property that each linear suborder has an upper bound has a maximal element.

Proof: Let $\leq$ be a partial order in which each linear suborder has a maximal element. Let P be the collection of all pairs $(\texttt{rng}(\leq_0), \{b\})$ where $\leq_0$ is a linear suborder of $\leq$ and b is an upper bound in $\leq$ for the domain of $\leq_0$, and a strict upper bound if there is one. Let C be a choice set for P. C is a function sending the range of each $\leq_0$ to the singleton of an upper bound for its domain which is strict if possible. Define $\theta(\{x\})$ as x. Define a map F from ordinals to double singletons of elements of the domain of $\leq$ as follows: choose $F(1)$ an arbitrary double singleton of an element of the domain of $\leq$ and otherwise define $F(\alpha) = \iota(C(\{\theta^2(F(\beta))) : \beta < \alpha\}))$, if the value of C computed is a strict upper bound, and otherwise undefined. The range of F (and of any restriction of F to an initial segment of the ordinals) is the range of a linear suborder (in fact, a sub-well-order) of $\leq^{\iota^2}$, as is evident by transfinite induction (and this linear suborder is the result of applying T^2 to a linear suborder of $\leq$). There must be a largest element ν in the range of F because application of C to the union of the union of the range of F must give an object which is an element of the range of F already computed (as $F(\nu)$, $F(\nu + 1)$ thus being undefined) and $\theta^2(F(\nu))$ must be a maximal element in $\leq$. Note that the kind of α here is the kind of an order type of a subset of $\kappa^0(\leq)$, and in fact in each instance of the recursive definition, the order type of the restriction of $\leq$ to $\{\theta^2(F(\beta)) : \beta <^* \alpha\}$ is α, so we are not going to have a failure of definition because we run out of ordinals.

Definition: We refer to a set which is not finite (whose cardinality is not a natural number) as infinite. We refer to a set with the cardinality of the set of natural numbers as countably infinite or just countable.

Theorem: Every set can be well-ordered.

Proof: Let A be a set. Apply Zorn's Lemma to the end extension order on well-orderings of subsets of A (an order is less than another order if

it is an initial segment of that order): a linear suborder of this order
has the union of its domain a well-ordering end extending all elements
of the domain of the linear suborder, which can be extended with one
more element if the union of its domain is not A. A maximal element
in this order is a well-ordering of A.

Theorem: Every infinite set has a countable subset.

Proof: Consider the inclusion order on finite subsets of an infinite set A.
Every linear suborder of this order is either infinite (but no more than
countably infinite) and will have countably infinite union, or finite,
and can be extended by adding another element to its maximum, if it
is finite. Apply Zorn's Lemma.

Corollary: $\kappa = \kappa + n$ for each infinite cardinal κ and natural number n.
This follows by considering a bijection fixing everything but a countable
subset of a set of size κ. $\kappa + \aleph_0 = \kappa$: this follows from the fact that
a countable set has the same cardinality as the union of two disjoint
countable sets.

Theorem: $\kappa + \kappa = \kappa$ for each infinite cardinal κ.

Proof: Let A be a set of size κ. Consider the inclusion order on bijections
from $(B \times \{x\}) \cup (B \times \{y\})$ [for fixed, distanct x, y] to B where $B \subseteq A$.
The union of a linear suborder in this order is an element of the domain
of the order: moreover, when $A \setminus B$ is infinite, this can be properly
extended by exploiting the fact that a countably infinite subset C of
$A \setminus B$ can be placed in bijection with $(C \times \{x\}) \cup (C \times \{y\})$. A maximal
element in this order gives a bijection between $(B \times \{x\}) \cup (B \times \{y\})$
and B when $A \setminus B$ is finite, witnessing $\kappa + \kappa = \kappa$.

Corollary: For infinite cardinals κ, λ, $\kappa + \lambda = \mathtt{max}(\kappa, \lambda)$.

Proof: If $\lambda \leq \kappa$, $\kappa \leq \kappa + \lambda \leq \kappa + \kappa = \kappa$.

Theorem: $\kappa \cdot \kappa = \kappa$, for each infinite cardinal κ.

Proof: Let $A \in \kappa$. Apply Zorn's Lemma to the inclusion order on bijec-
tions from $B \times B$ to B where $B \subseteq A$. The union of the domain of a
linear suborder of this order is an element of the domain of the order.
Moreover, if $A \setminus B$ is infinite we can properly extend an element of the

domain of this order by considering a countable subset C of $A \setminus B$. $B \cup C \times B \cup C$ breaks into four parts, $B \times B$, for which we have a bijection to B, $B \times C$ and $C \times B$, for which we can construct bijections because $B \times C$ injects to $B \times B$, but also $B \times B$ injects to B which injects to $B \times C$, and $C \times C$ is countable. We already know that $|B| + |B| + |B| + \aleph_0 = |B|$, so we can in fact contruct a bijection from $(B \cup C) \times (B \cup C)$ to $B \cup C$ by previous results. A maximal element in the order will be a bijection from $B \times B$ to B with $A \setminus B$ finite, which witnesses $\kappa \cdot \kappa = \kappa$.

Corollary: For infinite cardinals κ, λ, $\kappa \cdot \lambda = \texttt{max}(\kappa, \lambda)$.

Proof: If $\lambda \leq \kappa$, $\kappa \leq \kappa \cdot \lambda \leq \kappa \cdot \kappa = \kappa$.

The arithmetic of addition and multiplication is vastly simplified by the assumption of choice. It is worth noting that the arithmetic results stated above hold for cardinals of well-orderable sets in any case. In our final remarks in this section we are not assuming choice, though we comment on its consequences.

Theorem: The order on cardinals of well-orderable sets is a well-ordering. Of course, if the axiom of choice holds, the order on all cardinals is a well-ordering.

Proof: $|A| \leq |B| \leftrightarrow \texttt{init}(|A|) \leq \texttt{init}(|B|)$.

Definition: Define $\aleph$ as the natural order on infinite cardinals of well-orderable sets. Then $\aleph_\alpha$ is defined for any (not too large) ordinal α by the ordinal indexing convention. Let $\beth$ be the order on cardinals in the smallest set of cardinals containing $\aleph_0$ and closed under $(\kappa \mapsto \kappa^+)$ and suprema. Again, $\beth_\alpha$ is defined for any (not too large) α by the ordinal indexing convention. The cardinals in the domain of $\beth$ are not necessarily cardinals of well-orderable sets, but it should be clear that they are well-ordered by the natural order on cardinals, so the indexing works. We do not have strong assumptions about how long these orders are.

15 Modelling our set theory in itself; a parallel consideration of ways to strengthen the theory; ambiguity considered

In this section, we first discuss modelling the theory in itself, and consequences which seem to follow in the most natural such models which are not provable in our theory, then we discuss two incompatible ways in which the theory can be extended.

15.1 Natural models

We discuss the project of building a structure in our theory which is a model of the theory. Such a structure would have to have a domain (which we call M for model) and support equality (which we allow equality restricted to M to interpret) and membership, which we suppose interpreted by a relation E. Though conceptually it is nice to have a name M for the domain of the model, in fact it is definable as $\mathtt{dom}(E)$: every object is a member of something.

Definition: We call a relation E *extensional* iff for any elements x, y of $\mathtt{rng}(E)$ with the property that $E^{-1}\text{``}x = E^{-1}\text{``}y$ we have $x = y$.

This amounts to a description of what has to be true of E if it is to model the axiom of extensionality for nonempty sets.

Definition: We call a sorted extensional relation E *sorted* iff for every $x \in \mathtt{dom}(E)$ there is a $K(x) \in E\text{``}\{x\}$ such that

$$E^{-1}\text{``}K(x) = \{y \in \mathtt{dom}(E) : E\text{``}\{x\} \cap E\text{``}\{y\} \neq \emptyset\};$$

further, for any $x, y \in \mathtt{dom}(E)$, if $K(K(x)) = K(K(y))$ then $K(x) = K(y)$.

This expresses the axiom of kinds and the axiom of diversity in internal terms.

We provide for the axiom of separation in internal terms by making a very strong assumption. The models we describe in this way are called *natural* models. The idea of a natural model is that *every* subset of what the model

thinks of as a kind is implemented in the model. A more careful analysis of what is needed for a model of separation, which is much less, will be given below.

Definition: We call an extensional relation E *complete* iff for each $a \in \mathtt{rng}(A)$, and for each nonempty $B \subseteq E^{-1}``A$, there is $b$ such that $b \, E \, K(a)$ and $E^{-1}``b = B$. We further provide that every $K(a)$ contains at least one element c such that there is no $d \, E \, c$.

We could in addition provide a function N such that $N(K(x)) \, E \, K(K(x))$ and for all y, $\neg y \, E \, N(K(x))$. This would implement empty sets in each appropriate kind.

Finally, we provide for pairing in the model.

Definition: We call a sorted extensional relation E *coupled* iff each $K(a)$ has at least two elements and there is an injective map P whose domain is the union of all $K(a) \times K(a)$ for $a \in \mathtt{dom}(E)$ such that

$$P``(K(a) \times K(a)) \subseteq K(a)$$

for each $a \in \mathtt{dom}(E)$.

A natural model of our theory in our theory is determined by a coupled complete sorted extensional relation. It should be clear that all primitive notions and axioms of our theory translate into terms of these models. We haven't provided for choice but it could easily be expressed (and will in fact hold in natural models if assumed in the theory itself).

Our theory does not prove that there are natural models of itself. An elegant argument for this is to show that if there is a natural model of the theory, there is a natural model which itself thinks there are no natural models.

Notice that for any $a \in M = \mathtt{dom}(E)$, $|K(K(a))| \geq 2^{|K(a)|}$ by completeness. So the existence of a natural model of our theory implies the existence of a sequence of infinite cardinals κ_i such that $\kappa_{i+1} \geq 2^{\kappa_i}$. This implies that each $\beth_i$ for i a natural number exists. Choose an element A_i of each $\beth_i$ (in such a way that these sets are pairwise disjoint) and choose an injection $f_i : \mathcal{P}(A_i) \to \iota``A_{i+1}$. Define a relation E on $\bigcup_{i \in \mathbb{N}} A_i$: $x \, E \, y$ iff $(\exists i : x \in f_i^{-1}(\{y\}))$. It is straightforward to see that this models our theory (showing that this E is a coupled relation is an exercise) and evident that

in this model of our theory there is no natural model of our theory, because any kind in the sense of this model is of size $\beth_n$ for a natural number n and so cannot support a sequence of infinite cardinals κ_i such that $\kappa_{i+1} \geq 2^{\kappa_i}$.

Intuitively, the reason that our theory cannot see natural models of itself in general is that all sets in a natural model must be in the same kind in the theory we are working in, and there is no guaranteed that a single kind in our world holds large enough sets to serve as kinds in the sense of a natural model.

Our theory does not prove that there are individuals (showing this will require extra work). But every natural model contains at least one kind of individuals. The reason for this is that $\aleph(|K(a)|) \leq \aleph(\exp^2(|K(a)| \cdot K(a)|)) \leq \aleph(|K^3(a)|)$. There must be a smallest cardinal $\aleph(|K(a)|)$ even if choice is not assumed, and by the inequality this $K(a)$ cannot be a $K^3(a)$, so there must be a type of individuals in the natural model.

It is impossible to express in the language of our theory the idea that there is a single sequence $\{\kappa^i(x)\}_{i \geq 1}$ containing all kinds, but the most natural construction of a natural model leads to this picture. Choose a sequence of infinite cardinals κ_i such that $\kappa_{i+1} \geq 2^{\kappa_i}$ and $\kappa_i^2 = \kappa_i$. Choose an element A_i of each κ_i (in such a way that these sets are pairwise disjoint) and choose an injection $f_i : \mathcal{P}(A_i) \to \iota``A_{i+1}$ (and choose pairing functions on each A_i). Define a relation E on $\bigcup_{i \in \mathbb{N}} A_i$: $x \, E \, y$ iff $(\exists i : x \in f_i^{-1}(\{y\})$. Nothing prevents us from choosing many such sequences and having disjoint indexed hierarchies of kinds, but nothing is gained by it in terms of the mathematics we can do.

15.2 Extending the theory to internalize the hierarchy of kinds

The simplest axioms which give the theory the ability to talk about the iterated hierarchy of kinds are motivated by things we have actually done in our mathematical development.

Axiom of Iterated Singletons: We postulate for each object x and for each natural number n (not necessarily of the same kind as x) an object $\iota^n(x)$, with $\iota^0(x) = x$ and $\iota^{n+1}(x) = \{\iota^n(x)\}$. This internalizes a definition we are already using as an axiom. The power of introducing this is that it can be used in comprehension.

Definition: $\kappa^{n+1}(x) = \kappa(\iota^n(x))$ [it is provable by induction on n that $\kappa^{n+1}(x) = \kappa^{n+1}(y)$ for $x \sim y$.] $\iota^n {}^{\text{\textquotedblleft}}A = \{\iota^n(a) \in \kappa^n(A) : a \in A\}$. $T^n(x) = |\iota^n {}^{\text{\textquotedblleft}}A|$. Of course we have already established these facts, but we are now able to prove by induction statements we could only give metatheoretical arguments for beforehand.

Discussion: Some other iterations such as $\mathcal{P}^n(A)$ and $\bigcup^n(A)$ seem to present more difficulties, but can be defined cleverly in terms of these operations. In fact, iteration can be managed. For example if $F(x)$ is a definable object in $\kappa^2(x)$, we can define $F^n(x)$ as the term a_n of a finite sequence a such that $a_0 = \iota^n(x)$ and for each i, for some u, $a_i = \iota^{n-i}(u)$ and $a_{i+1} = \iota^{n-(i+1)}(F(u))$. Other differentials in kind between input and output of an operation can be handled similarly, and certainly the iterations just mentioned are manageable in this way.

Axiom of Hierarchy: There is an x such that for every y there is an n such that $y \in \kappa^n(x)$. This axiom neatly confines the hierarchy to the minimal configuration, an iterated construction of kinds over a kind of individuals.

Observation: It is worth noting that this theory with the Axiom of Iterated Singletons and without Hierarchy proves that there is a type of individuals, using the same argument that $T^2(\aleph(|\kappa(x)|)) < \aleph(|\kappa^3(x)|)$ indicated above, and then remarking that for any x there must be a smallest cardinal of the form $T^{-n}(\aleph(\kappa^{1-n}(x)))$, so there must be a type of individuals $\kappa(u)$ and a natural number m so that $\kappa^m(u) = x$.

Discussion: The key thing here is actually being able to express general statements about the relationships between kinds in a hierarchy, which the basic language of our theory cannot do. The extension of our language described here is very natural and true to the mathematics as we have actually developed it. It allows some simplifications of mathematical terminology. For example, natural numbers can be taken to be the natural numbers of the lowest kind in which there are natural numbers in the general sense. We can say that two sets A, B are equinumerous iff there is a bijection from $\iota^m {}^{\text{\textquotedblleft}}A$ to $\iota^n {}^{\text{\textquotedblleft}}B$ for some natural numbers m, n and define $|A|$ as the set of all sets B equinumerous with A in this sense and not equinumerous with any set in a kind of lower index than $\kappa(B)$. In this way sets in all kinds are counted by the same cardinals.

Something similar can be done with ordinals. For many purposes one would still want to talk about cardinals and ordinals of a fixed kind, but this device is amusing and philosophically satisfying.

It is still impossible to say anything about the relationship between kinds which are not in a hierarchy with each other using the iterated singleton approach.

We outline a quite different seeming structural axiom with similar consequences.

Axiom of Ordered Pairs (2): We postulate for any objects x, y whatsoever an object (x, y): we postulate the axioms $(x, y) = (z, w) \rightarrow x = z \wedge y = w$ and $x \sim z \wedge y \sim w \rightarrow (x, y) \sim (z, w)$.

Discussion: This axiom allows definition of equinumerousness between sets of any kinds, and implies the existence of a kind of individuals for the usual cardinality reasons. A further refinement of this which allows a limited amount of dependent typing is to assert that $y \in z \rightarrow (x, y) \sim (x, z)$.

15.3   The temptation of ambiguity

In the main development of this theory, except in the previous subsection, there is a good deal of relativity of kinds. Mathematical objects are defined relative to a base kind $\kappa(x)$ to the extent that we felt the need for a special convention to simplify references to the base kind. The set of natural numbers, for example, is defined in each type $\kappa^2(x)$. Anything that we can actually prove about a kind $\kappa(x)$ or define over $\kappa(x)$ [often in some $\kappa^n(x)$ for $n > 1$] we can prove or define over any other kind.

This suggests the following

Axiom of Ambiguity: For any statement $P(x)$ in which x is the only free variable, $(\forall xy : P(\kappa(x)) \leftrightarrow P(\kappa(y)))$ holds.

This says that all the kinds look the same in any way we can express. Notice that it implies immediately that there are no individuals, so it is not consistent with the axiom of iterated singletons.

The theory still proves that the kinds are distinct, but it makes us quite indifferent to what kind we are talking about. It encourages our use of notations with bold face numeral subscripts replacing kind subscripts, as long as [as the convention requires in any case] all the boldface subscripts hide reference to the same kind.

With some logical finesse, we can further use our theory with the axiom of ambiguity to describe a world in which the kinds are all the same.