Set Theory

Robin Truax

April 2020

Contents

1 First Definitions

In these notes, we will discuss set theory using the language of ZFC. Set theory, alongside basic logic, is commonly seen as the basis of modern mathematics. However, in some sense, we have to "believe" that this system works – there is no way to find any reasonable consistent axiomatic system which is either complete or provably consistent. Following are informal descriptions of Gödel's famous incompleteness theorems, which formalize this sentiment:

Theorem 1.0.1 (Gödel's first incompleteness theorem): Let T be a set of axioms expressed in a formal language $\mathcal{L}$ such that (i) T is consistent, (ii) there is an effective algorithm that decides for an arbitrary sequence of the language $\mathcal{L}$ whether it is in T or not, and (iii) T contains the arithmetic of the natural numbers. Then there is a sentence φ of the language $\mathcal{L}$ such that neither φ nor its negation $\neg\varphi$ can be deduced from the axioms T.

Theorem 1.0.2 (Gödel's second incompleteness theorem): Let T be a set of axioms with the same properties as above. Then there is a sentence φ of the language $\mathcal{L}$ that encodes the statement "T is consistent", but φ is not a consequence of the axioms T.

We do not shy away from this failure: the author considers worrying about possible esoteric contradictions akin to solipsism. It is pointless to worry about whether or not we live in a perfectly made simulation, since we cannot prove it one way or another. Thus, unless we get an actual example of a contradiction that irreparably ruins ZFC set theory, we will continue to make use of the powerful tool that is set theory.

1.1 Primitive Notions

We begin with two primitive notions: the *set* (which is either the empty set or an object containing other sets) and the *class*, which is an intuitive concept meant to give us a term for those objects "too large" to be sets.

The *class* is not a part of ZFC, and we will not use it in any rigorous or formal capacity (though one can in other axiomatic systems such as NBG), but it is a helpful linguistic tool.

Here, we will usually use English letters (fonts include a, A, $\mathcal{A}$, or $\mathscr{A}$) to denote sets, and Greek letters (such as φ or γ) to denote formulas. Furthermore, we will freely use the following logical symbols:

- $\forall$ means "for all", $\exists$ means "there exists" (and $!\exists$ means "there exists a unique").
- $\in$ means "is a member of" and $\subseteq$ means "is a subset of"
- $\Rightarrow$ means "implies", $\Leftarrow$ means "is implied by", and $\Leftrightarrow$ means "is logically equivalent to".
- "And", "or", and "not" are all standard logical operators, and we may replace them with the symbols $\wedge$, $\vee$, and $\neg$, respectively.

1.2 The Zermelo-Fraenkel Axioms

Axiom 1 (The Axiom of Extensionality)**.** *Two sets are equal if they have the same members.*

$$\forall A\, \forall B\, [\forall x(x \in A \Leftrightarrow x \in B) \Rightarrow A = B] \tag{1}$$

Axiom 2 (The Pairing Axiom)**.** *Given two sets u and v, there exists a set $\{u, v\}$.*

$$\forall u \forall v\, \exists B\, \forall x(x \in B \Leftrightarrow x = u \textit{ or } x = v) \tag{2}$$

Axiom 3 (The Union Axiom)**.** *Given any set A, there is a set B whose elements are the members of the members of A. In this case, we write $B = \bigcup A$.*

$$\forall A\, \exists B\, \forall x(x \in B \Leftrightarrow \exists b \in A(x \in b)) \tag{3}$$

Definition 1.1 (Power Set)**.** The *power set* of a set S is the collection of subsets of S and is denoted $\mathcal{P}(S)$. Clearly, if S has finitely many elements, $|\mathcal{P}(S)| = 2^{|S|}$.

Axiom 4 (The Axiom of the Power Set)**.** *The power set $\mathcal{P}(S)$ of any set S is a set.*

$$\forall a\, \exists B(x \in B \Leftrightarrow x \subseteq a) \tag{4}$$

Axiom 5 (The Axiom Schema of Specification)**.** *For each formula* ___ *not containing* B*,*

$$\forall t_1 \dots \forall t_n \, \forall c \, \exists B \, \forall x (x \in B \Leftrightarrow x \in c \textit{ and } \_\_\_) \tag{5}$$

is an axiom.

Definition 1.2 (Successor, Inductive Sets)**.** Let the *successor* of a set a, denoted a^+, be $a \cup \{a\}$. We call a set S *inductive* if $\varnothing \in S$ and $\forall(a \in A)a^+ \in A$.

Axiom 6 (The Axiom of Infinity)**.** *There exists an inductive set. Formally:*

$$\exists A[\varnothing \in A \textit{ and } \forall(a \in A)a^+ \in A] \tag{6}$$

Axiom 7 (The Axiom of Replacement)**.** *If* $\varphi(x, y)$ *is a formula not containing* B*, the following is an axiom:*

$$\forall A[(\forall x \in A)\forall y_1, \forall y_2(\varphi(x, y_1) \textit{ and } \varphi(x, y_2) \Rightarrow y_1 = y_2) \Rightarrow \exists B \forall y(y \in B \Leftrightarrow (\exists x \in A)\varphi(x, y))] \tag{7}$$

In English, this axiom says that if we have a *function-class* $\mathbf{H} = \{\langle x, y\rangle \mid x \in A \text{ and } \varphi(x, y)\}$, then $B = \mathbf{H}|_A$ (the image of B under $\mathbf{H}$) is a set.

Axiom 8 (The Regularity Axiom)**.** *Every nonempty set* A *has a member* m *with* $m \cap A = \varnothing$*.*

Theorem 1.1 (Russell's Paradox)**.** *There is no set to which every set belongs.*

Proof. Let A be such a set. Then, take $B = \{x \mid x \in A \text{ and } x \notin x\}$. By the construction of B, $B \in B \Leftrightarrow B \in A$ and $B \notin B$. Since it cannot be the case that $B \in B$ and $B \notin B$, $B \notin A$, a contradiction. □

Theorem 1.2 (Intersections)**.** *For any nonempty set* A*, there is a unique set* B *such that* $x \in B \Leftrightarrow \forall a \in A, x \in a$*. We write* $B = \bigcap A$ *and say* B *is the intersection of all sets in* A*.*

Theorem 1.3 (Empty Set)**.** *The empty set exists and is unique.*

Proof. Consider the set A defined in the axiom of infinity. Then consider the subset $B \subseteq A$ given by all elements $x \in A$ such that $x \neq x$. Clearly, no such set exists, so B contains no elements. The uniqueness of B follows from the axiom of extensionality. □

Thus, we may denote the empty set by $\varnothing$. On the other hand, there are infinitely many sets with one element, which we call the "singletons".

Definition 1.3 (Difference and Symmetric Difference)**.** The *difference* $A - B$ is the set of all elements in A not in B. The *symmetric difference* of two sets A and B is denoted by $A \triangle B$ and defined as the union $(A - B) \cup (B - A)$.

Theorem 1.4 (Foundation)**.** *No set is a member of itself.*

Theorem 1.5 (Laws on the Algebra of Sets)**.**

1. $A \cap B = B \cap A$ *and* $A \cup B = B \cup A$ *(commutativity)*
2. $A \cup (B \cup C) = (A \cup B) \cup C$ *and* $A \cap (B \cap C) = (A \cap B) \cap C$ *(associativity)*
3. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ *and* $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ *(distributivity)*
4. $C - (A \cup B) = (C - A) \cap (C - B)$ *and* $C - (A \cap B) = (C - A) \cup (C - B)$ *(de Morgan's laws)*
5. $A \cup \varnothing = A$*,* $A \cap \varnothing = \varnothing$*, and* $A \cup (C - A) = \varnothing$ *(identities on* $\varnothing$*)*

Proof: The proof of these rules is left as an exercise for the reader. A hint: use Venn diagrams to visualize each rule and then formalize your results.

1.3 Ordered Tuples, Relations, and Functions

Definition 1.4 (Ordered Pair)**.** An *ordered pair* of sets x, y (denoted $\langle x, y\rangle$) is the set $\{x, \{x, y\}\}$. Notice that $\langle x, y\rangle = \langle u, v\rangle$ if and only if $x = u$ and $y = v$.

Definition 1.5 (Cartesian Product)**.** The *Cartesian product* of two sets A and B is the set of all ordered pairs of elements $\langle x, y\rangle$ with $x \in A, y \in B$. Formally, we write $A \times B = \{\langle x, y\rangle \mid x \in A, y \in B\}$.

To prove that $A \times B$ is actually a set, notice that $C \times C$ is contained in the set $\mathcal{P}(\mathcal{P}(C))$ and then realize that $A \times B$ is contained in the set $\mathcal{P}(\mathcal{P}(A \cup B))$.

Definition 1.6 (Relation)**.** A *relation* is a set of ordered pairs. Instead of writing $\langle x, y\rangle \in R$, we write xRy.

For example, let ω be the set of natural numbers (ignore that we haven't defined that yet) and $<$ be the relation $\{\langle x, y\rangle \in \omega \times \omega \mid x \text{ is less than } y\}$, Then, $\langle 2, 3\rangle \in <$ so we write $2 < 3$.

Definition 1.7 (Properties of Relations)**.** Following are some possible properties of relations:

- A relation is *symmetric* if $xRy \Leftrightarrow yRx$ and *anti-symmetric* if xRy and $yRx \Rightarrow x = y$.
- A relation is *reflexive* if xRx always holds, and *irreflexive* if xRx *never* holds.
- A relation is *connected* if for any x, y such that $x \neq y$, either xRy or yRx holds.
- A relation is *tricohotomous* (or *satisfies trichotomy*) if, for any $x, y \in R$, exactly one of $xRy, x = y, yRx$ holds.

Definition 1.8 (Domain, Range, Field)**.** The *domain* of R, denoted $\operatorname{dom} R$, are those x such that $\exists y \, \langle x, y\rangle \in R$. Conversely, the *range* of R, denoted $\operatorname{ran} R$, are those y such that $\exists x \, \langle x, y\rangle \in R$. Finally, the *field* of R, $\operatorname{fld} R$, is $\operatorname{dom} R \cup \operatorname{ran} R$.

Again, we cannot immediately assume that $\operatorname{dom} R$ and $\operatorname{ran} R$ are actually sets, but it is not hard to prove that they are, as Lemma 1.6 implies that $\operatorname{dom} R$ and $\operatorname{ran} R$ are contained in a set.

Lemma 1.6. *If $\langle x, y\rangle \in R$, then x, y belong to $\bigcup\bigcup A$.*

Definition 1.9 (n-Tuples)**.** Let $\langle x\rangle := x$ be a 1-tuple and then define $\langle x_1, \ldots, x_n\rangle := \langle\langle x_1, \ldots, x_{n-1}\rangle, x_n\rangle$ inductively as an n-tuple. For example, $\langle x, y, z\rangle = \langle\langle x, y\rangle, z\rangle$. For each n, two n-tuples are equal if and only if each of their ith entries, for every i, are equal.

Definition 1.10 (n-ary Relation)**.** An *n-ary relation* on a set A is a set of ordered n-tuples with all elements in A. In other words, a binary relation on A is a subset of $A \times A$, a ternary relation on A is a subset of $(A \times A) \times A$, and so on.

Definition 1.11 (Function)**.** A *function* is a relation F such that, for any x in $\operatorname{dom} F$, there is exactly one y such that $\langle x, y\rangle \in F$. In this case, we call y the *value of F at x*, and write $F(x) = y$.

Definition 1.12 (Properties of Functions)**.**

- We say that F is a function from A to B (and write $F : A \to B$) if $\operatorname{dom} F = A$ and $\operatorname{ran} F \subseteq B$. In this case, we call B the *codomain* of F.
- Furthermore, if $\operatorname{ran} F = B$, we call the function *onto* or *surjective* on B, and say F is a function from A *onto* B.
- If, for every $y \in \operatorname{ran} F$, there is only one x such that $F(x) = y$, we call the function *one-to-one* or *injective.* We generalize this to the concept of *single-rooted relations* when F is not a function.
- A function which is surjective and injective is called *bijective* or a *one-to-one correspondence.*

Theorem 1.7 (Objects Formed From Relations)**.** *For any relation F (not just functions), the following objects are sets:*

- *The set* $F^{-1} = \{\langle u, v\rangle \mid vFu\}$*, called the inverse of* F*.*
- *The set* $F \circ G = \{\langle u, v\rangle \mid \exists t(uGt \;\&\; tFv)$*, called the composition of* F *and* G*.*
- *The set* $F|_A = \{\langle u, v\rangle \in F \mid u \in A\}$*, called the restriction of* F *to* A*.*
- *The set* $\operatorname{ran}(F|_A)$*, called the image of* A *under* F*.*

Theorem 1.8 (Properties of These Objects)**.**

- *For any relation* F*,* $\operatorname{dom} F^{-1} = \operatorname{ran} F$ *and* $\operatorname{ran} F^{-1} = \operatorname{dom} F$*.*
- *For any relation* F*,* $\left(F^{-1}\right)^{-1} = F$*.*
- *For any relation* F*,* F^{-1} *is a function if and only if* F *is single-rooted.*
- *If* F *is a one-to-one function,* $x \in \operatorname{dom} F \Rightarrow F^{-1}F(x)) = x$ *and* $y \in \operatorname{ran}(F) \Rightarrow F(F^{-1}(y)) = y$*.*
- *If* F *and* G *are functions,* $F \circ G$ *is a function with domain* $\{x \in \operatorname{dom} G \mid G(x) \in \operatorname{dom} F\}$*.*
- *For any relations* F, G*,* $(F \circ G)^{-1} = G^{-1} \circ F^{-1}$*.*

1.4 Equivalence Relations, Unions, and Products

Definition 1.13 (Equivalence Relation)**.** We call a relation R an *equivalence relation on* A if R is a binary relation on A that is *reflexive*, *symmetric*, and *transitive*.

Definition 1.14 (Equivalence Class)**.** The set $[x]_R$ is defined by $[x]_R = \{t \mid xRt\}$, and is called the *equivalence class of* x *modulo* R (if R is an equivalence relation).

Lemma 1.9. *If* R *is an equivalence relation on* A *with* $x, y \in A$*, then* $[x]_R = [y]_R$ *iff* xRy*.*

Definition 1.15. A *partition* Π of a set A is a set of nonempty subsets of A that is *disjoint* and *exhaustive*. In other words, no two distinct sets in Π have any common elements, and each element of A is in some set of Π.

Theorem 1.10 (Partitions and Equivalence Relations)**.** *There is a natural bijective correspondence between partitions and equivalence relations, namely:*

- *If* R *is an equivalence relation on* A*, then the set* $\{[x]_R \mid x \in A\}$ *is a partition of* A*.*
- *If* Π *is a partition of* A*, then the relation* $R = \{\langle x, y\rangle \mid \exists p \in \Pi, x, y \in p\}$ *is an equivalence relation on* A*.*

Definition 1.16 (Quotient Set and Natural Map)**.** If R is an equivalence relation on A, then we can define the *quotient set* $A/R = \{[x]_R \mid x \in A\}$ whose members are the equivalence classes. The expression A/R is read A modulo R, and gives us the *natural map* or *canonical map* $\varphi : A \to A/R$ given by $\varphi(x) = [x]_R$.

Definition 1.17 (Compatibility)**.** A function $F : A \to A$ is compatible with an equivalence relation R on A if, for all x and y in A, $xRy \Rightarrow F(x)RF(y)$. In this case, there exists a unique $\tilde{F} : A/R \to A/r$ such that $\tilde{F}([x]_R) = [F(x)]_R$.

Definition 1.18 (Indexing an Infinite Union)**.** Let I be a set, called an *index set*. Further let F be a function whose domain includes I. Then we define

$$\bigcup_{i\in I} F(i) = \bigcup \{F(i) \mid i \in I\} \tag{8}$$

Definition 1.19 (General Cartesian Products)**.** Let I be our index set and let H be a function whose domain includes I. Then, we define:

$$\prod_{i\in I} H(i) = \{f \mid f \text{ is a function with domain } I \text{ and } (\forall i \in I) f(i) \in H(i)\} \tag{9}$$

Thus the members of $\prod_{i\in I} H(i)$ are $|I|$-tuples for which the ith coordinate is in $H(i)$.

1.5 Orderings and Isomorphisms

Definition 1.20 (Partial Ordering)**.** A *partial ordering* on a set S is a relation R such that R is transitive, reflexive, and antisymmetric.

Definition 1.21 (Linear Ordering)**.** Let A be any set. A *linear ordering* on A (also called a *total ordering on* A) is a binary relation R on A which is transitive and trichotomous on A. Any linear ordering is irreflexive and connected.

Theorem 1.11. *Any partial ordering R on X can be extended to a linear ordering R' on X such that $R \subseteq R'$.*

It should be clear why we call such a ordering *linear* – we can never go in a circle without contradicting trichotomy.

Definition 1.22 (Structure)**.** A *structure* is a pair $\langle A, R\rangle$ consisting of a set A and a binary relation $R \subseteq A \times A$ on A. In particular, if R is a partial ordering, we call the structure a *partially ordered structure* or *poset*, and if it is a linear ordering, we call the structure a *linearly ordered structure* or *loset.*

Definition 1.23 (Minimal/Maximal/Least/Greatest Elements)**.** A *minimal element* in a set S with a partial order $<$ is an element m such that there is no $x \in S$ with $x < m$. In this case, we say that m is $<$-minimal on S. Similarly, we call m a *least element* if $m \leq x$ for all $x \in S$.

A *maximal element* in a set S with a partial order $<$ is an element m such that there is no $x \in S$ with $m < x$. In this case, we say that m is $<$-maximal on S. Similarly, we call m a *greatest element* if $x \leq m$ for all $x \in S$.

Clearly, a least/greatest element is also minimal/maximal, but if our relation is not linear, then "least/greatest" is a stronger condition.

Definition 1.24 (Upper/Lower Bound, Supremum, Infimum)**.** An *upper bound* on a subset $A \subseteq S$ with partial ordering $<$ is an element $s \in S$ such that $x \leq b$ for all $x \in A$. A *supremum* is a *least upper bound* – i.e. a least element of the set of all upper bounds for C (if it exists). We analogously define *lower bound* and *greatest lower bound/infimum.*

Definition 1.25 (Well-Founded)**.** A relation R is *well-founded* on D if every nonempty set D contains an R-minimal element. Clearly if $D \not\subseteq \operatorname{fld} R$, then any $m \in D - \operatorname{fld} R$ is an R-minimal element, so we can restrict our attention to subsets of $\operatorname{fld} R$.

For example, the regularity axiom implies that for any set S, the membership relation $\in_S = \{\langle x, y\rangle \in S \times S \mid x \in y\}$ is well-founded on S. In fact, one can prove that the regularity axiom is *equivalent* to the theorem that any nonempty set has an $\in$-minimal element.

Theorem 1.12 (Well-Founded $\Leftrightarrow$ No Descending Chains)**.** *A relation R is well-founded iff there is no function f with domain ω such that $f(n^+)Rf(n)$ – i.e. it has no infinitely descending chains.*

Proof. If R is not well founded, then there exists a nonempty set A without an R-minimal element. Thus, define $f(0)$ to be any element in A, and $f(n^+)$ to be the element y such that $yRf(n)$ (which we know exists by the lack of a minimal element), giving us an infinitely descending chain. The opposite direction is similarly easy. □

Definition 1.26 (Well Ordering)**.** A *well ordering* on S is a well-founded linear ordering on S.

Definition 1.27 (Isomorphism)**.** An *isomorphism* from a structure $\langle A, R\rangle$ onto $\langle B, S\rangle$ is a bijection f from $A \to B$ such that xRy if and only if $f(x)Sf(y)$. In this case, we call the two structures *isomorphic.* This is an equivalence relation on the class of sets.

Definition 1.28 (Chain)**.** A subset A of a partially ordered set $\langle P, \leq \rangle$ is called a *chain* if $\langle A, \leq |_A \rangle$ is a totally ordered set. Equivalently, $A \subseteq P$ is a chain if, for any $a, b \in A$, either $a \leq b$ or $b \leq a$.

If no partial ordering is explicitly mentioned (as in the case of formulation 5 of Theorem 1.13), then we assume that the partial ordering being mentioned is the inclusion ordering $\subseteq$.

1.6 The Axiom of Choice

Axiom 9 (The Axiom of Choice)**.** *Let A be a family of nonempty disjoint sets. Then there exists a set C containing exactly one element from each member of A.*

Theorem 1.13 (Equivalent Formulations of the Axiom of Choice)**.** *All of the following theorems are equivalent to the axiom of choice (assuming the other ZF axioms).*

1. *For any relation R there is a function $F \subseteq R$ with* $\operatorname{dom} F = \operatorname{dom} R$.
2. *The Cartesian product of nonempty sets is nonempty.*
3. *For any set A there is a function F (a "choice function" for A) such that the domain of F is the set of nonempty subsets of A and $F(B) \in B$ for each nonempty $B \subseteq A$.*
4. *Zorn's Lemma: If $\langle P, \leq \rangle$ is a partially ordered set such that every chain in P has an upper bound then P has a maximal element m.*
5. *Hausdorff Maximal Principle: If $\mathcal{F}$ is a nonempty family of subsets of a set A such that the union of every chain $S \subseteq \mathcal{F}$ is also in $\mathcal{F}$, then $\mathcal{F}$ has a maximal subset.*
6. *The Well Ordering Theorem: Any set can be well-ordered.*

Theorem 1.14 (Inverses of Functions)**.** *Suppose that $F : A \to B$ is a function with $A, B \neq \varnothing$. Then*

1. *There exists a function $G : B \to A$ such that $G \circ F$ is the identity function I_A on A (concisely, we say G is a left inverse of F) iff F is injective.*
2. *There exists a function $H : B \to A$ such that $F \circ H$ is the identity function I_B on B (concisely, we say H is a right inverse of F) iff F is surjective.*

Corollary 1.14.1 (Two-Sided Inverses of Bijections)**.** *A function $F : A \to B$ is bijective iff it has a two-sided inverse (a function G such that $F \circ G = I_B$ and $G \circ F = I_A$).*

Warning: This is not perfectly trivial from Theorem 1.14, as you must prove that the left and right inverses are actually the same.

Definition 1.29 (Set of Functions)**.** For sets A and B we can form the *set of functions F from A into B*, denoted B^A. Notice that if A and B are finite sets with $|A| = a, |B| = b$, then $|B^A| = b^a$.

2 Important Constructions

At first glance, numbers may seem like completely different entities than sets. However, akin to how the letter A is a label for a set, so too is the number 0 or the number 5. In this section, we will explore exactly what this means and define each of the usual number classes we encounter in modern math.

2.1 The Natural Numbers and Peano Systems

Recall the definition of the *successor* a^+ of a set from Definition 1.2. Further recall the definition of *inductive sets* from Definition 1.2 and Axiom 6 which states that such a set exists.

Definition 2.1 (Natural Number)**.** A *natural number* is a set that belongs to every inductive set.

There is a set of all natural numbers (as it must be a subset of the set specified in the Axiom of Infinity), which we will denote ω here (though in other areas of math, it is usually called $\mathbb{N}$). Clearly, $\omega = \{0, 0^+, 0^{++}, 0^{+++}, \dots\}$, since this is an inductive set but any inductive set must contain this set. We have shorthand for these symbols, of course: 1 is defined as 0^+, 2 is defined as 0^{++}, and so on.

Theorem 2.1 (Finite Induction)**.** *Any inductive subset S of ω coincides with ω. In particular, if one lets the subset $S \subseteq \omega$ be the set of natural numbers with a certain property P, and one proves that $0 \in S$ and $n \in S \Rightarrow n^+ \in S$, then all natural numbers satisfy P.*

Definition 2.2. A *Peano system* is a triple $\langle N, S, e\rangle$ consisting of a set N, a function $S : N \to N$, and a member $e \in N$ such that the following three conditions are met:

1. $e \notin \operatorname{ran}(S)$
2. S is injective.
3. Any subset A of N that contains e and is closed under S (that is, if $x \in A$, $S(x) \in A$) is N itself.

The first two conditions force our chain $e, S(e), S^2(e), \dots$ to be a never-ending "line", and the third condition is an analogue of the induction postulate that we proved for natural numbers. From this alone we can build up the arithmetic of the natural numbers – called Peano arithmetic.

Notice that if σ denotes the restriction of the successor operation to ω, then $\langle \omega, \sigma, 0\rangle$ is a Peano system. That is, the natural numbers initialized from 0 and inductively defined via succession form a Peano system. Of course, this should not be surprising: the natural numbers were the model for the definition of a Peano system.

Definition 2.3 (Transitive)**.** A set A is said to be *transitive* if every member of a member of A is a member of A. Formally, $x \in a \in A \Rightarrow x \in A$.

Notice that every natural number is a transitive set (since $n = \{0, 1, 2, \dots, n-1\}$). Thus, the set ω, the union of them, is a transitive set. Furthermore, notice that if a is a transitive set (not just a natural number), then $\bigcup(a^+) = a$.

Theorem 2.2 (The Recursion Theorem on ω)**.** *If A is a set with $a \in A$, and F is a function $A \to A$, then there exists a unique function $h : \omega \to A$ such that $h(0) = a$ and $h(n^+) = F(h(n))$.*

Proof. Call a function ν "acceptable" if $\operatorname{dom} \nu \subseteq \omega, \operatorname{ran} \nu \subseteq A$, and:

1. If $0 \in \operatorname{dom} \nu$, then $\nu(0) = a$.
2. If $n^+ \in \operatorname{dom}(\nu)$, then $n \in \operatorname{dom}(\nu)$ and $\nu(n^+) = F(\nu(n))$.

Let $\mathscr{H}$ be the set of all acceptable functions, and let $\mathcal{H} = \bigcup \mathscr{H}$. Thus $n\mathcal{H}y$ iff $\nu(n) = y$ for some acceptable ν. We first show that $\mathcal{H}$ is a function, which amounts to showing that two acceptable functions always agree with each other when both are defined. Then we show that h is acceptable and has domain ω, and finally demonstrate that h is unique.

All of these steps follow fairly easily from creating a subset for which the claim we want to be true holds, and then showing that said subset is inductive. □

Theorem 2.3 (Peano Systems Model Natural Numbers)**.** *Let $\langle N, S, e\rangle$ be a Peano system. Then $\langle \omega, \sigma, 0\rangle$ is isomorphic to $\langle N, S, e\rangle$ – that is, there is a bijective function h from ω to N which preserves the successor operation ($h(\sigma(n)) = S(h(n))$) and the zero element ($h(0) = e$).*

Proof. Use Theorem 2.2 to prove that such a function exists and then prove that it is bijective. □

Thus, up to isomorphism, $\langle \omega, \sigma, 0\rangle$ is the unique number system satisfying the Peano axioms.

2.2 Arithmetic

Definition 2.4 (Binary Operation)**.** A *binary operation* on a set A is a function $A \times A \to A$.

We now want to find a way to define addition. To do this, consider a function $A_m : \omega \to \omega$ where $A_m(n)$ should be m added to n. For this to be the case, then $A_m(0)$ must surely equal m, and furthermore $A_m(n^+) = A_m(n)^+$. But then, by Theorem 2.2, there is a unique such function!

Definition 2.5 (Addition)**.** *Addition* $(+)$ is the binary operation on ω such that for any m and n in ω, $m + n = A_m(n)$.

Theorem 2.4 (Properties of Addition)**.**

1. $m + 0 = 0 + m = m$
2. $m + n^+ = (m + n)^+$
3. $(m + n) + k = m + (n + k)$
4. $m + n = n + m$

We now want to define multiplication, and follow our previous example by considering the unique function M_m such that $M_m(0) = 0$ and $M_m(n^+) = M_m(n) + m$.

Definition 2.6 (Multiplication)**.** *Multiplication* is the binary operation $\cdot$ on ω such that $\forall m, n \in \omega$, $m \cdot n = M_m(n)$.

Theorem 2.5 (Properties of Multiplication)**.**

1. $m \cdot 0 = 0 \cdot m = 0$ *and* $m \cdot 1 = 1 \cdot m = m$
2. $m \cdot n^+ = m \cdot n + m$
3. $(m \cdot n) \cdot k = m \cdot (n \cdot k)$
4. $m \cdot n = n \cdot m$
5. $m \cdot (n + p) = m \cdot n + m \cdot p$

Similarly, it is not difficult to define the exponentiation operation and the subtraction/divisors operations (where they exist).

Definition 2.7 (Order on ω)**.** We associate with ω the order given by $x < y$ if $x \in y$. Similarly, we define $x \leq y$ if $x \in y$ or $x = y$.

Theorem 2.6 (Properties of Ordering)**.**

1. *Notice $p < k^+$ if and only if $p \leq k$, and $m < n$ if and only if $m^+ < n^+$.*
2. *Notice $m \not< m$ for any $m \in \omega$ (by the earlier result, this would imply by induction that $\varnothing < \varnothing$, but clearly $\varnothing \notin \varnothing$, a contradiction).*
3. *For natural numbers n and m, exactly one of the three (1) $n < m$, (2) $n = m$, (3) $n > m$ holds.*
4. *Order is respected by addition and multiplication.*
5. *This is a well ordering (see Def. 1.26) on ω*

Theorem 2.7. *Let A be a subset of ω with the property that, for every $n \in \omega$, if all the numbers less than n are in A, then $n \in A$. Then $A = \omega$.*

2.3 Integers and Rational Numbers

Definition 2.8 (Integers)**.** *Let $\sim$ be the equivalence relation on $\omega \times \omega$ where $\langle m, n\rangle \sim \langle p, q\rangle$ if $m + q = p + n$. Then we define the integers $\mathbb{Z}$ to be the set $\omega \times \omega / \sim$.*

We define addition by $\langle m, n\rangle + \langle p, q\rangle = \langle m + p, n + q\rangle$ and multiplication by $\langle m, n\rangle \cdot \langle p, q\rangle = \langle mp + nq, mq + np\rangle$. We also notice that we can identify $\mathbb{Z}$ with the set $\{\dots, -1, 0, 1, \dots\}$. From here, any important basic properties of the integers (including arithmetic, the fact that $\mathbb{Z}$ is a ring, and the fact that $\mathbb{Z}$ is an initial object in **Ring**) can be easily deduced.

Definition 2.9 (Rationals)**.** *Let $\sim$ be the equivalence relation on $\mathbb{Z} \times \mathbb{Z}$ where $\langle m, n\rangle \sim \langle p, q\rangle$ if $mq = np$. Then we define the rational numbers $\mathbb{Q}$ to be the set $\mathbb{Z} \times \mathbb{Z} / \sim$.*

We define addition by $\langle m, n\rangle + \langle p, q\rangle = \frac{mq+np}{nq}$, and multiplication by $\langle m, n\rangle \cdot \langle p, q\rangle = \langle mp, nq\rangle$. Again, we can derive any important basic properties of the rational numbers.

2.4 Defining the Real Numbers

We simply state two equivalent definitions of the real numbers here, and leave it to the reader (if they are curious) to use the properties of the rationals to prove the basic rules of arithmetic for real numbers.

Definition 2.10 (Cauchy Sequences)**.** *A Cauchy sequence is a function $s : \omega \to \mathbb{Q}$ such that $|s_m - s_n|$ is arbitrarily small for sufficiently large m, n. Formally, $s : \omega \to \mathbb{Q}$ is Cauchy if for any $\varepsilon > 0$, there exists an N such that if $m, n > N$, $s_m - s_n < \varepsilon$.*

Definition 2.11 (Cauchy Equivalence)**.** *Two Cauchy sequences r and s are equivalent under the equivalence relation $\sim$ if they have the same limit (formally, if $|r_m - s_m| < \varepsilon$ for sufficiently large m for any $\varepsilon > 0$).*

Definition 2.12 (Reals, Def. 1)**.** *The set of real numbers is the set of all Cauchy sequences C modulo the equivalence relation $\sim$ described in the above definition.*

Definition 2.13 (Dedekind Cut)**.** *A Dedekind cut is a subset x of $\mathbb{Q}$ such that (1) $\varnothing \neq x \neq \mathbb{Q}$, (2) x is "closed downward" ($q \in x$ and $r < q \Rightarrow r \in x$), and (3) x has no largest number.*

Definition 2.14 (Reals, Def. 2)**.** *The set of real numbers is the set of all Dedekind cuts.*

From here, it is not hard to satisfactorily define multiplication, division, addition, subtraction, the usual linear ordering, and so on.

3 Ordinal Numbers

Definition 3.1 (Initial Segment)**.** *If $<$ is an ordering on A and $t \in A$, then the set $\operatorname{seg} t = \{x \mid x < t\}$ is called the initial segment up to t. For example, with ω ordered by $\in$, $\operatorname{seg} n = \{x \mid x \in n\} = n$.*

Theorem 3.1 (Transfinite Induction)**.** *Assume that $<$ is a well ordering on A. Further assume that $B \subseteq A$ has the special property that for every $t \in A$,*

$$\operatorname{seg}(t) \subseteq B \Rightarrow t \in B. \tag{10}$$

Then $B = A$.

Proof. If B is a proper subset of A, then $A - B$ has a least element m. By the leastness, $y \in B$ for each $y < m$: but then $\operatorname{seg} m \subseteq B$, so $m \in B$ by the construction of m, a contradiction. Thus $A = B$. □

Theorem 3.2 (Transfinite Recursion Theorem)**.** *For any formula $\gamma(x, y)$, the following is a theorem:*

Let A be a set well-ordered by $<$. Assume that for any f there is a unique y such that $\gamma(f, y)$. Then, there exists a unique function F with domain A such that $\gamma(F|_{\operatorname{seg} t}, F(t))$ for all $t \in A$.

Proof. For $t \in A$, we declare a function ν to be γ-constructed up to t if $\operatorname{dom} \nu = \{x \mid x \leq t\}$ and, for any $x \in \operatorname{dom} \nu$,

$$\gamma(\nu|_{\operatorname{seg} x}, \nu(x)). \tag{11}$$

First, we prove (as in Theorem 2.2), that if ν_1 and ν_2 are γ-constructed up to t, then $\nu_1 = \nu_2$. Then consider the set:

$$\mathscr{X} = \{v \mid (\exists t \in A)\ \nu \text{ is a function } \gamma\text{-constructed up to} t\} \tag{12}$$

Now let F be $\bigcup \mathscr{X}$. Thus $\langle x, y\rangle \in F \Leftrightarrow \nu(x) = y$ for some ν in $\mathscr{X}$. This means that F is indeed a function. Then, we prove that for any $x \in \operatorname{dom} F$, $\gamma(F|_{\operatorname{seg} x}, F(x))$ by construction. Similarly, $\operatorname{dom} F = A$ and F is unique, which proves the theorem. □

3.1 Ordinal Numbers and Ordinal Arithmetic

Definition 3.2 ($\in$-images and Ordinals)**.** Consider a well-ordered set $\langle A, <\rangle$ and let $\gamma(x, y)$ mean that $y = \operatorname{ran}(x)$. The transfinite recursion theorem states that there is a unique function E with domain A such that for any $t \in A$,

$$E(t) = \operatorname{ran}(E\ |_{\operatorname{seg} t}) = \{E(x) \mid x < t\} \tag{13}$$

Then the range of E is some set α, which we call the $\in$-*image* (the "epsilon image") of A. We also call α an *ordinal number* if it is the $\in$-image of some set A – specifically, we call it the *ordinal number of the set A.*

Theorem 3.3 (A is isomorphic to its $\in$-image)**.** *Suppose that $\langle A, <\rangle$ is a well-ordered set and let α denote the $\in$-image of α. Then:*

1. *E is a bijection between A and α.*
2. *For any $s, t \in A$, $s < t$ if and only if $E(s) \in E(t)$.*

Indeed, if we define the binary relation $\in_\alpha = \{\langle x, y\rangle \in \alpha \times \alpha \mid x \in y\}$, then the well-ordered structure $\langle A, <\rangle$ is isomorphic to $\langle \alpha, \in_\alpha\rangle$.

Corollary 3.3.1. *Any ordinal number α is well-ordered by $\in_\alpha$ and it is a transitive set.*

Theorem 3.4 ($\in$-images are unique)**.** *Two well-ordered structures $\langle A, <_A\rangle$ and $\langle B, <_B\rangle$ are isomorphic if and only if they have the same $\in$-image (not just isomorphic ones).*

One can view this as a statement about the category of structures, where one direction (they are isomorphic if they have the same $\in$-image) is trivial and the other (they have the same $\in$-image if they are isomorphic) states that any isomorphism between $\langle A, <_A\rangle$ and $\langle B, <_B\rangle$ factors through $\langle \alpha, \in_\alpha\rangle$.

Theorem 3.5 (Putting the "order" in ordinals)**.** *For any two well-ordered structures, either they are isomorphic to each other or one is isomorphic to an initial segment of the other.*

Theorem 3.6 (Classifying Ordinals)**.** *Suppose that α is a transitive set such that $\langle \alpha, \in_\alpha\rangle$ is a well-ordered structure. Then α is an ordinal number.*

Theorem 3.7 (Properties of Ordinal Numbers)**.**

- *Any member of an ordinal number is itself an ordinal number.*

- *For any two ordinal numbers α, β, exactly one of $\alpha \in \beta, \alpha = \beta, \beta \in \alpha$ is true.*
- *Any transitive set of ordinal numbers is itself an ordinal number.*
- 0 *is an ordinal number.*
- *If α is any ordinal number, then α^+ is also an ordinal number. Indeed, it is the least ordinal larger than α.*
- *If A is any set of ordinal numbers, then $\bigcup A$, the supremum of A, is also an ordinal number.*

Thus, we finally have an idea of what the ordinals look like: $\{0, 1, 2, \ldots, \omega, \omega^+, \omega^{++}, \ldots\}$. Of course, we can put the dots here, even though this set is not countable, because we are speaking *transfinitely.*

3.2 Rank

We want to characterize the totality of sets. V_0 will be the empty set, and in general, V_α is the set of sets whose elements are drawn from all lower-rank sets. Formally, $V_\alpha = \bigcup\{\mathcal{P}(V_\beta \mid \beta < \alpha)\}$. By transfinite induction, we know this defines a set for all ordinals α.

Definition 3.3 (Types of Ordinals)**.** 0 is *the zero ordinal*, and any ordinal α such that $\alpha = \beta^+$ for some β is called a *successor ordinal.* The ordinals left over, such as ω, are called the *limit ordinals.*

For a successor ordinal α^+, it is not hard to see that $V_{\alpha^+} = \mathcal{P}(V_\alpha)$.

Theorem 3.8 (Every Set Has a Rank)**.** *For a given set A, the smallest ordinal α such that $A \in V_\alpha$ is called the rank of A. For example, the rank of ω is ω (in general, the rank of an ordinal is itself).*

4 Cardinal Numbers

Definition 4.1 (Equinumerosity)**.** A set A is *equinumerous* to a set B if there is a bijection between A and B. Clearly, this is an equivalence relation on the class of sets, so we write $A \approx B$ if A and B are equinumerous.

For example, the set ω is is equiumerous with the set $\omega \times \omega$. An example bijection from $\omega \times \omega \to \omega$ is $J(m, n) = \frac{1}{2}((m + n)^2 + 3m + n)$.

Theorem 4.1 (Cantor's Diagonal Argument Pt. 1)**.** *The set ω is not equinumerous to the set of all real numbers $\mathbb{R}$.*

Proof. Suppose there is a function $f : \omega \to \mathbb{R}$. Then we arrange the list of values of f in a list (which we can do by induction): $f(0) = r_0, f(1) = r_1, f(2) = r_2, \ldots$. Now define q as the real number whose first digit is r_0's incremented by 1, whose second digit is r_1's incremented by 1, and so on. Then $q \notin \operatorname{ran}(f)$ so f is not bijective. The result follows. □

Theorem 4.2 (Cantor's Diagonal Argument Pt. 2)**.** *There is no set A such that S is equinumerous to its power set $\mathcal{P}(A)$.*

Proof. Let f be a function $A \to \mathcal{P}(A)$. Consider the set $B = \{x \in A \mid x \notin f(x)\}$. Clearly $B \subseteq A$ so $B \in \mathcal{P}(A)$. Now consider any $x \in A$: either $x \in f(x)$ or $x \notin f(x)$. In the former case, $f(x)$ cannot equal B because $x \in f(x)$. In the latter case, $f(x)$ cannot equal B because $x \notin f(x)$ by assumption and $x \in B$ by the construction of B. □

Definition 4.2 (Finite and Infinite)**.** A set is *finite* if it is equinumerous with a natural number. Otherwise, the set is *infinite.*

Theorem 4.3 (Properties of Finite and Infinite Numbers)**.**

1. *Pigeonhole Principle: No natural number is equinumerous with a proper subset of itself.*

2. *No finite set is equinumerous with a proper subset of itself. Indeed, the property of being infinite is equivalent to the property of being equinumerous to a proper subset of itself. Thus, for example,* $\omega \approx \omega - \{0\}$ *is infinite.*
3. *Any subset of a finite set is finite.*

Definition 4.3 (Dominant Sets and Countability)**.** A set A is *dominated* by a set B (we write $A \preceq B$) if there is an injective function $A \to B$. If a set is dominated by ω, we call it *countable* (otherwise, it is *uncountable*).

Definition 4.4 (Omega)**.** We call the smallest uncountable ordinal Ω.

Theorem 4.4 (Burali-Forti Theorem)**.** *There is no set of all ordinal numbers.*

Proof. Such a set would be transitive and well ordered by epsilon and thus an ordinal number. But this implies that some set contains itself, which is a contradiction with the Axiom of Regularity. □

Theorem 4.5 (Numeration Theorem)**.** *Any set A is equinumerous to some ordinal number.*

Proof. Use the Well Ordering Theorem to well-order A according to some relation $<$. Then consider the $\in$-image of $\langle A, < \rangle$: it is an ordinal equinumerous to A. □

Definition 4.5 (Cardinality)**.** For any set A, define the cardinal number of A ($\operatorname{card} A$) to be the least ordinal equinumerous to A. In particular, any natural number's cardinality is itself, and the cardinality of ω (and any countably infinite set) is ω.

Theorem 4.6 (Schröder-Bernstein Theorem)**.** *If $A \preceq B$ and $B \preceq A$, then $A \approx B$. Similarly, if* $\operatorname{card} A \leq \operatorname{card} B$ *and* $\operatorname{card} B \leq \operatorname{card} A$, *then* $\operatorname{card} A = \operatorname{card} B$.

Corollary 4.6.1. $\mathbb{R} \approx 2^\omega$. *That is,* $\operatorname{card} \mathbb{R} = \operatorname{card} \mathcal{P}(\omega)$.

Proof. The map $2^\omega \to \mathbb{R}$ given by $\{a_1, a_2, \dots\} \to 0.a_1 a_2 \dots$ is injective, so $2^\omega \preceq \mathbb{R}$. Similarly, the map $[0,1] \approx \mathbb{R} \to 2^\omega$ given by $0.a_1 a_2 \dots$ (where each a_i is 0 or 1 and this is all in binary) to $\{a_1, a_2, \dots\}$ is injective, so $\mathbb{R} \preceq 2^\omega$. Thus $\mathbb{R} \approx 2^\omega$. □

4.1 The Alephs ($\aleph_0, \aleph_1, \dots, \aleph_\omega, \dots$)

Theorem 4.7. *The class of all cardinals is not a set.*

Proof. Notice that any unbounded class of ordinals is not a set (this is because its union is the entire class of ordinal – which would be a contradiction with Theorem 4.4). The class of all cardinals is an unbounded class of ordinals, and thus it is not a set. □

Definition 4.6. Let $\aleph_0$ be the least infinite cardinal. Then there is a least infinite cardinal larger than $\aleph_0$, which we call $\aleph_1$. Similarly, we define $\aleph_\alpha$ for any ordinal α as the least cardinal number such that $\aleph_\alpha > \aleph_\beta$ for all $\beta < \alpha$.

Theorem 4.8 (Classification of the Cardinals)**.** *If $\alpha \in \beta$, then $\aleph_\alpha < \aleph_\beta$. From this, one may deduce that every cardinal is of the form $\aleph_\alpha$ for some cardinal α.*

4.2 Order Type and Ordinal Arithmetic

Definition 4.7 (Isomorphism Type)**.** Let R be a binary relation on A. The *isomorphism type* it $\langle A, R \rangle$ of the structure $\langle A, R \rangle$ is the set of all structures $\langle B, S \rangle$ such that:

1. $\langle A, R \rangle \cong \langle B, S \rangle$, and
2. no structure of rank less than rank $\langle B, S \rangle$ is isomorphic to $\langle B, S \rangle$.

Theorem 4.9. *Structures $\langle A, R\rangle$ and $\langle B, S\rangle$ are isomorphic iff they have the same isomorphism type.*

Definition 4.8 (Order Type)**.** An *order type* is the isomorphism type of some linearly ordered structure. Any member of an order type ρ is said to be a linearly ordered structure of *type* ρ.

Definition 4.9 (Order Type Arithmetic)**.** Suppose we have two structures $\langle A, R\rangle$ and $\langle B, S\rangle$, with order types ρ and σ, respectively. Then we define $R \oplus S = R \cup S \cup (A \times B)$ and $R * S$ to be the "reverse lexicographic order" on $A \times B$: $\langle a_1, b_1\rangle\,(R * S)\,\langle a_2, b_2\rangle$ if either $b_1 S b_2$ or $b_1 = b_2$ and $a_1 R a_2$. Furthermore, we define $\rho + \sigma = \mathrm{it}\,\langle A \cup B, R \oplus S\rangle$ and $\rho \cdot \sigma = \mathrm{it}\,\langle A \times B, R * S\rangle$

There is work to be done to prove that everything here is well-defined, but it is not that difficult.

Definition 4.10 (Order Types of Common Structures)**.** The order type of ω is denoted $\overline{\omega}$ (indeed, the order type of any ordinal $\langle \alpha, \in_\alpha\rangle$ is denoted $\overline{\alpha}$). We let μ denote the order type of $\mathbb{Q}$ and λ denote the order type of $\mathbb{R}$.

Finally, for any linearly ordered structure $\langle A, R\rangle$ of type ρ, then define ρ^* to be the isomorphism type of it $\left\langle A, R^{-1}\right\rangle$. Then $\mathbb{Z}$ has order type $\overline{\omega}^* + \overline{\omega} \neq \overline{\omega} + \overline{\omega}^*$. Notice that the addition of order types is not commutative!

Theorem 4.10 (Properties of Order Types Arithmetic)**.**

1. $(\rho + \sigma) + \tau = \rho + (\sigma + \tau)$ *and* $(\rho \cdot \sigma) \cdot \tau = \rho \cdot (\sigma \cdot \tau)$
2. $\rho \cdot (\sigma + \tau) = \rho \cdot \sigma + \rho \cdot \tau$.
3. $\rho + \overline{0} = \overline{0} + \rho = \rho$
4. $\rho \cdot \overline{1} = \overline{1} \cdot \rho = \rho$
5. $\rho \cdot \overline{0} = \overline{0} \cdot \rho = \overline{0}$

Definition 4.11. Let α and β be ordinal numbers. Then define $\alpha + \beta$ to be the unique ordinal γ such that $\overline{\alpha} + \overline{\beta} = \overline{\gamma}$ and similarly let $\alpha \cdot \beta = \gamma$ if $\overline{\alpha} \cdot \overline{\beta} = \overline{\gamma}$. Thus, all properties of order type arithmetic (associativity, distributivity, and the behavior of identity elements) are inherited by ordinal numbers.

There are many other properties that we could cover here, but the most important are the subtraction, division, and logarithm theorems:

Theorem 4.11 (Subtraction Theorem)**.** *If $\alpha \in \beta$ (for ordinal numbers α and β), then there exists a unique ordinal number δ (their difference) such that $\alpha + \delta = \beta$.*

Theorem 4.12 (Division Theorem)**.** *Let α and δ be ordinal numbers with δ nonzero. Then there is a unique pair of ordinal numbers β and γ such that $\alpha = \delta \cdot \beta + \gamma$ and $\gamma \in \delta$.*

4.3 Cardinal Arithmetic

Definition 4.12 (Addition and Multiplication of Cardinal Numbers)**.** If κ and λ are cardinal numbers with associated sets K and L, we define $\kappa + \lambda = \mathrm{card}(K \cup L)$, $\kappa \cdot \lambda = \mathrm{card}(K \times L)$, and $\kappa^\lambda = \mathrm{card}\,{}^L K$.

Theorem 4.13 (Properties of Cardinal Arithmetic)**.** 1. $\kappa + \lambda = \lambda + \kappa$ *and* $\kappa \cdot \lambda = \lambda \cdot \kappa$

2. $\kappa + (\lambda + \mu) = (\kappa + \lambda) + \mu$
3. $\kappa \cdot (\lambda + \mu) = \kappa \cdot \lambda + \kappa \cdot \mu$
4. $\kappa^{\lambda + \mu} = \kappa^\lambda \cdot \kappa^\mu$, $(\kappa \cdot \lambda)^\mu = \kappa^\mu \cdot \lambda^\mu$, *and* $(\kappa^\lambda)^\mu = \kappa^{\lambda \cdot \mu}$

Theorem 4.14. *The countable union of countable sets is countable. More generally, if every member of a set $\mathscr{A}$ has cardinality κ or less, then* $\mathrm{card} \bigcup \mathscr{A} \leq \mathrm{card}\, A \cdot \kappa$.

Theorem 4.15 (The Absorption Law of Cardinal Arithmetic)**.** *Let κ and λ be cardinal numbers, at least one infinite and both nonzero. Then* $\kappa + \lambda = \kappa \cdot \lambda = \max(\kappa, \lambda)$

4.4 The Continuum Hypothesis

So far, every single infinite set we have created either has cardinality $\aleph_0$ or cardinality at least $2^{\aleph_0}$. Thus, we ask the question: "is there a set with cardinality κ such that $\aleph_0 < \kappa < 2^{\aleph_0}$"? The *continuum hypothesis* is the hypothesis that no such set exists: equivalently, any set of real numbers is equinumerous to either the natural numbers or the real numbers.

The *generalized continuum hypothesis* is the theory that for each κ, there is no cardinal number strictly between κ and 2^κ. Gödel proved in 1939 that neither the generalized or normal continuum hypothesis is disprovable in ZFC, and Paul Cohen also proved that even the normal continuum hypothesis is not provable in ZFC. Thus, we say the continuum hypothesis is *independent* of ZFC. How this was proven will be discussed later in our coverage of forcing.

4.5 Cofinality

Notice that any limit ordinal is the supremum of the set of all smaller ordinals. In other words, $\lambda = \bigcup \lambda$ for any limit ordinal λ. However, sometimes it is unnecessary to take *all* smaller ordinals: sometimes we can take a proper subset of λ.

Definition 4.13 (Cofinality)**.** The *cofinality* of a limit ordinal λ, denoted $\operatorname{cf} \lambda$, is the smallest cardinal κ such that λ is the supremum of κ smaller ordinals. The cofinality of a nonlimit ordinal is defined to be 0 for 0 and 1 for any successor ordinal.

A set S for which $\lambda = \sup S$ is called *cofinal* in λ. Of course, there are many such S – in particular, since λ is cofinal in itself, $\operatorname{cf} \lambda \leq \operatorname{card} \lambda$.

Definition 4.14 (Regular and Singular Cardinals)**.** A cardinal κ is called *regular* if $\operatorname{cf} \kappa = \kappa$ and *singular* if $\operatorname{cf} \kappa < \kappa$.

Theorem 4.16. *$\aleph_{\alpha+1}$ is a regular cardinal for every ordinal α. In contrast, for any limit ordinal λ,* $\operatorname{cf} \aleph_\lambda = \operatorname{cf} \lambda$.

The second part of the above result can be used to show that $\operatorname{cf} \aleph_\omega = \operatorname{cf} \omega = \aleph_0$ and $\operatorname{cf} \aleph_\Omega = \operatorname{cf} \Omega = \aleph_1$, so both are singular. We cannot prove that $\aleph_\lambda$ is regular for any limit ordinal λ within our axioms (if $\aleph_\lambda$ is, we call it *weakly inaccessible*).

Theorem 4.17. *For any ordinal λ,* $\operatorname{cf} \lambda$ *is a regular cardinal.*

Theorem 4.18. *If λ is an infinite cardinal, then* $\operatorname{cf} \lambda$ *is the least cardinal number κ such that λ can be decomposed into the union of κ sets each having cardinality less than λ.*

4.6 Inaccessible Cardinals

Consider a formula σ defined on the entire class of sets $\mathbf{V}$. Given some set M, we convert this σ to a new formula σ^M (called the *relativization of σ to M*) by replacing $\forall x$ and $\exists x$ by $\forall x \in M$ and $\exists x \in M$, respectively. If σ^M is true, we say that σ is true in M and then M is a *model* of σ.

Suppose that λ is any limit ordinal greater than ω (the smallest example being $\omega \cdot 2$). Then V_λ is actually a model of all the ZFC axioms except replacement. As an interesting side note, we can write $V_{\omega\cdot 2} = V_\omega \cup \mathcal{P}(V_\omega) \cup \mathcal{PP}(V_\omega) \dots$. This set contains the real numbers, all functions from reals to reals, and so on. Thus, it contains almost all the sets we actually care about.

Theorem 4.19. *Not all of the replacement axioms are true in $V_{\omega\cdot 2}$.*

Theorem 4.20. *Not all of the replacement axioms are theorems of the other ZFC axioms.*

Proof. Any theorem of the Zermelo axioms must be true in any model of those axioms, such as $V_{\omega \cdot 2}$. Thus, by the preceding result, not all replacement axioms are theorems of the other ZFC axioms. □

This is a great example of the use of models to prove the independence of various axioms/theorem, but we do discuss models in greater detail later.

Definition 4.5.3: A cardinal number κ is called *inaccessible* if:

1. κ is greater than $\aleph_0$.
2. For any cardinal $\lambda < \kappa$, $2^\lambda < \kappa$.
3. κ is regular (see Def. 4.14)

One can prove that for any inaccessible cardinal number κ, V_κ is a complete model of the ZFC axiom. However, one of the corollaries of Gödel's second incompleteness theorem is that we cannot prove from our axioms that such a model exists. Thus, the existence of inaccessible cardinals cannot be proven using ZFC. Some people, like Alfred Tarski, proposed the following axiom in response: "for any cardinal number there is a larger inaccessible cardinal number". This is an example of a so-called "large cardinal axiom", of which there are many variants.

5 The Unreasonable Effectiveness of Set Theory

Set theory is not just pursued in pure mathematics for its own sake – it has numerous applications in other areas of pure math. Here, we will list a few of the most accessible, to inspire your research.

5.1 Zorn's Lemma Applied to Algebra

Here we give an example of the usual way one applies Zorn's Lemma. Other results that follow from the extremely useful lemma include (1) every proper ideal of a commutative ring (unital, of course) is contained in a maximal ideal, (2) the intersection of all prime ideals in a commutative ring is the set of nilpotent elements in the ring, and so on.

Theorem 5.1. *If S_0 is a linearly independent subset of a vector space V over K, then there exists a basis $\mathcal{B}$ of V that contains S_0. In particular, since the empty set is linearly independent, every vector space has a basis.*

Proof. Let

$$\mathcal{F} = \{S \subset V \mid S_0 \subset S \text{ and } S \text{ is linearly independent in } V\} \tag{14}$$

Then $\mathcal{F}$ satisfies the assumptions of the Hausdorff maximality principle: clearly $\mathcal{F}$ is nonempty and the union of any chain of linearly independent sets is linearly independent (since an infinite set is linearly independent if any finite set is linearly independent).

Thus there exists a maximal element $\mathcal{B}$ in $\mathcal{F}$. If a vector v is not in the span of $\mathcal{B}$, then it is trivial to show that $\mathcal{B} \cup \{v\}$ is linearly independent and thus an element containing $\mathcal{B}$ in $\mathcal{F}$, a contradiction. Thus $\mathcal{B}$ is spanning and linearly independent, so it is a basis for V. □

5.2 Strange Subsets of $\mathbb{R}^n$

We simply list some of the most interesting results of transfinite induction:

Theorem 5.2. *There exists a subset A of the plane with every horizontal section $A^y = \{x \in \mathbb{R} \mid (x, y) \in A\}$ being dense in $\mathbb{R}$ and with every vertical section $A^x = \{y \in \mathbb{R} \mid (x, y) \in A\}$ having precisely one element.*

Theorem 5.3. *There exists a subset A of the plane $\mathbb{R}^2$ that intersects every straight line in exactly two points.*

Theorem 5.4. *The space $\mathbb{R}^3$ is a union of disjoint circles, but the plane $\mathbb{R}^2$ is not.*

Theorem 5.5. *There is a countable partition $\{S_i : i < \omega\}$ of $\mathbb{R}^2$ such that the distance between any two different points of the same set S_i is irrational.*

Theorem 5.6. *Every family $\mathcal{U}$ of pairwise disjoint open subsets of $\mathbb{R}^n$ is countable. Similarly, every discrete subset S of $\mathbb{R}^n$ (that is, a set where each point in S has a neighborhood around it containing no other point in S) is countable.*

Proof. Define $f : \mathcal{U} - \varnothing \to \mathbb{Q}^n$ by choosing $f(U) \in U \cap \mathbb{Q}^n$ for every $U \in \mathcal{U}$. Then f is one-to-one, so $\mathcal{U} \leq |\mathbb{Q}^n| + 1 = \omega$. Notice that a counterexample to the second part of the theorem would also contradict the first, so by contraposition both follow. □

5.3 Strange Real Functions

Theorem 5.7 (The Intermediate Value Theorem)**.** *Suppose we have any continuous function f. Then, for every $a < b$ and every number y between $f(a)$ and $f(b)$ there is an $x \in (a, b)$ such that $f(x) = y$.*

Definition 5.1 (Darboux)**.** Functions satisfying the above property are called *Darboux.* If furthermore for every $a < b$ and every number y there is an $x \in (a, b)$ such that $f(x) = y$, we call f *strongly Darboux.*

Theorem 5.8. *There exists a strongly Darboux function $f : \mathbb{R} \to \mathbb{R}$ which is Darboux but everywhere discontinuous.*

Theorem 5.9. *Let $\mathcal{G}$ be a family of real functions $\mathcal{G} \subset \mathbb{R}^{\mathbb{R}}$ with $|\mathcal{G}| \leq \mathfrak{c}$. Then there exists a function $f : \mathbb{R} \to \mathbb{R}$ such that $f + g$ is strongly Darboux for every $g \in \mathcal{G}$.*