

Ordinal Numbers

Ryan Tay, MA213 Second Year Essay

17 April 2023

Abstract

It is common to introduce both ordinal and cardinal numbers in any text covering the construction and generalisation of natural numbers, however we shall restrict ourselves to only ordinal numbers here. Ordinal numbers can be thought of as generalisations of natural numbers, where we prioritise position and order over cardinality. Through the approach of axiomatic set theory, we will develop ordinal numbers and explore how we may use them.

This essay is broken into three sections: a preliminaries section covering the basic theory of sets and relations (which the reader may skip), a section on the construction of ordinal numbers, and a section on some applications of ordinal numbers.

Contents

1	Introduction and Preliminaries	2
1.1	ZFC Set Theory	2
1.1.1	The Axioms, Part 1	3
1.1.2	Developing Basic Set Theory, Part 1	4
1.1.3	The Axioms, Part 2	6
1.1.4	Developing Basic Set Theory, Part 2	6
1.2	Relations	7
1.2.1	Binary Relations	7
1.2.2	Functions	7
1.2.3	Well-Ordering	8
2	Ordinal Numbers	9
2.1	The Structure of Ordinal Numbers	9
2.2	The Successor Operator	11
2.3	Natural Numbers	12
3	Using Ordinal Numbers	14
3.1	Transfinite Induction and Transfinite Recursion	14
3.2	Well-Ordering Theorem	17
3.3	Basic Ordinal Arithmetic	19
3.4	Goodstein's Theorem	21

1 Introduction and Preliminaries

This essay will largely follow the approach taken in *Axiomatic Set Theory* by Patrick Suppes [16]. In particular, we shall assume the existence of urelements (objects which are not sets), and the notion of classes will not be developed. For a more modern approach, we recommend *Set Theory* by Thomas Jech [7].

This section is devoted to defining notations and developing the tools which we shall use throughout the essay. If the reader is comfortable with the basic theory of sets and relations the reader may skip ahead to the heart of the essay, from Section 2 onwards, and may revisit this section if needed.

The development of mathematical logic and the notion of equality, where we write $x = y$ to mean “the objects x and y are equal”, deserves extensive discussion and is fully developed in Whitehead and Russell 1910 [18]. We shall assume all results from mathematical logic as well as the properties of equality. The symbols we require are defined below for logical statements P and Q .

Notation	Meaning
$P \wedge Q$	P and Q
$P \vee Q$	P or Q
$\neg P$	P is false
$P \implies Q$	If P then Q
$P \iff Q$	P if and only if Q
$(\forall x)P$	For all objects x , P
$(\exists x)P$	There exists an object x such that P
$(\exists!x)P$	There exists a unique object x such that P
$(\nexists x)P$	There does not exist an object x such that P

We also adopt the notation $x := y$ to mean “the object x is defined to satisfy $x = y$ ”, and the notation $x \neq y$ to mean $\neg(x = y)$.

1.1 ZFC Set Theory

The ZFC axioms are the axioms of Zermelo-Fraenkel set theory together with the axiom of choice. They are the most widely accepted axiomatic system among mathematicians [4, p. 4], and we will adopt them as laid out by Suppes 1960 [16].

We start with two special symbols: \in and \emptyset . We treat \emptyset as a special constant, which we shall later define, and we write $x \in y$ to mean “the object x is a member of the object y ”. We will write $x \notin y$ if and only if $\neg(x \in y)$. With these, we define the notion of a *set*.

Definition 1.1.1 (Sets). An object x is a *set* if and only if $(\exists y)(y \in x) \vee (x = \emptyset)$.

Remark. From Definition 1.1.1, we observe that \emptyset is a set.

To avoid the repetition of the predicate “is a set”, we will use capital Latin alphabets (A, B, C, \dots) to denote sets, and lower case Latin alphabets (x, y, z, \dots) to denote general objects. To reduce the number of parentheses, we establish the following order of precedence for logical symbols, listed in order of priority:

1. $\forall, \exists, \exists!$, and \neg ;
2. $=, \neq, \in$, and \notin ;
3. \wedge and \vee ;
4. \implies and \iff .

For example, “ $x \in A \vee x \neq B \iff P$ ” means “ $((x \in A) \vee (x \neq B)) \iff P$ ”. We also adopt the following conventions:

- “ $(\exists x \in A)P$ ” means “ $(\exists x)(x \in A \wedge P)$ ”,
- “ $(\forall x \in A)P$ ” means “ $(\forall x)(x \in A \implies P)$ ”,
- “ $x, y \in A$ ” means “ $(x \in A) \wedge (y \in A)$ ”, and
- “ $x, y, z \in A$ ” means “ $(x \in A) \wedge (y \in A) \wedge (z \in A)$ ”.

1.1.1 The Axioms, Part 1

Our first five axioms are as follows.

Axiom 1 (Axiom of Extensionality). Let A and B be sets. Then

$$(\forall x)(x \in A \iff x \in B) \implies A = B.$$

Remark. We can establish the converse by using the properties of equality.

Axiom 2 (Axiom Schema of Separation). Let A be a set, and $\varphi(x)$ be a logical statement depending on the variable x . Then

$$(\exists B)(\forall x)(x \in B \iff x \in A \wedge \varphi(x)).$$

Remark. Notice that, by our convention of using capital letters for sets, we require that B is a set. Moreover, by the axiom of extensionality, this set B is unique.

Axiom 3 (Axiom of Pairing). Let x and y be objects. Then

$$(\exists A)(\forall z)(z \in A \iff z = x \vee z = y).$$

Remark. As before, A is a set and, by the axiom of extensionality, is unique. We define the notation $\{x, y\} := A$, where A satisfies the axiom of pairing for x and y .

Axiom 4 (Axiom of Union). Let A be a set which may contain sets. Then

$$(\exists B)(\forall x)(x \in B \iff (\exists C \in A)(x \in C)).$$

Remark. Again, B is the unique set with this property. We define the notation $\bigcup A := B$, where B satisfies the axiom of union for the set A .

Axiom 5 (Axiom Schema of Replacement). Let A be a set and $\varphi(p, q)$ be a logical statement depending on two variables p and q . Suppose that

$$(\forall x \in A)(\forall y)(\forall z)(\varphi(x, y) \wedge \varphi(x, z) \implies y = z).$$

Then

$$(\exists B)(\forall y)(y \in B \iff (\exists x \in A)(\varphi(x, y))).$$

Remark. This set B is, again, unique.

1.1.2 Developing Basic Set Theory, Part 1

Before introducing the remaining axioms, we first develop some set theory to introduce some notations in set theory. The aim of this section is just to define basic operations in set theory, and not to derive their properties. We leave those developments to Suppes 1960 [16, pp. 14–89].

Theorem 1.1.2. *There exists a unique set A such that $(\forall x)(x \notin A)$, and $A = \emptyset$.*

Proof. From the axiom schema of separation, there exists a set B such that

$$(\forall x)(x \in B \iff x \in \emptyset \wedge x \neq x)$$

since \emptyset is a set. Now notice that $(\forall x)(x \notin B)$, since $x \neq x$ is always false. Since B is a set, Definition 1.1.1 requires $B = \emptyset$. Thus we obtain

$$(\forall x)(x \in \emptyset \iff x \in \emptyset \wedge x \neq x),$$

which reduces to

$$(\forall x)(x \in \emptyset \implies x \neq x).$$

Therefore $(\forall x)(x \notin \emptyset)$. So the existence condition is satisfied by \emptyset . Uniqueness then follows from the axiom of extensionality. \square

Remark. This justifies naming \emptyset *the empty set*.

The theorem below shows that a “universal set” cannot exist in ZFC set theory.

Theorem 1.1.3. *There does not exist a set V such that $(\forall x)(x \in V)$.*

Proof. Suppose, for a contradiction, such a set V exists. Then, from the axiom schema of separation,

$$(\exists B)(\forall x)(x \in B \iff x \in V \wedge x \notin x).$$

Since we have $(\forall x)(x \in V)$, this reduces to

$$(\exists B)(\forall x)(x \in B \iff x \notin x).$$

Now, taking $x = B$, this yields the absurd result $B \in B \iff B \notin B$. \square

We now introduce notation for sets of all objects with a specified property.

Definition 1.1.4. Let $\varphi(x)$ be a logical statement depending on x . We define $\{x : \varphi(x)\} := A$ where A is the set which satisfies

$$(\forall x)(x \in A \iff \varphi(x)) \vee (A = \emptyset \wedge (\nexists B)(\forall x)(x \in B \iff \varphi(x))).$$

This definition is purposely clunky to avoid the construction of sets in which paradoxes such as Russell’s paradox [13] may arise. Due to Theorem 1.1.3, we require $\{x : x = x\} = \emptyset$.

We also introduce the following more flexible notations.

Definition 1.1.5. Let A be a set and $\varphi(x)$ be a logical statement depending on x . We define

$$\{x \in A : \varphi(x)\} := \{x : x \in A \wedge \varphi(x)\}.$$

Definition 1.1.6. Let $\varphi(x_1, \dots, x_n)$ be a logical statement depending on the variables x_1, \dots, x_n , and let $\tau(x_1, \dots, x_n)$ be a *term* depending on x_1, \dots, x_n (which is any valid construction of objects from the variables x_1, \dots, x_n). We define

$$\{\tau(x_1, \dots, x_n) : \varphi(x_1, \dots, x_n)\} := \{y : (\exists x_1) \dots (\exists x_n)(y = \tau(x_1, \dots, x_n) \wedge \varphi(x_1, \dots, x_n))\}.$$

We now start to develop basic tools in set theory. The definitions of a set being a *subset* or a *proper subset* of another set are straightforward.

Definition 1.1.7 (Subsets). Let A and B be sets. We write $A \subseteq B$ if and only if $(\forall x)(x \in A \implies x \in B)$. Furthermore, we write $A \subset B$ if and only if $A \subseteq B$ and $A \neq B$.

Before defining the *union*, *intersection*, and *set difference* of two sets, we must first prove that such sets exist and are unique.

Theorem 1.1.8. Let A and B be sets. Then $(\exists! C)(\forall x)(x \in C \iff x \in A \vee x \in B)$.

Proof. From the pairing and union axioms, $\bigcup\{A, B\}$ exists. Then for all x ,

$$x \in \bigcup\{A, B\} \iff (\exists D \in \{A, B\})(x \in D) \iff x \in A \vee x \in B.$$

So the existence claim is satisfied by $\bigcup\{A, B\}$. For uniqueness, suppose for some set C' we have $(\forall x)(x \in C' \iff x \in A \vee x \in B)$. Then for all x ,

$$x \in C' \iff x \in A \vee x \in B \iff x \in \bigcup\{A, B\},$$

and thus by the extensionality axiom $C' = \bigcup\{A, B\}$. \square

Theorem 1.1.9. Let A and B be sets. Then $(\exists! C)(\forall x)(x \in C \iff x \in A \wedge x \in B)$.

Proof. The existence of such C follows from the axiom schema of separation, with $\varphi(x)$ as $x \in B$. Uniqueness is proved similarly as in Theorem 1.1.8. \square

Theorem 1.1.10. Let A and B be sets. Then $(\exists! C)(\forall x)(x \in C \iff x \in A \wedge x \notin B)$.

Proof. Similarly as in Theorem 1.1.9. \square

Theorems 1.1.8, 1.1.9, and 1.1.10 justify the following three definitions.

Definition 1.1.11 (Unions, Intersections, and Set Differences). Let A and B be sets. We define

$$\begin{aligned} A \cup B &:= \{x : x \in A \vee x \in B\}, \\ A \cap B &:= \{x : x \in A \wedge x \in B\}, \\ A \setminus B &:= \{x : x \in A \wedge x \notin B\}. \end{aligned}$$

We can also construct of sets by simply listing their elements. Note that due to the axiom of pairing, for any objects x and y , there exists a set $\{x, y\}$ containing only x and y .

Definition 1.1.12. For an object x, y, z, w , define $\{x\} := \{x, x\}$;

for objects x, y, z , define $\{x, y, z\} := \{x, y\} \cup \{z\}$;

for objects x, y, z, w , define $\{x, y, z, w\} := \{x, y, z\} \cup \{w\}$;

and so on.

Definition 1.1.13. Let A be a set, and let $\tau(x)$ be a term depending on the variable x . Then

$$\bigcup_{x \in A} \tau(x) := \bigcup\{y : (\exists x \in A)(y = \tau(x))\}.$$

We have the following theorem justifying the existence of this set.

Theorem 1.1.14. Let A be a set, and let $\tau(x)$ be a term depending on the variable x . Then for all y ,

$$y \in \bigcup_{x \in A} \tau(x) \iff (\exists x \in A)(\exists B)(B = \tau(x) \wedge y \in B).$$

Proof. Define the logical statement $\varphi(x, y)$ to be true if and only if $x \in A$ and $y = \tau(x)$. If $\varphi(x, y)$ and $\varphi(x, z)$ are both true, then we easily conclude that $y = z$. Hence by the axiom schema of replacement on A and φ , we obtain

$$(\exists C)(\forall z)(z \in C \iff (\exists x \in A)(z = \tau(x))).$$

Using the axiom of union on C then gives the desired result. \square

1.1.3 The Axioms, Part 2

We now complete our collection of the axioms of Zermelo-Fraenkel set theory. The axiom of choice will be introduced in Section 1.2.2, after developing functions.

Axiom 6 (Axiom of Power Set). Let A be a set. Then

$$(\exists B)(\forall C)(C \subseteq A \implies C \in B).$$

Remark. We can establish the converse of the axiom of power set as follows. Due to the axiom schema of separation, $(\exists D)(\forall x)(x \in D \iff x \in B \wedge x \subseteq A)$. Now due to the power set axiom, this reduces to the desired $(\exists D)(\forall C)(C \in D \iff C \subseteq A)$. Furthermore, by the axiom of extensionality, this set D is unique. Thus we define the *power set* of A as $\mathcal{P}(A) := D$.

Axiom 7 (Axiom of Regularity). Let A be a set such that $A \neq \emptyset$. Then

$$(\exists x \in A)(\forall y \in x)(y \notin A).$$

Remark. Intuitively, this says $(\exists x \in A)(x \cap A = \emptyset)$. The purpose of the axiom being formulated this way is for the case when x is not a set, and so intersections with x are not defined.

Axiom 8 (Axiom of Infinity). There exists a set A such that

$$\emptyset \in A \wedge (\forall B \in A)(B \cup \{B\} \in A).$$

1.1.4 Developing Basic Set Theory, Part 2

Definition 1.1.15. For objects x and y , the *ordered pair* (x, y) is defined as $(x, y) := \{\{x\}, \{x, y\}\}$.

Definition 1.1.16 (Cartesian Products). Let A and B be sets. We define

$$A \times B := \{(x, y) : x \in A \wedge y \in B\}.$$

Due to Definitions 1.1.4 and 1.1.6, we require the following theorem to avoid $A \times B$ simply being the empty set.

Theorem 1.1.17. Let A and B be sets. Then for all z , we have $z \in A \times B$ if and only if

$$(\exists x \in A)(\exists y \in B)(z = (x, y)). \quad (1)$$

Proof. Due to Definition 1.1.4, it suffices to show that there exists a set C such that for all z , we have $z \in C$ if and only if (1) holds.

Now, from the axiom schema of separation, there exists a set C such that for all z , we have that $z \in C$ if and only if $z \in \mathcal{P}(\mathcal{P}(A \cup B))$ and (1) holds. So if we show that (1) implies $z \in \mathcal{P}(\mathcal{P}(A \cup B))$, then we are done.

Suppose (1) holds, so $z = \{\{x\}, \{x, y\}\}$ for some $x \in A$ and $y \in B$. Now it is clear that $\{x\}, \{x, y\} \subseteq A \cup B$, therefore $\{x\}, \{x, y\} \in \mathcal{P}(A \cup B)$. This then gives us $\{\{x\}, \{x, y\}\} \subseteq \mathcal{P}(A \cup B)$, and hence $\{\{x\}, \{x, y\}\} \in \mathcal{P}(\mathcal{P}(A \cup B))$. \square

It is also worth noting that the axiom of regularity gives us the following two important theorems.

Theorem 1.1.18. For all sets A , we have $A \notin A$.

Proof. Suppose, for a contradiction, that $(\exists A)(A \in A)$. Clearly we have $A \in \{A\}$, and so $A \in A \cap \{A\}$. However, from the axiom of regularity, $(\exists x \in \{A\})(x \cap \{A\} = \emptyset)$ and thus $A \cap \{A\} = \emptyset$ since A is the only element of $\{A\}$. This gives us the absurd result $A \in \emptyset$. \square

Theorem 1.1.19. For all sets A and B , we have $\neg(A \in B \wedge B \in A)$.

Proof. Suppose, for a contradiction, that $(\exists A)(\exists B)(A \in B \wedge B \in A)$. Then $A \in \{A, B\} \cap B$ and $B \in \{A, B\} \cap A$. Now by the axiom of regularity, $(\exists x \in \{A, B\})(x \cap \{A, B\} = \emptyset)$. This requires $x = A$ or $x = B$. However we then obtain $\{A, B\} \cap B = \emptyset$ or $\{A, B\} \cap A = \emptyset$, giving the contradiction $A \in \emptyset$ or $B \in \emptyset$. \square

1.2 Relations

1.2.1 Binary Relations

Definition 1.2.1 (Relations). Let A and B be sets. A set R is a *(binary) relation* on $A \times B$ if and only if $R \subseteq A \times B$. For a relation R , we write xRy if and only if $(x, y) \in R$.

Definition 1.2.2 (Domains and Images of Relations). The *domain* and *image* of a relation R are defined, respectively, as

$$\begin{aligned}\text{Dom}(R) &:= \{x : (\exists y)(xRy)\}, \\ \text{Im}(R) &:= \{y : (\exists x)(xRy)\}.\end{aligned}$$

Remark. The proofs that the sets $\text{Dom}(R)$ and $\text{Im}(R)$ are not simply the empty set follow a similar argument as in Theorem 1.1.17, using the axiom schema of separation on the set $\bigcup \bigcup R$.

Definition 1.2.3 (Converse Relations). The *converse* of a relation R is $R^T := \{(y, x) : xRy\}$.

Remark. The proof that R^T is not simply the empty set also follows a similar argument to Theorem 1.1.17, using the axiom schema of separation on the set $\text{Im}(R) \times \text{Dom}(R)$.

Definition 1.2.4 (Restrictions of Relations). Let R be a relation and A be a set. Then

$$R|_A := R \cap (A \times \text{Im}(R)).$$

It is clear that R^T and $R|_A$ above are relations.

1.2.2 Functions

Definition 1.2.5 (Functions). A relation f is a *function* if and only if

$$(\forall x)(\forall y)(\forall z)(xfy \wedge x fz \implies y = z).$$

Definition 1.2.6. For a function f and an object x , we write $f(x) := y$ where y satisfies

$$((\exists! z)(xfz) \wedge xfy) \vee ((\nexists z)(xfz) \wedge y = \emptyset).$$

Definition 1.2.7 (Function Spaces). Let A and B be sets. Then

$$\mathcal{F}(A, B) := \{f : f \text{ is a function} \wedge \text{Dom}(f) = A \wedge \text{Im}(f) \subseteq B\}.$$

Remark. By the axiom schema of separation on the set $A \times B$, we establish the existence of the desired set of all functions from A to B similarly as in Theorem 1.1.17. The conventional notation for $f \in \mathcal{F}(A, B)$ is $f : A \rightarrow B$, but we will not be using that notation in this essay due to the confusions which may arise when working with sets of functions.

Definition 1.2.8. Let A and B be sets, and let $f \in \mathcal{F}(A, B)$. We say f is *injective* if and only if $f^T \in \mathcal{F}(B, A)$. Furthermore, we say f is *bijective* if and only if f is injective and $\text{Im}(f) = B$.

Finally, with the notion of functions defined, we introduce the axiom of choice.

Axiom 9 (Axiom of Choice). Let A be a set such that $(\forall x \in A)(x \text{ is a set} \wedge x \neq \emptyset)$. Then

$$\left(\exists f \in \mathcal{F} \left(A, \bigcup A \right) \right) (\forall B \in A)(f(B) \in B).$$

1.2.3 Well-Ordering

Definition 1.2.9. Let A be a set, and let R be a relation on $A \times A$.

- We say R is *asymmetric* in A if and only if $(\forall x, y \in A)(xRy \implies \neg(yRx))$.
- We say R *connects* A if and only if $(\forall x, y \in A)(x \neq y \implies xRy \vee yRx)$.
- We say $x \in A$ is an *R -least element* of A if and only if $(\forall y \in A)(x \neq y \implies xRy)$.
- We say R *well-orders* A if and only if R is asymmetric in A and connects A , and

$$(\forall B)(B \subseteq A \wedge B \neq \emptyset \implies B \text{ has an } R\text{-least element}).$$

Theorem 1.2.10. Let a relation R well-order a set A , and let $B \subseteq A$. Then R well-orders B .

Proof. If $B = \emptyset$ or $B = \{x\}$ for some $x \in A$, then B is trivially well-ordered. So suppose $(\exists x, y \in A)(x \neq y \wedge x, y \in B)$.

The asymmetry of R in B is carried over from the asymmetry of R in A . For any $x, y \in B$ such that $x \neq y$, we have $x, y \in A$. Then $xRy \vee yRx$ since R connects A , and so R connects B . Now for any set C such that $C \subseteq B \wedge C \neq \emptyset$, we have $C \subseteq A \wedge C \neq \emptyset$. Therefore C has an R -least element. \square

Theorem 1.2.11. Let A be a set such that $A \neq \emptyset$, and let a relation R well-order A . Then A has a unique R -least element.

Proof. Since $A \neq \emptyset$, the well-ordering of A requires that A has an R -least element. Now suppose $x, y \in A$ are both R -least elements of A with $x \neq y$. Then xRy and yRx . Now, if we consider the subset $A_{x,y} := \{x, y\}$ of A , by Theorem 1.2.10, R well-orders $A_{x,y}$. But R being asymmetric in $A_{x,y}$ yields the contradiction $xRy \wedge yRx \wedge \neg(xRy) \wedge \neg(yRx)$. \square

2 Ordinal Numbers

The layman definition of an *ordinal number* is a number defining a thing's position in a series, such as “first” or “second” [15]. We will generalise this notion to infinity and beyond [9].

For convenience, any set defined using Definitions 1.1.4, 1.1.5 or 1.1.6 will henceforth be assumed to have a theorem justifying the existence of the desired set, like Theorem 1.1.17. For example, given a logical statement $\varphi(x)$ we will assume that the set $A := \{x : \varphi(x)\}$ satisfies $(\forall x)(x \in A \iff \varphi(x))$. Most of the content below come from Suppes 1960 [16], with some changes as laid out by Jech 2003 [7] to follow modern convention.

2.1 The Structure of Ordinal Numbers

Definition 2.1.1 (Transitive Sets). Let A be a set. We say A is *transitive* if and only if

$$(\forall x \in A)(x \text{ is a set} \wedge x \subset A).$$

Definition 2.1.2. Let A be a set. We define the relation \in_A on $A \times A$ by

$$\in_A := \{(x, y) : x, y \in A \wedge x \in y\}.$$

Definition 2.1.3 (Ordinals). Let α be a set. We say α is an *ordinal* (or an *ordinal number*) if and only if α is transitive and \in_α connects α .

A consequence of this definition is all ordinals α are well-ordered by the \in_α . Some texts define ordinals to be transitive and well-ordered, as in Jech 2003 [7, p. 19].

Lemma 2.1.4. *Let α be an ordinal. Then \in_α well-orders α .*

Proof. By definition, \in_α connects α .

Suppose we have $x, y \in \alpha$ such that $x \in_\alpha y$ and $y \in_\alpha x$. Then $x \in y$ and $y \in x$, contradicting Theorem 1.1.19. Hence \in_α is asymmetric in α .

Now suppose $A \subseteq \alpha$ is such that $A \neq \emptyset$. By the axiom of regularity and the transitivity of α ,

$$(\exists B \in A)(A \cap B = \emptyset),$$

and hence $(\forall x \in A)(x \notin B)$. Then because \in_α connects α , we must have $(\forall x \in A)(x \neq B \implies B \in x)$. Therefore B is an \in_α -least element of A . \square

The simplest example of an ordinal would be \emptyset . We will use this fact to later define the ordinal number 0 as \emptyset .

Lemma 2.1.5. *The empty set is an ordinal.*

Proof. We know that $(\forall x)(x \notin \emptyset)$. Therefore the statement “ $(x \text{ is a set} \wedge x \subset A)$ ” in Definition 2.1.1 trivially holds due to “ $x \in A$ ” being a false statement, and so \emptyset is transitive. A similar argument is used on Definition 1.2.9 to conclude that \in_\emptyset connects \emptyset . \square

Before introducing more examples of ordinals, we establish several theorems that reveal the structure of ordinals. We first show that every ordinal only consists of other ordinals.

Proposition 2.1.6. *Let α be an ordinal. Then $(\forall \beta \in \alpha)(\beta \text{ is an ordinal})$.*

Proof. Let $\beta \in \alpha$. The transitivity of α yields $\beta \subset \alpha$. From Theorem 1.2.10 we know that \in_α well-orders β , so \in_α connects β and every non-empty subset of β has an \in_α -least element. Since \in_α connects β ,

$$(\forall x, y \in \beta)(x \neq y \implies x \in_\alpha y \vee y \in_\alpha x).$$

Since $\beta \subset \alpha$, we may rewrite this as

$$(\forall x, y \in \beta)(x \neq y \implies x \in_\beta y \vee y \in_\beta x),$$

from which we conclude that \in_β connects β .

Now for any $x \in \beta$, the transitivity of α implies that x is a set. Then for all $y \in x$, we have $y \in_\alpha x \in_\alpha \beta$, so we obtain $y \in_\alpha \beta$ because \in_α connects α . Therefore $x \subset \beta$ and thus β is transitive. \square

We also have the following: for any ordinals α and β , we have $\alpha \in \beta$ if and only if $\alpha \subset \beta$. The forward implication is simply due to the transitivity of β . The converse of this statement is established below.

Proposition 2.1.7. *Let α and β be ordinals with $\alpha \subset \beta$. Then $\alpha \in \beta$.*

Proof. As $\alpha \subset \beta$, the set $\beta \setminus \alpha$ is non-empty and is thus well-ordered by \in_β , due to Theorem 1.2.10. So there is an \in_β -least element γ of $\beta \setminus \alpha$. We claim that $\gamma = \alpha$.

For any $x \in \gamma$, the transitivity of β and the minimality of γ imply that $x \in \alpha$. So we have the inclusion $\gamma \subseteq \alpha$. Now, for any $x \in \alpha$ satisfying $x \notin \gamma$, we cannot have that $x = \gamma$, otherwise $\gamma = x \in \alpha$ which contradicts γ being an (\in_β -least) element of $\beta \setminus \alpha$. So, as \in_β connects β , we must have $\gamma \in x$. However, the transitivity of α then implies that $\gamma \in \alpha$, again, a contradiction. Therefore every $x \in \alpha$ must satisfy $x \in \gamma$, giving $\alpha \subseteq \gamma$. \square

We can then establish the following important theorem: for any ordinals α and β , they are either equal or one must be a proper subset (and hence an element, by Proposition 2.1.7) of the other.

Theorem 2.1.8. *Let α and β be ordinals. Then $\alpha \subseteq \beta$ or $\beta \subseteq \alpha$.*

Proof. Define $\gamma := \alpha \cap \beta$. Arguing similarly as in Proposition 2.1.6, we conclude that γ is an ordinal.

We will show that $\gamma = \alpha$ or $\gamma = \beta$, from which the theorem will follow. Suppose, for a contradiction, that $\gamma \subset \alpha$ and $\gamma \subset \beta$. Then $\gamma \in \alpha$ and $\gamma \in \beta$ by Proposition 2.1.7, and so $\gamma \in \alpha \cap \beta = \gamma$, contradicting Theorem 1.1.18. \square

We can now define the notion of an ordinal number being “less than” another ordinal number.

Definition 2.1.9. Let α and β be ordinals. We write

- $\alpha < \beta$ if and only if $\alpha \in \beta$,
- $\alpha \leq \beta$ if and only if $\alpha < \beta \vee \alpha = \beta$.

The notations $\alpha > \beta$ and $\alpha \geq \beta$ are defined similarly.

Standard inequality results quickly follow. There is no ordinal α such that $\alpha < \emptyset$, due to Theorem 1.1.2. From Theorem 1.1.18, for any ordinal α we have $\neg(\alpha < \alpha)$. Due to the transitive property of ordinals, we also have for any ordinals α, β, γ , if $\alpha < \beta$ and $\beta < \gamma$ then $\alpha < \gamma$. Furthermore, Theorem 2.1.8 allows us to have a “law of trichotomy” for ordinal numbers, i.e. for any ordinals α and β , exactly one of the following holds: $\alpha < \beta$, or $\alpha = \beta$, or $\alpha > \beta$.

Putting all of the above together, we obtain the structure of all ordinals: as the set of all ordinals less than that ordinal.

Corollary. *Let α be an ordinal. Then*

$$\alpha = \{\beta : \beta \text{ is an ordinal} \wedge \beta < \alpha\}.$$

With all ordinals following a very specific structure, it would be natural to assume that we can have a “set of all ordinals”. However, as useful as it would be, the set of all ordinals does not exist.

Theorem 2.1.10. *There does not exist a set V such that $(\forall x)(x \in V \iff x \text{ is an ordinal})$.*

Proof. Suppose such a set V exists. We will show that V must be an ordinal, from which $V \in V$ follows, contradicting Theorem 1.1.18.

From Theorem 2.1.8, any $\alpha, \beta \in V$ such that $\alpha \neq \beta$ have the property $\alpha \in \beta$ or $\beta \in \alpha$. Hence \in_V connects V .

Now suppose $\gamma \in V$. For any $\delta \in \gamma$, Proposition 2.1.6 tells us that $\delta \in V$. Thus $\gamma \subset V$, which establishes the transitivity of V . \square

2.2 The Successor Operator

For any ordinal, we may think of its successor as the “next immediate ordinal”, analogous to “adding 1 to a natural number”. In fact, the successor operator will be used in the definition of addition of ordinal numbers, and will have the property of “adding 1”.

Definition 2.2.1 (Successor Operator). Let α be an ordinal. Then the *successor* of α is

$$S(\alpha) := \alpha \cup \{\alpha\}.$$

Remark. The successor operator is not a function. Functions are defined as relations between sets (see Section 1.2.2). The successor operator is defined on all ordinals, and a set of all ordinals does not exist due to Theorem 2.1.10.

We should expect that the successor of an ordinal α is also an ordinal. We observe that $S(\alpha) = \alpha \cup \{\alpha\} = \bigcup \{\beta : \beta \text{ is an ordinal} \wedge \beta \leq \alpha\}$, and so $S(\alpha)$ is a union of ordinals. Using the following theorem then shows that $S(\alpha)$ is an ordinal.

Proposition 2.2.2. *Let A be a set such that $(\forall x \in A)(x \text{ is an ordinal})$. Then $\bigcup A$ is an ordinal.*

Proof. From Proposition 2.1.6, elements of ordinals are ordinals. So for any $\alpha, \beta \in \bigcup A$ we conclude that α and β are ordinals. Thus if $\alpha \neq \beta$, then $\alpha \in \beta$ or $\beta \in \alpha$. Therefore $\in_{\bigcup A}$ connects $\bigcup A$.

Now suppose $\gamma \in \bigcup A$. As above, γ is an ordinal and is thus a set. Also, by definition, $(\exists \delta \in A)(\gamma \in \delta)$. Then, as δ is an ordinal, we have $\gamma \subset \delta$. We thus have

$$\gamma \subset \delta \subseteq \bigcup A$$

giving the desired $\gamma \subset \bigcup A$. Thus $\bigcup A$ is transitive. \square

Corollary. *Let α be an ordinal. Then $\bigcup \alpha \leq \alpha$.*

Definition 2.1.9 yields the intuitively obvious $\alpha < S(\alpha)$ for any ordinal α . We also expect the successor operator to have the property that there is no ordinal between α and $S(\alpha)$.

Theorem 2.2.3. *Let α be an ordinal. Then*

$$(\nexists \beta)(\beta \text{ is an ordinal} \wedge \alpha < \beta < S(\alpha)).$$

Proof. Suppose, for a contradiction, such a β exists. Then we have

$$\alpha \subset \beta \subset S(\alpha) = \alpha \cup \{\alpha\}$$

and so $(\exists x, y \in S(\alpha) \setminus \alpha)(x \in \beta \wedge y \notin \beta)$. However as $x, y \in S(\alpha) \setminus \alpha$, we require $x, y \in \{\alpha\}$ and thus $x = y = \alpha$, contradicting $(x \in \beta \wedge y \notin \beta)$. \square

Corollary. *Let α be an ordinal. Then $\bigcup S(\alpha) = \alpha$.*

We also have the “cancellation law” for the successor operator.

Proposition 2.2.4. *Let α and β be ordinals. If $S(\alpha) = S(\beta)$ then $\alpha = \beta$.*

Proof. Suppose $S(\alpha) = S(\beta)$. For any $\gamma \in \alpha$ we have $\gamma < S(\alpha)$, and so $\gamma < S(\beta)$. This gives $\gamma \leq \beta$. Suppose, for a contradiction, that $\gamma = \beta$. Then $S(\gamma) = S(\beta) = S(\alpha)$, contradicting the fact that $\gamma < \alpha < S(\alpha)$ (which requires $S(\gamma) < S(\alpha)$). Therefore $\gamma < \beta$, and hence $\alpha \subseteq \beta$. Similarly, we can establish that $\beta \subseteq \alpha$. \square

While we can “cancel” a successor operator as above, we cannot necessarily apply the inverse of the successor operator onto ordinals. More concretely, there are some ordinals which do not have predecessors. The ordinal \emptyset is a trivial example. A more interesting example is infinity, which we will soon formalise. We will introduce such examples properly when we define ordinal arithmetic later.

Definition 2.2.5 (Limit Ordinals). An ordinal α is a *limit ordinal* if and only if $\alpha \neq \emptyset$ and

$$(\nexists \beta)(\beta \text{ is an ordinal} \wedge S(\beta) = \alpha).$$

Lemma 2.2.6. *Let α be a limit ordinal. Then $\bigcup \alpha = \alpha$.*

Proof. We already know that $\bigcup \alpha \leq \alpha$. Now for any ordinal $\beta < \alpha$, since α is a limit ordinal, we have $\neg(S(\beta) = \alpha)$ and hence $S(\beta) < \alpha$. Therefore

$$\beta \in S(\beta) \subset \bigcup \{\gamma : \gamma \text{ is an ordinal} \wedge \gamma < \alpha\} = \bigcup \alpha,$$

which gives us $\alpha \leq \bigcup \alpha$. \square

2.3 Natural Numbers

Natural numbers are simply the finite ordinals, driven by the idea that a finite ordinal should have an “end”. Since all ordinals α are well-ordered by \in_α , all finite ordinals α should be well-ordered by \in_α^T . Intuitively they have an “ \in_α -greatest element”, and we can “count backwards” from a finite ordinal α down to \emptyset .

Definition 2.3.1 (Natural Numbers). Let n be an ordinal number. We say n is a *natural number* (or *finite ordinal*) if and only if \in_n^T well-orders n .

Using a similar argument to Lemma 2.1.5, we can show that \emptyset is a natural number. The following theorems then allow us to build up all the natural numbers.

Proposition 2.3.2. *Let n be a natural number. Then $S(n)$ is a natural number.*

Proof. We already know $S(n)$ is an ordinal, hence $\in_{S(n)}$ connects $S(n)$ and thus $\in_{S(n)}^T$ connects $S(n)$.

Now suppose a set B is such that $B \subseteq S(n)$ and $B \neq \emptyset$. If $n \notin B$ then $B \subseteq n$, and so B has an \in_n^T -least element because n is a natural number. Then since $n \subset S(n)$, this \in_n^T -least element is also an $\in_{S(n)}^T$ -least element of B . On the other hand, if $n \in B$ then $(\forall x \in B)(x \neq n \implies x \in_{S(n)} n)$ and so n is the $\in_{S(n)}^T$ -least element of B . \square

Proposition 2.3.3. *Let n be a natural number, and suppose m is an ordinal such that $m < n$. Then m is a natural number.*

Proof. Theorem 1.2.10 implies \in_n^T well-orders m . Then, since $m \subset n$, we conclude that \in_m^T well-orders m . \square

Proposition 2.3.4. *Let n be a natural number such that $n \neq \emptyset$. Then $n = S(m)$ for some natural number m .*

Proof. We already know that $\bigcup n \leq n$, so $\bigcup n$ is a natural number by Proposition 2.3.3. We will show that $\bigcup n < n$, so n is not a limit ordinal by Lemma 2.2.6 and hence $S(\bigcup n) = n$.

Let m be the \in_n^T -least element of n , so $(\forall x \in n)(x \neq m \implies x \in m)$. The transitivity of m then yields $(\forall x \in n)(x \neq m \implies x \subset m)$, hence we conclude that $m = \bigcup n$. Therefore $\bigcup n \in n$ as desired. \square

Remark. This shows that all natural numbers are not limit ordinals, and all limit ordinals are not natural numbers.

Putting together Propositions 2.3.2, 2.3.3, and 2.3.4 allows us to conclude that all the natural numbers form a long chain of ordinals connected by the successor operator S , and that there are no breaks in the chain. We can thus define the familiar natural numbers as follows.

Definition 2.3.5. We define the natural numbers $0, 1, 2, 3, \dots$ as follows:

$$0 := \emptyset, \quad 1 := S(0), \quad 2 := S(1), \quad 3 := S(2), \quad \dots$$

It is also convenient to have the set ω of all natural numbers, whose existence can be proven by using the axiom schema of separation on the set obtained from the axiom of infinity.

Definition 2.3.6. We define the set of all natural numbers as $\omega := \{n : n \text{ is a natural number}\}$.

Remark. It is clear that ω is an ordinal. In fact, ω is a limit ordinal. This is because if $S(\alpha) = \omega$ for some ordinal α then $\alpha < \omega$, and so $\alpha \in \omega$, meaning α is a natural number. But then Proposition 2.3.2 implies that ω is a natural number, and hence $\omega \in \omega$, contradicting Theorem 1.1.18. As a consequence, we conclude that ω is not a natural number. We call ordinals which are not natural numbers *transfinite ordinals*.

3 Using Ordinal Numbers

Now that we have developed ordinal numbers, we will see some of their applications in mathematics. We will develop the required tools to define arithmetic operations like addition by a recursion scheme. For example, we intend evaluate $3 + 2$ as follows:

$$3 + 2 = 3 + (1 + 1) = (3 + 1) + 1 = 4 + 1 = 5.$$

We will then see how ordinal numbers can be used to produce two startling results — the well-ordering theorem and Goodstein’s theorem.

To avoid the repetition of the predicate “is an ordinal”, we will use lower case Greek letters $(\alpha, \beta, \gamma, \dots)$ to denote ordinals.

3.1 Transfinite Induction and Transfinite Recursion

We begin by proving the principle of transfinite induction, which is a mathematical induction technique for statements about ordinal numbers. We provide several formulations of transfinite induction, as found in Suppes 1960 [16, pp. 195–197], each with their own quirk.

The first formulation of transfinite induction is most commonly used [16, pp. 205–224]. It consists of three parts: a base case, a weak inductive step, and a strong inductive step for limit ordinals.

Theorem 3.1.1 (Transfinite Induction: Formulation 1). *Let $\varphi(x)$ be a logical statement depending on the variable x , and suppose*

- $\varphi(0)$ is true,
- $(\forall \alpha)(\varphi(\alpha) \implies \varphi(S(\alpha)))$, and
- for all limit ordinals α , we have $(\forall \beta < \alpha)(\varphi(\beta)) \implies \varphi(\alpha)$.

Then $(\forall \alpha)(\varphi(\alpha))$.

Proof. Suppose, for a contradiction, $(\exists \alpha)(\neg \varphi(\alpha))$. Define

$$L_\alpha := \{\beta \in S(\alpha) : \neg \varphi(\beta)\}.$$

Observe that $\emptyset \neq L_\alpha \subseteq S(\alpha)$. Hence, by Theorem 1.2.10, $\in_{S(\alpha)}$ well-orders L_α and so by Theorem 1.2.11 there is a unique $\in_{S(\alpha)}$ -least element of L_α by Theorem 1.2.11. Let β_* be the $\in_{S(\alpha)}$ -least element of L_α .

Suppose β_* is not a limit ordinal. If $\beta = 0$ then this contradicts the first part of the hypothesis. So we must have $(\exists \gamma)(S(\gamma) = \beta_*)$. But now, due to β_* being the least element, we have $\varphi(\gamma)$. Therefore the second part of the hypothesis yields $\varphi(\beta_*)$, contradicting the definition of β_* .

On the other hand, if β_* is a limit ordinal, then for all ordinals $\delta < \beta_*$ we also have $\varphi(\delta)$. The third part of the hypothesis then yields $\varphi(\beta_*)$, also contradicting the definition of β_* . \square

The second formulation of transfinite induction is analogous to regular strong induction for natural numbers. That is, if we can perform the strong inductive step for any ordinal, then the inductive hypothesis holds for all ordinals. Its proof is identical to the proof for the limit ordinal case in the first formulation.

Theorem 3.1.2 (Transfinite Induction: Formulation 2). *Let $\varphi(x)$ be a logical statement depending on the variable x , and suppose*

$$(\forall \alpha)((\forall \beta < \alpha)(\varphi(\beta)) \implies \varphi(\alpha)).$$

Then $(\forall \alpha)(\varphi(\alpha))$.

The third formulation of transfinite induction is also similar to strong induction for natural numbers, however this time there is an “upper limit” to our inductive conclusion. This is useful because we can “stop” the induction once we reach a certain ordinal α , and the inductive hypothesis will be true for all ordinals $\beta < \alpha$. While this formulation is written in terms of sets instead of logical statements, the proof is similar to the limit ordinal case of the first formulation.

Theorem 3.1.3 (Transfinite Induction: Formulation 3). *Let A be a set, let α be an ordinal, and suppose*

$$(\forall \beta < \alpha)(\beta \subseteq A \implies \beta \in A).$$

Then $\alpha \subseteq A$.

We now introduce transfinite recursion, with which we will define ordinal arithmetic operations. As with transfinite induction, we will provide several formulations of transfinite recursion as found in Suppes 1960 [16, pp. 202–205]. We start with the first formulation, which recursively adds new mappings to a function based on the current function.

Theorem 3.1.4 (Transfinite Recursion: Formulation 1). *Let $\tau(x)$ be a term depending on the variable x , and let α be an ordinal. Then there exists a unique function F with $\text{Dom}(F) = \alpha$ satisfying*

$$(\forall \beta < \alpha)(F(\beta) = \tau(F|_\beta)).$$

Proof. For an ordinal $\beta < S(\alpha)$, define the logical statement $\varphi(\beta, f)$ to be true if and only if f is a function with $\text{Dom}(f) = \beta$ satisfying

$$(\forall \gamma < \beta)(f(\gamma) = \tau(f|_\gamma)).$$

Suppose $\varphi(\beta, f)$ and $\varphi(\beta, g)$ are both true, we will show by transfinite induction that $f = g$. Take $A := \{x \in \beta : f(x) = g(x)\}$, and suppose inductively that for any ordinal $\gamma < \beta$ we have $\gamma \subseteq A$. Then $f|_\gamma = g|_\gamma$ by the inductive hypothesis, and thus

$$f(\gamma) = \tau(f|_\gamma) = \tau(g|_\gamma) = g(\gamma),$$

giving us $\gamma \in A$. The third formulation of transfinite induction then yields $\beta \subseteq A$, and so we conclude $f = g$.

The above shows that φ satisfies the hypothesis of the axiom schema of replacement, so the axiom yields

$$(\exists B)(f \in B \iff (\exists \beta < S(\alpha))(\varphi(\beta, f))).$$

A corollary of the transfinite induction above is that for all $f, g \in B$, we have $f \subseteq g$ or $g \subseteq f$. Thus if we define $F := \bigcup B$, then F is a function with the property $(\forall \beta \in \text{Dom}(F))(F(\beta) = \tau(F|_\beta))$.

Now define $L_\alpha := \{\beta \in \alpha : \beta \notin \text{Dom}(F)\}$ and suppose, for a contradiction, that L_α is non-empty. Then, letting β_* be the \in_α -least element of L_α , there must be a function in B with the desired recursion property for all ordinals less than β_* , i.e.

$$(\exists f \in B)(\text{Dom}(f) = \beta_*).$$

But then $f \cup (\beta_*, \tau(f)) \in B$, which gives the contradiction $\beta_* \in \text{Dom}(F)$.

Therefore the function $F := \bigcup B$ satisfies the desired properties, proving the existence claim of the theorem. The uniqueness claim is then proven using third formulation of transfinite induction similarly as above. \square

The second formulation of transfinite recursion will recursively use the actual values of the function evaluated at smaller ordinals to define new mappings. The idea is to replace $F(\beta) = \tau(F|_\beta)$ with $F(S(\beta)) = \tau(F(\beta))$ in the first formulation. However, we must then create a separate case for limit ordinals in our recursive scheme.

Theorem 3.1.5 (Transfinite Recursion: Formulation 2). *Let $\tau(x)$ be a term depending on the variable x , let x_0 be any object, and let $\alpha \neq \emptyset$ be an ordinal. Then there exists a unique function F with $\text{Dom}(F) = \alpha$ satisfying*

- $F(0) = x_0$,
- for all ordinals β with $S(\beta) < \alpha$,

$$F(S(\beta)) = \tau(F(\beta)),$$

- for all limit ordinals β with $\beta < \alpha$,

$$F(\beta) = \bigcup_{\gamma < \beta} F(\gamma).$$

Proof. Define the function f_τ by

$$f_\tau := \{(\beta, \tau(\beta)) : \beta < \alpha\},$$

and let $I := \text{Im}(f_\tau)$. Then define the term $\sigma(x)$ by

$$\sigma(x) := \begin{cases} x_0 & \text{if } x = 0, \\ \tau(x(\beta)) & \text{if } \beta < \alpha \text{ and } x \in \mathcal{F}(S(\beta), I), \\ \bigcup_{\gamma < \beta} x(\gamma) & \text{if } \beta \text{ is a limit ordinal, } \beta < \alpha, \text{ and } x \in \mathcal{F}(\beta, I), \\ \emptyset & \text{otherwise.} \end{cases}$$

Then, by the first formulation of transfinite recursion, there exists a unique function F with $\text{Dom}(F) = \alpha$ satisfying

$$(\forall \beta < \alpha)(F(\beta) = \sigma(F|_\beta)).$$

Now, note that for all ordinals $\beta < \alpha$, $F|_\beta \in \mathcal{F}(\beta, I)$. Hence

- $F(0) = \sigma(F|_0) = x_0$,
- for all ordinals β with $S(\beta) < \alpha$

$$F(S(\beta)) = \sigma(F|_{S(\beta)}) = \tau(F|_{S(\beta)}(\beta)) = \tau(F(\beta)),$$

- for all limit ordinals β with $\beta < \alpha$,

$$F(\beta) = \sigma(F|_\beta) = \bigcup_{\gamma < \beta} F|_\beta(\gamma) = \bigcup_{\gamma < \beta} F(\gamma).$$

Thus F has the desired properties, proving the existence claim of the theorem.

We now use the first formulation of transfinite induction to prove the uniqueness claim. Suppose F and G are two functions with the desired properties. The base case of the induction holds because of the requirement $F(0) = G(0) = x_0$. Now suppose for all $\beta < \alpha$ we have $F(\beta) = G(\beta)$. Then, if $S(\beta) < \alpha$ (otherwise the weak inductive step vacuously holds), we have

$$F(S(\beta)) = \tau(F(\beta)) = \tau(G(\beta)) = G(S(\beta)).$$

Finally suppose for all limit ordinals $\beta < \alpha$ we have

$$(\forall \gamma < \beta)(F(\gamma) = G(\gamma)).$$

Then

$$F(\beta) = \bigcup_{\gamma < \beta} F(\gamma) = \bigcup_{\gamma < \beta} G(\gamma) = G(\beta).$$

Thus transfinite induction yields $F = G$. □

We should expect that the function F obtained from the second formulation of transfinite recursion is “independent” of the choice of the domain α . If we obtain functions F_1 and F_2 with different ordinal domains but the same term $\tau(x)$, then we should expect $F_1(\beta) = F_2(\beta)$ for all β that is in both domains. Transfinite induction easily affirms this.

Corollary. *Let $\tau(x)$ be a term depending on the variable x , and let α_1 and α_2 be ordinals. Suppose we obtain functions F_1 and F_2 from the second formulation of transfinite recursion on α_1 and α_2 respectively, but both using the same term $\tau(x)$ and starting object x_0 . Then for all $\beta < \alpha_1 \cap \alpha_2$, we have $F_1(\beta) = F_2(\beta)$.*

On the basis of this corollary and the second formulation of transfinite recursion, we obtain the following third formulation of transfinite recursion which we will subsequently use for all definitions of ordinal arithmetic operations. The reason this formulation uses terms instead of functions is because terms can be defined for all ordinals despite the set of all ordinals not existing, thus we would be allowed to define arithmetic operations on all ordinals.

Theorem 3.1.6 (Transfinite Recursion: Formulation 3). *Let $\sigma(\alpha_1, \dots, \alpha_n)$ and $\mu(\alpha_0, \alpha_1, \dots, \alpha_n)$ be terms depending on the ordinal variables $\alpha_1, \dots, \alpha_n$ and $\alpha_0, \alpha_1, \dots, \alpha_n$ respectively. Then the following definition of the term $\tau(\alpha_0, \alpha_1, \dots, \alpha_n)$ is well-defined:*

- $\tau(0, \alpha_1, \dots, \alpha_n) := \sigma(\alpha_1, \dots, \alpha_n)$,
- for all ordinals β ,

$$\tau(S(\beta), \alpha_1, \dots, \alpha_n) := \mu(\tau(\beta, \alpha_1, \dots, \alpha_n), \alpha_1, \dots, \alpha_n),$$

- for all limit ordinals β ,

$$\tau(\beta, \alpha_1, \dots, \alpha_n) := \bigcup_{\gamma < \beta} \tau(\gamma, \alpha_1, \dots, \alpha_n).$$

3.2 Well-Ordering Theorem

We can generalise the third formulation of transfinite induction to any well-ordered set, obtaining a general “strong principle of induction”.

Theorem 3.2.1 (Strong Principle of Mathematical Induction). *Let A and B be sets, let R be a relation which well-orders B , and suppose that*

$$(\forall y \in B)((\forall x \in B)(xRy \implies x \in A) \implies y \in A).$$

Then $B \subseteq A$.

A consequence of the axiom of choice is the well-ordering theorem, which states that every set can be well-ordered. In fact, the well-ordering theorem is equivalent to the axiom of choice [7, pp. 243 – 251]. In conjunction with the strong principle of mathematical induction, this shows that it is theoretically possible to perform strong induction on any set. That being said, such a relation may be incredibly difficult (or even impossible [10]) to explicitly construct within our framework, or be impractical to perform induction with. The following proof is adapted from both Jech 2003 [7, p. 48] and Suppes 1960 [16, pp. 241 – 242].

Theorem 3.2.2 (Well-Ordering Theorem). *For any set A , there exists a relation R on $A \times A$ which well-orders A .*

Proof. We will show that there exists an ordinal α and a function $f \in \mathcal{F}(\alpha, A)$ which is bijective, from which the theorem easily follows.

We already know that \emptyset is an ordinal, from Lemma 2.1.5. Thus \emptyset is well-ordered by \in_\emptyset . So let us assume that $A \neq \emptyset$.

Let $A_S := \mathcal{P}(A) \setminus \{\emptyset\}$ be the set of all non-empty subsets of A . Due to the axiom of choice, there exists a function $g \in \mathcal{F}(A_S, \bigcup A_S)$ such that

$$(\forall B \in A_S)(g(B) \in B).$$

Note that $\bigcup A_S = A$.

Now define the logical statement $\varphi(B, \beta)$ to be true if and only if β is an ordinal, $B \in A_S$, and

$$(\exists h \in \mathcal{F}(\beta, B)) (h \text{ is bijective} \wedge (\forall \gamma < \beta)(h(\gamma) = g(A \setminus \text{Im}(h|_\gamma)))).$$

If $\varphi(B, \beta)$ is true, and so such a function h exists, the first formulation of transfinite recursion says this h must be unique. Furthermore, notice that $h(0) = g(A)$. Together with the bijective requirement in φ , this yields

$$(\forall B \in A_S)(\forall \beta_1)(\forall \beta_2)(\varphi(B, \beta_1) \wedge \varphi(B, \beta_2) \implies \beta_1 = \beta_2).$$

Hence from the axiom schema of replacement,

$$(\exists C)(\forall \beta)(\beta \in C \iff (\exists B \in A_S)(\varphi(B, \beta))).$$

Now define $\alpha := \bigcup C$. We know that α is an ordinal, due to Proposition 2.2.2. Consider the function $f \in \mathcal{F}(\alpha, A)$ constructed using the first formulation of transfinite recursion with the recursion scheme

$$(\forall \delta < \alpha)(f(\delta) = g(A \setminus \text{Im}(f|_\delta))).$$

We will show that f is the desired bijection.

To show that f is injective, we proceed by the third formulation of transfinite induction. Let

$$I_f := \{\delta \in \alpha : (\nexists \varepsilon \in \alpha)(\varepsilon \neq \delta \wedge f(\delta) = f(\varepsilon))\}.$$

Suppose inductively that for all $\delta < \alpha$ we have $\delta \subseteq I_f$. Now for any ε such that $\delta < \varepsilon < \alpha$, we know that $f(\delta) \in \text{Im}(f|_\varepsilon)$. Hence

$$f(\varepsilon) = g(A \setminus \text{Im}(f|_\varepsilon)) \in A \setminus \text{Im}(f|_\varepsilon) \subseteq A \setminus \{f(\delta)\}.$$

Thus $f(\varepsilon) \neq f(\delta)$, giving us $\delta \in I_f$. Therefore transfinite induction yields $\alpha \subseteq I_f$, and so we conclude f is injective.

Finally, we show that $\text{Im}(f) = A$. Suppose, for a contradiction, that $A \setminus \text{Im}(f) \neq \emptyset$. Then $(\exists x \in A)(g(A \setminus \text{Im}(f)) = x)$. Now define the function $f' := f \cup \{(\alpha, x)\} \in \mathcal{F}(S(\alpha), A)$. Clearly f' , as a function in $\mathcal{F}(S(\alpha), \text{Im}(f'))$, is bijective. Hence $\varphi(\text{Im}(f'), S(\alpha))$ is true, and thus $S(\alpha) \in C$. However, since $\alpha = \bigcup C$, we obtain the absurd result $S(\alpha) \subseteq \alpha$.

We thus conclude that f is bijective. It is then clear that the relation R defined by

$$R := \{(f(\delta), f(\varepsilon)) : \delta < \varepsilon < \alpha\}$$

well-orders A . □

3.3 Basic Ordinal Arithmetic

Equipped with third formulation of transfinite recursion, we can now define the familiar basic arithmetic operations for ordinal numbers.

Definition 3.3.1 (Addition). For all ordinals α ,

- $\alpha + 0 := \alpha$,

- for all ordinals β ,

$$\alpha + S(\beta) := S(\alpha + \beta),$$

- for all limit ordinals β ,

$$\alpha + \beta := \bigcup_{\gamma < \beta} (\alpha + \gamma).$$

Definition 3.3.2 (Multiplication). For all ordinals α ,

- $\alpha \cdot 0 := 0$,

- for all ordinals β ,

$$\alpha \cdot S(\beta) := \alpha \cdot \beta + \alpha,$$

- for all limit ordinals β ,

$$\alpha \cdot \beta := \bigcup_{\gamma < \beta} (\alpha \cdot \gamma).$$

Remark. We will adopt the convention $\alpha\beta := \alpha \cdot \beta$.

Definition 3.3.3 (Exponentiation). For all ordinals α ,

- $\alpha^0 := 1$,

- for all ordinals β ,

$$\alpha^{S(\beta)} := \alpha^\beta \cdot \alpha,$$

- for all limit ordinals β ,

$$\alpha^\beta := \bigcup_{\gamma < \beta} (\alpha^\gamma).$$

If we restrict ourselves to natural numbers, we have all the usual arithmetic results for natural numbers. In particular, we may derive the Peano axioms [16, pp. 135–150] [11, pp. 27–70]. However, the story is different when we bring transfinite ordinals into the equation. For the complete development of ordinal arithmetic, we defer to Bachmann 1955 [2]. Several interesting ordinal arithmetic results are listed below, and are found in Suppes 1960 [16, pp. 205–224] and Jech 2003 [7, pp. 23–24]. We will omit their proofs, as most of them are simply proofs using the first formulation of transfinite induction.

Lemma 3.3.4. *Let α be an ordinal. Then $0 + \alpha = \alpha$ and $0 \cdot \alpha = 0$.*

Remark. Combined with the Definitions 3.3.1 and 3.3.2, we have $0 + \alpha = \alpha + 0 = \alpha$ and $0 \cdot \alpha = \alpha \cdot 0 = 0$ for all ordinals α .

Lemma 3.3.5. *Let α , β , and γ be ordinals. Then $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ and $(\alpha\beta)\gamma = \alpha(\beta\gamma)$. Furthermore, if $\beta < \gamma$ then $\alpha + \beta < \alpha + \gamma$ and $\alpha\beta < \alpha\gamma$.*

Lemma 3.3.6. *Let n be a natural number. Then $n + \omega = \omega$. Furthermore, if $n > 0$ then $n\omega = \omega$.*

Remark. From Definition 3.3.1, we obtain $\omega + 1 = S(\omega)$. Thus we have the intriguing result $1 + \omega = \omega < \omega + 1$. Also, from Definitions 3.3.2 and 3.3.1, we obtain $\omega \cdot 2 = \omega + \omega = \bigcup_{\gamma < \omega} (\omega + \gamma)$. Since $\omega \in \omega + 1$ we obtain $\omega \in \omega \cdot 2$. Thus we also have the intriguing result $2 \cdot \omega = \omega < \omega \cdot 2$. This shows that ordinal addition and multiplication is not commutative.

Lemma 3.3.7. *Let α , β , and γ be ordinals. If $\alpha + \beta = \alpha + \gamma$ then $\beta = \gamma$. Furthermore, if $\alpha > 0$ and $\alpha\beta = \alpha\gamma$ then $\beta = \gamma$.*

Remark. We can thus “cancel” addition and multiplication from the left, but not from the right due to Lemma 3.3.6.

Lemma 3.3.8. *Let α , β , and γ be ordinals. Then $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.*

Remark. This gives us the distributive law on the left. However, the distributive law on the right does not hold, since we know that $(1 + 1)\omega = 2\omega = \omega < \omega + \omega$.

Lemma 3.3.9. *Let α and β be ordinals with $\beta > 0$. Then there exist unique ordinals γ and δ such that $\delta < \beta$ and $\alpha = \beta\gamma + \delta$.*

Lemma 3.3.10. *Let α , β , and γ be ordinals. Then $\alpha^\beta \alpha^\gamma = \alpha^{\beta+\gamma}$ and $(\alpha^\beta)^\gamma = \alpha^{\beta\gamma}$. Furthermore, if $\alpha > 1$ and $\beta < \gamma$ then $\alpha^\beta < \alpha^\gamma$.*

We shall henceforth assume results from Peano and ordinal arithmetic, including the development of subtraction for natural numbers. The main result for this section is the following theorem, which states every ordinal number can be uniquely represented as a finite sum of multiples of powers of ω . The following proof is adapted from Jech 2003 [7, p. 24].

Theorem 3.3.11 (Cantor Normal Form). *Let α be an ordinal with $\alpha > 0$. Then there exist a unique natural number k , unique natural numbers n_1, \dots, n_k , and unique ordinals β_1, \dots, β_k such that*

- $k \geq 1$,
- $n_1, \dots, n_k > 0$,
- $\alpha \geq \beta_1 > \dots > \beta_k$, and
- $\alpha = \omega^{\beta_1} n_1 + \dots + \omega^{\beta_k} n_k$.

Remark. We call $\omega^{\beta_1} n_1 + \dots + \omega^{\beta_k} n_k$ the *Cantor normal form* of α .

Proof. We will use the second formulation of transfinite induction. If $\alpha = 0$ then the theorem holds vacuously. So suppose inductively that the theorem holds for all $\varepsilon < \alpha$, where $\alpha > 0$. Define $L_\alpha := \{\zeta \in \alpha : \omega^\zeta \leq \alpha\}$ and $\beta := \bigcup L_\alpha$, noting that $\beta \leq \alpha$.

If $\beta < \alpha$ then we must have $\beta \in L_\alpha$, otherwise if $\beta = L_\alpha$ then β is a limit ordinal and we obtain $\omega^\beta = \bigcup_{\varepsilon < \beta} \omega^\varepsilon \leq \alpha$ which yields the absurd $\beta \in L_\alpha = \beta$. On the other hand, if $\beta = \alpha$ then α is a limit ordinal and so $\omega^\alpha = \bigcup_{\varepsilon < \alpha} \omega^\varepsilon \leq \alpha$. In either case, we require $\omega^\beta \leq \alpha$.

From Lemma 3.3.9, there exist unique ordinals n and γ such that $\gamma < \omega^\beta$ and $\alpha = \omega^\beta n + \gamma$. Observe that we require $n \in \omega$, otherwise we can write $n = \omega + \delta$ for some ordinal δ and hence $\alpha = \omega^{\beta+1} + \omega^\beta \delta + \gamma$. However this implies $\beta + 1 \in L_\alpha$, contradicting the definition of β . Furthermore we require $n > 0$, otherwise we would obtain the contradiction $\alpha = \gamma < \omega^\beta$.

If $\gamma = 0$ then we have proven the existence claim of the theorem. Otherwise, we apply the inductive hypothesis onto γ to obtain its Cantor normal form $\gamma = \omega^{\beta_1} n_1 + \dots + \omega^{\beta_k} n_k$. Then noting that $\alpha \geq \beta > \beta_1$ completes the proof of the existence claim of the theorem.

For the uniqueness claim, suppose α has another Cantor normal form $\alpha = \omega^{\beta'_1} n'_1 + \dots + \omega^{\beta'_k} n'_k$. Suppose, for a contradiction and without loss of generality, that $\beta > \beta'_1$. Then, because both Cantor normal forms are finite sums, all n'_1, \dots, n'_k are natural numbers, and $\beta > \beta'_1 > \dots > \beta'_k$, we obtain the contradiction $\alpha = \omega^{\beta'_1} n'_1 + \dots + \omega^{\beta'_k} n'_k < \omega^\beta \leq \alpha$. Thus we require $\beta = \beta'_1$, which then clearly requires $n = n'_1$. Now as $\alpha = \omega^\beta n + \gamma$, applying the inductive hypothesis onto γ to obtain the uniqueness of its Cantor normal form completes the proof. \square

3.4 Goodstein's Theorem

This final subsection is devoted to stating and proving Goodstein's theorem [6], a theorem about natural numbers which is unprovable from the Peano axioms alone [8].

Theorem 3.3.11 can be generalised to to any base $b \geq 2$ instead of ω . This generalisation is found in Takeuti and Zaring 1982 [17, pp. 70–71], and its proof is similar to the proof of Theorem 3.3.11.

Theorem 3.4.1 (Basis Representation Theorem). *Let α and β be an ordinals with $\alpha > 0$ and $\beta \geq 2$. Then there exist a unique natural number k and unique ordinals $\gamma_1, \dots, \gamma_k, \delta_1, \dots, \delta_k$ such that*

- $k \geq 1$,
- $0 < \delta_1, \dots, \delta_k < \beta$,
- $\alpha \geq \gamma_1 > \dots > \gamma_k$, and
- $\alpha = \beta^{\gamma_1} \delta_1 + \dots + \beta^{\gamma_k} \delta_k$.

Remark. We call $\beta^{\gamma_1} \delta_1 + \dots + \beta^{\gamma_k} \delta_k$ the *base- β representation* of α .

We can recursively use Theorem 3.4.1 on the exponents of basis representations for natural numbers for a base b (where b is a natural number) to obtain a representation using only the numbers in $\{0, 1, \dots, b\}$.

Definition 3.4.2 (Hereditary Base- b Representation). Let m and b be natural numbers with $b \geq 2$. The *hereditary base- b representation* of m is a base- b representation of m with all its exponents (and their exponents, and so on) written in base- b representation.

Remark. We only convert all the digits appearing in the representation which are not in the set $\{0, 1, \dots, b\}$. Hence if 1 appears in the exponent, we do not convert it into b^0 .

Example 3.4.3. The base-3 representation of 250 is $250 = 3^5 + 3^1 \cdot 2 + 3^0$. The hereditary base-3 representation of 250 is $250 = 3^{3^1+3^0 \cdot 2} + 3^1 \cdot 2 + 3^0$.

Definition 3.4.4 (Hereditary Base Change Function). Let b and b' be natural numbers with $b, b' \geq 2$. Define the function $\text{HBC}_{b,b'} \in \mathcal{F}(\omega \setminus \{0\}, \omega \setminus \{0\})$ as follows: for all $1 \leq m < \omega$ define $\text{HBC}_{b,b'}(m)$ to be the natural number obtained by replacing all occurrences of b with b' in the hereditary base- b representation of m .

Example 3.4.5. $\text{HBC}_{3,5}(250) = \text{HBC}_{3,5} \left(3^{3^1+3^0 \cdot 2} + 3^1 \cdot 2 + 3^0 \right) = 5^{5^1+5^0 \cdot 2} + 5^1 \cdot 2 + 5^0 = 78136$.

Definition 3.4.6 (Goodstein Sequences). Let m be a natural number. Define the function $G_m \in \mathcal{F}(\omega, \omega)$ by

- $G_m(0) := 0$,
- $G_m(1) := m$, and
- for all $n \in \omega \setminus \{0\}$,

$$G_m(n+1) := \begin{cases} \text{HBC}_{n+1,n+2}(G_m(n)) - 1 & \text{if } G_m(n) > 0, \\ 0 & \text{if } G_m(n) = 0. \end{cases}$$

The sequence $(G_m(1), G_m(2), G_m(3), \dots)$ is the *Goodstein sequence* starting with m .

Remark. This function G_m exists and is unique due to the second formulation of transfinite recursion.

Preliminary calculations of some Goodstein sequences show that the terms in Goodstein sequences get very large very rapidly [3] [14]. For example, consider the Goodstein sequence starting with 16 [1].

n	$G_{16}(n)$
1	16
2	7625597484986
3	50973998591214355139406377
4	53793641718868912174424175024032593379100060
5	19916489515870532960258562190639398471599239042185934648024761145811
\vdots	\vdots

Goodstein's theorem is the startling result that any Goodstein sequence will eventually hit 0 (and subsequently stay at 0). The following proof is adapted from Rathjen 2015 [12, pp. 229–242].

Theorem 3.4.7 (Goodstein's Theorem). *Let m be a natural number. Then there exists a natural number $N \geq 2$ such that $G_m(N) = 0$.*

Proof. For natural numbers n and b with $b \geq 2$, define the term $\tau(b, n)$ to be the ordinal obtained by replacing b with ω in the hereditary base- b representation of n . Note that $\tau(b, n)$ acts as a well-defined version of $\text{HBC}_{b,\omega}(n)$, giving us ordinals in Cantor normal form.

Next, use the second formulation of transfinite recursion to define the function P_m with $\text{Dom}(P_m) = \omega$ by

- $P_m(0) := 0$, and
- for all $n \in \omega \setminus \{0\}$,

$$P_m(n) := \begin{cases} \tau(n+1, G_m(n)) & \text{if } G_m(n) > 0, \\ 0 & \text{if } G_m(n) = 0. \end{cases}$$

Let $n \in \omega \setminus \{0\}$ and suppose that $G_m(n) > 0$ and $G_m(n+1) > 0$. Notice that we have

$$P_m(n+1) = \tau(n+2, G_m(n+1)) < \tau(n+2, G_m(n+1) + 1). \quad (1)$$

Furthermore, notice that

$$P_m(n) = \tau(n+1, G_m(n)) = \tau(n+2, G_m(n+1) + 1). \quad (2)$$

Combining (1) and (2), we conclude that $P_m(n) > P_m(n+1)$ for all $n \in \omega \setminus \{0\}$ which satisfy $G_m(n) > 0$ and $G_m(n+1) > 0$.

Now suppose, for a contradiction, that $(\forall n \in \omega \setminus \{0\})(G_m(n) > 0)$. Then we must have $P_m(n) > P_m(n+1)$ for all $n \in \omega \setminus \{0\}$. Define the set

$$A := \{P_m(n) : n \in \omega \setminus \{0\}\}.$$

Observe that $\emptyset \neq A \subseteq S(P_m(2))$. Thus by Theorem 1.2.10 and Theorem 1.2.11 we know that A has a unique $\in_{S(P_m(2))}$ -least element, say α_0 . Then $(\exists n \in \omega \setminus \{0\})(\alpha_0 = P_m(n))$. However we then obtain $P_m(n+1) < \alpha_0$, which contradicts α_0 being the $\in_{S(P_m(2))}$ -least element of A . \square

References

- [1] Goodstein sequence starting with 16. The On-Line Encyclopedia of Integer Sequences. URL <https://oeis.org/A059933>.
- [2] H. Bachmann. *Transfinite Zahlen*. Springer-Verlag, 1955.
- [3] A. Caicedo. Goodstein's function. *Revista Colombiana de Matemáticas*, 41(2):381–391, 2007.
- [4] K. Ciesielski. *Set Theory for the Working Mathematician*. Cambridge University Press, 1997.
- [5] A. Fraenkel. Zu den grundlagen der cantor-zermeloschen mengenlehre. *Mathematische Annalen*, 86:230–237, 1953.
- [6] R. L. Goodstein. On the restricted ordinal theorem. *The Journal of Symbolic Logic*, 9(2):33–41, 1944.
- [7] T. Jech. *Set Theory*, volume 14. Springer, 3rd edition, 2003.
- [8] L. Kirby and J. Paris. Accessible independence for peano arithmetic. *Bulletin of London Mathematical Society*, 14(4):285–293, 1982.
- [9] J. Lasseter. Toy story. United States of America: Pixar Animation Studios, 1995.
- [10] C. Mummert (https://math.stackexchange.com/users/630/carl_mummert). Is there a known well ordering of the reals? Mathematics Stack Exchange. URL <https://math.stackexchange.com/q/6504>. URL : <https://math.stackexchange.com/q/6504> (version: 2010-10-11).
- [11] G. Peano. *Formulario Mathematico*. Fratres Bocca, 5th edition, 1908.
- [12] M. Rathjen and R. Kahle. *Gentzen's Centenary: The Quest for Consistency*. Springer, 2015.
- [13] B. Russell. *The Principles of Mathematics*. W. W. Norton and Company, 2nd edition, 1938.
- [14] S. G. Simpson. Unprovable theorems and fast-growing functions. *Contemporary Mathematics*, 65:359–394, 1987.
- [15] A. Stevenson and M. Waite. *Concise Oxford English Dictionary*. Oxford University Press, 12th edition, 2011.
- [16] P. Suppes. *Axiomatic Set Theory*. D. Van Nostrand Company, 1st edition, 1960.
- [17] G. Takeuti and W. M. Zaring. *Introduction to Axiomatic Set Theory*. Springer-Verlag New York Inc., 2nd edition, 1982.
- [18] A. Whitehead and B. Russell. *Principia Mathematica*, volume 1. Cambridge University Press, 1910.
- [19] E. Zermelo. Untersuchungen über die grundlagen der mengenlehre. i. *Mathematische Annalen*, 65(2):261–281, 1908.