

Type theory without types

Randall Holmes (expanding on Thomas Forster)

March 15, 2021

This is a presentation of the simple typed theory of sets as an unsorted theory.

The initial observation is due to Thomas Forster. In a context which is unsorted in the logical sense, we can recognize that two objects are of the same type if they belong to a common set.

Definition: $x \sim_\tau y$ is defined as $(\exists z : x \in z \wedge y \in z)$. We read this as “ x has the same type as y ”, though we do need to justify that it is an equivalence relation.

Axiom of Reflexivity: For every x , $x \sim_\tau x$. This amounts to the assertion that every object belongs to some set, and so that “ x has the same type as x ”.

Axiom of Separation: For every x and for each formula ϕ , there is an object $\{y \sim_\tau x : \phi\}$ such that $\{y \sim_\tau x : \phi\} \sim_\tau \{y \sim_\tau x : y = y\}$ and $(\forall y : y \in \{y \sim_\tau x : \phi\} \leftrightarrow y \sim_\tau x \wedge \phi)$. We define $\tau(x)$ as $\{y \sim_\tau x : y = y\}$ and note that we can equivalently write $\{y \sim_\tau x : \phi\}$ as $\{y \in \tau(x) : \phi\}$. In other words, for any property we can collect the objects with that property of the same type as a given x . The type of x , $\tau(x)$, is the set of all objects of the same type as x : we reiterate in different terminology that $\{y \in \tau(x) : \phi\} \sim_\tau \tau(x)$, or $\{y \in \tau(x) : \phi\} \in \tau^2(x)$.

We can now establish that \sim_τ is a symmetric and transitive. That

$$x \sim_\tau y \leftrightarrow y \sim_\tau x$$

is simply a truth of logic (given the definition of \sim_τ). It follows from Separation that $x \in \tau(y) \leftrightarrow y \in \tau(x)$.

Now suppose that $x \sim_\tau y$ and $y \sim_\tau z$. It follows that $x \in \tau(y)$ and $z \in \tau(y)$, whence $x \sim_\tau z$.

Thus we have established that \sim_τ is an equivalence relation.

We have one more axiom controlling types.

Axiom of Level: If $x \in y, u \in v$, and $y \in_\tau v$ then $x \sim_\tau u$. It is worth noticing that in the presence of the other axioms, the axiom of level is equivalent to the usual Axiom of Union asserting the existence for each A of $\{x : (\exists y : x \in y \wedge y \in A)\}$ or an axiom of binary union asserting that if $x \sim_\tau y$, then $\{z : z \in x \vee z \in y\}$ exists.

We state identity criteria for sets.

Axiom of Extensionality:

$$(\forall xyz : z \in x \wedge (\forall u : u \in x \leftrightarrow u \in y) \rightarrow x = y)$$

Nonempty objects with the same elements are equal.

Axiom of Null Sets: If $(\forall y \sim_\tau x : \neg \phi)$ and $x \sim_\tau x'$ then

$$\{y \sim_\tau x : \phi\} = \{u \sim_\tau x' : u \neq u\},$$

We define $\emptyset_{\tau(x)}$ as $\{u \sim_\tau x : u \neq u\}$.

We can define a notion of sethood at this point, and state some convenient alternative extensionality theorems.

Definition: We define “ x is a set”, written as $\mathbf{set}(x)$, as

$$(\exists y : y \in x \vee x = \emptyset_{\tau(y)}).$$

Theorem: $(\forall xy : \mathbf{set}(x) \wedge \mathbf{set}(y) \wedge x \sim_\tau y \wedge (\forall z : z \in x \leftrightarrow z \in y) \rightarrow x = y)$

Theorem: $\{y \sim_\tau x : \phi\} = \{y \sim_\tau u : \psi\}$ iff $x \sim_\tau u$ and $(\forall y : \phi \leftrightarrow \psi)$.

Note that nothing here prevents the existence of non-sets with no elements, in addition to distinct null sets in each type which contains sets.

We define local hierarchies of types.

Definition: We define $\tau^1(x)$ as $\tau(x)$ and $\tau^{n+1}(x)$ as $\tau(\tau^n(x))$ for each concrete natural number n . We will see that we cannot associate these indices sensibly with natural numbers inside the theory. We will certainly never quantify over these indices.

Observation: If $x \in y$, we have $x \in \{z \sim_\tau x : z \in y\} \sim_\tau \tau(x)$ by Separation. $\{z \sim_\tau x : z \in y\}$ is nonempty and has the same elements as y [note that any element of y is of the same type as x by definition of \sim_τ], so is equal to y , so $y \sim_\tau \tau(x)$, so $y \in \tau^2(x)$.

Observe that $\tau(u) = \tau(v)$ iff $u \sim_\tau v$. If $\tau^2(u) = \tau^2(v)$ we have $\tau(u) \sim_\tau \tau(v)$ whence by the axiom of levels and $u \in \tau(u), v \in \tau(v)$, we have $u \sim_\tau v$ whence we have $\tau(u) = \tau(v)$.

So if we have $x \in y$ and $y \in \tau^n(u)$ with $n \geq 2$, we have $\tau^2(x) = \tau^n(u)$, whence $\tau(x) = \tau^{n-1}(u)$.

Clearly if $x = y$, it follows that $\tau(x) = \tau(y)$.

A type $\tau(u)$ is called a base type iff there is no $\tau(x) \in \tau(u)$. Notice that a non-base type will contain only one type as a member: a $\tau(u)$ in which u is not a type can be rewritten in just one way as $\tau^2(x)$ (up to the identity of $\tau(x)$, not the identity of x).

Further observations: For any x , $\{x\} = \{y \sim_\tau x : y = x\} \in \tau^2(x)$. If we define $\iota(x)$ as $\{x\}$ and define iteration in the natural way, $\iota^n(x) \in \tau^{n+1}(x)$. We now show that $\tau^2(x) \neq \tau(x)$: the key observation is that $\{y \in \tau(x) : y \notin y\}$ belongs to $\tau^2(x)$ but cannot belong to $\tau(x)$ for the usual Russellian reasons. More generally $R^n = \{\iota^{n-2}(y) \in \tau^{n-1}(x) : \iota^{n-2}y \notin y\} = \{z \in \tau^{n-1}(x) : (\exists y : z = \iota^{n-2}(y) \wedge z \notin y)\}$ will cause disaster if $\tau^n(x) = \tau(x)$. This will certainly belong to $\tau^n(x)$ by separation. The point is that $\iota^{n-2}(R^n) \in R^n$ exactly if $\iota^{n-2}(R^n) \notin R^n$ and $\iota^{n-2}(R^n) \in \tau^{n-1}(x)$, and if $\tau(x) = \tau^n(x)$ it will be true that $\iota^{n-2}(R^n) \in \tau^{n-1}(x)$, leading to contradiction.

Equivalence to the usual simple theory of types? We have given enough information to make it clear that this theory interprets TSTU (in a version with a notion of sethood). The translation is achieved by fixing an object x and interpreting each type n variable as a variable restricted to $\tau^{n+1}(x)$.

This theory certainly has more expressive power than TSTU, but we argue that TSTU interprets it as well. Take a model of TSTU and add the assertions making the types disjoint. The only axioms that present any difficulty are unstratified instances of Separation. Each such axiom can be converted to a well-typed statement (or to many well-typed statements). Start by assigning a type to each free variable (which will specify the type $\tau(x)$ to which the set constructed is restricted). Then work from the outside in, replacing each universal quantifier with the conjunction of quantified statements restricted to specific types and each existential quantifier with the disjunction of quantified statements restricted to specific types. Whenever an atomic sentence becomes explicitly ill-typed, replace it with the False. Only finitely many types need be considered when expanding quantified sentences (there is no need to consider type $\tau^n(x)$ or a type y with $\tau^n(y) = \tau(x)$ for n greater than the number of variables present, and any statement about variables of types not connected to the type of x in one of these ways can be reduced to a truth value). This process will terminate with well-typed instances of the separation axiom scheme of TSTU for each assignment of types to the free variables of the original instance of separation.

This theory is much vaguer about what types there are than is TSTU. There may be many disjoint series of types, each of the order type of \mathbb{N} or of \mathbb{Z} . There is no way whatsoever to talk about relationships between disconnected type systems.

An ambiguity axiom may be formulated: For each formula $\phi(x)$ in which no variable occurs free other than x , $(\forall xy : \phi(\tau(x)) \leftrightarrow \phi(\tau(y)))$. We note that if our language includes a Hilbert symbol, this axiom will lead very naturally to an interpretation of NFU.

We define some further frequently used concepts.

Definition: If u is of the form $\tau^2(v)$, we define $\tau^{-1}(u)$ as $\tau(v)$. Otherwise this notation is undefined.

Note that we can use this to define the usual notation for sets derived by separation: $\{x \in A : \phi\}$ may be taken hereinafter to abbreviate $\{x \in \tau^{-1}(\tau(A)) : x \in A \wedge \phi\}$, when A is a set and $A \in \tau^2(a)$ for some a .

Definition: Define $x \subseteq y$ as $x \sim_\tau y \wedge \text{set}(x) \wedge \text{set}(y) \wedge (\forall z : z \in x \rightarrow z \in y)$.
 For a set A , define $\mathcal{P}(A)$, the power set of A , as $\{B \in \tau(A) : B \subseteq A\}$.
 Notice that $\mathcal{P}(A) \in \tau^2(A)$.

We do some sample mathematics, introducing the natural numbers as represented in this theory.

Definition: We define $\{x\}$ as $\{y \sim_\tau x : y = x\}$ as above. Note that $\{x\} \in \tau^2(x)$.

Definition: If x and y belong to $\tau^2(u)$, we define $x \cup y$ as

$$\{z \sim_\tau u : z \in x \vee z \in y\},$$

noting that this set will contain all elements of x , all elements of y , and no other elements, as expected.

Definition: We define $0_{\tau(x)}$ as $\{\emptyset_{\tau(x)}\}$. Note that $0_{\tau(x)} \in \tau^3(x)$ is the set of all sets of the same type as $\tau(x)$ with 0 elements.

Definition: We define $\sigma(A)$ as $\{x \cup \{y\} \in \tau^{-1}(\tau(A)) : x \in A \wedge y \notin x\}$.
 Recall that $\tau^{-1}(\tau^2(x))$ is defined as $\tau(x)$, and $\tau^{-1}(u)$ is not defined if u is not the type of a type. This can be read “the successor of A ”.

Definition schema: Define $1_{\tau(x)}$ as $\sigma(0_{\tau(x)})$ and more generally define $(n+1)_{\tau(x)}$ as $\sigma(n_{\tau(x)})$. Notice that $1_{\tau(x)}$ will be the set of all singleton sets of the same type as $\tau(x)$ and in general $n_{\tau(x)}$ for a natural number n will be the set of all sets of the same type as $\tau(x)$ with n elements. Note that $n_{\tau(x)} \in \tau^3(x)$. We reserve the right to omit type subscripts where they can be understood from context.

Definition: We call a set I a -inductive iff $a \in A \wedge (\forall x : x \in I \rightarrow \sigma(x) \in I)$.
 We define \mathcal{I}_a as the set of all a -inductive sets.

Definition: We define $\bigcup A$ as $\{x \in \tau^{-2}(\tau(A)) : (\exists y : x \in y \wedge y \in A)\}$. We define $\bigcap A$ as $\{x \in \tau^{-2}(\tau(A)) : (\forall y : y \in A \rightarrow x \in y)\}$.

Definition: We define $\mathbb{N}_{\tau(x)}$ as $\bigcap \mathcal{I}_{0_{\tau(x)}}$. We thus define $\mathbb{N}_{\tau(x)}$ as the intersection of all sets which contain $0_{\tau(x)}$ and are closed under successor. It is useful to note that there is a set which satisfies this closure condition: $\tau^3(x)$. $\mathbb{N}_{\tau(x)} \in \tau^4(x)$ is useful to note.

Definition: We define the set $\mathcal{F}_{\tau(x)}$ of finite subsets of $\tau(x)$ as $\bigcup \mathbb{N}$.

Axiom of Infinity: We assert that for every x , $\tau(x) \notin \mathcal{F}_{\tau(x)}$. Each type is an infinite set.

Observation: Notice that we do not quantify over numerals n in the context $n_{\tau(x)}$ any more than we do in the context $\tau^n(x)$. We certainly can quantify over all $n \in \mathbb{N}_{\tau(x)}$, but that is not (officially) the same thing. The natural numbers as we know them before we start developing our type theory are not actually part of our subject matter.

Theorems: The Peano axioms for arithmetic hold for our implementation(s):

1. $0_{\tau(x)} \in \mathbb{N}_{\tau(x)}$
2. For each $n \in \mathbb{N}_{\tau(x)}$, $\sigma(n) \in \mathbb{N}_{\tau(x)}$.
3. For each $n \in \mathbb{N}_{\tau(x)}$, $\sigma(n) \neq 0_{\tau(x)}$.
4. For each $m, n \in \mathbb{N}_{\tau(x)}$, $\sigma(m) = \sigma(n) \rightarrow m = n$ (the proof of axiom 4 requires the Axiom of Infinity, unlike the others).
5. (mathematical induction) Any $0_{\tau(x)}$ -inductive set includes all elements of $\mathbb{N}_{\tau(x)}$: for any predicate P , if $P(0_{\tau(x)})$ and

$$(\forall n \in \mathbb{N}_{\tau(x)} : P(n) \rightarrow P(\sigma(n))),$$

it follows that $(\forall n \in \mathbb{N}_{\tau(x)} : P(n))$, because under these conditions, $\{x \in \mathbb{N}_{\tau(x)} : P(x)\}$ will be inductive.

We can at once define an operation on sets generalizing addition of natural numbers.

Definition: Define $x \cap y$ (where $x \sim_{\tau} y$) as $\{u \in x : u \in y\}$. Define $A + B$ as $\{a \cup b \in \tau^{-1}(\tau(A \cup B)) : a \in A \wedge b \in B \wedge a \cap b = \emptyset_{\tau^{-2}(\tau(A \cup B))}\}$. We can expect intuitively that for familiar numerals m, n , $m_{\tau(x)} + n_{\tau(x)} = (m + n)_{\tau(x)}$, but we do not have a way of saying this internally to this theory.

Theorems: The first-order Peano axioms for addition hold for this definition of addition.

1. $(\forall m, n \in \mathbb{N}_{\tau(x)} : m + n \in \mathbb{N}_{\tau(x)})$

2. $(\forall m \in \mathbb{N}_{\tau(x)} : m + 0_{\tau(x)} = m)$
3. $(\forall m, n \in \mathbb{N}_{\tau(x)} : m + \sigma(n) = \sigma(m + n))$

There is an obvious correlation between natural numbers in related types.

Definition: Define $\iota^{\text{“}}A$ as $\{\{a\} \in \tau(A) : a \in A\}$.

Theorem: $(\forall A : A \in \mathcal{F}_{\tau(x)} \rightarrow \iota^{\text{“}}A \in \mathcal{F}_{\tau^2(x)})$

Definition: For $n \in \mathbb{N}_{\tau(x)}$, define $T(n)$ as the unique $m \in \mathbb{N}_{\tau^2(x)}$ which contains $\iota^{\text{“}}A$ for each $A \in n$. We firmly believe of course that $T(n_{\tau(x)}) = n_{\tau^2(x)}$ for each familiar natural number n in the pre-type-theoretic sense, but we do not allow ourselves a way to say this. More generally, we surely believe that $T^i(n_{\tau(x)}) = n_{\tau^{i+1}(x)}$

As usual, we define $T^1(n) = T(n)$; $T^{i+1}(n) = T(T^i(n))$ for each concrete integer i .

Theorem: For every $n \in \mathbb{N}_{\tau(x)}$, $T(n)$ is defined, and for every $n \in \mathbb{N}_{\tau^2(x)}$, there is a unique m such that $T(m) = n$, which we call $T^{-1}(n)$.

We define $T^{-(i+1)}(n) = T^{-1}(T^{-i}(n))$ for each concrete positive integer i .

Definition: For $m, n \in \mathbb{N}_{\tau(x)}$, we define $m \cdot n$ as the set of all unions $\bigcup A$ where $A \in T(n)$, $A \subseteq m$, and for each $a, b \in A$, $a \cap b = \emptyset_{\tau(x)}$. This is basically the elementary school definition of multiplication with some finicky attention to types.

Theorems: The first order Peano axioms for multiplication hold in our interpretation(s):

1. $(\forall m, n \in \mathbb{N}_{\tau(x)} : m \cdot n \in \mathbb{N}_{\tau(x)})$
2. $(\forall m \in \mathbb{N}_{\tau(x)} : m \cdot 0_{\tau(x)} = 0_{\tau(x)})$
3. $(\forall m, n \in \mathbb{N}_{\tau(x)} : m \cdot \sigma(n) = (m \cdot n) + n)$

The T operation commutes with operations of arithmetic, as one would expect.

Theorem: For any $m, n \in \mathbb{N}_{\tau(x)}$, $T^i(\sigma(m)) = \sigma(T^i(m))$, $T^i(m + n) = T^i(m) + T^i(n)$, and $T^i(m \cdot n) = T^i(m) \cdot T^i(n)$, for any concrete positive integer i and for any concrete negative integer i for which either side is defined.

Now we introduce an ordered pair in order to commence discussion of relations and functions.

Definition: For $x \sim_{\tau} y$, define $\{x, y\}$ as $\{z \in \tau(x) : z = x \vee z = y\}$. Define $\{x_1, x_2, \dots, x_n\}$ as $\{x_1\} \cup \{x_2, \dots, x_n\}$. Define $[x, y]$, the Kuratowski ordered pair of x and y , as $\{\{x\}, \{x, y\}\}$. Note that $[x, y] \in \tau^2(x) = \tau^2(y)$.

Theorem: For any x, y, z, w all of the same type, $[x, y] = [z, w] \rightarrow x = z \wedge y = w$.

For us, the Kuratowski pair is a nonce construction which we use in order to introduce another ordered pair construction almost immediately.

Definition: $[A \times B] = \{[a, b] \in \tau^2(A) : a \in A \wedge b \in B\}$.

Definition: A Kuratowski relation is a set of Kuratowski ordered pairs. We write $x[R]y$ for $[x, y] \in R$. Notice that for this to be true we must have $\tau^4(x) = \tau^4(y) = \tau^2([x, y]) = \tau(R)$. This type differential eventually proves inconvenient and we make an adjustment. But the adjusted theory is difficult to motivate without the initial development using this pair.

Definition: if R is a Kuratowski relation, $[R]^{-1}$ is defined as $\{[y, x] : x[R]y\}$. We write just $x[R]^{-1}y$ instead of $x[[R]^{-1}]y$. We define $\text{dom}[R]$ as

$$\{x \in \tau^{-3}(R) : (\exists y : x[R]y)\}.$$

We define $\text{rng}[R]$ as $\text{dom}[[R]^{-1}]$.

Definition: A Kuratowski function f is a Kuratowski relation such that for all x, y, z , $x[f]y \wedge x[f]z \rightarrow y = z$. We define $f : [A \rightarrow B]$ as asserting that f is a function, $\text{dom}[f] = A$, and $\text{rng}[f] \subseteq B$.

Definition: A Kuratowski function f is injective iff $[f]^{-1}$ is a Kuratowski function. A Kuratowski function is a Kuratowski bijection from A to B iff $f : [A \rightarrow B]$ and $[f]^{-1} : [B \rightarrow A]$.

Definition: We define $A \sim B$ (A is the same size as B) as holding iff there is a Kuratowski bijection from A to B . We define $|A|$, the cardinality of A , as $\{B \in \tau(A) : B \sim A\}$. Notice that $|A| \in \tau^2(A)$. We refer to any set $|A|$ as a cardinal number.

Theorems: Natural numbers are cardinal numbers (for any finite set A , $A \in n \in \mathbb{N}_{\tau^{-1}(\tau(A))} \leftrightarrow |A| = n$). \sim is an equivalence relation.

Definition: We define $T(|A|) = |\iota "A|$. This extends the identification of natural numbers with natural numbers in the next higher type to cardinal numbers. We will see below that T^{-1} is not total on general cardinals: the T operation is one-to-one, though.

***Definition:** We define $|A| + |B|$ as $T^{-2}(|[A \times \{x\}] \cup [B \times \{y\}]|)$ where x and y are distinct elements of $\tau^{-1}(\tau(A))$. We define $|A| \cdot |B|$ as $T^{-2}(|[A \times B]|)$. There is some evidence here of the inconvenience of the type differential in the definition of the Kuratowski ordered pair.

The inconvenient character of the proposed definitions of addition and multiplication of cardinals can be taken to motivate our immediate introduction of our

Axiom of Ordered Pairs: For each x , there is a Kuratowski bijection

$$\pi_{\tau(x)} : [[\tau(x) \times \tau(x)] \rightarrow \iota^2 " \tau(x)].$$

Definition: We define (x, y) , the (official) ordered pair by $\pi_{\tau(x)}[[x, y]] = \{\{(x, y)\}\}$. Note that $\tau(x) = \tau(y) = \tau((x, y))$.

Theorem: If x, y, z, w are all of the same type $(x, y) = (z, w) \leftrightarrow x = z \wedge y = w$.

Definition: $A \times B = \{(a, b) \in \tau^{-1}(\tau(A \cup B)) : a \in A \wedge b \in B\}$.

Definition: A relation is a set of ordered pairs. We write xRy for $(x, y) \in R$. Notice that for this to be true we must have $\tau^2(x) = \tau^2(y) = \tau^2((x, y)) = \tau(R)$.

Definition: if R is a relation, R^{-1} is defined as $\{(y, x) : xRy\}$. We define $\text{dom}(R)$ as

$$\{x \in \tau^{-1}(\tau(R)) : (\exists y : xRy)\}.$$

We define $\text{rng}(R)$ as $\text{dom}(R^{-1})$.

Definition: A function f is a relation such that for all x, y, z , $xfy \wedge x fz \rightarrow y = z$. We define $f : A \rightarrow B$ as asserting that f is a function, $\text{dom}(f) = A$, and $\text{rng}(f) \subseteq B$.

Definition: A function f is injective iff f^{-1} is a function. A function is a bijection from A to B iff $f : A \rightarrow B$ and $f^{-1} : B \rightarrow A$.

Theorem: $A \sim B$ iff there is a bijection $f : A \rightarrow B$. So all notions and theorems about cardinality transfer to the new pair.

The underlying fact here is that there is a precise correspondence between functions f and Kuratowski functions

$$f^* = \{[x, f(x)] \in \tau^2(f) : x \in \text{dom}(f)\}.$$

Definition: We define $|A| + |B|$ as $|(A \times \{x\}) \cup (B \times \{y\})|$ where x and y are distinct elements of $\tau^{-1}(\tau(A))$. We define $|A| \cdot |B|$ as $|A \times B|$.

Theorems: The notions of addition and multiplication just defined extend the notions already defined for natural numbers.

The fact that we introduce our official ordered pair by axiom may require some comment. It is worth noting that the Axiom of Ordered Pairs in fact implies the Axiom of Infinity (this is not difficult: strictly speaking, the Axiom of Ordered Pairs implies that each type $\tau^2(x)$ is infinite; it is consistent with existence of base types with just one element). If we assumed the Axiom of Infinity and strengthened extensionality so that all objects are sets, we would be able to prove Ordered Pairs as a theorem (the definition of the Quine ordered pair gives an explicit $\pi_{\tau^5(x)}$). It is straightforward to produce an interpretation of our theory with Ordered Pairs from an interpretation with just Infinity, exploiting the fact that the definition of the Quine pair works on sets of sets in this theory without full extensionality. Further, our theory with Infinity and the Axiom of Choice also proves Ordered Pairs. Our view is that there is value in allowing non-sets in our theory and that the definition of the Quine ordered pair is remarkably complicated, and that results noted here show that the assumption of a type level ordered pair that we make here is not unreasonable.

We note that we seem to have given ourselves the power to choose a $\pi_{\tau(x)}$ for each of infinitely many types. It should be noted that unless significant power is added to the system, there isn't any mathematical content to this,

as we really have no way to consider more than a concrete finite collection of types at one time in any argument. It is also true that with full extensionality and Infinity, we really can do this, as the Quine pair is definable in the same way in every type with sufficiently high index. And the use of the Quine pair on sets of sets allows an interpretation of the Axiom of Ordered Pairs exactly as written in all types uniformly (assuming Infinity): the idea is that we restrict ourselves to sets of sets as our objects (in all types which contain sets of sets). The Quine pair on these is definable. We then restrict the membership relation so we consider only sets of sets of sets as the sets of the interpretation, and we recover the full theory (on a cleverly restricted domain) with a type level ordered pair.

For reference, we give the definition of the Quine pair, of which we make no use here.

Definition: Define $\sigma_0(a)$ as $\sigma(a)$ if $a \in \mathbb{N}_{\tau^{-2}(\tau(a))}$ and as a otherwise.

Define $\sigma_1(A)$ as $\{\sigma_0(a) \in \tau^{-1}(\tau(A)) : a \in A\}$.

Define $\sigma_2(A)$ as $\sigma_1(A) \cup \{0_{\tau^{-3}(A)}\}$

Define $\langle A, B \rangle$ as

$$\{c \in \tau^{-1}(\tau(A \cup B)) : (\exists a \in A : c = \sigma_1(a)) \vee (\exists b \in B : c = \sigma_2(b))\}.$$

Theorem: If $\langle A, B \rangle = \langle C, D \rangle$ then $A = C$ and $B = D$, if A, B, C, D are sets of sets belonging to a $\tau^5(x)$. It is also worth noting that every set of sets in such a type is a Quine pair.

We now look into function spaces and exponentiation.

Definition: We define B^A as $\{f \in \tau(A \cup B) : f : A \rightarrow B\}$. In a horrible pun, usual in set theory already, we define $|B|^{|A|}$ as $T^{-1}(|B^A|)$. As we will see below, this can be undefined, because T^{-1} , though injective (if $T(\kappa) = T(\lambda)$ then $\kappa = \lambda$) is partial. There are cardinals which are not outputs of the T operation. The use of T^{-1} is necessary if we are to have an exponentiation operation that we can manipulate as a mathematical function: if $A, B \in \tau^2(x)$ then $B^A \in \tau^3(x)$. We note that this definition does define the usual exponentiation operation on natural numbers.

Definition: We define $|A| \leq |B|$ as holding if there is an injection from A to B . We define $|A| < |B|$ as holding iff $|A| \leq |B|$ and $A \not\sim B$ (so $|A| \neq |B|$).

Theorem (Cantor-Schroder-Berstein): If $|A| \leq |B|$ and $|B| \leq |A|$ then $|A| = |B|$ (so $A \sim B$).

Theorem (Cantor): $T(|A|) < |\mathcal{P}(A)|$.

Proof of Cantor Theorem: The theorem is not stated in the usual form, because $|A| < |\mathcal{P}(A)|$ does not make sense, as elements of A and elements of $\mathcal{P}(A)$ are not of the same type. But $T(|A|)$ is in some intuitive sense the same cardinal as $|A|$: the mathematical content of the theorem is the same, and the proof, as we shall see, is very similar to the usual one.

Certainly $T(A) = |\iota“A| \leq |\mathcal{P}(A)|$. This is witnessed by the identity function restricted to $\iota“A$, which is an injection from $\iota“A$ to $\mathcal{P}(A)$.

So we need to show that $|\mathcal{P}(A)| \not\leq |\iota“A|$ (appealing here to Cantor-Schroder-Bernstein). Suppose we have an injection $f : \mathcal{P}(A) \rightarrow \iota“A$. Consider $R = \{x \in A : x \notin f^{-1}(\{x\})\}$. $f(R) = \{r\}$ for some $r \in A$. $r \in R$ iff $r \notin f^{-1}(\{r\}) = R$, which is a contradiction.

Corollary: $T^{-1}(|\tau^2(x)|)$ is undefined. This follows from $|\iota“\tau(x)| < |\mathcal{P}(\tau(x))| \leq |\tau^2(x)|$. So, as foreshadowed, the identification of cardinals in lower types with cardinals in higher types via the T operation is not surjective: the higher type contains new cardinals, as it were.

This also shows that the exponential map is partial: $2^{|A|} = T^{-1}(|2^A|) = T^{-1}(|\mathcal{P}(A)|)$, via the one to one correspondence of sets with their characteristic functions, and we have seen that $T^{-1}(|\mathcal{P}(\tau(x))|)$ is undefined.

It might be useful here to introduce the alternative notation $V_{\tau(x)}$ for the set $\tau(x)$: a type is a local universal set, and this would make some of what I write more familiar to readers of text about TSTU or NFU.

We discuss the Axiom of Choice, which is enormously useful in simplifying cardinal arithmetic, but which in type theory can be most naturally initially introduced as a formal device to simplify equivalence class constructions.

If R is an equivalence relation (reflexive, symmetric, and transitive) it determines a partition of $\text{dom}(R)$ into equivalence classes

$$[x]_R = \{y \in \tau^{-1}(\tau(R)) : yRx\}.$$

One frequently wants to identify all elements of an equivalence class for some mathematical purpose. In ordinary set theory, the object obtained from x by this identification can conveniently be taken to be the class $[x]_R$ itself, but here we see that $[x]_R \in \tau^2(x)$: these objects are of different types. So, it is also frequently some canonical element of $[x]_R$ which is taken to stand in for x in such constructions. The Axiom of Choice is the assertion that we can always choose a representative element from each compartment of a partition.

Definition: A partition is a set of nonempty sets P such that for any $A, B \in P$ either $A = B$ or $A \cap B = \emptyset_{\tau^{-2}(\tau(P))}$. A choice set for a partition P is a set $C \subseteq \bigcup P$ such that for each $A \in P$, $C \cap A$ has exactly one element.

Axiom of Choice: Every partition has a choice set.

Observation: If our language includes a Hilbert symbol which can occur essentially in instances of separation, then Choice is a theorem.

Theorems: The Axiom of Choice is very powerful and has many unexpected consequences. It is equivalent to the assertion that for any infinite cardinal κ , $\kappa \cdot \kappa = \kappa$, so Choice and Infinity together imply the Axiom of Ordered Pairs. The Axiom of Ordered Pairs does *not* imply choice, since it only requires this equation for the cardinals of types.

We have discussed infinite cardinal numbers. We now discuss infinite ordinals.

We use for the moment a representation of an order \leq by the set of its initial segments $\{y : y \leq x\}$. In this parlance, a well-ordering is represented by a collection of sets S with the property that S is linearly ordered by inclusion, for every $T \subseteq S$, $\bigcup T \in S$, and for every $s \in S$ other than $\bigcup S$, there is a set $s \cup \{x\} \in S$ with $x \notin s$. The order \leq_S on the domain $\bigcup S$ is defined by $x \leq_S y \leftrightarrow (\forall s \in S : y \in s \rightarrow x \in s)$. If $A \subseteq \bigcup S$ is nonempty, consider the union A' of all elements of S which do not meet A . This belongs to S and is not equal to $\bigcup S$, so there is a such that $A' \cup \{a\} \in S$, and this a must be the \leq_S -least element of A .

We say that sets S and S' representing well-orderings are isomorphic and write $S \approx S'$ iff there is a bijection f such that for each $s \in S$, $f"s \in S'$, and for each $s' \in S'$, $f^{-1}"s' \in S$.

Isomorphism is an equivalence relation. We define the order type of a representation S as the set of all $S' \approx S$, and write this $\text{ot}(S)$. We call a set an ordinal number iff it is the order type of some representation of a well-ordering.

We observe that if the elements of $\bigcup S$ are in $\tau(x)$, then elements of S are in type $\tau^2(x)$, $S \in \tau^3(x)$, and $\text{ot}(S) \in \tau^4(x)$.

We note that we could also define $\text{ot}(S)$ as the set of all relations $\leq_{S'}$ for $S' \approx S$, which would be in $\tau^3(x)$. In general, we could consider the well-orderings \leq_S directly as sets of ordered pairs. But the representation of well-orderings considered here is elegant and very useful for the following theorem.

Theorem (using Choice): For any set A , there is a set S representing a well-ordering such that $\bigcup S = A$.

Sketch of Proof: Consider the set M of all pairs (T, t) where T is a proper subset of A and t is a one-element subset of $A - T$. M admits a partition P into sets $M_t = \{(T, t') \in M : t' = t\}$. This partition admits a choice set C . Now define a nice representation as a set S' representing a well-ordering in which for each $s' \in S'$ other than $\bigcup S'$ the unique $s' \cup \{x'\} \in S'$ will satisfy $(s', \{x'\}) \in C$. The union of all nice representations is a representation S of a well-ordering of A . The idea is that C guides us in choosing the “next” element of A in the well-ordering represented by S at every stage.

If S represents a well-ordering and $x \in \bigcup S$, then $S_x = \{s \in S : x \notin s\}$ also represents a well-ordering, a proper initial segment of S . We define a strict partial order on ordinal numbers: $\text{ot}(T) < \text{ot}(S)$ iff $(\exists x \in \bigcup S : T \approx S_x)$. Of course we can define $\text{ot}(T) \leq \text{ot}(S)$ as $\text{ot}(T) < \text{ot}(S) \vee \text{ot}(T) = \text{ot}(S)$.

We define \mathbf{Ord} as the set of all ordinal numbers. We refer to the natural well-ordering on the ordinals as $W_{\mathbf{Ord}}$ (the set representing it), or $\leq_{\mathbf{Ord}}$ (the set of ordered pairs).

Theorem: $\{\{\beta \in \mathbf{Ord} : \beta < \alpha\} \in \mathcal{P}(\mathbf{Ord}) : \alpha \in \mathbf{Ord}\}$ represents a well-ordering of the ordinal numbers.

We can define T operation on ordinal numbers analogous to the operation on natural numbers and cardinals already defined. For $\alpha \in \mathbf{Ord}$, and any

$S \in \alpha$, $T(\alpha) = \text{ot}(\{\iota^{\text{“}}s : s \in S\})$, the order type of the well-ordering on $\iota^{\text{“}}\bigcup S$ induced by the order on $\bigcup S$ represented by S in the obvious way.

We observe that if $S \in \alpha$, the order type of the restriction of W_{ord} to ordinals less than α is $T^3(\alpha)$ (if we used equivalence classes of relations as order types, $T^2(\alpha)$). In some sense $T(\alpha)$ is the same order type as α translated up one type. We see that each well-ordering in a given type corresponds in this way to a proper initial segment of the natural well-ordering on the ordinals of that type (two or three types higher), and so we see that there are longer well-orderings and so larger ordinal numbers in higher types. We see that $T^{-3}(\text{ot}((W_{\text{ord}})_\alpha) = \alpha$, whence $T^{-3}(\text{ot}(W_{\text{ord}}))$ cannot be defined (it would have to be larger than every ordinal number of the type of α !)