Set Theory Notes

by Tyler Wright

github.com/Fluxanoia   fluxanoia.co.uk

These notes are not necessarily correct, consistent, representative of the course as it stands today, or rigorous. Any result of the above is not the author's fault.

These notes are marked as unsupported, they were supported up until January 2021.

Contents

1 The Axioms

1.1 Axiom of Extensionality

For two sets a and b, we have that $a = b$ if and only if for all x we have that:

$$x \in a \Longleftrightarrow x \in b.$$

For two classes A and B, we have that $A = B$ if and only if for all x we have that:

$$x \in a \Longleftrightarrow x \in b.$$

1.2 Axiom of Pair Sets

For any sets x and y, there is a set $z = \{x, y\}$. This is the (unordered) pair set of x and y.

1.3 Axiom of the Powerset

For each set x, there exists a set which is the collection of the subsets of x, the powerset $\mathcal{P}(x)$. We have the powerset defined as $\mathcal{P}(x) = \{z : z \subseteq x\}$.

1.4 Axiom of the Empty Set

There exists a set with no members, the empty set $\varnothing$. We have the empty set defined as $\varnothing = \{x : x \neq x\}$.

1.5 Axiom of Subsets

For some set x, we have that $\{y \in x : \Phi(y)\}$ is a set for some well-defined property of sets Φ.

1.6 Axiom of Infinity

There exists an inductive set.

1.7 Axiom of Unions (1.6)

We have the basic union of two sets x_1 and x_2:

$$x_1 \cup x_2 = \{y : y \in x_1 \text{ or } y \in x_2\},$$

but for when we want to unify the members of the sets in a set x, we define:

$$\bigcup x = \{y : \exists\, z \in x, y \in z\}.$$

This axiom states that for a set x, $\bigcup x$ is a set.

1.8 Intersections (1.8)

We have the basic intersection of two sets x_1 and x_2:

$$x_1 \cap x_2 = \{y : y \in x_1 \text{ and } y \in x_2\},$$

but for when we want to intersect the members of the sets in a set x, we define:

$$\bigcup x = \{y : \forall\, z \in x, y \in z\}.$$

This is a set by the Axiom of Subsets.

1.9 Axiom of Replacement

For a function F from V to itself and a set x, $F''x$ is a set.

1.10 Well-ordering Principle

For a set X, there is a well-ordering $\langle X, R\rangle$.

1.11 Axiom of Choice (5.1)

For a set of non-empty sets $\mathcal{G}$, there is a choice function F from $\mathcal{G}$ to $\bigcup \mathcal{G}$ such that for all X in $\mathcal{G}$, $F(X)$ is in X. This is equivalent to the Well-ordering Principle.

Proof. ($\Longrightarrow$) For an arbitrary set Y, it is sufficient to show Y has a well-ordering. We take $Y \neq \varnothing$ as otherwise Y is trivially well-ordered. We take $\mathcal{G} = \{X \subseteq Y : X \neq \varnothing\}$. By the Axiom of Choice, we have a choice function F_0 for $\mathcal{G}$. We take u to be any set not in Y and define F from V to V:

$$F(t) = \begin{cases} F_0(t) & \text{if } t \in \mathcal{G} \\ u & \text{otherwise.} \end{cases}$$

By the recursion theorem, we define H_0 from the ordinals to $Y \cup \{u\}$:

$$H_0(\xi) = F(Y \setminus \{H_0(\zeta) : \zeta < \xi\}).$$

We can see that:

$$\begin{aligned}
&H_0(0) = F_0(Y) \in Y,\\
&H_0(1) = F_0(Y \setminus \{F_0(Y)\}) \in Y \setminus \{F_0(Y)\},\\
&H_0(n) = F_0(Y \setminus \{F_k(Y) : k \in [n-1]_0\}) \in Y \setminus \{F_k(Y) : k \in [n-1]_0\}.
\end{aligned}$$

So, we can select distinct elements from Y recursively via our choice function on the subsets of Y. We want to show that there's some ordinal β such that $H_0(\beta) = u$. If we suppose there isn't, then H_0 is injective from the ordinals to Y, but we know that $\mathrm{ran}(H_0) \subseteq Y$ is a set by the Axiom of Replacement. Thus, H_0^{-1} is a surjection from $\mathrm{ran}(H_0)$ to the ordinals, which is a contradiction as the ordinals form a proper class. So, we take α to be the least ordinal such that $H_0(\alpha) = u$. We let $H = H_0 \restriction \alpha$, H is a bijection from α to Y, which gives us a well-ordering on Y via the well-ordering on α.

($\Longleftarrow$) For $\mathcal{G}$ any set of non-empty sets, we take $A = \bigcup \mathcal{G}$. By the well-ordering principle, there's a well-ordering $\langle A, R\rangle$. We can define a choice function as:

$$F(X) = R\text{-least element of } \langle X, R\rangle,$$

as required. $\square$

1.12 Chains

Any collection $\mathcal{G}$ of sets is called a chain if for all X and Y in $\mathcal{G}$, $X \subseteq Y$ or $Y \subseteq X$.

1.13 Zorn's Lemma (5.2)

For a set $\mathcal{F}$ such that for every chain $\mathcal{G} \subseteq \mathcal{F}$, $\bigcup \mathcal{G}$ is in $\mathcal{F}$, we have that $\mathcal{F}$ contains a maximal element Y where for all Z in $\mathcal{F}$:

$$Y \subseteq Z \Longrightarrow Y = Z.$$

This is equivalent to the Axiom of Choice and thus the Well-ordering Principle.

Proof. (ZL $\Longrightarrow$ AC) For a collection of non-empty sets $\mathcal{G}$, we want a choice function for $\mathcal{G}$. We define $\mathcal{F}$ to be the set of all choice functions that exist for subsets of $\mathcal{G}$, that is, for f in $\mathcal{F}$:

$$\mathrm{dom}(f) \subseteq \mathcal{G} \text{ and } \forall\, x \in \mathrm{dom}(f), f(x) \in x.$$

We know that $\mathcal{F}$ is non-empty as for some x in $\mathcal{G}$, x is non-empty so we choose any u in x and thus $\{\langle x, u\rangle\}$ is in $\mathcal{F}$. For any chain $\mathcal{H}$ in $\mathcal{F}$, $\mathcal{H}$ is a chain of partial choice functions on subsets of $\mathcal{G}$. We take $h = \bigcup \mathcal{H}$, so h is a function with $\mathrm{dom}(h) = \bigcup\{\mathrm{dom}(f) : f \in \mathcal{H}\} \subseteq \mathcal{G}$. Thus, f is a choice function so is in $\mathcal{F}$.

By Zorn's Lemma, there's a maximal m in $\mathcal{F}$ and we want to show that m is a choice function for $\mathcal{G}$. We know m must be a partial choice function so it's sufficient to show that $\mathrm{dom}(m) = \mathcal{G}$. We suppose that $\mathrm{dom}(m) \neq \mathcal{G}$, and take x in $\mathcal{G} \setminus \mathrm{dom}(m)$ which must be non-empty as it is in $\mathcal{G}$. For u in x, $m \cup \{\langle u, x\rangle\}$ is a partial choice function in $\mathcal{F}$ with domain $\mathrm{dom}(m) \cup \{u\}$ so $m \subset m \cup \{\langle u, x\rangle\}$. This is a contradiction of the maximality of m, so m is a choice function for $\mathcal{G}$.

(WP $\implies$ ZL) We take $\mathcal{F}$ to be a set such that for every chain $\mathcal{G} \subseteq \mathcal{F}$ we have that $\bigcup \mathcal{G}$ is in $\mathcal{F}$. By the Well-ordering Principle, $\mathcal{F}$ can be well-ordered by some relation R, we take an ordinal α such that $\langle \alpha, \in\rangle \cong \langle \mathcal{F}, R\rangle$ for some order isomorphism k. By recursion on the ordinals $\beta < \alpha$, we define a maximal chain $\mathcal{H}$ of $\mathcal{F}$. We start by putting $k(0)$ into $\mathcal{H}$, if $k(0) \subset k(1)$ then we add $k(1)$ too, if not, we move on, adding $k(\beta)$ if it contains the current maximal element of $\mathcal{H}$. This clearly forms a chain, and we will show that $Y = \bigcup \mathcal{H}$ is a maximal element of $\mathcal{F}$. By the definition of $\mathcal{F}$, as $\mathcal{H}$ is a chain, Y is in $\mathcal{F}$. If we suppose that there is some Z in $\mathcal{F}$ with $Y \subseteq Z$, then $k(\gamma) \subseteq Z$ for any γ such that $k(\gamma)$ is in $\mathcal{H}$. As Y is in $\mathcal{F}$, for some $\delta < \alpha$, $Z = k(\delta)$. But, by the definition of our recursion, at the stage γ, we decided that Z should be added to $\mathcal{H}$ so $Z \subseteq \bigcup \mathcal{H} = Y$ and as such, $Z = Y$ as required. □

1.14 Axiom of Foundation (6.4)

Every set x is well-founded, so if x is non-empty, there exists some y in x such that $x \cap y = \varnothing$. This is equivalent to saying there exists some α such that x is in V_α.

Proof. For a set x, we set $T = TC(X)$. If $T \subset V$, then for some α, $\rho''T \subseteq \alpha$ so $T \subseteq V_\alpha$. Thus, we are done for this case as $x \subseteq T \subseteq V_\alpha$ so $x \in V_{\alpha+1}$.

If we suppose that $T \setminus V \neq \varnothing$ and take y in $T \setminus V$ such that $(T \setminus V) \cap y = \varnothing$ by the Axiom of Foundation, then for any z in y, as z must be in T by the properties of TC. Also, z must be in V as $(T \setminus V) \cap y = \varnothing$. Hence, $y \subseteq V$. But, as in the first case, $\rho''y$ is a set of ordinals, with some strict upper bound β. As such, $y \subseteq V_\beta$ which implies y is in $V_{\beta+1}$ which is a contradiction of the definition of y. □

2 Relations

We will first state the significant properties relations can have. Taking a relation R on X with x, y, and z arbitrary in X:

Name	Property
Reflexive	xRx
Irreflexive	$\neg(xRx)$
Symmetric	$xRy \Rightarrow yRx$
Antisymmetric	$[xRy \text{ and } yRx] \Rightarrow [x = y]$
Connected	$[x = y]$ or $[xRy]$ or $[yRx]$
Transitive	$[xRy \text{ and } yRz] \Rightarrow [xRz]$

Equivalence relations must satisfy reflexivity, symmetry, and transitivity.

2.1 Partial Orderings (1.10)

We say that a relation $\prec$ on a set X is a (strict) partial ordering if it is irreflexive and transitive.

Similarly, we say that a relation $\preceq$ on a set X is a non-strict partial ordering if it is reflexive, antisymmetric, and transitive.

2.2 Bounding (1.11)

For a partially ordered set $(X, \prec)$, we take a non-empty subset Y of X:

- x is the infimum of Y if it's the $\prec$-greatest lower bound,
- x in Y is the minimum of Y if for all y in Y, $x \preceq y$,
- x in Y is minimal in Y if for all y in Y, $\neg(y \prec x)$,
- x is the supremum of Y if it's the $\prec$-least upper bound,
- x in Y is the maximum of Y if for all y in Y, $y \preceq x$,
- x in Y is maximal in Y if for all y in Y, $\neg(x \prec y)$.

2.3 Well-founded Orderings

A partial ordering $(X, \prec)$ is wellfounded if for any non-empty subset Y of X, Y has a $\prec$-least element.

2.4 Order Preserving Maps (1.12)

We say that f from $(X, \prec_1)$ to $(Y, \prec_2)$ is an order preserving map if for each x_1 and x_2 in X:

$$x_1 \prec_1 x_2 \Longrightarrow f(x_1) \prec_2 f(x_2).$$

Two orderings are (order) isomorphic if there is a bijective order preserving map between them.

2.5 Representation Theorem for Partially Ordered Sets (1.13)

For a partially ordered set $(X, \prec)$, there is a set $Y \subseteq \mathcal{P}(X)$ which is such that $(X, \preceq)$ is order isomorphic to $(Y, \subseteq)$.

Proof. For some x in X, we set $X^x = \{x' \in X : x' \preceq x\}$, and define φ from X to X^x by $\varphi(x) = X^x$. For x and y in X, as X^x contains x and X^y contains y, $x \neq y$ implies that $X^x \neq X^y$ by the Axiom of Extensionality so φ is injective. We have that φ is trivially surjective and:

$$x \preceq y \Longleftrightarrow X^x \subseteq X^y,$$

by our definition. Thus, φ is an order isomorphism. $\square$

2.6 Total Orderings (1.14)

A relation $\prec$ on a set X is a (strict) total ordering if it is a connected (strict) partial ordering.

Similarly, we say that a relation $\preceq$ on a set X is a non-strict total ordering if it is a connected non-strict partial ordering.

2.7 Well-orderings (1.15)

A relation $\prec$ on a set X is a well-ordering if it is a well-founded total ordering.

2.8 Ordered Pairs (1.17)

For x and y sets, the ordered pair of x and y is the set:

$$\langle x, y\rangle = \{\{x\}, \{x, y\}\}.$$

2.8.1 Uniqueness of Ordered Pairs (1.18)

For x, y, u, and v sets, we have that:

$$\langle x, y\rangle = \langle u, v\rangle \iff (x = u) \text{ and } (y = v).$$

Proof. ($\Longrightarrow$) If $x = y$ then $\langle x, y\rangle = \{\{x\}, \{x, x\}\} = \{\{x\}\}$ so $\langle u, v\rangle = \{\{u\}\}$. Hence $u = v$ and by the Axiom of Extensionality, we have that $x = u$ and so $y = x = u = v$.

If $x \neq y$, then $\langle x, y\rangle$ and $\langle u, v\rangle$ both have the two identical elements so $u \neq v$. We cannot have $\{x\} = \{u, v\}$ so $\{x\} = \{u\}$ which means $x = u$ by the Axiom of Extensionality. Thus, $\{u, v\} = \{x, y\} = \{u, y\}$ so $y = v$.

($\Longleftarrow$) The former holds trivially. $\square$

2.8.2 The Ordered k-tuple (1.20)

We define the k-tuple inductively. The 2-tuple is already defined in (2.8). For $k > 2$, we define the k-tuple as:

$$\langle x_1, x_2, \ldots, x_k\rangle = \langle\langle x_1, x_2, \ldots, x_{k-1}\rangle, x_k\rangle.$$

2.9 Cartesian Products (1.21)

For x and y sets, we define:

$$x \times y = \{\langle a, b\rangle : a \in x, b \in y\}.$$

For $x_1, x_2, \ldots, x_k$ sets, we define:

$$x_1 \times x_2 \times \cdots \times x_k = (x_1 \times x_2 \times \cdots \times x_{k-1}) \times x_k.$$

2.9.1 Indexed Cartesian Products (1.28)

For a set I with each i in I corresponding to a non-empty set A_i:

$$A = \bigcup\{A_i : i \in I\},$$
$$\prod_{i \in I} A_i = \{f \in {}^I A : \forall\, i \in I, f(i) \in A_i\}.$$

2.10 Binary Relations (1.22)

A binary relation R is a class of ordered pairs. We write $R^{-1} = \{\langle y, x\rangle : \langle x, y\rangle \in R\}$.

2.10.1 Domain and Range of Relations (1.24)

For a relation R, we define:

$$\begin{aligned} \mathrm{dom}(R) &= \{x : \exists\, y \text{ where } \langle x, y\rangle \in R\}, \\ \mathrm{ran}(R) &= \{y : \exists\, x \text{ where } \langle x, y\rangle \in R\}, \\ \mathrm{Field}(R) &= \mathrm{dom}(R) \cup \mathrm{ran}(R). \end{aligned}$$

2.11 Functions (1.25)

A relation F is a function if for all x in $\mathrm{dom}(F)$, there is a unique y in $\mathrm{ran}(F)$ with $\langle x, y\rangle$ in F. We say F is injective if and only if for all x and x':

$$(\langle x, y\rangle \in F \text{ and } \langle x', y\rangle \in F) \implies (x = x').$$

2.11.1 Range and Restriction of Functions (1.26)

For a function F from X to Y:

- $F''A = \{y \in Y : \exists\, x \in A \text{ such that } F(x) = y\}$,
- $F \restriction A = \{\langle x, y\rangle \in F : x \in A\}$.

We can see that $F''A = \mathrm{ran}(F \restriction A)$.

2.11.2 Sets of Functions (1.27)

For x and y sets, we have that ${}^{x}y$ is the set of functions from x to y.

3 Transitive and Inductive Sets

3.1 Transitive Sets (1.30)

A set x is transitive if and only if for all y in x, $y \subseteq x$. This is equivalent to $\bigcup x \subseteq x$.

3.2 The Successor Function (1.32-33)

For a set x, $S(x) = x \cup \{x\}$ is the successor of x. $S(x) = x$ is equivalent to saying x is transitive.

3.3 Transitive Closure (1.34)

For a set x, the transitive closure TC of x, is defined recursively as:

$$\bigcup^0 x = x,$$
$$\bigcup^{n+1} x = \bigcup \left(\bigcup^n x \right),$$

which we can write as:

$$TC(x) = \bigcup \left\{ \bigcup^n x : n \in \mathbb{N} \right\}.$$

The transitive closure of a set is always transitive.

3.3.1 Properties of Transitive Closure (1.35)

For a set x:

1. $x \subseteq TC(x)$,
2. $TC(x)$ is the smallest transitive set containing x,
3. $TC(x) = x$ if and only if x is transitive.

Proof. (1) This follows from $\bigcup^0 = x$.

(2) For a transitive set t with $x \subseteq t$, we have $\bigcup^0 x \subseteq t$ by definition. We proceed by induction taking $k > 0$, we see that:

$$A \subseteq B \text{ with } B \text{ transitive} \Longrightarrow \bigcup A \subseteq B,$$

so we deduce that $\bigcup^k x \subseteq t$. By induction we have that $TC(x) \subseteq t$ as required.

(3) If $TC(x) = x$, x is transitive. If x is transitive, $TC(x) \subseteq x$ by (2) and $x \subseteq TC(x)$ by (1). $\square$

3.4 Von Neumann Numerals

The von Neumann numerals are defined as:

$$\begin{aligned}
0 &= \varnothing, \\
1 &= \{\varnothing\} = \{0\}, \\
2 &= \{\varnothing, \{\varnothing\}\} = \{1, 2\}, \\
&\cdots \\
n+1 &= \{0, 1, \ldots, n\}.
\end{aligned}$$

3.5 Inductive Sets (2.1)

A set X is called inductive if $\varnothing$ is in X and for all x in X, $S(x)$ is in X.

3.6 Natural Numbers (2.2-4)

We say that x is a natural number if for all X:

$$X \text{ is an inductive set} \implies x \in X.$$

We define ω as the class of natural numbers, $\omega = \bigcap\{X : X \text{ is an inductive set}\}$. We have that ω is the smallest inductive set.

Proof. Let z be an inductive set (which exists by the Axiom of Infinity). We can define ω by the Axiom of Subsets:

$$\omega = \{x \in z : \forall Y, Y \text{ is inductive} \implies x \in Y\},$$

so ω is a set. We know that $\varnothing$ is in every inductive set by definition, so $\varnothing$ is in ω. For any x in ω, we know that for any inductive set Y that x is in Y and thus $S(x)$ is also in Y as Y is inductive. Thus, $S(x)$ is also in ω as Y was chosen arbitrarily. Hence, ω is an inductive set and the smallest such set by its definition. □

3.7 Principle of Mathematical Induction (2.5)

For a well-defined property of sets Φ, we have that:

$$\Big[\Phi(0) \text{ and } \forall\, x \in \omega, \Phi(x) \implies \Phi(S(x))\Big] \implies \Big[\forall\, x \in \omega, \Phi(x)\Big].$$

Proof. We take $Y = \{x \in \omega : \Phi(x)\}$, it suffices to show that Y is inductive as then $\omega \subseteq Y \subseteq \omega$ implying $\omega = Y$. As we assume $\Phi(0)$, we know that 0 is in Y. Then, by our assumption, Y is closed under the successor function. Thus, Y is inductive as required. □

3.8 Representation of Natural Numbers (2.6)

Every natural number is either 0 or $S(x)$ for some natural number x.

Proof. We take $Z = \{y \in \omega : y = 0 \text{ or } \exists x \in \omega \text{ such that } S(x) = y\}$, it suffices to show that Z is inductive as then $\omega \subseteq Z \subseteq \omega$ implying $\omega = Z$. Clearly, 0 is in Z. Taking z in Z, z must be in ω so $S(z)$ is also in ω as it is inductive. Thus, $S(z)$ is in Z, so Z is inductive as required. □

3.9 Transitivity of ω (2.7)

We have that ω is transitive.

Proof. We take $X = \{n \in \omega : n \subseteq \omega\}$, if $X = \omega$ then by definition ω is transitive so it suffices to show that X is inductive. Clearly, 0 is in X. For n in X, $\{n\} \subseteq \omega$ and $n \subseteq \omega$. Thus, $n \cup \{n\} \subseteq \omega$ so $S(n) \in X$ which means X is inductive, as required. □

3.10 Ordering on the Naturals (2.10-11)

For m and n in ω, we define:

$$m < n \iff m \in n,$$
$$m \leq n \iff m = n \text{ or } m \in n.$$

By definition, $n < S(n)$. We have that:

1. this ordering is transitive,
2. for all n in ω and for all m we have that $m < n$ if and only if $S(m) < S(n)$,
3. for all n in ω, $n \not< n$.

Proof. (1) This follows from the transitivity of set inclusion.

(2) ($\Longrightarrow$) We take $\Phi(k) = [(m < k) \Longrightarrow (S(m) < S(k))]$ and see that $\Phi(0)$ holds. We suppose $\Phi(k)$ holds for some k in ω. For $m < S(k)$, m is in $k \cup \{k\}$. If m is in k then by $\Phi(k)$ we have that $S(m) < S(k) < S(S(k))$. If $m = k$ then $S(m) = S(k) < S(S(k))$.

($\Longleftarrow$) We have that m is in $S(m) = m \cup \{m\}$ which is in $S(n) = n \cup \{n\}$. If $S(m) = n$, then m is in n so $m < n$. If $S(m)$ is in n then m is in n as n is transitive.

(3) We know that $0 \not< 0$ as $0 \notin 0$. For k in ω, $k \notin k$ then $S(k) \notin S(k)$ by (2). We have the result by induction. □

3.11 Total Ordering on the Naturals (2.12)

We have that $<$ is a (strict) total ordering on the naturals.

3.12 Well-ordering Theorem for ω (2.13)

For $X \subseteq \omega$, either $X = \varnothing$ or there is some n_0 in X such that for any m in X either $n_0 = m$ or $n_0 < m$.

Proof. If we suppose X has no least element and take $Z = \{k \in \omega : \forall\, n < k, n \notin X\}$. We want to show Z is inductive, meaning $Z = \omega$ and thus $X = \varnothing$. Vacuously, 0 is in Z. If we have k in Z, we take $n < S(k) = k \cup \{k\}$ and consider:

- if n is in k then n is not in X as $n < k \in Z$,
- if $n = k$ then n is not in X because if n was in X then it would be the least element of X, a contradiction.

Thus, $S(k)$ is in Z so Z is inductive, as required. □

3.13 Recursion Theorem on ω (2.14)

For any set A with a in A and f from A to A any function. There exists a unique function h from ω to A such that for any n in ω:

$$h(0) = a,$$
$$h(S(n)) = f(h(n)).$$

Proof. We will find h as a union of 'k-approximations' to h where we define a k-approximation u as a function with the following properties:

- $\mathrm{dom}(u) = k$,
- if $k > 0$ then $u(0) = a$,
- if $k > S(n)$ then $u(S(n)) = f(u(n))$.

We see that $\{\langle 0, a\rangle\}$ is a 1-approximation, if u is a k-approximation and $l \leq k$ then $u \restriction l$ is an l-approximation, and if $u(k-1) = c$ for some c, then $u' = u \cup \{\langle k, f(c)\rangle\}$ is a $(k+1)$-approximation.

Agreement on Domain If u is a k-approximation and v is a k'-approximation for some $k \leq k'$ then $v \restriction k = u$ (hence $u \subseteq v$).

Proof. We appeal to the contrary with $0 \leq m < k$ being the least natural such that $u(m) \neq v(m)$. We know that $m \neq 0$ as $u(0) = a = v(0)$. So, $m = S(m')$ for some m'. As m is chosen minimally, $u(m') = v(m')$. We can then see that $u(m) = f(u(m')) = f(v(m')) = v(m)$, a contradiction. $\square$

Uniqueness If h exists, it is unique.

Proof. Suppose h and h' are two different functions with domain ω satisfying the theorem. We take $0 \leq m < \omega$ to be the least natural such that $h(m) \neq h'(m)$ and apply the argument from the **Agreement on Domain** case. $\square$

Existence We take B to be the collection of u such that u is in B if and only if there exists k in ω such that u is a k-approximation. For any u and v in B either $u \subseteq v$ or vice-versa by our previous results. We take $h = \bigcup B$. We have that h is a function:

Proof. We appeal to the contrary, if $\langle n, c\rangle$ and $\langle n, d\rangle$ are in h with $c \neq d$, then we have u and v in B with $u(n) = c$ and $v(n) = d$ but this a contradiction by **Agreement on Domain**. $\square$

Domain We have that $\mathrm{dom}(h) = \omega$:

Proof. We appeal to the contrary and suppose $\varnothing \neq X = \{n \in \omega : n \notin \mathrm{dom}(h)\}$. By the definition of h this means that:

$$X = \{n \in \omega : \text{There's no } u\text{-approximation with } n \in \mathrm{dom}(u)\}.$$

We saw that there is a 1-approximation, so 0 is not the least element of X. We suppose $n_0 = S(m)$ is the least element of X. As m is not in X, there must be an n_0-approximation n with $n(m) = c$ for some c. But, we saw that we can extend k-approximations, so we can generate a $(n_0 + 1)$-approximation which is a contradiction. Thus, $X = \varnothing$. $\square$

Thus, we have that h exists and is a unique function as required. $\square$

3.14 Arithmetic (2.17)

For n and k in ω, we define the following arithmetic functions:

$$\begin{aligned} A_n(0) &= n, & A_n(S(k)) &= S(A_n(k)), \\ M_n(0) &= 0, & M_n(S(k)) &= M_n(k) + n, \\ E_n(0) &= 1, & E_n(S(k)) &= E_n(k) \cdot n. \end{aligned}$$

We have that addition is associative and commutative, multiplication is associative, distributive over addition, and commutative, and for m, n, and p in ω:

$$m^{n+p} = m^n \cdot m^p \text{ and } m^{np} = m^{n \cdot p}.$$

4 Well-orderings and Ordinals

4.1 The Principle of Transfinite Induction (3.3)

For a well-ordering $\langle X, \prec \rangle$, we have that:

$$\big[\forall\, x \in X, (\forall\, y \prec x, \Phi(y)) \implies \Phi(x)\big] \implies \big[\forall\, x \in X, \Phi(x)\big].$$

Proof. We appeal to the contrary and suppose that $\varnothing \neq Z = \{x \in X : \neg\Phi(x)\}$. As $\langle Z, \prec \rangle$, there is $\prec$-least element z_0. But then for all $x \prec z_0$, $\Phi(x)$ holds so $\Phi(z_0)$ holds, a contradiction. □

4.2 Initial Segments (3.4)

For a well-ordering $\langle X, \prec \rangle$, the $\prec$-initial segment of some element z in X is the set of predecessors of z, denoted by X_z. We note that X_z does not contain z.

4.3 Order Preserving Maps on Well-orderings (3.5)

For a well-ordering $\langle X, \prec \rangle$ with a function f from $\langle X, \prec \rangle$ to itself an order preserving map, we have that for all x in X, $x \preceq f(x)$.

Proof. We appeal to the contrary, that for some x in X, we have $f(x) \prec x$. As $\langle X, \prec \rangle$ is a well-ordering, there's a $\prec$-least x_0 in X with the property that $f(x_0) \prec x_0$. But, $f(f(x_0)) \prec f(x_0)$ as f is order preserving. Thus, a contradiction to the minimality of x_0. □

4.3.1 Uniqueness of Order Isomorphisms (3.6-7)

For well-orderings $\langle X, \prec_x \rangle$ and $\langle Y, \prec_y \rangle$ with an order isomorphism f from $\langle X, \prec_x \rangle$ to $\langle Y, \prec_y \rangle$. We have that f is unique.

Proof. If we suppose we have two such isomorphisms f and g, we have that $(f^{-1} \circ g)$ is also an order isomorphism. Taking x arbitrary in X by (4.3):

$$\begin{aligned} x \preceq_x (f^{-1} \circ g)(x) &\implies f(x) \preceq_y f(f^{-1} \circ g)(x) \\ &\implies f(x) \preceq_y g(x). \end{aligned}$$

By applying this argument with f and g swapped, we can also see that $g(x) \preceq_y f(x)$. Thus, $f(x) = g(x)$.

In particular, if $\langle X, \prec_x \rangle = \langle Y, \prec_y \rangle$ then this isomorphism is the identity map. □

4.3.2 Non-existence of Order Isomorphisms to Segments (3.8)

A well-ordered set is not order isomorphic to any segment of itself.

Proof. We appeal to the contrary and suppose there is such an order isomorphism on a well-ordering $\langle X, \prec\rangle$ to $\langle X_z, \prec\rangle$ for some z in X. But, we have that $x \preceq f(x)$ for any x in X by (4.3) and $f(z) \prec z$ as $f(z)$ is in X_z. Thus, we have that $z \preceq f(z) \prec z$, a contradiction. $\square$

4.3.3 Order Isomorphism to Set of Segments (3.9)

A well-ordered set $\langle X, \prec\rangle$ is order isomorphic to the set of its initial segments ordered by $\subset$.

Proof. We take $Y = \{X_a : a \in X\}$ and a function φ defined by $\varphi(a) = X_a$. For a and b in X:

$$\begin{aligned}\varphi(a) = \varphi(b) &\iff X_a = X_b \\ &\iff \{x \in X : x \prec a\} = \{x \in X : x \prec b\} \\ &\iff a = b,\end{aligned}$$

so we have that φ is injective and trivially surjective onto the set of initial segments of X. As $a \prec b \iff X_a \subset X_b$, the mapping is order preserving. $\square$

4.4 Ordinal Numbers (3.10-11)

We say that $\langle X, \in\rangle$ is an ordinal if and only if X is transitive and where $\langle X, \in\rangle$ is a well-ordering. We have that $\langle \omega, \in\rangle$ is an ordinal.

4.4.1 Segment and Element Equality (3.12)

For an ordinal $\langle X, \in\rangle$, every element z in X is identical to X_z. So, for any elements a, b of an ordinal:

$$a \in b \iff a \subset b \iff X_a \subset X_b.$$

Proof. We know that X is transitive and $\in$ well-orders X, we take z in X and see that:

$$\begin{aligned}w \in X_z &\iff w \in X \text{ and } w \in z \\ &\iff w \in z, &&\text{(as } z \subseteq X\text{)}\end{aligned}$$

thus, $X_z = z$ by the Axiom of Extensionality. $\square$

4.4.2 Ordinal Initial Segments (3.13)

For an ordinal $\langle X, \in \rangle$, any $\in$-initial segment of X is an ordinal.

Proof. We take some u in X and w in X_u. As $\in$ well-orders X, it well-orders any subset of X so $\langle X_u, \in \rangle$ is a well-ordering. We have that:

$$t \in w \in u \Longrightarrow t \in u = X_u,$$

thus X_u is transitive as required. □

4.4.3 Proper Subset Segments (3.14)

For an ordinal $\langle X, \in \rangle$ with $Y \subset X$, if $\langle Y, \in \rangle$ is also an ordinal, then Y is an $\in$-initial segment of X.

Proof. For a in Y, $Y_a = a$ as Y is an ordinal. As $Y \subset X$, a is in X so $X_a = a$. Thus, $X_a = Y_a$. As $Y \neq X$, we consider $c = \inf(\{z \in X : z \notin Y\})$ which exists as the set is non-empty and $\langle X, \in \rangle$ is a well-ordering. Hence, $Y = X_c$. □

4.4.4 The Intersection of Ordinals (3.15)

For ordinals X and Y, $(X \cap Y)$ is also an ordinal.

Proof. We know that $(X \cap Y)$ is transitive as X and Y are transitive. Any subset of X is a well-ordering under $\in$, in particular $(X \cap Y)$ is well-ordered by $\in$. □

4.5 Classification Theorem for Ordinals (3.16)

For ordinals X and Y, either $X = Y$ or one is an initial segment of the other (or equivalently a member).

Proof. We suppose that $X \neq Y$. We know that $(X \cap Y)$ is an ordinal by (4.4.4), so have two cases. If $X = (X \cap Y)$ or $Y = (X \cap Y)$, one must be an initial segment of the other by (4.4.2). If $(X \cap Y)$ is a proper subset of X and Y, it is an initial segment of X and Y simultaneously by (4.4.2). We set $(X \cap Y) = X_a = Y_b$ for some a in X and b in Y. But, we know that by (4.4.1), $a = X_a = Y_b = b$. However, this means $a = b \in (X \cap Y) = X_a$, but a is not in X_a, a contradiction. □

4.6 Equality under Isomorphisms (3.17)

For ordinals X and Y, if X is order isomorphic to Y then $X = Y$.

Proof. Suppose $X \neq Y$, then without loss of generality we take X to be an initial segment of Y. But, this would mean Y is order isomorphic to an initial segment of itself which is a contradiction by (4.3.2). □

4.7 Bound on Isomorphisms (3.18)

A well-ordering is order isomorphic to at most one ordinal.

Proof. If a well-ordering is isomorphic to more than one ordinal, then these ordinals are isomorphic to each other and thus, equal by (4.6). □

4.8 Criterion for Ordinals (3.19)

If every initial segment of a well-ordered set $\langle A, \prec \rangle$ is order isomorphic to some ordinal, $\langle A, \prec \rangle$ itself is order isomorphic to an ordinal.

Proof. Each initial segment must be order isomorphic to at most one ordinal (thus exactly one) by (4.7). We define a function F that assigns elements of A to unique ordinals such that $\langle F(b), \in \rangle \cong \langle A_b, \prec \rangle$. We take $Z = \mathrm{ran}(F)$ by the Axiom of Replacement and g_b to be the isomorphism from A_b to $F(b)$ noting that the isomorphism is unique by (4.3.1). If c and b are in A with $c \prec b$ then $A_c = (A_b)_c$ implying that $F(c) \neq F(b)$ by (4.3.2). Thus, F is injective and so bijective between A and Z. Continuing with $c \prec b$, we see that $(g_b \upharpoonright A_c)$ is an isomorphism from A_c to $(F(b))_{g_b(c)}$ and by (4.7), $(g_b \upharpoonright A_c) = g_c$ and $F(c) = (F(b))_{g_b(c)}$. Thus, $F(c)$ is in $F(b)$.

We know that Z is well-ordered by $\in$ as A is well-ordered by $\prec$ and F is an order isomorphism. So, for u in $F(b)$, as g_b is surjective, $u = g_b(c)$ for some $c \prec b$. As such, $u = F(b)_u = F(b)_{g_b(c)} = F(c)$ so u is in Z. Thus, Z is transitive so, Z is an ordinal. □

4.9 Representation Theorem for Well-orderings (3.20)

Every well-ordering is order isomorphic to exactly one ordinal.

Proof. We take $Z = \{v \in X : X_v \text{ is not isomorphic to an ordinal}\}$, and want to show it's empty as this will suffice by (4.3.1 and 4.8). We suppose the contrary, we take v_0 to be the $\prec$-least element of Z. We have that $\langle X_{v_0}, \prec \rangle$ is a well-ordering with $(X_{v_0})_w = X_w$ for each w in X_{v_0}. But, for each w in X_{v_0}, X_w is isomorphic to some ordinal by the minimality of v_0, as such X_{v_0} must be isomorphic to an ordinal by (4.8), a contradiction. Thus, Z is empty, as required. □

4.10 Order Type of Well-orderings (3.21)

For a well-ordering $\langle X, \prec \rangle$, the order type of $\langle X, \prec \rangle$ is the unique ordinal isomorphic to $\langle X, \prec \rangle$, written as $\mathrm{ot}(\langle X, \prec \rangle)$.

4.11 Classification Theorem for Well-orderings (3.22)

For well-orderings $\langle A, \prec_A \rangle$ and $\langle B, \prec_B \rangle$ we have that exactly one of the following holds:

- $\langle A, \prec_A \rangle \cong \langle B, \prec_B \rangle$,
- there exists b in B such that $\langle A, \prec_A \rangle \cong \langle B_b, \prec_B \rangle$,
- there exists a in A such that $\langle A_a, \prec_A \rangle \cong \langle B, \prec_B \rangle$.

Proof. We take $\langle X, \in \rangle$ and $\langle Y, \in \rangle$ to be the unique ordinals isomorphic to $\langle A, \prec_A \rangle$ and $\langle B, \prec_B \rangle$ respectively via the maps:

$$f : \langle X, \in \rangle \to \langle A, \prec_A \rangle,$$
$$g : \langle Y, \in \rangle \to \langle B, \prec_B \rangle.$$

We know that either these ordinals are order isomorphic or order isomorphic to an initial segment of the other. If the former is true, then we have that our well-orderings are isomorphic via f and g and their inverses. If the latter is true, we know that (without loss of generality) $\langle X, \in \rangle \cong \langle Y_y, \in \rangle$ for some y in Y. Thus:

$$f(\langle X, \in \rangle) \cong g(\langle Y_y, \in \rangle) \implies \langle A, \prec_A \rangle \cong \langle B_{g(y)}, \prec_B \rangle,$$

as required. $\square$

5 Ordinal Applications

We collate the properties of ordinals covered so far for some ordinals α, β, and γ:

- ordinals are transitive and well-ordered by $\in$ by definition,
- for x in α, x is an ordinal with $x = \alpha_x$,
- $\alpha \cong \beta$ implies that $\alpha = \beta$,
- we have exactly one of the following $\alpha = \beta$, α in β, or β in α.

5.1 Principle of Transfinite Induction on Ordinals (3.24)

For a well-defined property of ordinals Φ, we have that for all ordinals α:

$$\big[\forall\, \beta < \alpha, \Phi(\beta) \implies \Phi(\alpha)\big] \implies \Phi(\alpha). \tag{$*$}$$

Hence, the class of ordinals is well-ordered.

Proof. We take $C = \{\alpha \in \mathrm{On} : \neg\Phi(\alpha)\}$ and α_0 in C. If α_0 is not the least element of C, we have that $\varnothing \neq (\alpha_0 \cap C) \subseteq \alpha_0$ has an $\in$-least element α_1 as α_0 is an ordinal, which is well-ordered by $\in$. Thus, α_1 is the $\in$ least element of C. As we have a least element γ of C, we see that for all β in C with $\beta < \gamma$, we have $\Phi(\beta)$. But, our assumption implies that we have $\Phi(\gamma)$, a contradiction. Thus, $C = \varnothing$ as required. $\square$

5.2 The Class of Ordinals (3.25)

The class of ordinals is a proper class.

Proof. Suppose the class of ordinals is a set z. We have that $\langle z, \in\rangle$ is transitive and well-ordered by (5.1). Thus, z is an ordinal, as such z is in z. But, this contradicts the strict ordering of $\in$. $\square$

5.3 Sum of Orderings (3.26)

For strict total orderings $\langle A, R\rangle$ and $\langle B, S\rangle$ with $A \cap B$ empty, we define the sum ordering $\langle C, T\rangle$ as:

$$C = A \cup B,$$
$$xTy \iff \begin{cases} xRy & \text{for } x \text{ and } y \in A \\ xSy & \text{for } x \text{ and } y \in B \\ x \in A \text{ and } y \in B & \text{otherwise.} \end{cases}$$

We can avoid the disjoint constraint by taking the sum of $\langle A \times \{0\}, R\rangle$ and $\langle B \times \{1\}, S\rangle$. We name this operation $+'$ so for ordinals α and β:

$$\alpha +' \beta = \langle \mathrm{ot}((\alpha \times \{0\}) \cup (\beta \times \{1\})), T\rangle,$$
$$\langle \gamma, i\rangle T \langle \delta, j\rangle \iff (i = j \text{ and } \gamma < \delta) \text{ or } (i < j).$$

5.4 Product of Orderings (3.28)

For strict total orderings $\langle A, R\rangle$ and $\langle B, S\rangle$, we define the product of these orderings $\langle A, R\rangle \times \langle B, S\rangle$ to be the ordering $\langle C, U\rangle$:

$$C = A \times B$$
$$\langle x, y\rangle U \langle x', y'\rangle \iff (ySy') \text{ or } (y = y' \text{ and } xRx'),$$

defining an operation for ordinals, denoted by $\cdot'$.

5.5 Supremum of Ordinals (3.30, 3.32)

For a set of ordinals A, $\sup(A)$ is the least ordinal γ such that for all δ in A, $\delta \leq \gamma$. We also have the strict supremum $\sup^+(A)$ as the least ordinal γ^+ such that for all δ in A, $\delta < \gamma^+$. We have that $\sup(A) = \bigcup A$.

Proof. We know the supremum is well-defined as if we suppose there isn't an ordinal which is an upper bound for A, there's some δ in A such that $\delta > \gamma$ for each ordinal γ. However, this means $\bigcup A$ must be equal to On, which is a contradiction as $\bigcup A$ is a set by the Axiom of Unions.

We take $S = \sup(A)$ and u in $\bigcup A$, we know that there must be some a in A, such that $u < a < S$. Thus, u is in S as S is transitive, hence $\bigcup A \subseteq S$. Conversely, for s in S, $s < S$ so there is some a in A with $s < a \leq S$. Thus, s is in $\bigcup A$, so $S \subseteq \bigcup A$. Thus $S = \bigcup A$. □

5.6 Types of Ordinals (3.33)

We can consider three types of ordinals:

- the zero ordinal,
- successor ordinals, ordinals with immediate predecessors,
- limit ordinals, ordinals that are not of the other types.

5.7 Recursion Theorem on Ordinals (3.35)

For a function F from V to V, there exists a unique function H from the class of ordinals to V such that for all α:

$$H(\alpha) = F(H \restriction \alpha).$$

Proof. We define a function u to be a δ-approximation if $\mathrm{dom}(u) = \delta$ and for all $\alpha < \delta$, $u(\alpha) = F(u \restriction \alpha)$. For a δ-approximation u and $\delta > 0$, we see that $u(0) = F(u \restriction 0) = F(\varnothing)$ so a 1-approximation is equal to $\{\langle 0, F(\varnothing)\rangle\}$ with domain $\{0\} = 1$. Additionally, for some $\gamma < \delta$, $u \restriction \gamma$ is a γ-approximation. Furthermore, $u \cup \{\langle \delta, F(u)\rangle\}$ is a $(\delta + 1)$-approximation.

Agreement on Domain For a δ-approximation u and any γ-approximation v with $\delta < \gamma$, $u = v \restriction \delta$.

Proof. We appeal to the contrary and take τ be the least ordinal such that $u(\tau) \neq \gamma(\tau)$. Thus, $(u \restriction \tau) = (v \restriction \tau)$ but then:

$$u(\tau) = F(u \restriction \tau) = F(v \restriction \tau) = v(\tau),$$

which is a contradiction. □

Uniqueness If such H exists, it is unique.

Proof. We appeal to the contrary, taking H' to be some differing derivation of H. We consider the least τ such that $H(\tau) \neq H'(\tau)$ and apply the same argument as the **Agreement on Domain** case. □

Limits For some limit ordinal λ, if for all $\alpha < \lambda$ we have that u_α is an α-approximation, $\bigcup_{\alpha<\lambda} u_\alpha$ is a λ-approximation.

Proof. This union is of an increasing sequence of sets so:

$$\alpha < \beta < \lambda \Longrightarrow u_\alpha \subseteq u_\beta.$$

As each element is a function, and the functions agree on domain, the union is also a function and has domain λ. Thus, this union is a λ-approximation. □

Existence We define $H = \bigcup B$ which is a function with $\mathrm{dom}(H)$ being the set of ordinals.

Proof. We know that H is a function by the **Agreement on Domain** case. We take $C = \{\delta$: There's no δ-approximation$\}$ and suppose C is non-empty. By the Principle of Transfinite Induction on Ordinals, C has a least elememt ψ. We know that $\psi > 1$ as we defined a 1-approximation and by **Limits** it cannot be a limit ordinal. If $\psi = \mu + 1$ then there's a μ-approximation v by the minimality of ψ. However, we can extend v to a ψ-approximation u by setting $u(\mu) = F(v)$. This is a contradiction. □

Thus, we have that H exists and is a unique function as required. □

5.8 Recursion Theorem on Ordinals, Second Form (3.38)

For a in V, and functions F_0 and F_1 from V to V, there's a unique function H from the class of ordinals to V such that for an ordinal α and a limit ordinal λ:

$$\begin{aligned} H(0) &= a, \\ H(\alpha + 1) &= F_0(H(\alpha)), \\ H(\lambda) &= F_1(H \restriction \lambda). \end{aligned}$$

Proof. We define a function F from V to V by:

$$F(u) = \begin{cases} a & \text{for } u = \varnothing \\ F_0(u) & \text{if } u \text{ is a function with a successor domain} \\ F_1(u) & \text{if } u \text{ is a function with a limit domain} \\ \varnothing & \text{otherwise,} \end{cases}$$

and apply (5.7). □

5.9 Ordinal Addition (3.39)

We define ordinal addition A_α for some ordinals α and β, and a limit ordinal λ as:

$$\begin{aligned} A_\alpha(0) &= \alpha, \\ A_\alpha(\beta + 1) &= S(A_\alpha(\beta)), \\ A_\alpha(\lambda) &= \sup(\{A_\alpha(x) : x < \lambda\}). \end{aligned}$$

5.10 Ordinal Multiplication (3.39)

We define ordinal multiplication M_α for some ordinals α and β, and a limit ordinal λ as:

$$\begin{aligned} M_\alpha(0) &= 0, \\ M_\alpha(\beta+1) &= M_\alpha(\beta) + \alpha, \\ M_\alpha(\lambda) &= \sup(\{M_\alpha(x) : x < \lambda\}). \end{aligned}$$

5.11 Ordinal Exponentiation (3.39)

We define ordinal exponentiation A_α for some ordinals α and β, and a limit ordinal λ as:

$$\begin{aligned} E_\alpha(0) &= 1, \\ E_\alpha(\beta+1) &= E_\alpha(\beta) \cdot \alpha, \\ E_\alpha(\lambda) &= \sup(\{E_\alpha(x) : x < \lambda\}). \end{aligned}$$

5.12 Monotonicity of Ordinal Arithmetic (3.40-41)

For ordinals α, β, and γ with $\beta > 0$ and $\gamma > 1$, the functions A_α, M_β, and E_γ are strictly increasing and thus injective.

Proof. We take β, γ and δ to be ordinals and we proceed by induction, supposing that:

$$[\beta < \gamma] \implies [A_\alpha(\beta) < A_\alpha(\gamma)], \tag{$*$}$$

for all $\gamma \leq \delta$. The base case is trivial. For $\beta < \delta + 1$, if $\beta = \delta$, then:

$$A_\alpha(\delta) < S(A_\alpha(\delta)).$$

Otherwise, $\beta < \delta$ so by our hypothesis:

$$A_\alpha(\beta) < A_\alpha(\delta) < S(A_\alpha(\delta)) = A_\alpha(\delta + 1).$$

Now, we suppose $(*)$ holds for all $\gamma < \lambda$ for some limit ordinal λ. For $\beta < \lambda$, clearly $\beta < \beta + 1 < \lambda$ as λ has no immediate predecessor. By the hypothesis:

$$A_\alpha(\beta) < A_\alpha(\beta+1) \leq \sup(\{A_\alpha(\gamma) : \gamma < \lambda\}) = A_\alpha(\lambda),$$

as required. The arguments for M_α and E_α are similar. □

5.13 Remainders (3.43)

For α and β ordinals with $0 < \alpha \leq \beta$, there's a unique:

1. ordinal γ such that $\alpha + \gamma = \beta$,
2. pair of ordinals ζ and κ such that $\alpha \cdot \zeta + \kappa = \beta$ and $\kappa < \alpha$.

Proof. (1) As A_α is strictly increasing, we consider $Z = \{x : \alpha + x \geq \beta\}$ which must be non-empty as A_α is strictly increasing. We take $\gamma = \min(Z)$ and see that $\alpha + \gamma = \beta$ since if $\alpha + \gamma > \beta$ either:

- $\gamma = \delta + 1$ so $\alpha + \delta < \beta$ as δ is not in Z. But then, $\alpha + \gamma = \alpha + (\delta + 1) \leq \beta$, a contradiction,
- γ is a limit ordinal, $\alpha + \gamma = \sup(\{\alpha + \delta : \delta < \gamma\})$. But, as $\alpha + \gamma > \beta$ there's some $\delta < \gamma$ such that $\alpha + \delta \geq \beta$. This contradicts the minimality of γ.

(2) As M_α is strictly increasing, we choose the least ζ such that $\alpha \cdot \zeta \leq \beta < \alpha \cdot (\zeta + 1)$. We apply (1) to find some κ such that $\alpha \cdot \zeta + \kappa = \beta$. For some ζ' and κ' also satisfying (2), if $\zeta = \zeta'$ then by the uniqueness of (1), $\kappa = \kappa'$. We suppose $\zeta < \zeta'$ so $\zeta + 1 \leq \zeta'$:

$$\begin{aligned} \beta = \alpha \cdot \zeta + \kappa &< \alpha \cdot \zeta + \alpha \\ &= \alpha \cdot (\zeta + 1) \\ &\leq \alpha \cdot \zeta' \\ &\leq \alpha \cdot \zeta' + \kappa' \\ &= \beta, \end{aligned}$$

which is a contradiction. Hence, $\zeta = \zeta'$. $\square$

5.14 Ordinal Arithmetic (3.44)

We have that ordinal addition is associative, ordinal multiplication is distributive over addition and associative, and for ordinals α, β, and γ:

$$\alpha^{\beta + \gamma} = \alpha^\beta \cdot \alpha^\gamma.$$

6 Cardinality

6.1 Equinumerosity (4.1-2)

We say that two sets, A and B, are equinumerous if there is a bijection between them, written as $A \approx B$. We have that $\approx$ is an equivalence relation with equivalence classes the collections of all equinumerous sets of equal cardinality.

6.2 Finite Sets (4.3)

A set is finite if it is equinumerous with a natural number. Sets that are not finite are infinite.

6.3 Pidgeon-hole Principle (4.4-8)

No natural number is equinumerous to a proper subset of itself and thus:

- no finite set is equinumerous to a proper subset of itself,
- any set equinumerous to a proper subset of itself is infinite,
- any finite set is equinumerous to a unique natural number,
- ω is infinite.

Proof. We take Z to be the subset of ω such that for all z in Z, all injective functions f from z to z have $\mathrm{ran}(f) = z$. Trivially, Z contains 0. For n in Z, we consider an injective function f from $(n+1)$ to $(n+1)$.

Case 1 If $(f \restriction n)$ has domain and range n, by our inductive hypothesis, we have that $\mathrm{ran}(f \restriction n) = n$. Thus, $\mathrm{ran}(f) = n + 1$.

Case 2 If $f(m) = n$ for some $m < n$, as f is injective, we have that for some $k < n$, $f(n) = k$. We define g identically to f except $g(m) = k$ and $g(n) = n$ so that g is an injective function from $(n+1)$ to $(n+1)$. Thus, **Case 1** applies to g so $\mathrm{ran}(g) = n + 1 = \mathrm{ran}(f)$. □

6.4 Cantor's Diagonal Argument (4.9)

The natural numbers are not equinumerous with the real numbers.

Proof. We appeal to the contrary and suppose we have some injective map f from ω to $\mathbb{R}$. We can generate some x in $\mathbb{R}$ that is not in $\operatorname{ran}(f)$ by setting the i^{th} decimal place of x to the i^{th} decimal place of $f(i)$ mapped by:

$$k \mapsto \begin{cases} 1 & k \text{ even} \\ 2 & k \text{ odd.} \end{cases}$$

Thus, x would differ from every element of $\operatorname{ran}(f)$, a contradiction. □

6.5 Cantor's Theorem (4.10)

No set is equinumerous to its powerset.

Proof. We appeal to the contrary and suppose f from X to $\mathcal{P}(X)$ is a bijection for some set X. We set $Z = \{u \in X : u \notin f(u)\}$ and see that $Z \subseteq X$ so Z is in $\mathcal{P}(X)$. As such, $Z = f(u)$ for some u, but:

$$\begin{aligned} u \in Z &\Longrightarrow u \notin f(u) = Z, \\ u \notin Z &\Longrightarrow u \in f(u) = Z, \end{aligned}$$

which is a contradiction. □

6.6 Cantor-Schröder-Bernstein Theorem (4.11-12)

For sets X and Y, $X \preceq Y$ if there's an injection from X to Y and $X \prec Y$ if $X \preceq Y$ and $Y \npreceq X$. We have that $X \preceq Y$ and $Y \preceq X$ is equivalent to $X \approx Y$.

Proof. ($\Longrightarrow$) By our assumptions, there are some f from X to Y and g from Y to X both injective and want to form some bijection h from X to Y. We consider $C_0 = X \setminus \operatorname{ran}(g)$, the values suppressing the surjectivity of g. For n in $\mathbb{N}$, we define:

$$\begin{aligned} D_n &= f''C_n \\ C_{n+1} &= g''D_n = g''(f''C_n), \end{aligned}$$

$$h(v) = \begin{cases} f(v) & \text{if } v \text{ is in } C_n \text{ for some } n \\ g^{-1}(v) & \text{otherwise.} \end{cases}$$

To see that h is injective, we consider u and v in X, as f and g are injective, the only problems arise from u and v invoking differing cases of the definition of h. Without loss of generality, we suppose u in C_n for some n in $\mathbb{N}$ and v is not in any C_k for k in $\mathbb{N}$. Thus, for some m in $\mathbb{N}$:

$$\begin{aligned} h(u) &= f(u) \in D_p, \\ h(v) &= g^{-1}(v). \end{aligned}$$

We know that for all p in $\mathbb{N}$, $g^{-1}(v)$ is not in D_p as otherwise, $g(g^{-1}(v)) = v$ in C_{p+1} which is a contradiction. Thus, $u \neq v$ implies that $h(u) \neq h(v)$ and as such, h is injective. To see that h is surjective, we first note that $U = \bigcup_{m\in\mathbb{N}} D_m \subseteq \operatorname{ran}(h)$. We consider u in $Y \setminus U$, $g(u)$ is not in $C_0 = X \setminus \operatorname{ran}(g)$ and for n in $\mathbb{N}$, $g(u)$ is not in C_{n+1} because u is not in D_n and g is injective so there's no v in D_n such that $g(v) = g(u)$ As such, $h(g(u)) = g^{-1}(g(u)) = u$. So, h is surjective and as such, bijective.

($\Longleftarrow$) This is direct from the properties of bijections. □

6.7 Characteristic Function (4.13)

For a set X, we define the the characteristic function of any $Y \subseteq X$ to be χ_Y from X to 2 defined by:

$$\chi_Y(a) = \begin{cases} 1 & \text{if } a \text{ is in } Y \\ 0 & \text{if } a \text{ is in } X \setminus Y. \end{cases}$$

6.8 Countability (4.14-15)

A set X is countably infinite if $X \approx \omega$ and countable if $X \preceq \omega$. Subsets of countable sets are countable.

6.9 The Union of Countably Infinite Sets (4.16, 4.18)

The union of two countably infinite sets is also countably infinite. The countably infinite union of countably infinite sets is countably infinite.

Proof. The case for countably infinitely many sets follows from the Well-ordering Principle. □

6.10 Countably Infinite Subsets (4.17)

For an infinite set X with $\langle X, R\rangle$ a well-ordering, X has a countably infinite subset.

Proof. We take x_0 to be the R-least element of X and for n in ω, x_{n+1} is the R-least element of $X \setminus \{x_k : k \leq n\}$. Thus, $\{x_k : k < \omega\}$ is a countably infinite subset of X. □

6.11 Cardinality (4.20-21)

For a set X, the cardinality of X, $|X|$ is the least ordinal α such that $X \approx \alpha$. We have that for X and Y sets:

$$X \approx Y \iff |X| = |Y|,$$
$$X \preceq Y \iff |X| \leq |Y|,$$
$$X \prec Y \iff |X| < |Y|.$$

We note that the cardinality operation is a projection onto the ordinals.

6.12 Cardinal Numbers (4.22)

An ordinal α is a cardinal if $\alpha = |\alpha|$.

6.13 Cardinality Capture (4.23)

For α and γ ordinals, if $|\alpha| \leq \gamma < \alpha$ then $|\alpha| = |\gamma|$.

Proof. By our assumptions, there is a bijection f from α to $|\alpha|$, so $|a| = ||a||$. We know that $\gamma \subseteq \alpha$, so $f \upharpoonright \gamma$ is an injection from γ to $|\alpha|$ so $\gamma \preceq |\alpha|$. But, $|\alpha| \preceq \gamma$ by our assumption, so $|\alpha| \approx \gamma$ which implies that $|\gamma| = ||\alpha|| = |\alpha|$. □

6.14 Cardinal Addition and Multiplication (4.24)

For cardinals κ and λ, and sets K and L with cardinality κ and λ respectively, we define:

$$\kappa \oplus \lambda = |K \cup L|, \qquad \text{(for } K \text{ and } L \text{ disjoint)}$$
$$\kappa \otimes \lambda = |K \times L|.$$

We note that these operations are commutative and associative.

6.15 Confluence of Ordinal and Cardinal Arithmetic (4.25)

For ordinals m and n in ω, $m + n = m \oplus n$ and $m \cdot n = m \otimes n$.

Proof. This follows from induction on n. □

6.16 Hessenberg's Theorem (4.26)

For an infinite cardinal κ, there is a bijection from $\kappa \times \kappa$ to κ. Thus, $\kappa \otimes \kappa = \kappa$.

Proof. We already know that $\omega \times \omega \approx \omega$ and so $\omega \otimes \omega = |\omega \times \omega| = \omega$. We proceed by induction on $\kappa \geq \omega$. We assume for all infinite cardinals $\lambda < \kappa$ we have that $\lambda \otimes \lambda = \lambda$. We consider Gödel's ordering:

$$\begin{aligned}\big[\langle \alpha, \beta\rangle \lhd \langle \gamma, \delta\rangle\big] \iff & \big[(\max(\{\alpha, \beta\}) < \max(\{\gamma, \delta\})) \\ & \text{or } \big[(\max(\{\alpha, \beta\}) = \max(\{\gamma, \delta\})) \\ & \text{and } (\alpha < \gamma \text{ or } (\alpha = \gamma \text{ and } \beta < \delta))\big]\big],\end{aligned}$$

and we note that:

$$(\kappa \times \kappa)_{\langle \alpha, \beta \rangle} \subset \gamma \times \gamma,$$

where $\gamma = \max(\{\alpha, \beta\}) + 1$. So, as $\gamma < \kappa$, we have that $|\gamma| < \kappa$ as γ is an ordinal. Thus, $\gamma \otimes \gamma = |\gamma| \otimes |\gamma| = \gamma < \kappa$ by the inductive hypothesis and the fact that α and β must precede κ. As such, all initial segments must have order type preceding κ which means the order type of $\kappa \times \kappa$ is at most κ (Ex. 4.24). However, $\kappa \times \kappa$ must also have order type at least κ as $\langle \alpha, 0\rangle$ is in the initial segment for $\alpha < \kappa$. Thus, $\mathrm{ot}(\kappa \times \kappa, \lhd) = \kappa$. From this, we deduce that $\kappa \times \kappa \approx \kappa$ so $\kappa \otimes \kappa = \kappa$. □

6.17 Confluence of Addition and Multiplication (4.27)

For infinite cardinals κ and λ, $\kappa \oplus \lambda = \kappa \otimes \lambda = \max(\{\kappa, \lambda\})$.

Proof. Without loss of generality, we assume $\lambda \leq \kappa$ so $\max\{\kappa, \lambda\} = \kappa$. For X and Y disjoint with cardinality κ and λ respectively:

$$X \preceq X \cup Y \preceq (X \times \{0\}) \cup (X \times \{1\}) = X \times 2 \preceq X \times X.$$

So, in terms of cardinals we have:

$$\kappa \leq \kappa \oplus \lambda \leq \kappa \oplus \kappa = \kappa \otimes 2 \leq \kappa \otimes \kappa.$$

But, by Hessenberg's Theroem, $\kappa = \kappa \otimes \kappa$ which induces equality on all the above statements. As such, $\kappa = \kappa \oplus \lambda$ and similarly:

$$\kappa \leq \kappa \otimes \lambda \leq \kappa \otimes \kappa = \kappa,$$

we have that $\kappa = \kappa \otimes \lambda$, as required. □

6.18 Cardinality of a Countable Union of Infinite Cardinals (4.28-29)

For a set A, ${}^{<\omega}A = \bigcup_{n\in\omega} {}^{n}A$. For an infinite cardinal κ, $|{}^{<\omega}\kappa| = \kappa$.

6.19 Cardinal Exponentiation (4.30, 4.32)

For cardinals κ and λ, $\kappa^{\lambda} = |{}^{L}K|$ where K and L are sets of cardinality κ and λ respectively. For cardinals κ, λ, and μ, we have that:

$$\kappa^{\lambda\oplus\mu} = \kappa^{\lambda} \otimes \kappa^{\mu},$$
$$(\kappa^{\lambda})^{\mu} = \kappa^{\lambda\otimes\mu}.$$

6.20 Equinumerosity with Characteristic Functions (4.31)

For cardinals κ and λ with $\lambda \geq \omega$ and $2 \leq \kappa \leq \lambda$, then ${}^{\lambda}\lambda \approx {}^{\lambda}\kappa \approx {}^{\lambda}2 \approx \mathcal{P}(\lambda)$.

Proof. We know that ${}^{\lambda}2 \approx \mathcal{P}(\lambda)$ as we can assign characteristic functions to the subsets they identify. Then, using Hessenberg's Theorem:

$${}^{\lambda}2 \preceq {}^{\lambda}\kappa \preceq {}^{\lambda}\lambda \preceq \mathcal{P}(\lambda \times \lambda) \approx \mathcal{P}(\lambda) \approx {}^{\lambda}2,$$

inducing equinumerosity throughout. □

6.21 Class of Cardinals (4.34)

The class of cardinals is a proper class.

Proof. We suppose the class of cardinals is a set, as it's the union of ordinals, it's an ordinal τ. By Cantor's Theorem, $|\mathcal{P}(\tau)| > \tau$ which is a cardinal not in our set of cardinals, a contradiction. □

6.22 Unbounded Ordinals (4.35)

For any set x, there's an ordinal α with $\alpha \npreceq x$.

Proof. We take $\alpha = |\mathcal{P}(x)|$ and we are done by Cantor's Theorem. □

6.23 The $\aleph$ Cardinals (4.36-37)

For some ordinal α and a limit ordinal λ, we have the $\aleph$ cardinals:

$$\begin{aligned}\aleph_0 &= \omega_0 = \omega,\\ \aleph_{\alpha+1} &= \omega_{\alpha+1} = \omega_\alpha^+ = \text{the least ordinal containing } \omega_\alpha\\ \aleph_\lambda &= \omega_\lambda = \sup(\{\omega_\tau : \tau < \lambda\}).\end{aligned}$$

We have a function $F_\aleph$ from the ordinals to the $\aleph$ cardinals defined by:

$$F_\aleph(\alpha) = \omega_\alpha.$$

For an ordinal $\alpha > 1$, $F_\aleph(\alpha)$ is an uncountable cardinal, called a limit or successor cardinal, dependent on whether α is a limit or successor cardinal.

6.24 The $\beth$ Cardinals (4.39)

For some ordinal α and a limit ordinal λ, we have the $\beth$ cardinals:

$$\begin{aligned}\beth &= \omega,\\ \beth_{\alpha+1} &= 2^{\beth_\alpha}\\ \beth_\lambda &= \sup(\{\beth_\tau : \tau < \lambda\}).\end{aligned}$$

If the Generalised Continuum Hypothesis holds, then we have that $\beth_\alpha = \aleph_\alpha$ for all ordinals α.

6.25 The Continuum Hypothesis (4.38)

The hypothesis states that $2^{\omega_0} = \omega_1$ and for the general hypothesis, for all ordinals α, $2^{\omega_\alpha} = \omega_{\alpha+1}$. With our axioms, we can't prove the specific hypothesis is true or false. They are insufficient to this end.

7 The Universe of Sets

7.1 Classes

We have that classes are collection of objects, these could also be sets. Classes that are not sets are called proper classes.

7.2 Russell's Theorem (1.4)

We have that $R = \{x : x \notin x\}$ is a proper class.

Proof. Suppose we have a set z such that $z = R$, we consider the membership of z in R. If we suppose z is in R, by the definition of R, z is not in $z = R$, a contradiction. If we suppose z is not in R, by the definition of R, z is in $z = R$, a contradiction. Thus, z cannot be a set, so R is a proper class. □

7.3 The Universe of Sets (1.5)

We define the universe of sets as $V = \{x : x = x\}$. We have that V is a proper class.

Proof. If we suppose V is a set, we apply the Axiom of Subsets with $\Phi(x) = x \notin x$ and reach a contradiction via (7.2). □

7.4 The Well-founded Hierarchy of Sets (6.1)

For an ordinal α and a limit ordinal λ, we define the function V_α by transfinite recursion:

$$\begin{aligned}
V_0 &= \varnothing, \\
V_{\alpha+1} &= \mathcal{P}(V_\alpha), \\
V_\lambda &= \bigcup_{\alpha<\lambda} V_\alpha, \\
V &= \bigcup_{\alpha\in\mathrm{On}} V_\alpha.
\end{aligned}$$

7.5 Transitivity of V_α (6.2)

For any ordinal α, we have that V_α is transitive and for all $\beta < \alpha$, V_β is in V_α.

Proof. We proceed by induction, for $\alpha = 0$, $V_0 = \varnothing$ which trivially satisfies both statements. For $\alpha = \beta + 1$, we use the fact that if β is transitive, then $\mathcal{P}(\beta)$ is also. By the inductive hypothesis, $V_\alpha = \mathcal{P}(\beta)$ is transitive. As V_β is in $\mathcal{P}(V_\beta)$, we have that V_β is in V_α and if $\beta' < \beta$ then by the inductive hypothesis, $V_{\beta'}$ is in V_β and hence $V_{\beta'}$ is in V_α by transitivity. For a limit ordinal α, $V_\alpha = \bigcup_{\beta<\alpha} V_\beta$ is transitive by the inductive hypothesis. For $\beta < \alpha$, it must be that V_β is in V_α by the definition and transitivity of V_α. □

7.6 The Rank Function (6.3, 6.5)

For any x in V, $\rho(x)$ is the least τ such that $x \subseteq V_\tau$ (or rather, x is in $V_{\tau+1}$). We have that:

1. $V_\alpha = \{x \in V : \rho(x) < \alpha\}$,
2. For x in V, and all y in x, y is in V and $\rho(y) < \rho(x)$,
3. For x in V, $\rho(x) = \sup(\{\rho(y) + 1 : y \in x\}) = \sup^+(\{\rho(y) : y \in x\})$.

So, the relation on sets:

$$xRy \iff \rho(x) < \rho(y),$$

is a strict partial order that is well-founded, meaning there is a R-least element of every non-empty $X \subseteq V$.

Proof. (1) For x in V, then $\rho(x) < \alpha$ is equivalent to saying that there is some $\beta < \alpha$ such that $x \subseteq V_\beta$ or rather x is in $V_{\beta+1}$. This is then equivalent to saying x is in V_α as $V_{\beta+1} \subseteq V_\alpha$ by the transitivity of V.

(2) We take $\rho(x) = \alpha$ so $x \subseteq V_\alpha$ and as such y in x must be in V_α so $\rho(y) < \alpha$.

(3) We take $\alpha = \sup^+(\{\rho(y) : y \in x\})$ and y in x. By (2), $\rho(y) < \rho(y)+1 \leq \rho(x)$ so, $\alpha \leq \rho(x)$. By (1), $\rho(y) < \rho(y)+1 \leq \alpha$ so, y is in V_α. Thus, $x \subseteq V_\alpha$ so $\rho(x) \leq \alpha$. □

7.7 Rank and Ordinals (6.6)

For an ordinal α, $\rho(\alpha) = \alpha$ and $(\mathrm{On} \cap V_\alpha) = \alpha$.

Proof. The result is trivial for $\alpha = 0$, so we proceed by induction with $\alpha > 0$. By (7.6):

$$\begin{aligned} \rho(\alpha) &= \sup^+(\{\rho(\beta) : \beta < \alpha\}) \\ &= \sup^+(\{\beta : \beta < \alpha\}) && \text{(IH)} \\ &= \alpha. \end{aligned}$$

From this, we know that $\alpha \subseteq (\mathrm{On} \cap V_\alpha)$. We take β in $(\mathrm{On} \cap V_\alpha)$ so $\beta = \rho(\beta) < \alpha$. Thus, $(\mathrm{On} \cap V_\alpha) \subseteq \alpha$ and as such, $\alpha = (\mathrm{On} \cap V_\alpha)$. □

7.8 Principle of $\in$-induction (6.7)

For a well-defined and definite property of sets Φ:

$$\big[\forall\, z \in y, \Phi(z) \Longrightarrow \Phi(y)\big] \Longrightarrow \Phi(y), \tag{$*$}$$

and if x is a transitive set, we have $(*)$ for all y in x.

Proof. For a transitive set x, we take $Z = \{y \in x : \neg\Phi(y)\}$, supposing $Z \neq \varnothing$. By the Axiom of Foundation, we have y_0 in Z such that y_0 is $\in$-minimal (meaning $(y_0 \cap Z) = \varnothing$). For any u in y_0, u must be in x as x is transitive. By the minimality of y_0, it must be that $\Phi(u)$ holds as otherwise $(y_0 \cap Z)$ would contain u. As such, assuming the antecedent, we get $\Phi(y_0)$ which is a contradiction of the membership of y_0 in Z. For the case on classes, we just take $Z = \{y : \neg\Phi(y)\}$ and use the same argument. $\square$

7.9 Theorem of $\in$-recursion (6.8)

For a function G from V to V, there is exactly one function H from V to V such that for all x:

$$H(x) = G(H \restriction x) = G(\{\langle y, H(y)\rangle : y \in x\}).$$

Proof. The proof operates similarly to that on ordinals, but is omitted. $\square$