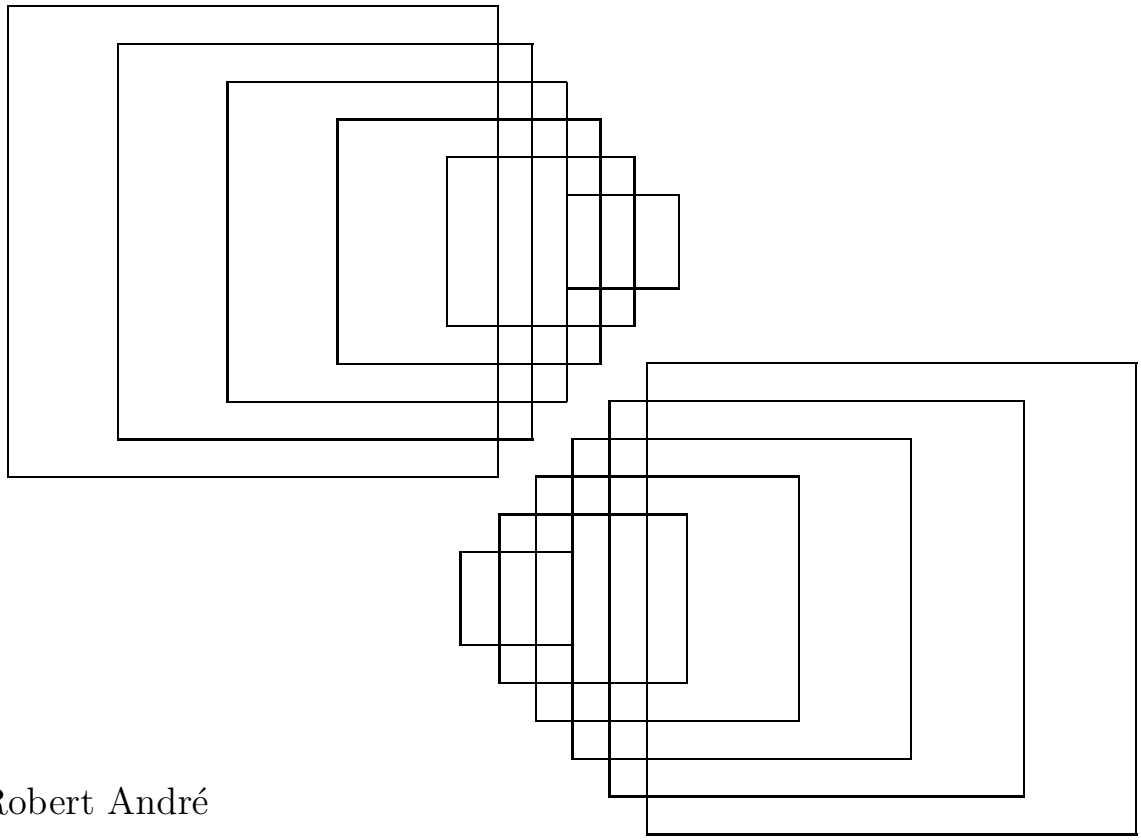
Axioms and Set Theory

A first course in Set Theory



Robert André

ISBN 978-0-9938485-0-6

Revised 21/07/06

À
Jinxia, Camille et Isabelle

"Everything has beauty, but not everyone sees it."

Confucius

Preface

A set theory textbook can cover a vast amount of material depending on the mathematical background of the readers it was designed for. Selecting the material for presentation in this book often came down to deciding how much detail should be provided when explaining concepts and what constitutes a reasonable logical gap which can be independently filled in by the reader. The initial chapters of this book will appeal to students who have little experience in proving mathematical statements, while the last chapters, significantly denser in subject matter, will appeal considerably more to senior undergraduate or graduate students. Choice of topics and calibration of the level of communication is based on the estimated mathematical fluency of the target students. The first part of this book (Chapters 1 to 21) is written for intermediate level math major students in mind. At this level, most students have not yet been exposed to the mathematical rigor normally found in most textbooks in set theory. The pace at which new concepts are introduced at the beginning is what some may subjectively consider as being quite "leisurely". The meaning of mathematical statements is explained at length and their proofs presented in great detail. As the student progresses through the course, he or she will develop a better understanding of what constitutes a correct mathematical proof. To help attain this objective, numerous examples of simple straightforward proofs are presented as models throughout the text.

The subject material is subdivided into ten major parts. The first few are themselves subdivided into "bite-size" chapters. Smaller sections allow students to test their understanding on fewer notions at a time. This will allow the instructor to better diagnose the understanding of those specific points which challenge the students the most, thus helping to eliminate obstacles which may slow down their progress later on.

Each chapter is followed by a list of *Concepts review* type questions. These questions highlight for students the main ideas presented in that section and help them deepen understanding of these concepts before attempting the exercises. The answers to all *Concept review* questions are in the main body of the text. Attempting to answer these questions will help the student discover essential notions which are often overlooked when first exposed to these ideas.

Textbook examples will serve as solution models to most of the exercise questions at the end of each section. Exercise questions are divided into three groups: A, B and C. The answers to the group A questions normally follow immediately from definitions and theorem statements presented in the text. The group B questions require a deeper understanding of the concepts, while the group C questions allow the students to deduce by themselves a few consequences of theorem statements presented in the text.

The course begins with an informal discussion of primitive concepts and a presentation of the ZFC axioms. We then discuss, in this order, operations on classes and sets, relations on classes and sets, functions, construction of numbers (beginning with the natural numbers followed by the rational numbers and real numbers), infinite sets, cardinal numbers and, finally, ordinal numbers. It is hoped that the reader will eventually perceive the ordinal

numbers as a natural logical extension of the natural numbers and as the "spine of set theory" − like many authors described them. Towards the end of the book we present a brief discussion of a few more advanced topics such as the *Well-ordering theorem*, *Zorn's lemma* (both proven to be equivalent forms of the *Axiom of choice*) as well as Martin's axiom. Finally, we briefly discuss the *Axiom of regularity* and a few of its implication. A brief and very basic presentation of ordinal arithmetic properties is then given.

The pace and level of abstraction increase considerably when *ordinals* are introduced. It is hoped that everything which is presented before this point will allow the students to master various proof techniques while simultaneously developing a feeling for what constitutes the essence of set theory. The format used in the book allows for some flexibility in how subject matter is presented, depending on the mathematical maturity of the audience or the pace at which the students can absorb new material. A determining factor may be the amount of practice that students require to understand and produce correct mathematical proofs. Some instructors may decide to use the first twenty chapters of the book as a text for an "Introduction to mathematical proofs" course.

Students who already possess a substantial amount of mathematical background may feel they can comfortably skip many chapters without loss of continuity, since these contain notions which are well-known to them. The following order sequence will allow readers with the required background to advance more quickly to the meat of the textbook: Chapter 1 on the topic of the *ZFC*-axioms can be immediately followed by chapters 13 and 14 on the topic of natural numbers, chapters 18 to 22 on the topic of infinite sets and cardinal numbers followed by chapters 26 to 29, 32 and 33, on ordinals, and finally, chapters 30 and 31 on the axiom of choice and the axiom of regularity.

As we all know, any textbook, when initially published, will contain some errors, some typographical, others in spelling or in formatting and, what is even more worrisome, some mathematical. Many readers of the text are required to help weed out the most glaring mistakes. If you happen to be a reader who has carefully studied a chapter or two of the book please feel free to communicate to me, by email, any errors you may have spotted, with your name and chapters reviewed. In the preface of further versions of the book, I will gladly acknowledge your help. This will be much appreciated by this writer as well as by future readers. It is always more pleasurable to study a book which is error-free.

Robert André
University of Waterloo, Ontario
randre@uwaterloo.ca

ISBN 978-0-9938485-0-6

Contents

Part I

Axioms and classes

1 / Classes, sets and axioms.

Summary. *In this section we discuss axiomatic systems in mathematics. We explain the notions of "primitive concepts" and "axioms". We declare as primitive concepts of set theory the words "class", "set" and "belong to". These will be the only primitive concepts in our system. We then present and briefly discuss the fundamental Zermelo-Fraenkel axioms of set theory.*

1.1 Contradictory statements.

When expressed in a mathematical context, the word "statement" is viewed in a specific way. A *mathematical statement* is a declaration which can be characterized as being either true or false. By this we mean that if a statement is not false, then it must be true, and vice-versa. There exists a predetermined set of rigorous logical rules which can be used to help determine the true or false value of such statements. Whether one does mathematics as an expert or as a beginner, these elementary rules of logic must always be respected. An argument which does not respect one of these rules is said to be "illogical". Then, by combining various mathematical statements whose true or false values are known, we can logically determine the true or false value of other mathematical statements. A rule of logic looks something like this:

> If Q is true whenever P is true, and T is true whenever Q is true, then T is true whenever P is true.

Such rules can be symbolically represented in a way that avoids the use of words. For example, the above statement is represented as:

$$[(P \Rightarrow Q) \wedge (Q \Rightarrow T)] \Rightarrow (P \Rightarrow T)$$

In this way, we can construct an elaborate system of mathematical statements each of which has been determined to be true or false. The logical steps which help us determine the true or false value of a statement is called a "mathematical proof". Most readers have previously been exposed to this particular way of thinking in various courses such as calculus and linear algebra. Basic rules of logic are normally not taught explicitly in such courses. It is however expected that a student who has sufficiently been exposed to rigorous mathematical arguments and has often enough attempted to formulate correct mathematical proofs − sometimes more successfully than others − progressively learns to distinguish valid logical arguments from ones that are flawed. Like learning to speak any language, formulating correct mathematical arguments is a skill that is developed with practice.

If the truth of a mathematical statement is logically deduced by combining statements previously known to be true, then clearly there had to be, at some point, a set of statements whose true-false values were not derived from previous statements. That is,

the process must start somewhere, with some initial statements whose true-false value were unknown. Such initial statements are not "deduced" but simply *declared* to be true based on nothing more than "common sense". For example, one may declare the statement: "Distinct parallel lines cannot intersect" as being *self-evident* or being so "elementary" that it cannot be proved. Once we give ourselves a set $\mathscr{A}$ of *self-evident statements* and a list of rules that can be used to determine the true-false value of other statements then the universe $\mathscr{U}_{\mathscr{A}}$ of all possible true statements derived from $\mathscr{A}$ is determined. This determined universe $\mathscr{U}_{\mathscr{A}}$ of statements constitutes a mathematical theory which is ours to explore, or discover, one statement at a time.

But what if the choice of our original set $\mathscr{A}$ of statements was not a wise one? "How can it not be a wise one if based on common sense?" one might ask. Imagine this scenario:

> Say that from a set $\mathscr{A}$ of initial self-evident statements, a statement A has been shown to be true, and given that A is true it is deduced that statement B must be true, and from B we deduce that P is true. On the other hand it is shown that given A, statement D must be true and that from D we show that P is false. Hence, from A we have deduced that the statement P is both true and false.

A statement which has been determined to be both true and false is referred to as a "contradictory statement" or a *paradox*. If a contradictory statement logically flows from what was assumed to be a paradox-free system, then the foundation of this system, as well as the methods used to determine the true or false value of statements, must be carefully scrutinized to determine the incorrect assumption(s) which allowed this "renegade" statement to emerge. In this book we will explore a specific mathematical system. It is hoped that in the process, the reader will be able to appreciate the skill and ingenuity required for the construction of this impressive mathematical structure. This system is called the "*theory of sets*" or more simply "*set theory*".

1.2 Sets.

Most people are familiar with the notion of a *set* and its *elements*. "Sets" are viewed as collections of things while "elements" are viewed as those things which belong to sets. Normally, a set is defined in terms of certain properties shared by its elements. These properties must be well described, with no ambiguities, so that it is always clear whether a given element belongs to a given set or not. Being a "set" can also be an element property; so sets whose elements are sets exist. For example, the set S of all teams in a particular hockey league. The elements of the set S are sets of hockey players.

Let us consider a few examples of entities we may consider to be sets.

a) Let T denote the set of all straight lines in the Cartesian plane. For example, the set $A = \{(x, y) : y = 2x + 3\}$ belongs to T while the set $B = \{1, 2, 3\}$ does not. We easily see that T is not an element of T since T is not a line in the Cartesian plane.

b) Let U denote the set of all sets which contain infinitely many elements. This set is well-sdefined since we can easily distinguish those elements which belong to U from those that do not belong to U. For example, the subset $\{-2, 0, 100\}$ is not an element of U since it contains only three elements. We ask the question: Is the set U an element of U? To help answer this question, witness the sets

$$\begin{aligned} A_0 &= \{0, 1, 2, 3, \cdots\} \\ A_1 &= \{-1, 0, 1, 2, 3, \cdots\} \\ A_2 &= \{-2, -1, 0, 1, 2, 3, \cdots\} \\ &\vdots \\ A_n &= \{-n, -(n-1), -(n-2), \cdots, -2, -1, 0, 1, 2, 3, \cdots\} \\ &\vdots \end{aligned}$$

Every element in the set $\{A_0, A_1, A_2, A_2, \ldots\}$ belongs to U. Hence, U contains infinitely many elements. We conclude that U is an element of U.

c) Define S to be the set of all "sets that are not elements of themselves". For example, the set T described in example a) is in S since T is not an element of itself. The set U described in example b) does not belong to S since U is an element of itself.

We now look more closely at the three sets described above. Other than the fact that it is an extremely large set, there is nothing extravagant about the set T described in example a). On the other hand, the set U discussed in example b) also appears to be well-defined since a set which is infinite can easily be distinguished from one that is not. But the fact that this set is an element of itself makes one wonder whether we should allow sets to satisfy this property. On the other hand, it is difficult to express what could possibly go wrong with such sets. Let us now look closely at the "set" described in the example c): A set belongs to S only if it does not belong to itself. We wonder whether, like the U in example b), the set S is an element of itself. But S cannot belong to S since no element in S can belong to itself. So S is not an element of itself. Then, by definition of S, S would then be an element of S. This doesn't make sense. There is obviously a problem with the "set" described in example c). Even if it was fairly easy to detect the contradiction which follows from example c) specifically determining the source of this contradiction can be more difficult.

Example c) nicely illustrates what is called a *paradox*. As we mentioned earlier, a paradox consists of two contradictory statements both of which logically flow from what was thought to be a well understood and clearly defined concept. In this case, to say that the statement "S is an element of S" is true means that the statement "S is an element of S" is false, and vice-versa. For many, this is just a play on words; it may seem harmless enough. But for mathematicians this was not a trivial matter. The discovery of this particular paradox by Bertrand Russell was equivalent to the uncovering of a malicious virus lying dormant in the heart of the operating system of a microcomputer. It cannot be ignored. The way sets are defined along with their universally accepted properties form the foundation of modern mathematics. Paradoxes such as this undermines the confidence we have in mathematics, a discipline which prides itself on its clarity of thought, a discipline which sees itself as a seeker of irrefutable truths. Once this flaw was exposed it was important to understand why we did not see this before. Mathematicians also wondered whether any other cracks in the foundation of mathematics existed, requiring immediate attention.

Paradoxes are usually the result of some "erroneous assumptions". In example c), we are making an erroneous assumption of some kind. But it is not obvious what the erroneous assumption is. Should we allow ourselves to talk of a "set that contains itself as an element"? Or maybe the erroneous assumption is to say that there exists a "set that contains all sets". How do we decide which assumptions are acceptable ones and which are not? This problem motivated mathematicians to determine as clearly as possible what are acceptable properties for "sets". Since most of modern mathematics can be derived from the notion of "sets" this question was labeled "High priority" in the early 1900's. It is in this period that particular attention was given to developing a reliable *axiomatic system* which could serve as a foundation of modern mathematics. The modern set-theoretic axiomatic system which evolved as a result of these efforts is the main topic of this book.

1.3 Axiomatic systems.

An axiomatic system is normally set up by first declaring some *primitive concepts* or *undefined notions*. These *primitive concepts* carry no intrinsic meaning although the symbols or words used to represent them often convey some intuitive concept in the mind of the reader. That is, the words which represent this undefined notion are such that the user will more easily understand the properties which will be prescribed for this concept. Specific rules and properties which declare how these concepts relate to each other are then formulated; these rules and properties must allow mathematical constructs which are viewed as being important in our mathematical system. These properties and rules are called the "axioms".

Euclid's axiomatic system. Euclid provided us with a useful model for constructing an axiomatic system. He is the first person known to apply the axiomatic method to

study the field of geometry. In his axiomatic system, the words "line" and "point" are primitive concepts. He instinctively recognized that some undefined terms would be required. He then described properties of "lines" and "points". These properties are his *axioms*. These axioms are statements whose true-false values are not logically deduced from statements previously shown to be true. They are simply assumed to be true. The important point here is that he *explicitly* states what these "assumed to be true properties" are. The proposed primitive concepts and these axioms, when gathered together, constitute the foundation of the "Euclidean axiomatic universe" more commonly referred to as *Euclidean geometry*. Euclid justified his choice of axioms by saying that these point and line properties were "self-evident". Note that Euclid's primitive concepts and axioms differ entirely from the set-theoretic axioms we will be studying. But the axiomatic method he used to study geometry has served as a valuable model for others who wanted to develop different mathematical systems.

Euclid then used deductive reasoning to show that various geometric statements were true. The assumptions made were limited to

1) the stated axioms, along with
2) other statements previously shown to be true.

In this way, Euclidean geometry came to be. In spite of Euclid's best efforts, careful scrutiny of his work revealed that Euclid erred in certain ways. He unknowingly made assumptions which were neither stated as axioms nor previously proven to be true. In 1899, the mathematician David Hilbert revised the Euclidean axiomatic system by proposing three primitive concepts: point, straight line, plane. He also proposed 21 axioms. In 1902, one axiom was shown to be redundant and so was eliminated from the list. These primitive concepts along with 20 axioms are now widely accepted as forming a firm logical footing for Euclidean geometry.

1.4 The Zermelo-Fraenkel axiomatic system.

The axiomatic system of set theory as we know it today was in large part developed in the period of 1908 to 1922 by Ernst Zermelo and Abraham Fraenkel. Mathematicians T. Skolem and John Von Neumann made slight modifications to these a few years later. These axioms are now referred to as the *ZF-axioms* which stands for the "Zermelo-Fraenkel axiomatic system".

In what follows we will strive to develop an intuitive understanding of what axiomatic set theory is all about. We will avoid *logical formalism*, a treatment of axiomatic set theory based entirely on symbols normally reserved for more advanced courses. Our approach to set theory is referred to as "naive set theory" in the sense that we use ordinary language (words, sentences) to better understand what the basic axioms actually mean. We will see how these axioms are used to construct well-known sets such as the natural numbers and the real numbers. Finally, we will see how and why the chosen axioms serve as a widely accepted foundation of modern mathematics.

Primitive concepts and notation. What makes a set of primitive concepts and axioms suitable for a particular theory? Most will agree that these must satisfy the following conditions:

1) The number of undefined terms and axioms must be as few as possible.
2) Normally, an axiom should not be logically deducible from other axioms. (If one is deducible from the others this should be explicitly expressed.)
3) We should be able to prove from these axioms and concepts most of what we consider to be interesting or useful mathematics. Often, parts of the logical universe which is determined from the axioms are preconceived, in the sense that axioms are introduced so that certain mathematical statements will turn out to be true. This, of course, can turn out to be a "dangerous game". But there has to be some motivation for choosing one statement as an axiom rather than choosing another.
4) These axioms must not lead to any paradoxes. An axiomatic system which contains contradictions is either modified to one in which these contradictions do not occur or some axioms are simply discarded and replaced with others if needed.

The primitive concepts in our theory. There will be three undefined notions in our axiomatic system. They are the words:

"class"
"set"
"belongs to"

The expression "*belongs to*" is often stated as "*is a member of* " or "*is an element of* "; it is usually abbreviated by the symbol, $\in$.
All objects in our theory are *classes*. There is nothing else. We will soon distinguish special kinds of classes. Once we have discussed a few axioms, we will define a *set* as being a special kind of *class*. A *class* which is not a *set* will be called a *proper class*. Some axioms will help us distinguish between those *classes* which are *sets* and those which are *proper classes*. *Classes* will be represented either by lower-case or upper-case letters. So we can write "Let x and A be two classes".

The expression

$$x \in A$$

is to be read as "the class x belongs to the class A", or "the class x is in the class A" or "x is an element of A". However, no class will be representable by a lower-case letter, x, unless it is known that $x \in B$ for some class B. Those classes which can be represented by a lower-case letter, say x, will be given a special name:

> *If a class A is such that $A \in B$ for some class B, then we will refer to the class A as being an "element".*

Elements are still classes; but they are special classes, since they "belong to" another class. So an element can be represented by both a lower-case or an upper-case letter. For example, if we write $x \in y$ or $A \in B$ this means that x, y and A are *elements*, while B may or may not be an *element*.

Why is "element" not an undefined notion? The reader may find surprising that the object *element* is not expressed as an undefined notion. After all, we are accustomed to distinguishing *elements* from *sets*. Introducing a fourth undefined term was eventually seen as being superfluous. This became clear when we realized that we often view sets as being "elements" of other sets.[1] Witness:

- Points (a, b) in the Cartesian plane are actually two-element sets $\{a, b\}$ of real numbers stated in a particular order.
- Rational numbers a/b can be described as the set of all two-element sets $\{a, b\}$ of integers in a particular order where b is not 0.
- Irrational numbers can be viewed as infinite sequences of rational numbers converging to a non-rational number, again a set.

1.5 The axioms of set theory.

We now give a "preview" of what the set-theoretic axioms are, keeping in mind that a full understanding of what they mean will only be developed when we actually invoke each of these in various situations where they are required. These are called the *ZF*-axioms. The reader will see how surprisingly few are required. At this point, much of this will look like gibberish, but as we prod through the subject matter, we will step by step develop a better understanding of what they mean.

Primitive concepts: *"class"*, *"set"* and *"belongs to"*.

Axiom A1 (Axiom of extent): For the classes x, A and B, $[A = B] \Leftrightarrow [x \in A \Leftrightarrow x \in B]$

Axiom A2 (Axiom of class construction): Let $P(x)$ designate a statement about x which can be expressed entirely in terms of the symbols $\in, \vee, \wedge, \neg, \Rightarrow, \forall$, brackets and variables x, y, z, ..., A, B, ... Then there exists a class C which consists of all the elements x which satisfy $P(x)$.

Axiom A3 (Axiom of pair): If A and B are sets, then the doubleton $\{A, B\}$ is a set.

[1] There exists a branch of set theory in which mathematical entities which are neither sets nor classes are considered. These are referred to as "urelements". We will not consider these in this text.

Axiom A4 (Axiom of subsets): If S is a set and ϕ is a formula describing a particular property, then the class of all sets in S which satisfy this property ϕ is a set. More succinctly, every subclass of a set of sets is a set.[1]

Axiom A5 (Axiom of power set): If A is a set, then the power set $\mathscr{P}(A)$ is a set.

Axiom A6 (Axiom of union): If $\mathscr{A}$ is a set of sets, then $\bigcup_{C\in\mathscr{A}} C$ is a set.

Axiom A7 (Axiom of replacement): Let A be a set. Let $\phi(x, y)$ be a formula which associates to each element x of A an element y in such a way that whenever both $\phi(x, y)$ and $\phi(x, z)$ hold true, $y = z$. Then there exists a set B which contains all elements y such that $\phi(x, y)$ holds true for some $x \in A$.[2]

Axiom A8 (Axiom of infinity): There exists a non-empty class A called a *set* that satisfies the condition: "$X \in A$" $\Rightarrow$ "$X \cup \{X\} \in A$". (A set satisfying this condition is called a *successor set* or an *inductive set*.)

Axiom A9 (Axiom of regularity): Every non-empty set A contains an element x whose intersection with A is empty.

Another "special" axiom is usually stated separately from the other nine axioms above. It is viewed by many as being different in nature. It was also, at least initially, quite controversial. It is called the *Axiom of choice*.

Axiom of choice: For every set $\mathscr{A}$ of non-empty sets there is a function f which associates to every set A in $\mathscr{A}$ an element $a \in A$.

In this text we will refer to this set of nine axioms viewed together with the *Axiom of choice* as "ZF + Choice" or simply by *ZFC*.[3]

The Axiom of choice. The controversy surrounding the Axiom of choice requires some explanation. The *Axiom of choice* is an axiom which was added after most of the *ZF*-axioms were widely accepted as a foundation for modern mathematics. It is so subtle a concept that many early mathematicians unknowingly invoked it in their proofs. That is, it was invoked without stating it explicitly as an assumption. Some

[1] This axiom is more often expressed as the Axiom of comprehension, Axiom of Specification or Axiom of separation. It is in fact many axioms (which, when viewed together, are referred to as a *schema*) each differing only by the formula ϕ it refers to. So to be more precise, given a formula ϕ in set theory language, we would refer to it as axiom A4(ϕ) rather than A4.

[2] This axiom is more often expressed as the *Replacement axiom schema* since it is in fact many axioms each differing only by the formula ϕ it refers to. So to be more precise, given a formula ϕ in set theory language, we would refer to it as axiom A7(ϕ) rather than A7. It essentially allows us to confirm that if the domain A of a set f is a set, then the image $f[A]$ is a set.

[3] Note that some of the ZF axioms listed have been shown to follow from the others. So some set theory texts may omit one or more of these from their formal list of ZFC axioms. Since most of these axioms are non-controversial we will adopt, for this text, this list of 10 axioms as the ZFC-axioms. The reader should simply be alerted to the fact that the list of the ZFC axioms may vary from text to text.

mathematicians publicly questioned this assumption, asking openly whether the word "obviously" was sufficient justification for using it. These questions could not be ignored. Numerous attempts at proving this axiom from the *ZF*-axioms failed. In 1963, it was finally proven that neither the Axiom of choice, nor its negation, can be proven from the *ZF*-axioms. This implied that we are free to state it as an axiom, along with the other *ZF*-axioms, without fear of producing a contradiction. A lengthy debate on whether this statement should be included with the other "fundamental" *ZF*-axioms followed. Some described it as "the most interesting and, in spite of its late appearance, the most discussed axiom of mathematics, second only to Euclid's axiom of parallels which was introduced more than two thousand years ago" (Fraenkel, Bar-Hillel & Levy 1973). Eventually, it was felt that "not accepting" the Axiom of choice closes the door to many fundamentally important results of modern mathematics. One could say that the Axiom of choice had already been used so extensively that it was deeply ingrained in the modern mathematical fabric; we were "addicted" to the Axiom of choice, so to speak.

Even though proofs that invoke the Axiom of choice are widely viewed as being acceptable, it is often felt that a correct proof that does not invoke the Axiom of choice is preferable to a simpler proof which invokes it. This is because it assumes the existence of something that can neither be seen nor constructed. It is viewed somewhat pejoratively by some as the "magic wand" that magically opens closed doors. For this reason, when proving a statement, it is customary to point out explicitly the steps where the Axiom of choice is invoked. Actually, there is a general consensus on one point: The Axiom of choice should not be listed with the *ZF*-axioms; it should be set apart in a category of its own. This is why we refer to this group of axioms as "*ZF*+Choice", or simply *ZFC*. One can view this as some sort of compromise.

1.6 A few more words on these axioms.

Even though the words *class* and *set* are undefined, the axioms will allow us to perceive them (once we can decode them) as "collections of objects". It is still too early to extract the full meaning of the axioms stated above. But the reader will feel more at ease if we interpret at least certain aspects of these immediately.
The axiom A1,

$$\text{"For the classes } x, A \text{ and } B, [A = B] \Leftrightarrow [x \in A \Leftrightarrow x \in B]\text{."}$$

is an axiom which states that "a class is defined by its *elements*". If two classes are equal, then they have the same elements. Conversely, two classes which have the same elements are the same class.

We now examine more closely the axioms A2, A9 and the *Axiom of choice*, in random order.

Axiom A2: (Class construction)

"Let $P(x)$ designate a statement about x which can be expressed entirely in terms of the symbols $\in, \vee, \wedge, \neg, \Rightarrow, \forall$, brackets and variables x, y, z, ..., A, B, ... Then *there exists* a class C which consists of all the elements x which satisfy $P(x)$."

states that we can use well-defined properties which can be expressed by the given symbols to construct *classes*. For example $A = \{X : u \in X\}$ and $B = \{X : X = X\}$ are different classes since the properties that characterize their elements are different. For the class A, $P(x)$ is the property $u \in X$ while for the class B, $P(x)$ is the property "$X = X$".

Axiom A8 (Infinity):

"There exists a non-empty class A called a *set* that satisfies the condition:

$$\text{"}X \in A\text{"} \Rightarrow \text{"}X \cup \{X\} \in A\text{"}."$$

says that there exists a class called a *set* which is infinite in size. (This axiom also guarantees that at least one class called a set exists.) It essentially allows us to define the "natural numbers", $0, 1, 2, 3, \ldots,$.

Axiom of choice:

"For every set $\mathscr{A}$ of non-empty sets there is a rule f which associates to every set A in $\mathscr{A}$ an element $a \in A$."

says that given a set of non-empty sets, *there exists* a certain type of function. But it does not show how to construct or find such a function.

Note that axioms A1 and A2 refer only to classes while all the other axioms (Axioms A3 to A9 and the Axiom of choice) are "set axioms". The set axioms determine what kind of objects exist in the universe of all sets.

Axioms A2, A3, A4, A5, A6 and A7 are "constructive" axioms since A2 gives us a way to construct a class by referring to a property. Axioms A3 to A7 provide a method to construct new sets from ones that are known to exist.

Axiom A9, the Axiom of regularity, is sometimes referred to as the "useless" axiom by some. Others don't consider it as a basic axiom since most of mathematics which is based on set theory does not require it. It will be invoked only in the last chapter of this book. Although it is not obvious, just from reading it, this axiom actually states that "those non-empty classes which don't have a least element are not sets". It is in fact an axiom which does not allow certain types of sets to exist in the universe of sets. It is of an exclusionary nature. The other axioms (except for axiom A1) increase the number of sets in the universe of sets.

The Axioms A4 (Subsets) and A7 (Replacement) each represent many axioms. We refer to such axioms as *schema*.[1] Axiom A4 speaks of a set S and a particular formula ϕ describing a property. For each property we have a different Axiom. Given ϕ, we could say the "Axiom A4 for ϕ". Axiom A7 speaks of a set A and a class B of sets along with a particular formula $\phi(x, y)$ which plays the role of a function (normally referred to as a *functional*). For each functional, $\phi(x, y)$, we have a different axiom.

1.7 Some things we may immediately wonder about.

As one may suspect, when formally expressed, axioms do not contain words. The *ZFC*-axioms are expressed using symbolism of first order logic. For example, when formally stated, an axiom may look like this:

$$\forall x \forall y \exists z (x \in z \vee y \in z)$$

This is the *Axiom of pair*. We use the words and sentences to develop an intuitive understanding of what this code means.

A second point one may wonder about: Are the *ZF*-axioms *consistent*? That is, do we know for sure that the *ZF*-axioms, as stated, will never yield some contradiction? If one day we actually encounter a contradiction that flows from these axioms, then we can answer: "No, the *ZF*-axioms are not consistent, since we have revealed a contradiction which flows from these axioms!" If such a contradiction is discovered we must tinker some more with the set-theoretic axioms to correct the flaw.

But as long as we do not encounter such a paradox, the answer to this question is: "We don't know for sure whether the *ZF*-axioms are consistent." It has been shown that using only the ZF-axioms, it is impossible to prove or disprove that the ZF-axioms are consistent. It is the "nature of the beast", so to speak. Since new forms of mathematics is uncovered every day, it is possible that next week, in a hundred years or in a thousand years someone will discover that ZF is inconsistent. "Set theory" is, as the words indicate, just a theory. By their very nature, all theories evolve to explain newly discovered previously unknown facts. The ZF-set-theoretic system is no different. As a foundation of modern mathematics, the ZF-set-theoretic system seems to serve its purpose well; it is the best theory we have today, even though some day we may discover significant ways of improving it.

Concepts review:

1. What is Russell's paradox?
2. Why do paradoxes occur?

[1] A dictionary describes *schema* as meaning "an underlying organizational pattern or structure".

3. What are three primitive concepts of set theory?
4. What is the difference between a *class*, a *set* and a *proper class*?
5. When is a class called an *element*?
6. Which classes can be represented by a lower case letter?
7. What does *ZFC* stand for?
8. How many axioms belong to *ZFC*?

2 / Constructing classes and sets.

Summary. *In this section we define the symbols "=" and "⊆" and discuss Axioms A1 to A5. We show how Axioms A1 and A2 are used to construct classes. Axioms A3 to A5 are used as tools to construct sets. We distinguish between classes, sets and elements by exhibiting a class which is not an element and a class which is not a set. We also show that all sets are elements. We introduce the concept of "power set of a set" as a set constructing tool.*

2.1 Basic statements, definitions and notation.

To discuss the axioms and some of their immediate consequences we first define a few words and symbols that will allow us to communicate certain ideas more efficiently. We first confirm that every class is equal to itself. This is not an axiom, since it is an immediate consequence of axiom A1.

Theorem 2.1 For any class C, $C = C$.

This follows from axiom A1: Since $x \in C \Rightarrow x \in C$ and $x \in C \Rightarrow x \in C$ then $C = C$.

If the statement "$A = B$" is false then we will write $A \neq B$.

Definition 2.2 If A and B are classes or sets we define $A \subseteq B$ to mean that every element of A is an element of B. That is,

$$A \subseteq B \text{ if and only if } x \in A \Rightarrow x \in B$$

If $A \subseteq B$ we will say that A is a *subclass* (*subset*) of B. If $A \subseteq B$ and $A \neq B$ we will say that A is a *proper subclass* (*proper subset*) of B; in this case we write $A \subset B$. So "$A \subset B$" is a shorter way of saying "$A \subseteq B$ but $A \neq B$".[1]

We restate the *Axiom of construction:*

> **Axiom A2**: If $P(x)$ is a property of an element x which can be expressed entirely in terms of the symbols $\in, \vee, \wedge, \neg, \Rightarrow, \forall$, brackets and variables x, y, z, ..., A, B, ..., then there exists a class C which consists of all the elements x which satisfy $P(x)$.

[1] Do not confuse $\subset$ with $\in$. When we say that "the class A belongs to the class B" we mean that $A \in B$, not $A \subset B$.

This axiom refers to properties of elements which can be expressed in terms of logical symbols. Many students may have used some or all of these symbols before; for completeness, we explicitly state how these symbols should be interpreted:

$\in$ is given the meaning "is an element of"

$\vee$ is given the meaning "or"

$\wedge$ is given the meaning "and"

$\neg$ is given the meaning "not"

$\Rightarrow$ is given the meaning "implies"

$\exists$ is given the meaning "there exists"

$\forall$ is given the meaning "for all"

The Axiom of construction allows us to construct a class by first defining a property P and then gathering together all the *elements* possessing this property to form the class

$$C = \{A : \ A \text{ possesses the property } P\} = \{x : P(x)\}$$

This class is succinctly expressed as: $\{x : P(x)\}$. The brackets { } refer to a "class". The symbol, $P(x)$, means "x possesses property P". The lower case symbol, x, refers to a "class which is an element of another class". Elements are the only classes that can be denoted by a lower case letter.

A word of caution: Axiom A2 allows us to gather together all the "*elements*" that possess a property P, **not** all "*classes*" that possess a property P. For if the word "element" is replaced with the word "class", then we easily obtain a paradox. Witness the class, C, defined as follows:

$$C = \{A : A \text{ is a class and } A \text{ satisfies the property } \neg(A \in A)\} = \{A : A \notin A\}$$

which leads to Russell's paradox, since neither $C \notin C \Leftrightarrow C \in C$ is both true and absurd.

Also, in expressions such as

$$\text{"the class } \{C : C = C\}\text{"}$$

it is understood that C must be an element. The expression $\{x : x = x\}$, means the same thing except it emphasizes that the classes it refers to are elements.

2.2 Properties of classes.

Axiom A1 allows classes to have the properties normally attributed to those things we call "collections of objects". After all, this is what we would like classes and sets to be. The axioms are developed with a preconceived idea of what the only objects

(classes and sets) in our set-theoretic universe are. The statements in the following theorem are all logical consequences of these axioms. They are easily seen to hold true, but it is good practice to explicitly write out the proofs.

Theorem 2.3 If C, D, and E are classes (sets) then:

a) $C = D \Rightarrow D = C$.

b) $C = D$ and $D = E \Rightarrow C = E$.

c) $C \subseteq D$ and $D \subseteq C \Rightarrow C = D$.

d) $C \subseteq D$ and $D \subseteq E \Rightarrow C \subseteq E$.

Proof:

a) Suppose $C = D$. Then by axiom A1, $x \in C \Rightarrow x \in D$ and $x \in D \Rightarrow x \in C$. Then $x \in D \Rightarrow x \in C$ and $x \in C \Rightarrow x \in D$. Hence, by definition of equality $D = C$.

Proofs of b) to d) are left as an exercise.

We said that all objects in our set-theoretic universe are classes. Some of those classes are elements provided these belong to another class. It is normal to ask whether there exists at least one class which is not an element. We answer this question in the following theorem.

Theorem 2.4 There exists a class which is not an *element*.

Proof:

Let C be the class

$$C = \{A : A \text{ is an element and } A \text{ satisfies } \neg(A \in A)\} = \{x : x \notin x\}$$

By Axiom A2, the class C is well-defined. Suppose C is an element. Then either C belongs to $\{x : x \notin x\}$ or it doesn't. Both of these options lead to a contradiction. Then C is not an element, as required.

2.3 The *universal class* and the *empty class*.

The class $\mathscr{U} = \{x : x = x\}$ is easily seen to be the class that contains precisely all elements. By theorem 2.3 part a), "Every element is equal to itself". The class $\mathscr{U}$ is called the *universal class*. In the proof of the theorem 2.4 we constructed a class C

which does not belong to $\mathscr{U}$. Hence, $\mathscr{U}$ does not contain all classes.

At this point, we have discussed only two of the primitive concepts: *class* and "*belongs to*". From these we have defined "*element*". The reader has maybe noticed that the word "*set*" has remained on the sidelines. We have not yet discussed this primitive concept, other than witnessing the very large set given to us "for free" by the Axiom of infinity. The Axiom A2 will allow us to construct a much smaller set. We start with the following definition.

Definition 2.5 The Axiom A2 authorizes us to call $C = \{x : x \neq x\}$ a *class*. Since we have proven above that every element is equal to itself, then this class contains no elements. We will call the class with no elements the *empty class* and denote it by $\varnothing$.

Theorem 2.6 For any class C, $\varnothing \subseteq C$.
Proof:

Let C be a class. To show that $\varnothing \subseteq C$ it suffices to show, by definition of "$\subseteq$", that $x \in \varnothing \Rightarrow x \in C$. Any element in $\varnothing$ belongs to C since $\varnothing$ contains no elements; then $\varnothing \subseteq C$, as required.[1]

2.4 Sets which are derived from other sets.

Axiom A1 "If $x = y$ and $x \in A$, then $y \in A$" is a statement about elements x, y and the class A. Since every set is a class, axioms referring to classes also refer to sets. Axiom A2 is a statement which shows how to construct classes by referring to some property, $P(x)$; it does not refer to properties specific to sets only. On the other hand, the axioms three and four are set-specific:

> **Axiom A3** (Axiom of pair): If A and B are sets, then the doubleton class $C = \{A, B\}$ is a set.
> **Axiom A4** (Axiom of subsets): Every subclass of a set is a set.

These two axioms alone will allow us to construct sets from those classes known to be "sets". Axiom A8 guarantees that at least one class called "set" exists: It contains the words "...there exists a class A called a *set* that...". We need not search any further.

The axiom A3 refers to the set $C = \{A, B\}$ as a "doubleton". We will use the word *doubleton* when referring to two sets A and B viewed together to form a collection $\{A, B\}$ of sets. For convenience, we will not put any restrictions on how the set B relates to A. For example, we can refer to the set $\{A, A\}$ as a doubleton even though it contains only one element.
The statements in the following theorem follow immediately from the axioms A3 and A4.

[1] Alternatively, the statement $x \notin C \Rightarrow x \notin \varnothing$ is the logical equivalent of $x \in \varnothing \Rightarrow x \in C$. If x is not an element of C, then x is not an element of $\varnothing$ since $\varnothing$ has no elements.

Theorem 2.7 Let S be a set. Then:

a) $\varnothing \subseteq S$ and so $\varnothing$ is a set.

b) The set S is an element. Hence, "*All sets are elements*".

Proof:

a) We are given that S is a set. We are required to prove that $\varnothing$ is a set. We can directly apply axiom A4: Since $\varnothing = \{x \in S : x \neq x\}$ and, by hypothesis, S is a set then, by A4, $\varnothing$ is a set as required.

b) We are given that S is a set. We are required to prove that S is an element. By axiom A3 (Axiom of pair), for any set S, $\{S, S\}$ is a set. Since $S \in \{S, S\}$, for all sets S, then, by definition, S is an element, as required.

In the proof above we discussed the set $\{S, S\}$ which contains the set S as an element. Since $\{S, S\} = \{S\}$ (Prove this!) this is a one element class. We call such sets *singleton sets*. The reader should note that according to our definition of *doubleton* above, every singleton set $\{A\}$ can be expressed as a doubleton $\{A, A\}$.

We can now verify that the universal class $\mathscr{U} = \{x : x = x\}$ is not a set. Suppose $\mathscr{U}$ is a set. See that the class $C = \{x \in \mathscr{U} : x \notin x\}$ is a subclass of $\mathscr{U}$. Since we assumed $\mathscr{U}$ to be a set, by the Axiom of subset, C must also be a set. But we showed in theorem 2.4 that the class C is not an element and so cannot be a set. We have a contradiction. Therefore the universal class is a proper class, as claimed.

2.5 Examples of sets which are non-empty.

At this point we have only exhibited one set, the empty set $\varnothing$. In the following examples we we use some axioms to construct other sets.

a) The set $\varnothing$ contains no elements. By axiom A3, the class $C = \{\varnothing, \varnothing\} = \{\varnothing\}$ is a set which contains exactly one element (the element $\varnothing$). Observe that $\varnothing \neq \{\varnothing\}$ since $\{\varnothing\}$ contains one element while $\varnothing$ does not.

b) Let $A = \varnothing$ and $B = \{\varnothing\}$. By axiom A3, $C = \{\varnothing, \{\varnothing\}\}$ is a set which contains exactly two elements (the element $\varnothing$ and the one element set $\{\varnothing\}$).

c) Let a, b, c be 3 sets. Then, by repeated applications of axiom A3, $\{a, \{a\}, \{\{a\}\}, \{a, b, c\}\}$ is a 4 element set.

d) Let c be an element (class). Then $A = \{c\}$ is a class with only one element since $A = \{x : x = c\}$ and so, by axiom A2, $A = \{c\}$ is a class. If c is known to be a set, $A = \{c\} = \{c, c\}$, so we can conclude that A is a set.

2.6 The class of all sets.

Recall that the word *set* is a primitive concept (along with the word *class*). We define the property symbol "set(x)" to mean "x is a set". Then by axiom A2, $\mathscr{S} = \{x : \text{set}(x)\}$ = "all elements which are sets" forms a class. Since every set is an element (by theorem 2.7), we can say $\mathscr{S} = \{x : \text{set}(x)\}$ is a subclass of the universal class, $\mathscr{U} = \{x : x = x\}$. We call $\mathscr{S}$ the *class of all sets*. From this we observe that:

Given any property P,

$$S = \{x : \text{set}(x) \wedge P(x)\} = \{x : x \text{ is a set and } x \text{ satisfies } P\}$$

is a class.

Axiom A2 said that $\{x : P(x)\}$ = "all elements which satisfy property P" is a class. Now it makes sense to talk about the "*class of sets satisfying a property P*".

Note that the class, $\mathscr{S}$, of all sets is a proper class. To see this, suppose $\mathscr{S}$ was a set. Let $D = \{x \in \mathscr{S} : x \notin x\}$. The class D cannot be a set, for if it was, then as previously shown, we would quickly obtain the contradiction, $D \in D$ and $D \notin D$. See that D is a subclass of $\mathscr{S}$. Since $\mathscr{S}$ was assumed to be a set, by the Axiom of subset, D must be a set. We have a contradiction. The source of the contradiction is our assumption that $\mathscr{S}$ is a set. So $\mathscr{S}$ is a proper class.

2.7 Power sets.

Axiom A3 allows us to construct new sets from known ones by forming doubleton sets, while Axiom A4 allows us to construct sets by taking subclasses of sets and calling them *subsets*. Axiom A5 will allow us to construct, from a known set A, what seems to be a larger set, $\mathscr{P}(A)$. It is called the *power set of A*. We define "power set".

Definition 2.8 If A is a set, then we define the *power set of A* as being the class $\mathscr{P}(A)$ of all subsets of A. It can be described as follows:

$$\mathscr{P}(A) = \{X : X \subseteq A\}$$

We verify the following facts:

- By axiom A4, "$X \subseteq A$" $\Rightarrow$ "X is a set" so all elements of $\mathscr{P}(A)$ are sets.
- By axiom A2, $\mathscr{P}(A)$ is a class. (The fact that "all sets are elements" is proved above).

No one has been able to prove that $\mathscr{P}(A)$ is a set. So if we want it to be a set, we must postulate this fact. Axiom A5 does precisely that. We will later see why this axiom plays a fundamental role in the mathematical universe we are exploring today.

Axiom A5: If A is a set, then the power set, $\mathscr{P}(A)$, is a set.

In formal language, the Axiom of power set actually reads as follows:

$$\forall A \exists P[B \in P \Longleftrightarrow B \subseteq A]$$

where A is specified to be a set. This expression includes the definition of the "power set $\mathscr{P}(A)$ of a set A". The interior of the square brackets specifies that "the elements B of the power set $\mathscr{P}(A)$ are precisely the subsets B of A". The sequence of symbols "$\forall A \exists P$" instructs the reader that given any set A, the class $\mathscr{P}(A)$ exists as a set. This axiom also guarantees that a set A is an element since A belongs to the set $\mathscr{P}(A)$. Declaring that the power set of any set is a set adds many sets to our universe of sets. There is, of course, some risk in doing this, since we may be allowing sets in our universe of sets which are so strange that we will not be sure whether we want them there or not. On the other hand, we will see that the *Power set axiom* is extremely useful for constructing sets we need. For example, the *Power set axiom* allows us to construct Cartesian products.

The four axioms A6, A7, A8 and A9 have not yet been discussed. We will study these only when we require them later on.

2.8 Examples.

We provide a few exercises which allow us to practice notions related to power sets.

1) *Power sets.* List the elements of the power set of
   a) the empty set, $\varnothing$.
   b) a singleton set.
   c) a doubleton set.

   Solution:

   a) The power set of the empty set: $\mathscr{P}(\varnothing) = \{X : X \subseteq \varnothing\}$. If $X \in \mathscr{P}(\varnothing)$, then $X \subseteq \varnothing$. Thus, $X = \varnothing$. So
   $$\mathscr{P}(\varnothing) = \{\varnothing\}$$
   b) The power set of a singleton set $\{x\}$: For an element x, $\mathscr{P}(\{x\}) = \{\varnothing, \{x\}\}$. Note that $x \nsubseteq \{x\}$ since the elements of x are not in $\{x\}$, a single element set.
   c) The power set of a doubleton set $\{x, y\}$: For the elements x and y, $\mathscr{P}(\{x, y\}) = \{\varnothing, \{x\}, \{y\}, \{x, y\}\}$.

2) Consider the 3-element class $C = \{x, \{x, y\}, \{z\}\}$. Determine which of the following statements are true and which are false.
   a) We can write $x \in C$.
   b) We can write $x \subseteq C$.

c) We can write $\{x\} \subseteq \mathscr{P}(C)$.
d) We can write $\{\{z\}\} \in \mathscr{P}(C)$.
e) We can write $z \in \mathscr{P}(C)$.
f) We can write $\{z\} \subseteq C$.

Solution:

a) True. We can write $x \in C$ since x is an element explicitly listed as a class in C.
b) False. We cannot write $x \subseteq C$ since this does not satisfy the definition of $\subseteq$. To write $x \subseteq C$ is to say that every element in x is an element in C. But the contents of x is unknown. So there is no basis to state that $x \subseteq C$.
c) True. We can write $\{x\} \subseteq \mathscr{P}(C)$ since every element in $\{x\}$ is also an element of C.
d) True. We can write $\{\{z\}\} \in \mathscr{P}(C)$ since $\{\{z\}\} \subseteq C$. The only element in $\{\{z\}\}$ is in C.
e) False. We cannot write $z \in \mathscr{P}(C)$ since the element z does not appear as an element of C.
f) False. We cannot write $\{z\} \subseteq \mathscr{P}(C)$ since $\{z\}$ contains only one element z. This element is not a subset of C.

It was shown above that "*all sets satisfying a property P*" is a class. In the following example we say something similar. But there are subtle differences in the statement. See if you can detect these differences.

3) Let A be a set and P denote some property. Show that the class

$$S = \{x : (x \subseteq A) \wedge P(x)\}$$

is a set.

Solution:

We are given that A is a set and $S = \{x : (x \subseteq A) \wedge P(x)\}$. We are required to show that S is a set. Since A is a set, and, for every $x \in S$, $x \subseteq A$ then every $x \in S$ is a set (by axiom A4). By axiom A5, $\mathscr{P}(A)$ is a set. Since the class $S \subseteq \mathscr{P}(A)$, then S is a set (by axiom A4). This is what we were required to prove.

Concepts review:

1. What does it mean to say "the class A is equal to the class B", $A = B$?
2. What does it mean to say "the class A is contained in the class B", $A \subseteq B$?
3. What does it mean to say the class A is a *proper subclass* of the class B?

4. How should we read the expression $C = \{x : P(x)\}$?
5. Is it true that $\varnothing \notin \varnothing$? Why?
6. State a class that is not an *element*.
7. What is the *universal class*?
8. What is the *empty class* $\varnothing$?
9. Is a *set* an *element*?
10. Given a set A, what is the *power set*, $\mathscr{P}(A)$, of A? How do we know that $\mathscr{P}(A)$ is a set?
11. If B is a set and $A \subseteq B$, what can we say about A? Why?
12. Why is $\varnothing$ a set?

EXERCISES

A.

1. Suppose A is a proper class and $A \subseteq B$. Show that B is a proper class.
2. Prove the following.
   a) $\{c, d, e\} = \{c, d\}$ if and only if $e = c$ or $e = d$.
   b) $c = \{d\} \Rightarrow d \in c$
3. If $x = \{u, v\}$, $y = \{v, \{w\}\}$ and $z = \{x, y\}$ write out explicitly the elements of the following classes: $\mathscr{P}(x)$, $\mathscr{P}(y)$, $\mathscr{P}(\mathscr{P}(x))$ and $\mathscr{P}(z)$
4. Prove parts c), d) and e) of theorem 2.3.
5. For sets S and T show that:
   a) $S \subseteq T$ if and only if $\mathscr{P}(S) \subseteq \mathscr{P}(T)$
   b) $S = T$ if and only if $\mathscr{P}(S) = \mathscr{P}(T)$
6. Show all the elements in the set $\mathscr{P}(\mathscr{P}(\varnothing))$.
7. Show all the elements in the set $\mathscr{P}(\mathscr{P}(\mathscr{P}(\varnothing)))$.
8. Show that $[S \subset T)] \wedge [T \subseteq V)] \Rightarrow (S \subset V)$.
9. Show that $[S \subseteq T)] \wedge [T \subset V)] \Rightarrow (S \subset V)$.
10. Show that $c = d$ if and only if $\{c\} = \{d\}$.
11. Show that $c \in d$ if and only if $\{c\} \subseteq d$.

B.

12. Show that $(S \subseteq \varnothing) \Rightarrow (S = \varnothing)$.
13. Using the empty set, $\varnothing$, construct a set containing 7 elements.
14. If $A = \{\varnothing, \varnothing, \varnothing, \{\varnothing, \varnothing\}\}$ and $B = \{\varnothing, \{\varnothing\}\}$ show that $A = B$.

15. Show that the statement "For any set S, $\mathscr{P}(S) \subseteq S$" is a false statement.
16. Suppose U and V are sets. Determine whether the statement $\mathscr{P}(U) \cup \mathscr{P}(V) = \mathscr{P}(U \cup V)$ is true or false. Justify your answer.
17. Suppose U and V are sets. Determine whether the statement $\mathscr{P}(U) \cap \mathscr{P}(V) = \mathscr{P}(U \cap V)$ is true or false. Justify your answer.

Part II

Class operations

3 / Operations on classes and sets.

Summary. *In this section we define operations on classes that will allow us to better see how classes relate to each other. The concept of unions, intersections and complements of classes and sets are defined. The axiom A6 (Axiom of union) is discussed. This axiom is used to determine when the union of sets is a set. We show how Venn diagrams can serve as a guide when interpreting operations on sets. Some basic laws for the union, intersection and complements of classes and sets are presented in the form of theorems. De Morgan's laws are also stated.*

3.1 Unions, intersections and complements of classes.

The union and intersection of classes is a method for constructing new classes and sets from old ones. We begin by defining these formally.

Definition 3.1 Let A and B be classes (sets). We define the *union*, $A \cup B$, of the class A and the class B as

$$A \cup B = \{x : (x \in A) \vee (x \in B)\}$$

That is, $x \in A \cup B$ if and only if $x \in A$ or $x \in B$. If $\mathscr{A}$ is a non-empty class of classes then we define the *union of all classes in* $\mathscr{A}$ as

$$\bigcup_{C \in \mathscr{A}} C = \{x : x \in C \text{ for some } C \in \mathscr{A}\}$$

That is, $x \in \bigcup_{C \in \mathscr{A}} C$ if and only if there exists $C \in \mathscr{A}$ such that $x \in C$.

Definition 3.2 Let A and B be two classes (sets). We define the *intersection*, $A \cap B$, of the class A and the class B as

$$A \cap B = \{x : (x \in A) \wedge (x \in B)\}$$

That is, $x \in A \cap B$ if and only if $x \in A$ and $x \in B$. If $\mathscr{A}$ is a non-empty class of classes, then we define the *intersection of all classes in* $\mathscr{A}$ as

$$\bigcap_{C \in \mathscr{A}} C = \{x : x \in C \ \ \forall \ \ C \in \mathscr{A}\}$$

That is, $x \in \bigcap_{C \in \mathscr{A}} C$ if and only if $x \in C$ for every class C in $\mathscr{A}$.

Observe that

a) if $\mathscr{A} = \{D, E\}$ then $D \cup E = \bigcup_{C \in \mathscr{A}} C$.
b) if $\mathscr{A} = \{D, E\}$ then $D \cap E = \bigcap_{C \in \mathscr{A}} C$.

Also see that the axiom A2, *Axiom of construction*, guarantees that both $\bigcup_{C \in \mathscr{A}} C$ and $\bigcap_{C \in \mathscr{A}} C$ are classes. If the class $\mathscr{A}$ contains no elements, then by definition of "union" and "intersection" the union and intersection of all elements in $\mathscr{A}$ is $\varnothing$.

Definition 3.3 We will say that two classes (sets) C and D are *disjoint* if the two classes have no elements in common. That is, the classes C and D are disjoint if and only if $C \cap D = \varnothing$.

Definition 3.4 The *complement*, C', of a class (set) C consists of all *elements* which do not belong to C. That is, if C is a class, then

$$C' = \{x : x \notin C\}$$

Hence, $x \in C'$ if and only if $x \notin C$. Again, the axiom of construction (A2) guarantees that C' is a class. Given two classes (sets) C and D, the *difference* $C - D$, of C and D, is defined as

$$C - D = C \cap D'$$

This is also a class. The *symmetric difference*, $C \triangle D$, is defined as (the class)

$$C \triangle D = (C - D) \cup (D - C)$$

3.2 Unions and intersections referring specifically to sets.

Observe that $\bigcap_{C \in \mathscr{A}} C \subseteq C$ for all $C \in \mathscr{A}$ since, by definition, every element in $\bigcap_{C \in \mathscr{A}} C$ belongs to every $C \in \mathscr{A}$.

- If $\mathscr{A}$ is a non-empty class of sets then $\bigcap_{C \in \mathscr{A}} C$ is a subclass of every set C. So, by axiom A4, $\bigcap_{C \in \mathscr{A}} C$ is a set.
- Hence, we can say, in a general way, that the intersection of arbitrarily large classes $\mathscr{A}$ of sets is always a set.

If $\mathscr{A}$ is a non-empty class of sets is $\bigcup_{C \in \mathscr{A}} C$ necessarily a set? What about the special case where $\mathscr{A}$ is a set of sets? None of the axioms A1 to A5, nor any previously proven statement resulting from these help to answer this question. It will turn out to be useful if we can answer "yes" to the second question. However, we have been unable to prove this. Then, we need an axiom that will postulate this to be true. The axiom A6, *Axiom of union*, declares when a union of sets is a set. We restate it here:

Axiom 6: If $\mathscr{A}$ is a non-empty set of sets then $\bigcup_{C\in\mathscr{A}} C$ is a set.

Thus, axiom A6 says "The union of all sets in a set of sets is a set". We should also be clear about what axiom A6 *does not* say: "The union of all sets in a class of sets is a set." If we make the mistake of assuming this to be true it will lead to a contradiction, as the following example shows.

Example: Suppose $\mathscr{A}$ is the class $\mathscr{A} = \{x : x \text{ is a set and } x \notin x\}$. Show that $D = \bigcup_{x\in\mathscr{A}} x$ is **not** a set.

Solution:

What is given: $\mathscr{A} = \{x : x \text{ is a set and } x \notin x\}$
What we are required to show: $D = \bigcup_{x\in\mathscr{A}} x$ is **not** a set.
Suppose $D = \bigcup_{x\in\mathscr{A}} x$ is a set. Then by axiom A5, $\mathscr{P}(D)$ is also a set. We know that for every $x \in \mathscr{A}$, $x \subseteq D = \bigcup_{x\in\mathscr{A}} x$. (Make sure you see why. If not look at $A \subseteq A \cup B$.) So, for every $x \in \mathscr{A}$, $x \in \mathscr{P}(D)$. Hence, $\mathscr{A} \subseteq \mathscr{P}(D)$.
Since $\mathscr{A}$ is a subclass of a set, then, by axiom A4, $\mathscr{A}$ is a set.
We now argue as in theorem 2.4: Since $\mathscr{A}$ is a set, "$\mathscr{A} \notin \mathscr{A}$" $\Rightarrow$ "$\mathscr{A} \in \mathscr{A}$" and "$\mathscr{A} \in \mathscr{A}$" $\Rightarrow$ "$\mathscr{A} \notin \mathscr{A}$". This is a contradiction. So $D = \bigcup_{x\in\mathscr{A}} x$ cannot be a set. This is what we were required to show.

Then the statement "The union of all sets in a class of sets is a set" is **not** a true statement, in general. Even though, in set theory, both *class* and *set* intuitively represent a "collection of objects", freely substituting the word *set* with the word *class* may lead to some nasty consequences.

3.3 Venn diagrams.

We often use *Venn diagrams* as a tool to visualize how sets relate to others. Venn diagrams should not be substitutes for proofs of statement; but they are helpful when used to guide our intuition. We represent here Venn diagrams representing some relations defined above.

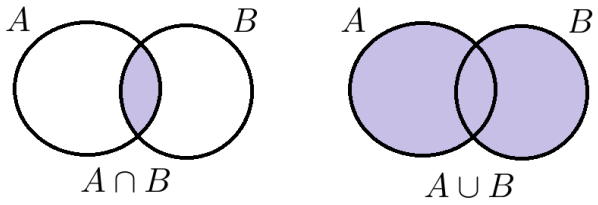


Figure 1: Intersection and union of two sets

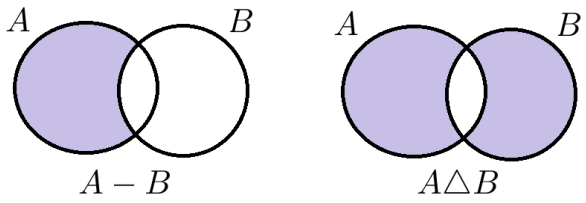


Figure 2: Difference and symmetric difference

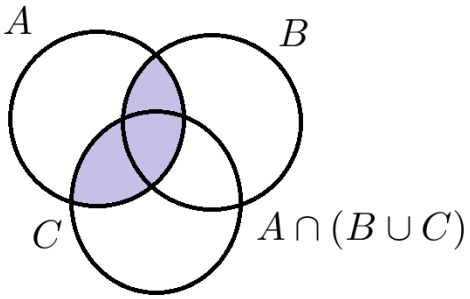


Figure 3: Intersection distributing over a union.

3.4 Basic laws for operations on classes and sets.

We now prove a few fundamental properties of sets. The proofs of less common properties will be left as exercises.

Theorem 3.5 Let C and D be classes (sets). Then

a) $C \subseteq C \cup D$

b) $C \cap D \subseteq C$

Proof:

a) Let $x \in C$. It suffices to show that $x \in C \cup D$.

$$\begin{aligned} x \in C &\Rightarrow x \in \{x : x \in C \text{ or } x \in D\} \\ &\Rightarrow x \in C \cup D \end{aligned}$$

b) The proof is left as an exercise.

Theorem 3.6 Let C and D be classes (sets). Then

a) $C \cup (C \cap D) = C$
b) $C \cap (C \cup D) = C$

Proof:

a) What we are given: C and D are classes (sets).
What we are required to show: $C \cup (C \cap D) = C$

$$\begin{aligned} C &\subseteq C \cup (C \cap D) && \text{(By theorem 3.5 a))} \\ C \cap D &\subseteq (C \cap D) \cup C && \text{(By theorem 3.5 a))} \end{aligned}$$

$$\begin{aligned} x \in (C \cap D) \cup C &\Rightarrow x \in C \cap D \text{ or } x \in C \\ &\Rightarrow x \in C \text{ or } x \in C && \text{(Since theorem 3.5 b) says } C \cap D \subseteq C) \\ &\Rightarrow x \in C \\ &\Rightarrow (C \cap D) \cup C \subseteq C \end{aligned}$$

So $C \subseteq C \cup (C \cap D)$ and $(C \cap D) \cup C \subseteq C$ implies $C \cup (C \cap D) = C$ (Def. of equal classes).

b) The proof is left as an exercise.

Theorem 3.7 Let C be a class (a set). Then $(C')' = C$.

Proof:

$$\begin{aligned} x \in (C')' &\Rightarrow x \notin C' && \text{(By definition 3.4)} \\ &\Rightarrow x \in C && \text{(By definition 3.4)} \\ &\Rightarrow (C')' \subseteq C \end{aligned}$$

$$\begin{aligned} x \in C &\Rightarrow x \notin C' && \text{(By definition 3.4)} \\ &\Rightarrow x \in (C')' && \text{(By definition 3.4)} \\ &\Rightarrow C \subseteq (C')' \end{aligned}$$

$$\left\{ \begin{array}{ccc} (C')' & \subseteq & C \\ C & \subseteq & (C')' \end{array} \right. \Rightarrow (C')' = C \quad \text{(By definition of =)}$$

Theorem 3.8 *(De Morgan's laws)* Let C and D be classes (sets). Then

a) $(C \cup D)' = C' \cap D'$

b) $(C \cap D)' = C' \cup D'$

Proof:

a) Given: C and D are classes (sets).

$$\begin{aligned} x \in (C \cup D)' &\Rightarrow x \notin C \cup D \quad \text{(By definition 3.4)} \\ &\Rightarrow x \notin C \text{ and } x \notin D \quad \text{(For if } x \in C \text{ or } \in D, \text{ then } x \in C \cup D) \\ &\Rightarrow x \in C' \cap D' \\ &\Rightarrow (C \cup D)' \subseteq C' \cap D' \end{aligned}$$

$$\begin{aligned} \text{Next } x \in C' \cap D' &\Rightarrow x \in C' \text{ and } x \in D' \\ &\Rightarrow x \notin C \text{ and } x \notin D \quad \text{(By definition 3.4)} \\ &\Rightarrow x \notin C \cup D \quad \text{(For if } x \in C \cup D, \text{ then } x \in C \text{ or } x \in D) \\ &\Rightarrow x \in (C \cup D)' \quad \text{(By definition 3.4)} \\ &\Rightarrow C' \cap D' \subseteq (C \cup D)' \end{aligned}$$

$$\left\{ \begin{array}{rcl} (C \cup D)' & \subseteq & C' \cap D' \\ C' \cap D' & \subseteq & (C \cup D)' \end{array} \right. \Rightarrow (C \cup D)' = C' \cap D'$$

b) The proof is left as an exercise.

Theorem 3.9 Let C, D and E be classes (sets). Then

a) $C \cup D = D \cup C$ and $C \cap D = D \cap C$ (*Commutative laws*)
b) $C \cup C = C$ and $C \cap C = C$ (*Idempotent laws*)
c) $C \cup (D \cup E) = (C \cup D) \cup E$ and $C \cap (D \cap E) = (C \cap D) \cap E$ (*Associative laws*)
d) $C \cup (D \cap E) = (C \cup D) \cap (C \cup E)$ and $C \cap (D \cup E) = (C \cap D) \cup (C \cap E)$ (*Distribution*)

Proof: The proofs of a) to d) are left as an exercise.

Theorem 3.10 Let A be a class and $\mathscr{U}$ denote the class of all elements.

a) $\mathscr{U} \cup A = \mathscr{U}$
b) $A \cap \mathscr{U} = A$
c) $\mathscr{U}' = \varnothing$
d) $\varnothing' = \mathscr{U}$
e) $A \cup A' = \mathscr{U}$

Proof:

a) By theorem 3.5 a), $\mathscr{U} \subseteq \mathscr{U} \cup A$. If $x \in \mathscr{U} \cup A$, then $x \in A$ or $x \in \mathscr{U}$. In either case x is an element and so $x \in \mathscr{U}$. Thus, $\mathscr{U} \cup A \subseteq \mathscr{U}$. Hence, $\mathscr{U} \cup A = \mathscr{U}$.
Parts b) to e) are left as an exercise.

3.5 Generalized distributive laws and De Morgan's laws.

The distributive law and the De Morgan's laws generalize to arbitrarily large unions and intersections.

Theorem 3.11 Let $\mathscr{A}$ be a non-empty class (set).

a) $\left(\bigcup_{C\in\mathscr{A}} C\right)' = \bigcap_{C\in\mathscr{A}} C'$
b) $\left(\bigcap_{C\in\mathscr{A}} C\right)' = \bigcup_{C\in\mathscr{A}} C'$

Proof:
a)

$$\begin{aligned} x \in \left(\bigcup_{C\in\mathscr{A}} C\right)' &\Leftrightarrow x \notin \bigcup_{C\in\mathscr{A}} C \\ &\Leftrightarrow x \notin C \text{ for all } C \in \mathscr{A} \\ &\Leftrightarrow x \in C' \text{ for all } C \in \mathscr{A} \\ &\Leftrightarrow x \in \bigcap_{C\in\mathscr{A}} C' \end{aligned}$$

Part b) is left as an exercise.

Theorem 3.12 Let D be a class and $\mathscr{A}$ be a non-empty class (set) of classes.

a) $D \cap \left(\bigcup_{C\in\mathscr{A}} C\right) = \bigcup_{C\in\mathscr{A}} (D \cap C)$
b) $D \cup \left(\bigcap_{C\in\mathscr{A}} C\right) = \bigcap_{C\in\mathscr{A}} (D \cup C)$

Proof:
a)

$$\begin{aligned} x \in D \cap \left(\bigcup_{C\in\mathscr{A}} C\right) &\Leftrightarrow x \in D \text{ and } x \in \bigcup_{C\in\mathscr{A}} C \\ &\Leftrightarrow x \in D \text{ and } x \in C \text{ for some } C \in \mathscr{A} \\ &\Leftrightarrow x \in D \cap C \text{ for some } C \in \mathscr{A} \\ &\Leftrightarrow x \in \bigcup_{C\in\mathscr{A}} (D \cap C) \end{aligned}$$

Parts b) is left as an exercise.

Theorem 3.13 Let $\{B_{(i,j)} : i = 1, 2, 3, \ldots,\ j = 1, 2, 3, \ldots\}$ be a set of sets Then

$$\cup_{i=1}^{\infty}(\cap_{j=1}^{\infty} B_{(i,j)}) = \cap_{j=1}^{\infty}(\cup_{i=1}^{\infty} B_{(i,j)})$$

Proof:

$$\begin{aligned} x \in \cup_{i=1}^{\infty}(\cap_{j=1}^{\infty} B_{(i,j)}) &\Leftrightarrow x \in \cap_{j=1}^{\infty} B_{(k,j)} \text{ For some } k. \\ &\Leftrightarrow x \in B_{(k,j)} \text{ For some } k \text{ and all } j. \\ &\Leftrightarrow x \in \cup_{i=1}^{\infty} B_{(i,j)} \text{ For all } j. \\ &\Leftrightarrow x \in \cap_{j=1}^{\infty}(\cup_{i=1}^{\infty} B_{(i,j)}) \end{aligned}$$

Concepts review:

1. If $\mathscr{A}$ is a class of classes how should we interpret the class $\bigcup_{C\in\mathscr{A}} C$?
2. If $\mathscr{A}$ is a class of classes how should we interpret the class $\bigcap_{C\in\mathscr{A}} C$? How do we know that this is indeed a class?
3. What does it mean to say that two classes A and B are *disjoint*?
4. What is the complement, C', of a class C?
5. What is the *difference*, $C - D$, of the two classes C and D? What is the *symmetric difference* $C \triangle D$?
6. If $\mathscr{A}$ is a non-empty set of sets how do we know that the union $\bigcup_{C\in\mathscr{A}} C$ is a set?
7. What do *De Morgan's laws* say in reference to two classes C and D?
8. Let $\mathscr{A}$ be a class of classes. Can we generalize De Morgan's laws to $\{C : C \in \mathscr{A}\}$?
9. Is it true that the union of sets C in a class $\mathscr{A}$ is a set?
10. List the ZF-axioms that refer specifically to sets and were invoked at least once up to now?
11. In algebra, we know about the distributive property of "multiplication over sums and differences". Is there a similar property which refers to "unions distributing over intersections" and "intersections distributing over unions"?

EXERCISES

A. 1. Prove or disprove that if $D \in \mathscr{A}$, then it is always true that $D \subseteq \bigcup_{C \in \mathscr{A}} C$.
2. Show that if $\mathscr{A}$ is a class of sets, then $\mathscr{A} \subseteq \mathscr{P}(\bigcup_{x \in \mathscr{A}} x)$.

B. 3. Show that $\mathscr{P}(A) \cap \mathscr{P}(B) = \mathscr{P}(A \cap B)$.
4. Show that $\mathscr{P}(A) \cup \mathscr{P}(B) = \mathscr{P}(A \cup B)$.
5. Show that $C \cap D = \varnothing$ if and only if $\mathscr{P}(C) \cap \mathscr{P}(D) = \{\varnothing\}$.
6. If A is a set show that $\cup_{C \in \mathscr{P}(A)} C = A$.
7. If A is a set show that $\cap_{C \in \mathscr{P}(A)} C = \varnothing$.
8. If A is a class show that $A = \bigcup_{x \in A} \{x\}$.
9. Prove the following statements.
   a) Part b) of theorem 3.5.
   b) Part b) of theorem 3.6.
   c) Part b) of theorem 3.8.
   d) Part a) to d) of theorem 3.9.
   e) Part b) to e) of theorem 3.10.
   f) Part b) of theorem 3.11.
   g) Part b) of theorem 3.12.

C. 10. If A and B are sets show that $\mathscr{P}(A) \in \mathscr{P}(B)$ implies $\mathscr{P}(A) \subseteq B$ and so $A \in B$.

4 / Cartesian products.

Summary. *In this section we define the notion of "ordered pairs" in terms of classes and sets. We then define the Cartesian product of two classes (sets). We also present a few of the basic properties of Cartesian products.*

4.1 Ordered pairs.

The notion of "ordered pairs" is an important one since it is involved in most areas of mathematics. Most students are familiar with the idea of *ordered pairs* since they have learned early on that when given an ordered pair of numbers, the order in which the numbers appear conveys a particular meaning.

For example, say 120 desks in an exam room are arranged in a rectangular grid of 10 rows with each row containing 12 desks. Suppose each of the 120 students writing an exam in this room is given an ordered pair, (a, b), to be interpreted as follows: You will write your exam on a^{th} desk in the b^{th} row where the first row is the one in the front of the room and the first desk in this row is the one which is closest to the door as you enter the room". The student understands that the desk labeled (2, 3) is not the same desk as the one labeled (3, 2). The order in which the numbers are presented has meaning.

We know that functions can also be represented by ordered pairs. For example, the function $f(x) = x^2$ with domain $\mathbb{R}$ can be represented by the set $S = \{(x, x^2) : x \in \mathbb{R}\}$. So we can say that the ordered pair (5, 25) is an element of this function while the pair (25, 5) is not, since the second entry is not the square of the first entry.

We would now like to formally define "ordered pairs" in our set-theoretic axiomatic system. Our first step will be to remind ourselves of the way "ordered pairs" as we know them are defined. Someone may attempt to define an ordered pair as follows:

> Given two elements a and b, an *ordered pair*, (a, b), is a doubleton $\{a, b\}$ where one element a is labeled as the "first" while the other element b is labeled as the "second". The element labeled "first" must be listed first. The round brackets "( )" are used to indicate that (a, b) is not a simple doubleton but rather a doubleton where the order in which the elements a and b appear has a particular meaning.

This is a bit wordy. Also, it is not clear what the words "first" and "second" mean. We have not defined these in our set-theoretic universe. Can we define ordered pairs without using the words "first" and "second"? That is, can we obtain an equivalent definition of "ordered pairs" by avoiding these two words entirely? Let us consider the following definition and then see if it works.

Definition 4.1 (*Kuratowski definition*) Given a pair of sets c and d, we can construct the class

$$\{\{c\}, \{c, d\}\}$$

The doubleton $\{\{c\}, \{c, d\}\}$ is called an *ordered pair*. The sets c and d need not be distinct. Ordered pairs are denoted as $(c, d) = \{\{c\}, \{c, d\}\}$.

First, we should verify that there are no inherent ambiguities in this definition. We verify immediately that if c and d are sets, then $\{\{c\}, \{c, d\}\}$ is a set:

$$\begin{aligned} c \text{ and } d \text{ are sets} &\Rightarrow \{c\} \text{ and } \{d\} \text{ are sets. (Axiom of pair)} \\ &\Rightarrow \{c\} \cup \{d\} = \{c, d\} \text{ is a set. (Axiom of union)} \\ &\Rightarrow \{\{c\}, \{c, d\}\} \text{ is a set. (Axiom of pair)} \end{aligned}$$

Now that this has been established, we should make sure that the doubleton defined above satisfies the essential "*ordered pairs* property", $[(a, b) = (c, d)] \Leftrightarrow [a = c \text{ and } b = d]$.

Theorem 4.2 Let a, b, c and d be sets. Then $(a, b) = (c, d)$ if and only if $a = c$ and $b = d$.

Proof:

($\Leftarrow$) That $a = c$ and $b = d$ implies $(a, b) = \{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\} = (c, d)$ is immediate.

($\Rightarrow$) What we are given: $(a, b) = (c, d)$. What we are required to show: $a = c$ and $b = d$.

$$(a, b) = (c, d) \quad \Rightarrow \quad \{\{a\}, \{a, b\}\} = \{\{c\}, \{c, d\}\}$$

$$\begin{aligned} \text{Case 1: } a \neq b &\Rightarrow \{a, b\} \neq \{c\} \text{ hence } \{a, b\} = \{c, d\} \\ &\Rightarrow \{a\} = \{c\} \\ &\Rightarrow a = c \\ \{a, b\} = \{c, d\} \text{ and } a = c &\Rightarrow b = d \end{aligned}$$

$$\begin{aligned} \text{Case 2: } a = b &\Rightarrow \{\{a\},\{a,b\}\} = \{\{a\},\{a,a\}\} = \{\{a\}\} \\ &\Rightarrow \{a\} = \{c\} \text{ and } \{a\} = \{c,d\} \\ &\Rightarrow a = c \\ \{a\} = \{c,d\} \text{ and } a = c &\Rightarrow \{a\} = \{a,d\} \\ &\Rightarrow \{a,d\} = \{a,a\} \\ &\Rightarrow a = d \end{aligned}$$

Then $(a,b) = (c,d)$ if and only if $a = c$ and $b = d$.

From this theorem we deduce that:

$$[a \neq b] \Rightarrow [(a,b) \neq (b,a)]$$

Given the distinct sets a and b, we see that the sets $\{a,b\}$ and (a,b) are indeed different mathematical "creatures". The definition of ordered pair simply shows that (a,b) is constructed from a and b in a way that guarantees that the "ordered pair" property holds true. It says that (a,b) is a doubleton where one of its elements is a singleton while the other is itself a doubleton. Since the two elements of this set have different characteristics it allows us to decide which is the first entry and which is the second. We can say that "the singleton is the first entry", while "the doubleton is the second entry".

Having defined an ordered pair of sets, we can conveniently define, in a similar way, an *ordered triple* (x,y,z) as an ordered pair where the "first" entry is itself an ordered pair:

$$(x,y,z) = ((x,y),z)$$

Alternate definitions of ordered pairs. There are other possible definitions of ordered pairs (c,d) in terms of sets. Many readers may find the following definition intuitively preferable. This definition is a slight variation of the one put forward by Felix Hausdorff (1868 - 1942), so we will label it as being Hausdorff's.

Definition 4.3 (*Hausdorff*) If c and d are sets, the expression (c,d) is defined as follows:

$$(c,d) = \{\ \{c,\varnothing\},\{d,\{\varnothing\}\}\ \}$$

The main reason why this definition may be intuitively appealing to some is that it looks more like we are indexing the two elements c and d with the symbols ϕ and $\{\phi\}$. It allows one to visualize the ordered pair as follows:

$$(c,d) = \{\ \{c,\varnothing\},\{d,\{\varnothing\}\}\ \} = \left\{c_{\varnothing}, d_{\{\varnothing\}}\right\} = \{c_0, d_1\}$$

This in fact resembles more the way we will be viewing ordered pairs once we define "functions" and the "natural numbers". We will defer the proof which guarantees that this definition satisfies the essential property of ordered pairs to the end of this section. In this text, we will adopt the more commonly used Kuratowski definition.

4.2 Cartesian products

Now that we have defined ordered pairs of classes we can construct new classes with old ones. Recall how, from two known sets, say $\mathbb{N}$ and $\mathbb{R}$ we can construct a new set $\mathbb{N}\times\mathbb{R} = \{(n,x) : n\in\mathbb{N}, x\in\mathbb{R}\}$. This is what we want to do with classes. Any two classes (sets) C and D can be used to construct another class (called a *Cartesian product*) whose elements are *ordered pairs*.

Definition 4.4 Let C and D be two classes (sets). We define the *Cartesian product*, $C\times D$, as follows:

$$C\times D = \{(c,d) : c\in C \text{ and } d\in D\}$$

We could of course also write $C\times D = \{\ \{\{c\},\{c,d\}\} : c\in C \text{ and } d\in D\}$. Since we are particularly interested in constructing new sets from old ones, we should first make sure that as long as C and D are sets, then $C\times D$ is a set. We will do this by first proving the following lemma.

Lemma 4.5 Let C and D be two classes (sets). Then the *Cartesian product*, $C\times D$, of C and D satisfies the property $C\times D\subseteq \mathscr{P}(\mathscr{P}(C\cup D))$.

Proof:

What is given: That C and D are two classes (sets).
What we are required to show: $C\times D\subseteq \mathscr{P}(\mathscr{P}(C\cup D))$.
Let $c\in C$ and $d\in D$. It will suffice to show that $(c,d)\in \mathscr{P}(\mathscr{P}(C\cup D))$.

$$\begin{aligned} \{c\} \in \mathscr{P}(\{c,d\}) \text{ and } \{c,d\} \in \mathscr{P}(\{c,d\}) &\Rightarrow \{\{c\},\{c,d\}\} \subseteq \mathscr{P}(\{c,d\}) \\ &\Rightarrow (c,d) \subseteq \mathscr{P}(\{c,d\}) \\ &\Rightarrow (c,d) \in \mathscr{P}(\mathscr{P}(\{c,d\})) \\ \mathscr{P}(\mathscr{P}(\{c,d\})) \subseteq \mathscr{P}(\mathscr{P}(C \cup D))^{\dagger} &\Rightarrow (c,d) \in \mathscr{P}(\mathscr{P}(C \cup D)) \end{aligned}$$

Hence, $C \times D \subseteq \mathscr{P}(\mathscr{P}(C \cup D))$, as required.

† To see this verify that $S \subset T \Rightarrow \mathscr{P}(S) \subset \mathscr{P}(T)$

Theorem 4.6 If C and D are classes, then the *Cartesian product*, $C \times D$, is a class. If C and D are sets, then $C \times D$ is a set.

Proof:

To show that $C \times D$ is a class we can express $C \times D$ as

$$C \times D = \{x : x \in \mathscr{P}(\mathscr{P}(C \cup D)) \text{ and } x = (c,d) \text{ for some } c \in C \text{ and some } d \in D\}$$

and invoke axiom of construction A2 which declares it to be a class.

Given that C and D are sets, then $C \cup D$ is a set (by the Axiom of union A6) which implies that $\mathscr{P}(\mathscr{P}(C \cup D))$ is a set (by the Axiom of power set A5). Since $C \times D \subseteq \mathscr{P}(\mathscr{P}(C \cup D))$, then $C \times D$ is a set (by the axiom of subset A4).

We can then write that if C and D are sets, $C \times D$ is the set of all those specific elements u in $\mathscr{P}(\mathscr{P}(C \cup D))$ which are of the form $u = (c,d)$ for some c in C and d in D.

Once we have defined the Cartesian product of two classes C and D, referring to our definition of ordered triples $(c,d,e) = ((c,d),e)$, we can define the Cartesian product of three classes C, D and E as follows:

$$\begin{aligned} C \times D \times E &= \{(c,d,e) : c \in C, d \in D, e \in E\} \\ &= \{((c,d),e) : c \in C, d \in D, e \in E\} \\ &= (C \times D) \times E \end{aligned}$$

4.3 A few properties of Cartesian products

The following theorem illustrates properties of Cartesian products involving the symbols intersections, "$\cap$", and union, "$\cup$".

Theorem 4.7 Let C, D, E and F be classes. Then

a) $C \times (D \cap E) = (C \times D) \cap (C \times E)$
b) $C \times (D \cup E) = (C \times D) \cup (C \times E)$
c) $(C \cap E) \times D = (C \times D) \cap (E \times D)$
d) $(C \cup E) \times D = (C \times D) \cup (E \times D)$
e) $(C \cup D) \times (E \cup F) = (C \times E) \cup (D \times E) \cup (C \times F) \cup (D \times F)$
f) $(C \cap D) \times (E \cap F) = (C \times E) \cap (D \times E) \cap (C \times F) \cap (D \times F)$

Proof:

a)
$$\begin{aligned} (c,d) \in C \times (D \cap E) &\Leftrightarrow c \in C \text{ and } d \in (D \cap E) \\ &\Leftrightarrow c \in C \text{ and } d \in D \text{ and } d \in E \\ &\Leftrightarrow (c,d) \in C \times D \text{ and } (c,d) \in C \times E \\ &\Leftrightarrow (c,d) \in (C \times D) \cap (C \times E) \end{aligned}$$

Hence, $C \times (D \cap E) = (C \times D) \cap (C \times E)$ (by axiom A1).

Proofs of parts b) to f) are left as an exercise.

Theorem 4.8 If $C \subseteq D$ and $E \subseteq F$, then $C \times E \subseteq D \times F$

Proof:

By definition $C \times E = \{(c,e) : c \in C \text{ and } e \in E\}$ and

$D \times F = \{(d,f) : d \in D \text{ and } f \in F\}$.

$$\begin{aligned} (c,e) \in C \times E &\Rightarrow c \in C \text{ and } e \in E \\ &\Rightarrow c \in D \text{ and } e \in F \quad \text{(Since } C \subseteq D \text{ and } E \subseteq F\text{)} \\ &\Rightarrow (c,e) \in D \times F \end{aligned}$$

Hence, $C \times E \subseteq D \times F$.

The following theorem shows that there is a one-to-one correspondence between the elements of $S \times (U \times V)$ and the elements of $(S \times U) \times V$.

Theorem 4.9 Given three classes (sets) S, U and V there is a one-to-one correspondence between the two classes (sets) $S \times (U \times V)$ and $(S \times U) \times V$.

Proof:

Let $\phi : S \times (U \times V) \to (S \times U) \times V$ be defined as: $\phi((s,(u,v))) = ((s,u),v)$. We will prove that ϕ maps distinct elements in $S \times (U \times V)$ to distinct elements in $(S \times U) \times V$. We can prove this by invoking theorem 4.2 as follows:

$$\begin{aligned}
(s,(u,v)) = (a,(b,c)) &\Leftrightarrow s = a \text{ and } (u,v) = (b,c) \\
&\Leftrightarrow s = a \text{ and } u = b \text{ and } v = c \\
&\Leftrightarrow (s,u) = (a,b) \text{ and } v = c \\
&\Leftrightarrow ((s,u),v) = ((a,b),c) \\
&\Leftrightarrow \phi(s,(u,v)) = \phi((a,(b,c))
\end{aligned}$$

4.4 Proof for the Hausdorff definition of *ordered pairs*

We end this section with a proof showing that the alternate of

$$(c,d) = \{\ \{c,\varnothing\},\{d,\{\varnothing\}\}\ \}$$

satisfies the essential property of "ordered pairs" and so can also be used to represent ordered pairs.

Theorem 4.10 For classes c, d, e and f, if $(c,d) = \{\{c,\varnothing\},\{d,\{\varnothing\}\}\}$ and $(e,f) = \{\{e,\varnothing\},\{f,\{\varnothing\}\}\}$, then $(c,d) = (e,f)$ if and only if $c = e$ and $d = f$.

Proof:

($\Leftarrow$) That $c = e$ and $d = f$ implies $(c,d) = (e,f)$ is immediate.

($\Rightarrow$) What we are given:
- $(c,d) = \{\{c,\varnothing\},\{d,\{\varnothing\}\}\}$
- $(e,f) = \{\{e,\varnothing\},\{f,\{\varnothing\}\}\}$
- $(c,d) = (e,f)$

What we are required to show: $c = e$ and $d = f$.
We first consider the case where c is the empty class.

$$\begin{aligned}
(\varnothing,d) = (e,f) &\Rightarrow \{\{\varnothing,\varnothing\},\{d,\{\varnothing\}\}\} = \{\{e,\varnothing\},\{f,\{\varnothing\}\}\} \\
&\Rightarrow \{\{\varnothing\},\{d,\{\varnothing\}\}\} = \{\{e,\varnothing\},\{f,\{\varnothing\}\}\} \\
&\Rightarrow \{\{\varnothing\},\{d,\{\varnothing\}\}\} = \{\{\varnothing,\varnothing\},\{f,\{\varnothing\}\}\} \quad \text{(Since } \{f,\{\varnothing\}\} \text{ can never equal } \{\varnothing\}.\text{)} \\
&\Rightarrow \{\{\varnothing\},\{d,\{\varnothing\}\}\} = \{\{\varnothing\},\{f,\{\varnothing\}\}\} \\
&\Rightarrow \{d,\{\varnothing\}\} = \{f,\{\varnothing\}\} \\
&\Rightarrow d = f
\end{aligned}$$

Thus, $(\varnothing, d) = (e, f) \Rightarrow e = \varnothing$ and $d = f$ as required. We now consider the case where $c \neq \varnothing$.

$$(c, d) = (e, f) \quad \Rightarrow \quad \{\{c, \varnothing\}, \{d, \{\varnothing\}\}\} = \{\{e, \varnothing\}, \{f, \{\varnothing\}\}\}$$

If $\{c, \varnothing\} = \{e, \varnothing\} \quad \Rightarrow \quad c = e$ and it quickly follows that $d = f$. (Check the details.)

$$\begin{array}{rcl} \text{If } \{c, \varnothing\} \neq \{e, \varnothing\} & \Rightarrow & \{c, \varnothing\} = \{f, \{\varnothing\}\} \text{ and } \{d, \{\varnothing\}\} = \{e, \varnothing\} \\ \{c, \varnothing\} = \{f, \{\varnothing\}\} & \Rightarrow & f = \varnothing \quad (\text{Since } \varnothing \neq \{\varnothing\} \text{ this forces } f = \varnothing.) \\ & \Rightarrow & \{\varnothing\} = c \\ \{d, \{\varnothing\}\} = \{e, \varnothing\} & \Rightarrow & d = \varnothing \text{ and } e = \{\varnothing\} \quad (\text{For the same reasons as above.}) \\ c = \{\varnothing\} = e \text{ and } d = \varnothing = f & \Rightarrow & c = e \text{ and } d = f. \end{array}$$

Note that the two different representations of ordered pairs (a, b), $\{\{a\}, \{a, b\}\}$ and $\{\{a, \varnothing\}, \{b, \{\varnothing\}\}\}$ do not form equal sets. These two classes only share the fundamental property of ordered pairs.

Concepts review:

1. What is the (Kuratowski) definition of the ordered pair (c, d)?
2. Given two classes C and D what is the definition of $C \times D$?
3. If C and D are sets is it true that $C \times D \subseteq \mathscr{P}(\mathscr{P}(C \cup D))$? Why?
4. Is it generally true that $C \times D = D \times C$? If so, why? If not give a counterexample.
5. Is it generally true that $(C \times D) \cup (E \times F) = (C \cup E) \times (D \cup F)$? If so, why? If not give a counterexample.

EXERCISES

A.
1. Prove that $C \times D = \varnothing$ if and only if $C = \varnothing$ or $D = \varnothing$.
2. Show that for classes C, D and E, $(C \times D) \cap (C' \times E) = \varnothing$.
3. Show that $A \subseteq B \Rightarrow A \times C \subseteq B \times C$.

B.

4. If $(c, d) \in C \times D$ is it necessarily true that $\{c\} \in C$ and $\{c, d\} \in D$? If so, why? If not, give a counterexample.
5. Suppose A, B and C are sets. Show that $A \times C \subseteq B \times C \Rightarrow A \subseteq B$, that is, the converse of the statement in question 3) holds true.
6. Prove parts b) and c) of theorem 4.7 on page 39.
7. Let $S = \{x\}$ be a set. Show that $(S \times S) \times S \neq S \times (S \times S)$.
8. Describe each of the following classes. But first explain why each of these classes is a set.
   a) $\varnothing \times \{\varnothing\}$
   b) $\{\varnothing\} \times \varnothing$
   c) $\varnothing \times \varnothing$
   d) $\{\varnothing\} \times \{\varnothing\}$
   e) $\{\varnothing \times \varnothing\}$

C.

9. Show that $C \times (D - E) = (C \times D) - (C \times E)$.
10. Is the statement "$C \times D = E \times F$ if and only if $C = E$ and $D = F$" always true? If there are situations where it fails to be true, state which ones.
11. Show that if a and b are sets, then $\{\{a, \varnothing\}, \{b, \{\varnothing\}\}\}$ is a set.

Part III

Relations

5 / Relations on a class or set.

Summary. *In this section we define a relation R on a class (a set) S. For a relation R on a set S we define the inverse R^{-1} of the relation R. We also define the domain and the image of R. The composition of two relations R and T is defined and some of their properties are given.*

5.1 Relations on a class or set.

Recall that the symbol $\mathscr{U}$ denotes the "Universal class", $\{x : x = x\}$. Since $\mathscr{U}$ is a class, we can then construct the Cartesian product, $\mathscr{U} \times \mathscr{U}$, itself a class (as we have seen). Recall that the elements of $\mathscr{U} \times \mathscr{U}$ are ordered pairs.

Definition 5.1 a) We will call any subset R of ordered pairs in $\mathscr{U} \times \mathscr{U}$ a *binary relation.*[1]

b) We will say that R is a *binary relation on a class* C if R is a subclass (subset) of $C \times C$. In such cases we will simply say that R is a *relation in* C (or *on* C).

c) If A and B are classes (sets) and R is a subclass (subset) of $A \times B$ then R can be viewed as a relation on $A \cup B$.

From this definition we see that any Cartesian product $C \times D$ is a relation. But a relation need not be a Cartesian product. For example, the smallest Cartesian product which contains the set $G = \{(a, c), (b, d)\}$ is $A \times B$ where $A = \{a, b\}$ and $B = \{c, d\}$. But $G \neq A \times B$ since $(a, d) \in (A \times B) - G$.

We will be referring to specific kinds of relations on a class or set. Given a class C, a relation on C is usually expressed by the symbol, R, although other capital letters are used whenever it is necessary to distinguish between two relations on a same class. Suppose R is a relation on a class C and that $(x, y) \in R$. Then, by definition, both x and y belong to C. Common ways of expressing that (x, y) belong to the relation R are:

- $(x, y) \in R$
- xRy holds true
- x is related to y under R.

[1†]The word *binary* refers to fact that the elements of R are doubletons (pairs). We can also speak of a *ternary relation* when considering subclasses of the Cartesian product $\mathscr{U} \times \mathscr{U} \times \mathscr{U}$. Unless we specify otherwise, all relations in this text are assumed to be "binary" and so the word *relation* will be used to abbreviate the words *binary relation.*

Remark: Recall from lemma 4.5 that $C \times D \subseteq \mathscr{P}(\mathscr{P}(C \cup D)$; hence, if R is a relation subset of $C \times D$ then $R \subseteq \mathscr{P}(\mathscr{P}(C \cup D)$.

5.2 Examples of relations on a class.

We provide a few examples of relations in $\mathscr{U}$.

1) We will define a relation, R_1, in $\mathscr{U}$ as follows: $(x, y) \in R_1$ if and only if $x \in y$.
   This says that "any x is *related* precisely to those classes (sets) y which contain it.". We can also write
   $$R_1 = \{(x, y) : x \in y\}$$
   For example, we can write $(a, \{a, b\}) \in R_1$ or, if one prefers, $aR_1\{a, b\}$ "holds true". On the other hand, we can write $(b, \{c, d\}) \notin R_1$. Also, $(\varnothing, \varnothing) \notin R_1$.

2) We define a relation, R_2, in $\mathscr{U}$ as follows: $(x, y) \in R_2$ if and only if $x \subset y$ where "$x \subset y$" means "$x \subseteq y$ and $x \neq y$".
   This says that a class (a set) x is related precisely to those non-empty classes (sets) $y \neq x$ which contain the elements of x. Since $x = x$, then $x \not\subset x$; hence, $(x, x) \notin R_2$. Assuming that $a \neq \{b\}$, we see that $(a, \{a, b\}) \notin R_2$, but that $(\{a\}, \{a, b\}) \in R_2$. The statement $xR_2\varnothing$ is false for all classes x including the class $\varnothing$.

3) We define a relation R_3 in $\mathscr{U}$ as follows: $(x, y) \in R_3$ if and only if $x = y$.
   This says that a class (a set) x is related only to itself and no other class. We can write $R_3 = \{(x, y) : x = y\}$. We see that $(a, \{a\}) \notin R_3$ but that $(\{a\}, \{a\}) \in R_3$. The statement $\varnothing R_3 \varnothing$ is true.

4) We define a relation R_4 in $\mathscr{U}$ as follows: $R_4 = \{(x, y) : x \subseteq y\}$.
   This means two elements x and y are related only if x is contained in y. Notice that if $x \neq y$ and $(x, y) \in R_4$, then it is impossible for (y, x) to belong to R_4. The ordered pair, $(\{\varnothing\}, \{\varnothing, \{\varnothing\}\})$, belongs to R_4 but $(\{\varnothing, \{\varnothing\}\}, \{\varnothing\})$ does not. We also see that the pair $(x, \{x\}) \notin R_4$.

Definition 5.2 Let C be a class (a set).

a) The relation
   $$\in_C = \{(x, y) : (x, y) \in C \times C, x \in y\}$$
   is called the *membership relation on C*.

b) The relation
   $$Id_C = \{(x, y) : (x, y) \in C \times C,\ x = y\}$$
   is called the *identity relation on C*.

We see that the only elements in the *identity relation* $Id_{\mathscr{U}}$ on $\mathscr{U}$ are those of the form (x, x).

5.3 The *domain* and the *image* of a relation.

The reader may see some similarities between the concept of a *relation on* C and what is known to be a "function from a set C to C":

- Both relations and functions are collections of ordered pairs.
- In both cases, the first entry should not be confused with the second entry. The second entry is often defined in terms of the first entry: That is, a rule states why the second entry is related to the first. This rule is the mechanism which allows us to determine which ordered pairs belong to the relation or the function and which don't.

Definition 5.3 Let R be a relation on a class (set) C. The *domain of* R is the class, $\operatorname{dom} R = \{x : x \in C \text{ and } (x, y) \in R \text{ for some } y \in C\}$. The *image of* R is the class, $\operatorname{im} R = \{y : y \in C \text{ and } (x, y) \in R \text{ for some } x \in C\}$. The word *range of* R is often used instead of "the image of R". If $R \subseteq A \times B$ is viewed as a relation on $A \cup B$, then $\operatorname{dom} R \subseteq A$ and $\operatorname{im} R \subseteq B$.

Example: Suppose R is the membership relation on the set S where S is defined as,

$$S = \{\ a, b, c, \{a\}, \{a, b\}, \{\{c\}\}, \varnothing, \{\varnothing\}\}$$

That is, $R = \{(x, y) : x \in y\}$. Find $\operatorname{dom} R$ and $\operatorname{im} R$.

Solution:

To find the domain and the image of R we will write out the elements of R explicitly:

$$R = \{\ (a, \{a\}),\ (a, \{a, b\}),\ (b, \{a, b\}),\ (\varnothing, \{\varnothing\})\ \}$$

The domain, $\operatorname{dom} R$, is

$$\operatorname{dom} R = \{a, b, \varnothing\}$$

while the image, $\operatorname{im} R$, is

$$\operatorname{im} R = \{\{a\}, \{a, b\}, \{\varnothing\}\}$$

5.4 The inverse of a relation R on a set S.

Just as for one-to-one functions, we can speak of the inverse of a relation R on a set S. However, a relation need not be "one-to-one" to have an inverse. "One-to-many" relations are quite common. We will begin by formally defining what we mean by the inverse of a relation.

Definition 5.4 Let C be a class (a set) and let R be a relation defined in C. The inverse, R^{-1}, of the relation R is defined as follows:

$$R^{-1} = \{(x, y) : (y, x) \in R\}$$

In the example above we defined a relation on the set

$$S = \{\ a, b, c, \{a\}, \{a, b\}, \{\{c\}\}, \varnothing, \{\varnothing\}\}$$

as $R = \{(x, y) : x \in y\}$.

Since R was found to be:

$$R = \{\ (a, \{a\}),\ (a, \{a, b\}),\ (b, \{a, b\}),\ (\varnothing, \{\varnothing\})\ \}$$

then the inverse relation, R^{-1}, is

$$R^{-1} = \{\ (\{a\}, a),\ (\{a, b\}, a),\ (\{a, b\}, b),\ (\{\varnothing\}, \varnothing)\ \}$$

The inverse of R can also be expressed in the more succinct form
$R^{-1} = \{(x, y) : y \in x\}$.

5.5 The composition of two relations R and T.

Just like pairs of functions f and g, a pair of relations R and T on a class C can be combined to obtain a new relation. Other than the fact that the first entries of a relation can be associated to many values in the image of R, compositions of relations work exactly like the composition of functions. We define the composition of two relations as follows.

Definition 5.5 Let C be a class (a set) and let R and T be two relations in C. We define the relation $T \circ R$ as follows:

$$T \circ R = \{(x, y) : (z, y) \in T \text{ where } (x, z) \in R\}$$

Suppose, for example that the relations R and T on the set

$$S = \{\ a, b, c, \{a\}, \{a, b\}, \{\{c\}\}, \varnothing, \{\varnothing\}\}$$

are defined as:

$$\begin{aligned} R &= \{(x, y) : x \in y\} \\ T &= \{(x, y) : x = \{y\}\} \end{aligned}$$

Then the relation R can be described as:

$$R = \{\ (a, \{a\}),\ (a, \{a, b\}),\ (b, \{a, b\}),\ (\varnothing, \{\varnothing\})\ \}$$

The relation T is:

$$T = \{\ (\{a\}, a\}), (\{\varnothing\}, \varnothing\})\}$$

For the relation $T \circ R$ we obtain:

$$T \circ R = \{(a, a), (\varnothing, \varnothing)\}$$

For the relation $R \circ T$ we obtain:

$$R \circ T = \{(\{a\}, \{a\}), (\{\varnothing\}, \{\varnothing\})\}$$

Concepts review:

1. Given a class C, what is a *relation* on C?
2. Given a relation R on a class C, what does the expression xRy mean?
3. Given a class C what is the *membership relation, $\in_C$, on C*?
4. Given a class C what is the *identity relation, Id_C, on C*?
5. Given a relation R on a class C what is the domain, dom R, of R and the image, im R, of R?
6. Given a relation R on a class C what is the inverse, R^{-1}, of the relation R? Is R^{-1} a relation on C?
7. Does a relation R on C have to be "one-to-one" for R^{-1} to be a relation?
8. If R and T are two relations on a class C, what does $R \circ T$ mean?

EXERCISES

A.
1. Suppose R, S and T are three relations on a set A. Prove that $T \circ (R \circ S) = (T \circ R) \circ S$.
2. Suppose R is a relation on a set A. Prove that $(R^{-1})^{-1} = R$.
3. Let $R = \{(a, a), (a, c), (c, c), (c, d)\}$ and $T = \{(a, b), (c, a), (d, c)\}$. Describe:
   a) R^{-1} and T^{-1}.
   b) $R \circ T$ and $T \circ R$.
   c) $R \circ T^{-1}$

4. If R is a relation on a set S show that $\operatorname{dom} R^{-1} = \operatorname{im} R$.

B. 5. Let $C = \{\varnothing, \{\varnothing\}\}$ and $D = \mathscr{P}(C)$.
   a) Write out explicitly the elements of the membership relation, $\in_D$, on D.
   b) Write out explicitly the elements of the identity relation, Id_D, on D.
   c) Describe the $\operatorname{dom} \in_D$ and $\operatorname{im} \in_D$.
   d) List all elements of $D = \mathscr{P}(C)$.
   e) List all possible relations on C.

C. 6. Let R, S and T be three relations on a set A.
   a) Prove that $(R \cup S) \circ T = (R \circ T) \cup (S \circ T)$.
   b) Prove that $(R \cap S)^{-1} = R^{-1} \cap S^{-1}$.
   c) Prove that $R \subseteq T$ implies $R^{-1} \subseteq T^{-1}$.
   d) Prove that $\operatorname{dom}(S \cup T) = \operatorname{dom} S \cup \operatorname{dom} T$.

6 / Equivalence relations and order relations.

Summary. *In this section we define special types of relations on a class or set: reflexive, symmetric, antisymmetric, asymmetric and transitive relations. Equivalence relations on a class S will be defined as those relations which are simultaneously reflexive, symmetric and transitive. We also define "partial order relations" and "strict order relations" on sets and provide examples of these.*

6.1 A few special types of relations on a class S.

Although we would not normally think of it as a relation, the empty set is a relation on any non-empty set S:

$$\varnothing = \varnothing \times S = \{(x, y) : x \in \varnothing, y \in S\} \subseteq S \times S$$

Its main properties are:

i) $\operatorname{dom} \varnothing = \{x \in S : (x, y) \in \varnothing\} = \varnothing$,
ii) $\varnothing^{-1} = \{(x, y) : (y, x) \in \varnothing\} = \varnothing$,
iii) $\operatorname{im} \varnothing = \{y \in S : (x, y) \in \varnothing\} = \varnothing$.

We define other special types of relations below.

Definition 6.1 Let S be a class and R be a relation on S.

a) We say that R is a *reflexive* relation on S if, for every $x \in S$, $(x, x) \in R$.

b) We say that R is a *symmetric* relation on S if, whenever $(x, y) \in R$, then $(y, x) \in R$.

c) We say that R is an *antisymmetric* relation on S if, whenever $(x, y) \in R$ and $(y, x) \in R$, then $x = y$.

d) We say that R is an *asymmetric* relation on S if, whenever $(x, y) \in R$, then $(y, x) \notin R$.

e) We say that R is a *transitive* relation on S if, whenever $(x, y) \in R$ and $(y, z) \in R$, then $(x, z) \in R$.

f) We say that R is an *irreflexive* relation on S if, for every $x \in S$, $(x, x) \notin R$.

g) If, for every $x, y \in S$ where $x \neq y$, either $(x, y) \in R$ or $(y, x) \in R$, then we say "any two elements a and b in S are *comparable* under the relation R".

It follows from these definitions that a relation which is both antisymmetric and irreflexive must be asymmetric. We illustrate these relation properties with the following four examples.

1) Let G denote the set of all individuals in Gotham City. Consider two relations on G:

$$\begin{aligned} R &= \{(f,b) : f \text{ is a female and } b \text{ is a brother of } f\} \\ T &= \{(x,y) : \ x \text{ and } y \text{ are distinct siblings.}\} \end{aligned}$$

- R is irreflexive since a female inhabitant of this city cannot be her own brother. The relation T is also irreflexive since a person cannot be two children of the same biological parents.
- R is an asymmetric relation on G since if $(f,b) \in R$, then $(b,f) \notin R$ since b is male. However, x and y are distinct siblings in whichever order we consider them. So T is symmetric.
- Transitivity: Since both (a,b) and (b,c) cannot belong to R (since b is male), then we will say that R is "vacuously" transitive. The relation T is transitive since if x and y are siblings and y and z are siblings, then x and z are siblings.
- If we assume that Gotham City contains more than one family, then there are pairs of individuals which are not comparable under both R an T.

2) Let $S = \{a,b,c,d\}$. Consider the relation $R_1 = \{(a,a),(b,b),(c,c),(d,d),(a,b)\}$ on S.

- R_1 is a reflexive relation on S since R_1 contains (x,x) for each $x \in S$.
- R_1 is *not* a symmetric relation on S since R_1 contains (a,b) but not (b,a).
- R_1 is "vacuously" antisymmetric on S since R_1 contains (a,b) and that (b,a) is not in R_1.[1]
- R_1 is not asymmetric since R_1 contains (a,a).
- R_1 is transitive on S since R_1 contains (a,b) and (b,b) and also contains (a,b).

3) Let $S = \{a,b,c,d\}$. Consider the relation $R_2 = \{(a,a),(b,b),(d,d),(a,b)\}$.

- R_2 is not a reflexive relation on S since R_2 does not contain (c,c). It is not irreflexive since it contains (a,a).
- R_2 is not a symmetric relation on S since R_2 contains (a,b) but not (b,a).
- R_2 is vacuously antisymmetric since R_2 contains (a,b) and (b,a) is not in R^2.
- R_2 is not asymmetric since R_2 contains (a,a).
- R_2 is transitive since R_2 contains (a,b) and (b,b) and also contains (a,b).

4) Let $S = \{a,b,c,d\}$. Consider the relation

$$R_3 = \{(a,a),(b,b),(c,c),(d,d),(a,b),(b,a),(b,c),(c,b)\}$$

- We see that R_3 is a reflexive relation on S.

[1]Note that the statement "whenever (a,b) and (b,a) are in S, then $a = b$" holds true.

- Since R_3 contains both pairs $\{(a,b),(b,a)\}$ and $\{(b,c),(c,b)\}$, then R_3 is a symmetric relation on S.
- Since R_3 contains $\{(a,b),(b,a),(a,a)\}$ and $\{(b,c),(c,b),(b,b)\}$, then R_3 is antisymmetric.
- Since R_3 contains (a,a), then R_3 is not asymmetric.
- Since R_3 contains the triples $\{(a,b),(b,a),(a,a)\}$ and $\{(b,c),(c,b),(b,b)\}$, then R_3 is transitive on S.

A word of caution: Some readers may conclude that a relation R which is both symmetric and transitive on a class S is automatically reflexive based on the following reasoning:

> Symmetric R says "$(a,b) \in R$ implies $(b,a) \in R$" while transitive R says "(a,b) and (b,a) in R implies $(a,a) \in R$ ". So "symmetric + transitive $\Rightarrow$ reflexive".

This conclusion is however not correct.
Consider the relation $R = \{(a,a),(b,b),(d,d),(a,b),(b,a)\}$ on the set $S = \{a,b,c,d\}$. It is symmetric and transitive and yet (c,c) is not in R and so R is not reflexive. We should remember that if a relation R is to be reflexive on S we must have $(x,x) \in R$ *for all* $x \in S$

6.2 Equivalence relations on a class S.

Equivalence relations are important types of relations on classes and sets. Students who study other fields of mathematics will frequently encounter sets equipped with this type of relation.

Definition 6.2 Let S be a class and R be a relation on S. We say that R is an *equivalence relation* on S if R is simultaneously reflexive, symmetric and transitive on S.

Examples: Let S and T be two non-empty classes.

a) Recall that Id_S denotes the *identity relation* on S:

$$(x,y) \in \mathrm{Id}_S \text{ if and only if } x = y$$

In this relation an element is only related to itself and no other. It is easily seen that Id_S is reflexive, symmetric and transitive on S. Thus, the identity relation, Id_S, is an equivalence relation on S.

b) Let $D = S \times T$. We define a relation, R, on D as follows:

$$((a,b),(c,d)) \in R \text{ if and only if } a = c$$

This means all ordered pairs in $S \times T$ with the same first entry are related under R. Since

i. $((a,b),(a,b)) \in R)$ for all $(a,b) \in D$ so R is reflexive on D.

ii. $((a,b),(c,d)) \in R \Rightarrow a = c \Rightarrow ((c,d),(a,b)) \in R$ so R is symmetric.

iii. $((a,b),(c,d)) \in R$ and $((c,d),(e,f)) \in R \Rightarrow a = c = e \Rightarrow ((a,b),(e,f)) \in R$ so R is transitive.

We conclude that R is an equivalence relation on D.

c) We will refer to the first example presented on page 52. If G denotes all the inhabitants of Gotham City and H is the relation on G defined as

$$H = \{(x,y) : \ x \text{ and } y \text{ are siblings or the same person}\}$$

then H is reflexive, symmetric and transitive and so forms an equivalence relation on G.

6.3 Order relations on a class S.

We now discuss another very important type of relation called "order relation". An order relation will be either *strict* or *non-strict*. In each of those two categories an order relation can be a *partial ordering* or a *linear ordering*. These terms are defined below.

Definition 6.3 Let S be a class.

a) *Non-strict order relation.* The relation R is a non-strict order relation on S if it is simultaneously reflexive (aRa holds true for any a in S), antisymmetric (if aRb and bRa then $a = b$) and transitive (aRb and bRc implies aRc) on S. A non-strict order relation, R, on S is said to be a *non-strict linear order relation* if, for every pair of elements a and b in S, either $(a,b) \in R$, $(b,a) \in R$ or $a = b$. That is, every pair of elements are comparable under R.[1] A non-strict ordering, R, on S which is not linear is said to be a *non-strict partial ordering relation* on S.[2]

[1] A class on which is defined a linear ordering R is also said to be *fully ordered* or *totally ordered* by R. In certain branches of mathematics, "linearly ordered set" is abbreviated as *l.o.set* or simply called *loset*.

[2] Again in certain branches of mathematics, "partially ordered set" is abbreviated as *p.o.set* or simply called a *poset*

b) *Strict order relation.* The relation R is a strict order relation if it is simultaneously irreflexive ($(a, a) \notin R$), asymmetric ($(a, b) \in R \Rightarrow (b, a) \notin R$) and transitive on S. If every pair of distinct elements, a and b, in S are comparable under a strict order relation, R, then R is a *strict linear ordering* on S. Those strict orderings which are not linear are called *strict non-linear orderings* or, more commonly, *strict partial ordering relation.*
A non-strict partial order R on S always induces a strict partial order R^* by defining $aR^*b \Rightarrow [aRb \text{ and } a \neq b]$. Similarly, a strict partial order R on S always induces a non-strict partial order $R^\dagger$ by defining $aR^\dagger b \Rightarrow [aRb \text{ or } a = b]$.

Note that none of the relations defined above are equivalence relations since a *partial ordering relation* is normally not symmetric, while the *strict ordering relation* is not reflexive.

At first glance, the reader may find it difficult to distinguish one relation from the other or remember precisely what they mean. Studying the following few examples carefully will help construct a mental representation of the structure these relations provide to sets.

Example 1.

Mortimer is constructing a chart in which he will list all of his ancestors. The set of all his ancestors is represented by S. He defines an order relation R on S as follows: If a and b are two ancestors, $(a, b) \in R$ only if b is an ancestor of a − equivalently, a is a descendant of b. We list some properties of the relation R:

- We see that R is transitive since, "a is a descendant of b" and "b is a descendant of c" implies "a is a descendant of c".
- Since an element a cannot be an ancestor of a, R is irreflexive.
- Finally, if a is a descendant of b, then clearly b cannot be a descendant of a. So R is asymmetric.

We conclude that R is a strict order relation on S. Instead of writing $(a, b) \in R$, we will write $a < b$ with the understanding that "$<$" is only to be interpreted as "a is a descendant of b". We list a few more properties of R:

- It is clear that R does *not* linearly order S since one parent of Mortimer cannot be an ancestor of the other parent (excluding cases where something highly unnatural is going on). Hence, there exists pairs of ancestors a and b such that $a \not< b$ and $b \not< a$.
- Let's assume that Mortimer has included himself in the set S and is represented by the letter M. Then $M < a$ for all $a \in S$. We will say that M is the *minimum element* of S with respect to the ordering "$<$".

– Beginning with M, Mortimer can trace different paths upwards forming *chains* of inequalities each in the form $M < a < b < c < \cdots < \cdots$. Such chains are linearly ordered subsets of S since for any two elements a, b in such chains either $a < b$ or $b < a$. So, not only is M the minimum element of S, M is also the *minimal element* of each chain.

To allow us to illustrate in this example as many properties of ordered relations as possible, let's assume that Adam and Eve were "spontaneously generated" and so were the most ancient of Mortimer's ancestors (assuming Adam and Eve are the only human beings which were spontaneously generated). Say that in the set S, Adam is represented by A and Eve by E. We add another few properties of the set S when equipped with the given order relation R:

– We see that there are numerous chains of elements (linearly ordered subsets of S) each of which begins with the minimal element M and finishes with either A or E.

– If a chain C linearly links M to A, then A is a *maximal* element of this chain in the sense that all elements of S which are comparable to A are "below" A. Similarly, E is the maximal element of all chains which link M to E.

– However S has no maximum element since A is not an ancestor of E and E is not an ancestor of A. The elements A and E are simply not comparable under R. In this sense, we can say that S has no *maximum* element and only two *maximal* elements.

We now formalize some of the concepts illustrated in this example with the following definitions.

Definition 6.4 Let S be a class and R be an order relation on S. If R is a non-strict order relation, "$(a, b) \in R$" is represented as "$a \leq b$", and if R is a strict order relation, "$(a, b) \in R$" is represented as "$a < b$". If $a \leq b$ and $a \neq b$, we will simply write $a < b$.

a) A subset of S which is linearly ordered by R is called a *chain* in S. If R linearly orders S, then S is a linearly ordered subset of itself and therefore is a chain.

b) An element, M, of S is called a *maximal* element of S with respect to $\leq$ if there does not exist an element b in S such that $M < b$. An element m of S is called a *minimal* element of S with respect to $\leq$ if there does not exist an element b in S such that $b < m$.

c) Let M and m be elements which are comparable (with respect to $\leq$) to all elements of the set S. The element, M, in S is called the *maximum element of* S with respect to $\leq$ if there does not exist an element a in S such that $M < a$. The element m in S is called the *minimum element of* S if there does not exist an element a in S such that $a < m$.

Note that a *maximum* or *minimum* element of an ordered set S must be comparable to all elements of S. This is not necessarily the case for a *maximal* or *minimal* element of S. Referring to the example on "ancestors" of Mortimer above, S has two maximal elements A and E; it has no maximum element, since no single element in S is strictly larger than all other elements. But not much is needed to give to S a maximum element: Suppose that only Adam was "spontaneously generated" and Eve was somehow formed from one of Adams ribs; then Adam would be Eve's only ancestor and so A would be the maximum element of S.

Example 2.

Let S denote the set of all molecules constructed from the atoms listed in the periodic table of elements. In this case, molecules are viewed as sets whose elements are atoms. (We exclude crystals.) The simplest molecules are those that contain only one atom. We define the relation R on S as follows: $(a, b) \in R$ if $a \neq b$ and all atoms in molecule a are contained in molecule b and any atom which appears n times in a also appears n times in b. If $(a, b) \in R$, we will write $a \subset b$ (or say that a is a proper subset of b). We describe the structure of S when equipped with this particular relation R.

- By definition, R is irreflexive.
- If molecule a is a proper subset of molecule b, then b cannot be a proper subset of molecule a and so R is asymmetric.
- If $a \subset b$ and $b \subset c$, then $a \subset c$ and so R is transitive.

We conclude that the relation R strictly orders the set S. The set S is however not linearly ordered since well-known molecules such as Cl_2 (gaseous chlorine) and H_2 (hydrogen gas) are not comparable under "$\subset$". We discuss a few more properties of S when equipped with R.

- Since a molecule made of a single atom cannot be properly contained in any other molecule, "single-atom molecules" are minimal elements of S. So S has as many minimal elements as there are atoms. Clearly, S does not contain a minimum element.
- There are atoms belonging to the family of noble gases (helium (He), argon (Ar), krypton (Kr), etc.) that are non-reactive and so do not tend to bond with other elements to form molecules. In S, these elements will form one element chains. For example, the helium atom is not properly contained in any other molecule. It is both a maximal and minimal element of S. These particular atoms (noble gases) form in S what is called an *antichain*. An *antichain* is a subset of an ordered set in which no two elements are comparable.
- Other molecules will join together to form new molecules. For example the carbon element, C, and hydrogen element, H, both belong to the molecule, CH_2, which will join with some other molecules containing carbon and oxygen atoms to form $C_2H_2O_4$.

– The order structure of S will then contain numerous chains of molecules. One suspects that at some point each of these chains may attain some extremely large molecule which is non-reactive or will be too unstable to form other lasting links. If this is the case, then such a molecule is a maximal element of S. However, S cannot have a maximum element since elements such as those belonging to the noble gases are not contained in any molecule.

Example 3.

Consider the set $\mathbb{Z}$ of all integers. If m and n are non-zero integers, we will say that "m divides n" if there exists a positive integer k such that $mk = n$. We will define the relation R on $\mathbb{Z}$ as follows:

$$R = \{(0, n) : n \in \mathbb{Z}\} \cup \{(m, n) : m \text{ divides } n\} \cup \{(m, n) : \ m < 0 \text{ and } n > 0\}$$

We first determine the most elementary properties of this relation.

– For every non-zero integer m, $m \times 1 = m$ so $(m, m) \in R$. Also $(0, 0) \in R$. Then R is a reflexive relation on $\mathbb{Z}$.

– Let m, n and k be non-zero integers in $\mathbb{Z}$. If $(0, m) \in R$ and $(m, n) \in R$, then $(0, n) \in R$. By considering each of the possible positive-negative cases for the values of m, n and k we see that if $(m, n) \in R$ and $(n, k) \in R$, then $(m, k) \in R$. Then R is a transitive relation on $\mathbb{Z}$.

– If m and n are both negative or both positive, then if m divides n and n divides m, then it must be the case that $m = n$. If m is negative and n is positive, then $(m, n) \in R$ but $(n, m) \notin R$. Finally $(0, m) \in R$ for all m, but $(m, 0) \in R$ only if $m = 0$. We conclude that R is antisymmetric.

We conclude that the relation R partially orders the set $\mathbb{Z}$. The expression "$a \leq b$" will be another way of stating that $(a, b) \in R$. We will write $a < b$ if $a \leq b$ and $a \neq b$. We see that R does not linearly order $\mathbb{Z}$ since neither "3 divides 5" nor "5 divides 3" holds true. So 3 and 5 are not comparable under R. We list a few more properties of the order structure imposed on $\mathbb{Z}$ by R.

– It is explicitly stated in the definition of the relation R that 0 is the *minimum* element of $\mathbb{Z}$. Given any positive integer n, $n + 1$ does not divide n and so n cannot be a maximum element of $\mathbb{Z}$. So no integer is a maximum element of $\mathbb{Z}$ with respect to the relation R.

– We will say that a chain C is a *maximal chain* if no larger chain in $\mathbb{Z}$ properly contains C. If p represents a positive prime number, maximal chains must begin with $0 < -1 < -p$, when listed in the order dictated by R. For example,

$$0 < -1 < -3 < -6 < -12 < -24 < \cdots < 1 < 7 < 14 < 28 < 56 < \cdots$$

is an example of a chain with respect to the order relation R. But many distinct chains may begin with $0 < -1 < -3$. For example,

$$0 < -1 < -3 < -9 < -27 < -54 < \cdots < 1 < 2 < 4 < 12 < 24 < \cdots$$

In the order structure defined by R all maximal chains contain the number one.

Example 4.

Let C denote the class

$$\{S : S \text{ is a set and } S \in S\}$$

It seems that up to now, we have avoided discussing sets S which satisfy the property "$S \in S$"; nor have we proved that such sets don't exist. We could put forward a set theory axiom which excludes such sets from the universe of sets, but a this point there seems to be no strong reason to deny the existence of such sets.[1] For now, we will assume that the class C is non-empty. Suppose T is a set, ordered by $\in$, which contains an element a such that $a \in a$. Let $U = \{a\}$ be a subset of T linearly ordered by $\in$. Suppose m is a minimum element of U. Then $m = a$; but we know that $a \in m = a$. This contradicts our definition of "$\in$-minimum". So U cannot contain a minimum element with respect to $\in$. The linearly $\in$-ordered set $U = \{a\}$ may appear a bit strange to many readers who wonder "How can a set which contains a single element not have a $\leq$-minimum element?". For now, we will let such sets be. If these become a source of important contradictions we will then see if a special axiom is required.

Concepts review:

1. What is the "empty relation" on a class S?
2. What does it mean to say that the relation R on S is *reflexive*?
3. What does it mean to say that the relation R on S is *irreflexive*?
4. What does it mean to say that the relation R on S is *symmetric*?
5. What does it mean to say that the relation R on S is *asymmetric*?
6. What does it mean to say that the relation R on S is *antisymmetric*?
7. What does it mean to say that the relation R on S is *transitive*?

[1] The Axiom of regularity will have something to say about this.

8. What does it mean to say that any two elements a and b in S are *comparable* under the relation R?
9. What is an *equivalence relation* R on a class S?
10. What is a *partially ordered class*?
11. Give an example of a partial ordering on $\mathscr{P}(S)$.
12. What is a partially ordered set?
13. What is a *strictly ordered class*?
14. Give an example of a strict ordering on $\mathscr{P}(S)$.
15. What is a *poset*?
16. If R is a partial order on the set S, what does it mean to say that a is a maximal element of S? What does it mean to say that a is a minimal element of S?
17. What is a chain in a partially ordered set?
18. What is the maximum element of a partially ordered set?
19. What is a linearly ordered set?

EXERCISES

A.
1. Suppose R is a reflexive relation on a class S. Show that $\mathrm{Id}_S \subseteq R$.
2. Show that if R is reflexive, then R^{-1} is reflexive.
3. Show that if R is symmetric, then R^{-1} is symmetric.
4. Show that if R is transitive, then R^{-1} is transitive.

B.
5. Show that if R is asymmetric, then $R \cap R^{-1} = \varnothing$.
6. Show that if R is an equivalence relation, then $R \circ R = R$.
7. Suppose T is a reflexive relation on a class S. Show that for every relation R on S, $R \subseteq T \circ R$ and $R \subseteq R \circ T$.

C.
8. Show that if R is a partial order relation on S, then $R \cap R^{-1} = \mathrm{Id}_S$ and $R \circ R = R$.
9. Show that if R is a partial order relation on S, then so is R^{-1}.
10. Suppose R is an equivalence relation on a class S. Show that if H and J are relations on S, then $R \subseteq (H \cap J) \Rightarrow R \subseteq H \circ J$.

11. Show that if a is a maximum element of a partially ordered set S, then it is the only maximal element of S.
12. Show that if a is a maximal element of a linearly ordered set S, then it is a maximum element of S.

---

7 / Partitions induced by equivalence relations.

Summary. *In this section we show that an equivalence relation R on a set S can be used to subdivide S into pairwise disjoint subsets. The subsets when viewed together are called a partition of S. We illustrate how such partitions are obtained.*

7.1 Subdividing a set S using an equivalence relation R on S.

Suppose R is an equivalence relation on the set S. We will show that this equivalence relation R, no matter how it is defined, can be used to subdivide the set S into a collection of non-empty pairwise disjoint[1] subsets. Furthermore, no element of S will be left out in this process (in the sense that every element of S will belong to one of these subsets).

The process is similar to subdividing a school's student population, S, into smaller subgroups $\{S_1, S_2, S_3, \ldots\}$ based on some predefined student characteristics; the characteristics are such that no student can belong to two subgroups.
Note: Although most of the results proven from here on can apply to classes we will prove these as applied to sets.

Notation 7.1 Let R be an equivalence relation on a set S and let $x \in S$. We define the set S_x as follows:

$$S_x = \{y : (x, y) \in R\}$$

That is, S_x is the set[2] of all elements y in S such that y is related to x under R.

The following five theorem statements show, step by step, how an equivalence relation, R, on S partitions the set S into subsets, $\{S_x : x \in S\}$. The reader should notice how all three relation properties which characterize equivalence relations are required to partition the set S in this way.

Theorem 7.2 Let R be an equivalence relation on a set S. Let x and y be two elements in S which are not related under R. Then any element z in S which is related to x cannot be related to y.

---

[1] A set $\mathscr{S}$ of sets whose elements are "pairwise disjoint" means that for any two sets A and B in $\mathscr{S}$, $A \cap B = \varnothing$.

[2] We justify that S_x a "set" as follows: We have that S_x is a subclass of the class S; given that S is declared to be a set, by axiom A4 (axiom of subset) S_x is a set.

Proof:

What we are given:

- R is an equivalence relation
- $x, y, z \in S$
- $(x, y) \notin R$ but $(x, z) \in R$

What we are required to show: That $(y, z) \notin R$.

Suppose $(y, z) \in R$. Then $(z, y) \in R$ (since R is an equivalence relation and so is symmetric). But $(x, z) \in R$ and $(z, y) \in R \Rightarrow (x, y) \in R$ (since R is an equivalence relation and so is transitive). This contradicts our hypothesis: $(x, y) \notin R$ Thus, $(y, z) \notin R$, as required.

Theorem 7.3 Let R be an equivalence relation on a set S. Let x and y be two elements in S which are not related under R. Then $S_x \cap S_y = \varnothing$.

Proof:

What we are given:

- R is an equivalence relation
- $x, y \in S$
- $(x, y) \notin R$

What we are required to show: $S_x \cap S_y = \varnothing$.

Suppose $z \in S_x \cap S_y$. Then z belongs to both S_x and S_y and so is related both to x and to y. Since R is an equivalence relation on S, it is transitive and so x must be related to y, a contradiction. So z cannot belong to $S_x \cap S_y$. So $S_x \cap S_y = \varnothing$ as required.

Theorem 7.4 Let R be an equivalence relation on a set S. Let x and y be two elements in S which are related under R. Then $S_x = S_y$.

Proof:

What we are given:

- R is an equivalence relation
- $x, y \in S$
- $(x, y) \in R$

What we are required to show: $S_x = S_y$.

We claim that $S_x \subseteq S_y$. Let $z \in S_x$. Then $(x, z) = (z, x) \in R$ (since R is symmetric). Since x is related to y, then $[(z, x) \in R$ and $(x, y) \in R] \Rightarrow [(z, y) \in R]$ (since R is transitive). So $z \in S_y$. Thus, $S_x \subseteq S_y$ as claimed.
We prove in a similar way that $S_y \subseteq S_x$. Thus, $S_x = S_y$ as required.

Theorem 7.5 Let R be an equivalence relation on a set S. For every $x \in S$, there exists some $y \in S$ such that $x \in S_y$.

Proof:

What we are given: That R be an equivalence relation and $x \in S$.
What we are required to show: That there exists some $y \in S$ such that $x \in S_y$.
Since R is reflexive, $(x, x) \in R$. Thus, $x \in S_x$. So it suffices to choose $y = x$. Hence, $x \in S_x$.

Theorem 7.6 Let R be an equivalence relation on a set S. Then $\bigcup_{x \in S} S_x = S$.

Proof:

What we are given: That R is an equivalence relation on S.
What we are required to show: That $\bigcup_{x \in S} S_x = S$.
Let $x \in S$. Then $x \in S_x$. Hence, $x \in \bigcup_{x \in S} S_x$. Thus, $S \subseteq \bigcup_{x \in S}$. Let $y \in \bigcup_{x \in S} S_x$. Then there exists some $x \in S$ such that $y \in S_x$. Since $S_x \subseteq S$ for all $x \in S$, then $y \in S$. Thus, $\bigcup_{x \in S} S_x \subseteq S$. We conclude that $\bigcup_{x \in S} S_x = S$ as required.

7.2 A partition of a set induced by an equivalence relation.

Given an equivalence relation R on a set S, we have seen how S can be expressed as a union, $\bigcup_{x \in S} S_x$, of pairwise disjoint subsets, S_x, of S. Let

$$\mathscr{S}_R = \{S_x : x \in S\}$$

denote the set of all the sets formed from the equivalence relation R. We verify that the class $\mathscr{S}_R$ is indeed a set: Since S is declared to be a set, then $\mathscr{P}(S)$ is a set (by the Axiom of power set); since $\mathscr{S}_R$ is a subclass of the set $\mathscr{P}(S)$, then, by the Axiom of subset, $\mathscr{S}_R$ is a set.

We examine the elements of $\mathscr{S}_R$ more carefully to see how they relate to each other.

- Having shown that $[(x, y) \in R] \Rightarrow [S_x = S_y]$, many of the sets in $\mathscr{S}_R$ are the same set. That is, identical sets may simply have different labels.
- If $(x, y) \notin R$, then $S_x \cap S_y = \varnothing$. So any two sets S_x and S_y in $\mathscr{S}_R$ are either the same set or $S_x \cap S_y = \varnothing$. No two subsets in $\mathscr{S}_R$ can overlap.
- Since R is reflexive, then no subset S_x is empty. (For any $x \in S$, $x \in S_x$.)
- On the other hand, if $x \in S$, since R is reflexive, $x \in S_x$ and so $x \in \bigcup_{x \in S} S_x$. Thus, $S \subseteq \bigcup_{x \in S} S_x$.

– We have already seen that $\bigcup_{x\in S} S_x = S$.

We summarize three important properties of $\mathscr{S}_R$:

1) $\bigcup_{x\in S} S_x = S$
2) $S_x \neq S_y \Rightarrow S_x \cap S_y = \varnothing$.
3) $S_x \neq \varnothing$ for all $x \in S$.

The three properties together describe what is called a *partition of a set S*.[1] A proper understanding of the method used to partition a set S in this way is important in a our study of set theory.

Finest and coarsest partitions. Let S be a non-empty set. The first example below illustrates the "finest" partitioning of S. The second example illustrates the "coarsest" partitioning of S.

Examples.

1) The identity relation, Id_S, on S is defined as follows: $(x, y) \in \mathrm{Id}_S$ if and only if $x = y$.
   a) Show that the relation Id_S is indeed an equivalence relation on S.
   b) If $\mathscr{S}_{\mathrm{Id}_S} = \{S_x : x \in S\}$, describe S_x for each $x \in S$.
   c) Compare the class of subsets in $\mathscr{S}_{\mathrm{Id}_S}$ to the class of subsets $\mathscr{M} = \{\{x\} : x \in S\}$. How do the elements of $\mathscr{M}$ compare to the elements of $\mathscr{S}_{\mathrm{Id}_S}$?
   d) In your opinion, is it possible to obtain an even "finer" partition of S? That is, is it possible to obtain a partition of S where each S_x is strictly smaller than the elements of $\mathscr{S}_{\mathrm{Id}_S}$?

2) Consider the relation R on the set S defined as follows: $(x, y) \in R$ if x and y both belong to S.
   a) Confirm that the relation R is an equivalence relation on S.
   b) If $\mathscr{S}_R = \{S_x : x \in S\}$, describe S_x for each $x \in S$.
   c) Describe the elements of $\mathscr{S}_R$. Which subsets of S are elements of $\mathscr{S}_R$? List all subsets of $\mathscr{S}_R$. Is this true or false: $\mathscr{S}_R = S$.
   d) In your opinion, is it possible to obtain an even "coarser" partition of S? That is, is it possible to obtain a partition of S where each S_x is strictly larger than the elements of $\mathscr{S}_R$?

Concepts review:

[1] If a set $\mathscr{S}_R = \{S_x : x \in S\}$ of subsets of S is such that $\bigcup_{x\in S} S_x = S$ we often say "$\mathscr{S}_R$ *covers* S" to express this fact.

1. Given an equivalence relation R on a set S and an element $x \in S$, what does the symbol S_x represent?

2. Given an equivalence relation R on a set S what does the symbol $\mathscr{S}_R$ represent?

3. Suppose $\mathscr{S}$ denotes a set of subsets of S. What does it mean to say that $\mathscr{S}$ partitions S?

4. Given an equivalence relation R on a set S what are the essential properties of $\mathscr{S}_R$?

5. What is the "finest" partition of S obtainable by an equivalence relation on S?

6. What is the "coarsest" partition of S obtainable by an equivalence relation on S?

EXERCISES

1. Let S be a set and let R and T be two equivalence relations on S. Is the relation $V = R \cup T$ necessarily an equivalence relation on S? Explain.
2. Suppose R is an equivalence relation on a set S and that $A \subseteq S$ where A is non-empty. We define the relation R_A on A as follows:

   $$R_A = \{(x, y) : (x, y) \in R \cap (A \times A)\}$$

   Show that R_A is an equivalence relation on A.
3. Suppose R is a partial order relation on a set S and that $A \subseteq S$ where A is non-empty. We define the relation R_A on A as follows:

   $$R_A = \{(x, y) : (x, y) \in R \cap (A \times A)\}$$

   Show that R_A is a partial order relation on A.

4. Let S be a set and let R and T be two equivalence relations on S. Show that $V = R \cap T$ is an equivalence relation on S.
5. Suppose R and T are two equivalence relations on a set S. For each $x \in S$, let ${}_RS_x = \{y : y \in S,\ (x, y) \in R\}$ and ${}_TS_x = \{y : y \in S,\ (x, y) \in T\}$. If, for each $x \in S$, ${}_RS_x \subseteq {}_TS_x$ show that $R \subseteq T$.
6. Suppose R and T are two equivalence relations on a set S. For each $x \in S$, let ${}_RS_x = \{y : y \in S,\ (x, y) \in R\}$ and ${}_TS_x = \{y : y \in S,\ (x, y) \in T\}$. If $R \subseteq T$, show that for each $x \in S$, ${}_RS_x \subseteq {}_TS_x$.

7. Let S be a set and let R and T be two equivalence relations on S. For each $x \in S$ let ${}_RS_x = \{y : y \in S,\ (x,y) \in R\}$ and ${}_TS_x = \{y : y \in S,\ (x,y) \in T\}$. Let $\mathscr{S}_R = \{{}_RS_x : x \in S\}$ and $\mathscr{S}_T = \{{}_TS_x : x \in S\}$. We have seen that $\mathscr{S}_R$ and $\mathscr{S}_T$ form sets of non-empty subsets of S which are pairwise disjoint and cover all of S. For each $x \in S$, let $S_x = {}_RS_x \cap {}_TS_x$. Show that $\mathscr{S} = \{S_x : x \in S\}$ forms a set of subsets of S satisfying the properties:

   a) $S_x \neq \varnothing$ for each x.
   b) Whenever $S_x \neq S_y$ then $S_x \cap S_y = \varnothing$.
   c) $\bigcup_{x \in S} S_x = S$.

8. Let S and T be sets. It has been shown that "$\subseteq$" constitutes a partial order relation on the set $\mathscr{P}(S)$. Consider the set $L = \mathscr{P}(S) \times \mathscr{P}(T)$. We define a relation R on L as follows: For (A,B) and $(C,D) \in \mathscr{P}(S) \times \mathscr{P}(T)$

$$((A,B),(C,D)) \in R \Leftrightarrow \begin{cases} A \subset C \\ \quad \text{or} \\ A = C \quad \text{and} \quad B \subseteq D \end{cases}$$

   Show that R is a partial ordering relation on L.[1]

[1] This relation is often called the *lexicographic ordering of a Cartesian product.*

8 / Equivalence classes and quotient sets.

Summary. *In this section we continue our discussion of partitions of a set S by an equivalence relation. When a set S is partitioned by an equivalence relation R the subsets in this partition are called "equivalence classes induced by R". These subsets together are called a "quotient set of S induced by R". We then show that any partition of S is induced by some equivalence relation R.*

8.1 More on partitions.

We have seen that an equivalence relation, R, on a set, S, subdivides S into pairwise disjoint subsets that cover all of S. Tools that allow us to systematically partition sets into subsets are important in mathematics. In set theory, this can be a method to construct a new set from a known set S by partitioning S into smaller pieces and forming a new set whose elements are those pieces. We have casually spoken of partitions with an intuitive understanding of what they are. We should formally define them before we go on.

Definition 8.1 Let S be a set. We say that a set of subsets $\mathscr{C} \subseteq \mathscr{P}(S)$ forms a *partition of S* if $\mathscr{C}$ satisfies the 3 properties:

1) $\bigcup_{A \in \mathscr{C}} A = S$

2) If A and $B \in \mathscr{C}$ and $A \neq B$, then $A \cap B = \varnothing$.

3) $A \neq \varnothing$ for all $A \in \mathscr{C}$.

Based on this definition, we can see that the set of subsets of S, $\mathscr{S}_R = \{S_x : x \in S\}$, formed by the equivalence relation, R, is a *partition of the set S*.[1] The set, $\mathscr{S}_R$, and its elements are referred to in a particular way in various fields of mathematics. We formally define these below.

Definition 8.2 Let S be a set on which an equivalence relation, R, is defined.

a) Each element S_x of $\mathscr{S}_R = \{S_x : x \in S\}$ is called an *equivalence class of x under R* or an *equivalence class induced by the relation R*.[2]

[1] We can say *$\mathscr{S}_R$ partitions S*.

[2] The expression "*equivalence class of x modulo R*" is sometimes used.

b) The set, $\mathscr{S}_R = \{S_x : x \in S\}$, of all equivalence classes induced by the relation, R, is called the *quotient set of S induced by R*. The set, $\mathscr{S}_R$, is more commonly represented by the symbol, S/R. So

$$S/R = \{S_x : x \in S\}$$

From here on we will use the more common notation, S/R.

Note that if S is a set, then by the Axiom of subset, the equivalence classes in $S/R = \mathscr{S}_R = \{S_x : x \in S\}$ are in fact "sets". But if S is a proper class, then it may occur that the elements of $S/R = \{S_x : x \in S\}$ are equivalence classes which are also proper classes.

8.2 Examples of quotient sets induced by an equivalence relation R.

In the following examples, S and T are non-empty sets.

a) Recall that Id_S denotes the *identity relation* on S:

$$(x, y) \in \mathrm{Id}_S \text{ if and only if } x = y$$

We have seen on page 53 that Id_S is an equivalence relation on S. For each $x \in S$, the equivalent class of x induced by Id_x is

$$\begin{aligned} S_x &= \{y : (x, y) \in \mathrm{Id}_S\} \\ &= \{y : x = y\} \\ &= \{x\} \end{aligned}$$

So the quotient set of S induced by Id_S is

$$S/\mathrm{Id}_S = \mathscr{S}_{\mathrm{Id}_x} = \{\{x\} : x \in S\}$$

The set $\mathscr{S}_{\mathrm{Id}_x}$ is the largest possible quotient set on S induced by a relation. We also say that this quotient set is the "finest" partition of S.

b) Let R be the relation defined as follows: $(x, y) \in R$ if and only if x and y belong to S. Then $R = S \times S$ was shown to be an equivalence relation on S. For $x \in S$,

$$\begin{aligned} S_x &= \{y : y \in S, (x, y) \in R\} \\ &= \{y : y \in S\} \\ &= S \end{aligned}$$

Hence, for every $x \in S$, the equivalence class of x induced by R is $S_x = S$. So the quotient set of S induced by R only contains the element $S_x = S$ and no other. Since

$$S/R = \mathscr{S}_R = \{S\}$$

it is the smallest possible quotient set of S induced by a relation. We also say that this quotient set is the "coarsest" partition of S.

c) For non-empty sets S and T, let $D = S \times T$. We define a relation, R, on D as follows:

$$((a,b),(c,d)) \in R \text{ if and only if } a = c$$

We have seen on page 53 that R is an equivalence relation on D. Let $(a,b) \in D$. Then

$$\begin{aligned} S_{(a,b)} &= \{(x,y) : (x,y) \in D, x = a\} \\ &= \{(a,y) : y \in T\} \\ &= \{a\} \times T \end{aligned}$$

Thus, $\{a\} \times T$ is the equivalence class of (a,b) induced by R. The quotient set of D induced by R is

$$D/R = \mathscr{D}_R = \{\{x\} \times T : x \in S\}$$

8.3 Equivalence relations defined from a partition.

We have seen how any equivalence relation R on a set S can be used to partition S into pairwise disjoint subsets. We will now work the other way around: If we are given a partition, $\mathscr{C}$, of S, is there an equivalence relation, R, such that $\mathscr{S}_R = \mathscr{C}$? We will see if we can construct the required equivalence relation.

Let S be a set on which we have defined a partition $\mathscr{C}$.

This means that $\mathscr{C}$ is a class of non-empty pairwise disjoint subsets of S which covers all of S. Suppose we define a relation $R_{\mathscr{C}}$ on S in the following way:

$$(x,y) \in R_{\mathscr{C}} \text{ if and only if } \{x,y\} \subseteq C \text{ for some element } C \text{ in } \mathscr{C}$$

Thus, the only pairs of elements x and y of S which are related under $R_{\mathscr{C}}$ are those pairs that appear together in the same subset $C \in \mathscr{C}$. Is $R_{\mathscr{C}}$ an equivalence relation?

- We verify that $R_{\mathscr{C}}$ is reflexive: For every $x \in S$, x belongs to some C and so $\{x\} = \{x,x\} \subseteq C$ and so $(x,x) \in R_{\mathscr{C}}$.
- We verify symmetry of $R_{\mathscr{C}}$: If $\{x,y\} \subseteq C \in \mathscr{C}$, then $\{y,x\} \subseteq C$. So $(x,y) \in R_{\mathscr{C}} \Rightarrow (y,x) \in R_{\mathscr{C}}$.
- We verify transitivity of $R_{\mathscr{C}}$: If $\{x,y\} \subseteq C$ and $\{y,z\} \subseteq C$, then $\{x,z\} \subseteq C$. So $(x,y) \in R_{\mathscr{C}}$ and $(y,z) \in R_{\mathscr{C}} \Rightarrow (x,z) \in R_{\mathscr{C}}$.

The relation $R_{\mathscr{C}}$ is indeed an equivalence relation on S. We conclude that any partition $\mathscr{C}$ of a set S defines an equivalence relation $R_{\mathscr{C}}$ on S. This result deserves to be called a theorem.

Theorem 8.3 Let S be a set and $\mathscr{C}$ be a partition of S. Let $R_{\mathscr{C}}$ be the relation such that $(x, y) \in R_{\mathscr{C}}$ if and only if $\{x, y\} \subseteq C$ for some element, C, in $\mathscr{C}$. Then $R_{\mathscr{C}}$ is an equivalence relation on S.

8.4 Refining an equivalence relation[1].

Let $S = \{a, b, c, d, e\}$. Suppose R, T, K, Id_S and M are equivalence relations on S which are defined as follows:

$$\begin{aligned}
R &= \{(a,a),(b,b),(c,c),(d,d),(e,e),(a,b),(b,a),(d,c),(c,d)\} \\
T &= \{(a,a),(b,b),(c,c),(d,d),(e,e),(a,b),(b,a),(c,d),(d,c),(c,e),(e,c),(d,e),(e,d)\} \\
K &= \{(a,a),(b,b),(c,c),(d,d),(e,e),(a,c),(c,a),(a,b),(b,a),(c,b),(b,c),(d,e),(e,d)\} \\
\mathrm{Id}_S &= \{(x,y) : x \in S, y \in S \text{ and } x = y\} = \{(a,a),(b,b),(c,c),(d,d),(e,e)\} \\
M &= \{(x,y) : x \in S, y \in S\} = S \times S
\end{aligned}$$

Verify that R, T, K and M are indeed equivalence relations. (Note that it suffices to show that R, T and K partition S.)
We describe explicitly the elements of $\mathscr{S}_R$, $\mathscr{S}_T$, $\mathscr{S}_K$, $\mathscr{S}_{\mathrm{Id}_S}$ and $\mathscr{S}_M$. (Recall that these can also be expressed in the form S/R, S/T, S/K, S/Id_S, S/M):[2]

$$\begin{aligned}
\mathscr{S}_R &= \{{}_RS_a,\ {}_RS_b,\ {}_RS_c,\ {}_RS_d,\ {}_RS_e\} \\
&= \{\{a,b,\},\{c,d\},\{e\}\} \\
\mathscr{S}_T &= \{{}_TS_a,\ {}_TS_b,\ {}_TS_c,\ {}_TS_d,\ {}_TS_e\} \\
&= \{\{a,b,\},\{c,d,e\}\} \\
\mathscr{S}_K &= \{{}_TS_a,\ {}_TS_b,\ {}_TS_c,\ {}_TS_d,\ {}_TS_e\} \\
&= \{\{a,b,c\},\{d,e\}\} \\
\mathscr{S}_{\mathrm{Id}_S} &= \{{}_{\mathrm{Id}_S}dS_a,\ {}_{\mathrm{Id}_S}S_b,\ {}_{\mathrm{Id}_S}S_c,\ {}_{\mathrm{Id}_S}S_d,\ {}_{\mathrm{Id}_S}S_e\} \\
&= \{\{a\},\{b\},\{c\},\{d\},\{e\}\} \\
\mathscr{S}_M &= \{{}_TS_a,\ {}_TS_b,\ {}_TS_c,\ {}_TS_d,\ {}_TS_e\} \\
&= \{\{a,b,c,d,e\}\} = \{S\}
\end{aligned}$$

We now make a few observations:

a) We see that each pair in R belongs to the relation T. So we can write $R \subseteq T$. Similarly, we see that the Id_S is a subset of each of R, T and K. But this relationship doesn't hold true between R and K: The pair (c, d) is an element of R but not of K. So $R \not\subseteq K$.

b) Notice how every equivalence class under R is contained in an equivalence class under T:

$$\begin{aligned}
\{a,b,\} &\subseteq \{a,b,\} \\
\{c,d\} &\subseteq \{c,d,e\} \\
\{c\} &\subseteq \{c,d,e\}
\end{aligned}$$

[1] This section can be omitted without loss of continuity.
[2] Recall that ${}_RS_x = \{y : (x, y) \in R\}$ where R is an equivalence relation on S.

Similarly, every equivalence class under Id_S is a subset of some equivalence class under R. The same can be said for the relationship between the equivalence classes of Id_S and of K. But this relationship between the equivalence classes under R and the equivalence classes under K doesn't hold true. Witness:

$$\{c, d\} \not\subseteq \{a, b, c\}$$
$$\{c, d\} \not\subseteq \{d, e\}$$

c) In cases such as R and T above, we say that the equivalence relation R *refines* or *is a refinement of* the relation T. We can make the more general statement:

> *An equivalence relation R refines the equivalence relation T whenever $R \subseteq T$.*

This is the same as saying "R refines T if every equivalence class under R is a subset of some equivalence class under T". We see that Id_S will refine any equivalence relation R. We see that R does not refine K since $R \not\subseteq K$. Similarly, we see that T does not refine K since $T \not\subseteq K$.

d) Also note that no matter which equivalent relation R on S we consider, $R \subseteq M$, and so M is refined by any equivalence relation on S.

Concepts review:

1. Given a set S, what does it mean to say that the class $\mathscr{C}$ of subsets of S *partitions* S?
2. Given an equivalence relation R on a set S and an element $x \in S$, what is an *equivalence class of x under R*? How is it denoted?
3. Given an equivalence relation R on a set S, what is the *quotient set of S induced by R*? How is it denoted?
4. Given an equivalence relation R on a set S what do the expressions $\mathscr{S}_R$ and S/R mean?
5. Given a partition $\mathscr{C}$ of a set S can we define an equivalence relation R on S such that $S/R = \mathscr{C}$?
6. What does it mean to say that an equivalence relation R refines the equivalence relation T?
7. Is there an equivalence relation on a set S that refines all other equivalence relations?
8. Is there an equivalence relation on a set S that is refined by all other equivalence relations?

EXERCISES

1. Suppose R is an equivalence relation on a set S and that $A \subseteq S$ where A is non-empty. We define the relation R_A on A as follows:

$$R_A = \{(x, y) : (x, y) \in R \cap (A \times A)\}$$

   a) Show that R_A is an equivalence relation on A.
   b) If $S/R = \mathscr{S}_R = \{S_x : x \in S\}$ represents the quotient set of S induced by R, describe the elements of the quotient set, $\mathscr{S}_{R_A}$, of A induced by R_A.

2. Let S be a set and let R and T be two equivalence relations on S.
   a) Show that $V = R \cap T$ is an equivalence relation on S.
   b) If $S/R = \mathscr{S}_R = \{{}_RS_x : x \in S\}$ and $S/T = \mathscr{S}_T = \{{}_TS_x : x \in S\}$ represent the quotient set of S induced by R and induced by T respectively, describe the elements of the quotient set of S, $\mathscr{S}_V$, induced by V.
3. Suppose R and T are two equivalence relations on a set S. For each $x \in S$, let $S/R = \mathscr{S}_R = \{{}_RS_x : x \in S\}$ and $S/T = \mathscr{S}_T = \{{}_TS_x : x \in S\}$ represent the quotient set of S induced by R and induced by T respectively. If $R \subseteq T$, show that for each $x \in S$,

$${}_TS_x = \bigcup_{y \in {}_TS_x} {}_RS_y$$

4. Let R and T be two equivalence relations on a set S. If $R \subseteq T$ we say that R is *finer* than T (or is a *refinement* of T). The choice of these expressions when comparing two equivalence relations is suggested by the result described in problem 3: The quotient sets of a finer equivalence relation all seem to fit neatly inside the quotient sets of a coarser equivalence relation. Give the finest possible equivalence relation on S. Give the coarsest possible equivalence relation on S.
5. Let $S = \{a, b, c, d, e, f\}$ be a set. Suppose R and T are equivalence relations on S defined as follows:

$$\begin{aligned} R &= \mathrm{Id}_S \cup \{(a, b), (b, a), (b, c), (c, b), (a, c), (c, a), (d, c), (c, d)\} \\ T &= \mathrm{Id}_S \cup \{(b, c), (c, b)\} \end{aligned}$$

   Write out explicitly the elements of the sets $\mathscr{S}_R$ and $\mathscr{S}_T$.

6. Let S be a set. Let R and T be two equivalence relations on S where $R \subseteq T$ (that is R is finer than T). For each $x \in S$, let $S/R = \mathscr{S}_R = \{{}_RS_x : x \in S\}$ and

$S/T = \mathscr{S}_T = \{{}_TS_x : x \in S\}$ represent the quotient set of S induced by R and induced by T respectively. We define the *quotient of T by R*, denoted T/R, as follows:

$$T/R = \{({}_RS_x,\ {}_RS_y) : (x, y) \in T\}$$

From this definition we see that T/R is a relation on $S/R = \mathscr{S}_R$. That is ${}_RS_x$ and ${}_RS_y$ are related under T/S if and only if x and y are related under T.

a) Show that T/R is an equivalence relation on $S/R = \mathscr{S}_R$.

b) Let K be another equivalence relation on S where $R \subseteq T \subseteq K$. Show that $T/G \subseteq K/R$.

c) Referring to the example on page 71 write out explicitly the elements of the equivalence relation T/R on $\mathscr{S}_R$.

d) Referring to the set S described in question 5 above, write out explicitly the elements of the equivalence relation T/R on $\mathscr{S}_R$.

Part IV

Functions

9 / Functions: A set-theoretic definition.

Summary. *In this section we formally define what we mean by a "function". This is done using only the set-theoretic concepts developed up to now. We introduce notation to simplify the discussion and definition of these concepts. Examples of simple functions such as the identity function, the characteristic function and constant functions are presented. Given a function, f, on a set A, we define the restriction of this function on a subset, D, of A. We state what we mean by "equal functions". The expressions "one-to-one", "injective", "onto", "surjective", "bijective functions" are also defined.*

9.1 A set-theoretic definition.

The concept of a function is not new to most readers. A standard definition goes something like this: "Given two sets A and B, a function is a rule, f, which associates to each element of A a single element of B".

To construct a function, first a property involving two elements x, y is defined. Say we represent this property by $\phi(x, y)$. The property ϕ is the blueprint which is used to construct a subset

$$f = \{(x, y) \in A \times B : x, y \text{ satisfies the property } \phi(x, y)\}$$

of the Cartesian product $A \times B$. So ϕ is the tool used to distinguish those ordered pairs (x, y) which belong to f from those that don't. This defines a *relation*, f. If it can be shown that "(x, y) and (x, z) belong to f implies $y = z$", then f is called a *function*. Since a function f is a subclass of a Cartesian product of sets, then functions are, by definition, sets. In practice, users often do not distinguish between the rule ϕ and the set f whose elements are determined by it (even though ϕ is just a formula, while f is a well-defined set and so is governed by axioms associated with sets). For example, let $A = \{1, 2\}$ and $B = \{2, 3, 4\}$; define the rule ϕ as "the second element is twice the first". The set which follows from ϕ is, $f = \{(1, 2), (2, 4)\} \subset A \times B$. Or, we could write, $f = \{(x, y) : x \in A,\ \phi(x, y) \text{ is satisfied}\} \subseteq A \times B$. For practical reasons, we normally just use the symbol f to represent both the rule and the set which flows from it. Opportunities to say more about the notion of a function abound in the following chapters in this text.

Our objective will be to define the concept of a function within the *ZFC*-universe, without adding any new primitive concepts to the three we already have: "class", "set", "belongs to". We must formulate this definition carefully so that it represents precisely what we want and understand it to be.

For most readers the notion of a "function" is intrinsically linked to those sets we call "numbers": natural numbers, integers, rational numbers and real numbers. So it may

appear strange to discuss the concept of a *function* before describing what *numbers* actually are in set theory. We have not yet shown how numbers can be constructed using our *ZFC* axioms. This is to come. In what follows, we will see that functions exist independently of *numbers*. *Functions* can be defined in terms of abstract sets. Studying functions in the absence of numbers will allow readers to better see, in essence, what they truly are.

Definition 9.1 A *function f mapping elements from a set A into a set B* is a triple $\langle f, A, B\rangle$ satisfying the following properties[1]:

1) $f \subseteq A \times B$

2) For every $a \in A$ there exists $b \in B$ such that $(a, b) \in f$.[2]

3) If $(a, b) \in f$ and $(a, c) \in f$, then $b = c$. Equivalently, if $(a, b) \in f$ and $(c, d) \in f$, then $(b \neq d) \Rightarrow (a \neq c)$.) [3]

From this definition we see that a function $f \subseteq A \times B$ is a special type of relation on $A \cup B$ with $\operatorname{dom} f \subseteq A$ and $\operatorname{im} f \subseteq B$. A function f can also be viewed as a particular element of $\mathscr{P}(A \times B)$.

9.2 Commonly used notation when discussing functions.

There is no reason why we should adopt functional notation which is different from the one we are accustomed to. We should however explain carefully how this notation is to be interpreted in set theory.

- Rather than represent a function as $\langle f, A, B\rangle$ we will write

$$f : A \to B$$

  and say "f maps elements of A into B". When we write, "$f : A \to B$", it will always be understood that A and B are sets.
- If $(x, y) \in f$, we will write

$$f(x) = y$$

  and say that *y is the image of x under f*.
- We will also say that x is a *preimage* or an *inverse image* of the element y.

[1] By "a triple $\langle f, A, B\rangle$" we mean that a function is characterized by three sets f , A and B with the described properties.

[2] Using logical symbols: $\forall x \in A\ \exists b \in B \,|\, (a, b) \in f$

[3] Using logical symbols: $[(a, b) \in f] \wedge [(a, c) \in f] \Rightarrow (b = c)$.

The definition of a function, f, states that f is a subset of $A \times B$ and therefore is a relation. If $A = B$ it is a relation on A with the extra condition: "$f(a) \neq f(b) \Rightarrow a \neq b$". Since a function is a relation we can then speak of its *domain* and its *image*.

- From the definition of "function", we see that for every $x \in A$, there is some $y \in B$ such that $(x, y) \in A \times B$. So, by definition of the domain of a relation (see definition 5.3),
$$A = \text{ dom } f$$
- It may be that not every $y \in B$ is such that $(x, y) \in f$ for some x in A. So we must be clear about what we mean by the image of f:
$$\text{im } f = \{y \in B : (x, y) \in f \text{ for some } x \in \text{ dom } f\}$$
If A is the domain of the function $f \subseteq A \times B$, then we will express the image, im f, of A under f as
$$\text{im } f = f[A]$$
- Note that im f is contained in B and need not be equal to B. To distinguish between $f[A]$ and B we will refer to B as being the *codomain* of A, abbreviated as codom f. The words "range of f", denoted as ran f, is often used instead of "image of f". In such cases you will read
$$\text{ran } f = f[A]$$

9.3 Restricting a function to a subset of its domain.

Suppose we are given a function, $f : A \to B$. Then f is mapping each element in its domain A into B. If $D \subseteq A$, then we may restrict the domain of f so that it only acts on the elements of D. We will show that $f : D \to C$ is also a function:

Since $f : A \to B$ is a function then by definition, for every $x \in A$, there exists $y \in B$ such that $f(x) = y$. Since $D \subseteq A$ then for every $x \in D$, $x \in A$; hence there exists $y \in B$ such that $f(x) = y$. Since the image (under f) of every $x \in A$ is unique, then the same is true for every $x \in D \subseteq A$. Thus, by definition, $f : D \to B$ is a function.

Notation to express the restriction of a function f to a subset of its domain will be useful. We introduce this now.

Definition 9.2 If $f : A \to B$ is a function and $D \subseteq A$, then we say that the function $f : D \to C$ is *a restriction of f to D*. In this case we will use the symbol, $f|_D$, to represent the restriction of f to D. Note that if $D \subseteq A$, then we can write, $f|_D \subseteq f$, since

$$f|_D = \{(x, y) : x \in D \text{ and } (x, y) \in f\} \subseteq f$$

We will now see that a function, $f : A \to B$, can always be expressed as the union of two functions, provided its domain contains more than one element.

Theorem 9.3 Let $f : A \to B$ be a function and suppose $A = C \cup D$, where neither C nor D is empty. Then $f = f|_C \cup f|_D$.

Proof:

Given: A function $f : A \to B$ is defined and $A = C \cup D$.

$$\begin{aligned}
f &= \{(x,y) : (x,y) \in A \times B \text{ and } (x,y) \in f\} \\
&= \{(x,y) : (x,y) \in (A \times B) \cap f\} \\
&= \{(x,y) : (x,y) \in [(C \cup D) \times B] \cap f\} \\
&= \{(x,y) : (x,y) \in [\ (C \times B) \cup (D \times B)\ ] \cap f\} \text{ (theorem 4.7)} \\
&= \{(x,y) : (x,y) \in [\ (C \times B) \cap f] \ \cup\ [\ (D \times B) \cap f\ ]\} \text{ (theorem 3.9)} \\
&= \{(x,y) : (x,y) \in (C \times B) \cap f\} \\
&\qquad \cup \{(x,y) : (x,y) \in (D \times B) \cap f\} \\
&= \{(x,y) : (x,y) \in (C \times B) \text{ and } (x,y) \in \cap f\} \\
&\qquad \cup \{(x,y) : (x,y) \in (D \times B) \text{ and } (x,y) \in \cap f\} \\
&= f|_C \cup f|_D
\end{aligned}$$

9.4 Equal functions.

We know that two sets are equal provided both sets contain the same elements. Since functions are defined as being sets of ordered pairs, then we can establish equality of two functions by comparing the elements of the sets they represent. If the function f and g contain the same ordered pairs, then we can write $f = g$.

Theorem 9.4 Two functions $f : A \to B$ and $g : A \to B$ are equal if and only if $f(x) = g(x)$ for all $x \in A$.

Proof:

$$\begin{aligned}
f = g \quad &\Leftrightarrow \quad \text{For any } x \in A,\ (x,y) \in f \text{ if and only if } (x,y) \in g \\
&\Leftrightarrow \quad \text{For any } x \in A,\ f(x) = y \text{ and } g(x) = y \\
&\Leftrightarrow \quad \text{For any } x \in A,\ f(x) = g(x)
\end{aligned}$$

9.5 Some particular types of functions.

We present a few elementary functions with particular properties often encountered in various fields of mathematics.

Definition 9.5 Let $f : A \to B$ be a function.

a) We say that "f maps A *onto* B" if $\operatorname{im} f = B$. We often use the expression "$f : A \to B$ is *surjective*" instead of the word *onto*.

b) We say that "f maps A *one-to-one into* B" if, whenever $f(x) = f(y)$, then $x = y$. We often use the expression "$f : A \to B$ is *injective*" instead of the words *one-to-one into* B.

c) If the function $f : A \to B$ is both one-to-one and onto B we then say that f is "one-to-one and onto". Another way of conveying this is to say that f is *bijective*, or f is a *bijection*. So "injective + surjective $\Leftrightarrow$ bijective".

d) Two classes (or sets) A and B for which there exists some bijective function $f : A \to B$ are said to be in *one-to-one correspondence*.

9.6 A few examples of simple functions.

a) *The constant function.* Let A and B be two sets and suppose $b \in B$. Define the function $f : A \to B$ as follows:

$$f(x) = b \text{ for all } x \in A$$

This function maps all elements of A to the same element of B. We call this a *constant function*.

- If $B \neq \{b\}$, then f is not "onto B" or surjective; it is just "into" B.
- However, if we were to write, $f : A \to \{b\}$ of course we could say that f is surjective.
- If A has only one element, say $A = \{a\}$, then $f : \{a\} \to \{b\}$ is a constant function which is bijective.

b) *The characteristic function.* Let C be the 2 element set $C = \{\varnothing, \{\varnothing\}\}$.[1] Let A be a set and D be a non-empty subset of A. We define a function denoted by $\chi_D : A \to C$ as follows[2]:

[1] We should justify that C is a set: Since $\varnothing$ is a subclass of any set, then by the Axiom A4 (Axiom of subset), $\varnothing$ is a set. Also $\{\varnothing\}$ is a set since $\{\varnothing, \varnothing\}$ is a set (by the axiom of pair). Again by the axiom of pair, $\{\varnothing, \{\varnothing\}\}$ is a set.

[2] The Greek letter χ is pronounced "kie" (like the word "pie").

$$\chi_D(x) = \begin{cases} \varnothing & \text{if} \quad x \notin D \\ \{\varnothing\} & \text{if} \quad x \in D \end{cases}$$

This is called the *characteristic function of D in A*.

- We see that χ_D is "onto C".
- The characteristic function is constant on D mapping all elements of D to the single element $\{\varnothing\}$. It is constant on $A - D$ mapping all elements of $A - D$ to the single element $\varnothing$.
- We can write

$$\begin{aligned} \chi_D &= \{(x, \varnothing) : x \in A - D\} \cup \{(x, \{\varnothing\}) : x \in D\} \\ &= (\chi_D)|_{A-D} \cup (\chi_D)|_D \end{aligned}$$

c) Recall that in theorem 4.9, we showed that the elements of the two classes $A \times (B \times C)$ could be matched one-to-one with the elements of $(A \times B) \times C$. The proof of this theorem shows that the function $f : A \times (B \times C) \to (A \times B) \times C$ defined as $f((a, (b, c))) = (a, (b, c))$ is a bijection between these two classes.

9.7 Class functions.

A "function" was formally defined as being a particular kind of subset of the Cartesian product of two sets. Suppose that X and Y are classes (possibly proper) and f is a subclass of $X \times Y$ which satisfies the conditions given in definition 9.1. In order to distinguish f from the notion of "function" as presented in definition 9.1, we will refer to f as a *class function*, keeping in mind that f may be a proper class.

Concepts review:

1. What is the definition of a function f from a set A to a set B?
2. Is it acceptable to view a function f from a set A to a set B as a set of ordered pairs?
3. Given a function f from a set A to a set B, what do each of the sets $\operatorname{dom} f$, $\operatorname{codom} f$, $\operatorname{im} f$ and $\operatorname{ran} f$ represent?
4. Given a function f from a set A to a set B and $y \in \operatorname{im} f$, what is the *preimage* or *inverse* of y?
5. Given a function f from a set A to a set B and a set D such that $D \subseteq A$ what does the symbol $f|_D$ mean?
6. Given a function f from a set A to a set B where $A = C \cup D$ is it true that $f = f|_C \cup f|_D$?

7. Given the two functions $f : A \to B$ and $g : A \to B$ what does it mean to say that the functions f and g are equal? How can we show that $f = g$?

8. Given a function f from a set A to a set B what does it mean to say that f is *onto* B?

9. Given a function f from a set A to a set B what does it mean to say that f is *one-to-one into* B?

10. Given a function f from a set A to a set B what does it mean to say that f is *injective*?

11. Given a function f from a set A to a set B what does it mean to say that f is *surjective*?

12. Given a function f from a set A to a set B what does it mean to say that f is *one-to-one and onto* B?

13. Given a function f from a set A to a set B what does it mean to say that f is *bijective* (or f is a bijection)?

14. Given a function f from a set A to a set B is it safe to say that if f is both injective and surjective, then it is bijective?

15. Given two sets A and B what does it mean to say that there is a one-to-one correspondence between A and B?

16. If D is a subset of A what is the characteristic function $\chi|_D$ of D in A? Describe $\chi|_D$ as a set of ordered pairs.

EXERCISES

A.

1. Suppose A is a set. Show that the set $\{(x, x) : x \in A\}$ is a function.
2. Let $f : A \to B$ be a function. Show that if $g \subseteq f$ and g is non-empty, then g is a function.
3. Suppose $f : A \to B$ and $g : A \to B$ are two functions each of which has the set A as domain. If $f \subseteq g$ show that $f = g$.
4. If C is as set, let $f : C \to I_C$ be defined as $f(a) = (a, a)$ for all $a \in C$
   a) Show that f satisfies the definition of a function.
   b) Show that f is one-to-one and onto and therefore is bijective

B.

5. Let D and E be two sets such that $D \cap E = \varnothing$. Let $g : D \to B$ and $h : E \to B$ be two functions. Let $f = g \cup h$.
   a) Show that $f : D \cup E \to B$ is a function.
   b) Show that $f|_D = g$ and $f|_E = h$.
6. Let $f : A \to B$ and $g : C \to D$ be two functions. We define $(f \times g) : A \times C \to B \times D$ as follows:
$$(f \times g)((x, y)) = (f(x), g(x)) \text{ for all } (x, y) \in A \times C$$
   a) Show that $(f \times g) : A \times C \to B \times D$ is a function.
   b) Show that if $f : A \to B$ and $g : C \to D$ are bijective, then $(f \times g) : A \times C \to B \times D$ is bijective.
7. Let $f : A \to B$ and $g : C \to D$ be two bijective functions where $A \cap C = \varnothing$ and $B \cap D = \varnothing$. Let $h : A \cup C \to B \cup D$ be defined as follows:
$$h(x) = \begin{cases} f(x) & \text{for all} \quad x \in A \\ g(x) & \text{for all} \quad x \in C \end{cases}$$
   a) Show that $h : A \cup C \to B \cup D$ is a function.
   b) Show that if $f : A \to B$ and $g : C \to D$ are bijective, then $h : A \cup C \to B \cup D$ is bijective.
8. Let S and T be sets and f be a function on $S \times T$ defined as: $f((x, y)) = x$ for all $(x, y) \in S \times T$.
   a) Verify that f is indeed a function.
   b) Describe the image of f.
   c) Verify whether f is one-to-one or not. If it is prove it, if it isn't show why not.
9. Let A be a set and $D \subseteq A$. Recall that $\chi|_D$ is the characteristic function mapping x to $\{\varnothing\}$ if $x \in D$ and x to $\varnothing$ if $x \notin D$. Show that $\operatorname{im} \chi|_D$ is a set.

C.

10. Let $f : A \to B$ and $g : C \to D$ be two bijective functions. Let $h : A \cup C \to B \cup D$ be defined as follows:
$$h(x) = \begin{cases} f(x) & \text{for all} \quad x \in A \\ g(x) & \text{for all} \quad x \in C \end{cases}$$
    a) Is h necessarily a function? If it isn't, give an example illustrating this.
    b) If h is a function is h necessarily bijective? If it isn't give an example illustrating this.
11. Let $f : A \to B$ be a function. Show that if $g \subseteq f$, then there exists some subset C of A such that $f|_C = g$.
12. Is $\varnothing$ a one-to-one function? Explain.

10 / Operations on functions.

Summary. *In this section we define the composition, $g \circ f$, of two functions $f : A \to B$ and $g : B \to C$. We view "composition of functions" as an operation "$\circ$" on two functions f and g. From this perspective we then discuss the main properties of composition of functions (such as non-commutativity and associativity). It is in this particular context that we describe the identity function and the inverse of a function. We also define the concept of "invertible function".*

10.1 Composition of functions: a set-theoretic definition.

Suppose $f : A \to B$ and $g : B \to C$ are two functions. A noticeable fact about these two functions is that the domain of the function g is the codomain of f. So for $x \in A$ we have $f(x) \in \operatorname{dom} g$. For such an element $x \in A$, the expression $g(f(x))$ is well-defined. This allows us to construct the set:

$$h = \{(x, y) : x \in A, y = g(f(x)) \in \operatorname{im} g\} \subseteq A \times C$$

By the *Axiom of construction* A2, h is a well-defined subset of $A \times C$. With these thoughts in mind, we formally define this notion of "composition of two functions".

Definition 10.1 Suppose $f : A \to B$ and $g : B \to C$ are two functions such that the codomain of the function f is the domain of the function g. Let

$$h = \{(x, z) \in A \times C : y = f(x) \text{ and } z = g(y) = g(f(x)) \}$$

Thus, $(x, z) \in h$ if and only if $(x, z) = (x, g(f(x))$. We will call h the *composition of g and f*, and denote it by $g \circ f$ where $(g \circ f)(x) = g(f(x))$.

Given the functions $f : A \to B$ and $g : B \to C$ and seeing that $g \circ f \subseteq A \times C$ we naturally suspect that $g \circ f : A \to C$ is a function. We will, of course, have to make sure that this is the case.

Theorem 10.2 Let $f : A \to B$ and $g : B \to C$ be two functions such that the codomain of the function f is the domain of the function g. Then the composition of g and f, $(g \circ f) : A \to C$, is a function.

Proof:

Given: $f : A \to B$ and $g : B \to C$ are two functions.
We are required to show that $g \circ f$ is a function.

1) By definition of $h = g \circ f$,
$$h = g \circ f \subseteq A \times C$$

2) Let $x \in A$. We are required to show that $h(x) \in C$.
$$\begin{aligned} x \in A &\Rightarrow f(x) \in B \quad \text{(Since } \operatorname{im} f \subseteq B\text{)} \\ &\Rightarrow f(x) \in \operatorname{dom} g \quad \text{(Since } \operatorname{im} f \subseteq \operatorname{dom} g\text{)} \\ &\Rightarrow g(f(x)) \in C \quad \text{(Since } g : B \to C \text{ is a function)} \\ &\Rightarrow h(x) \in C \quad \text{(Since } g(f(x)) = h(x)\text{)} \end{aligned}$$
Thus, $x \in A \Rightarrow h(x) = (g \circ f)(x) \in C$.

3) Suppose $h(a) = g(f(a)) \neq g(f(b)) = h(b)$.
$$\begin{aligned} h(a) = g(f(a)) &\neq g(f(b)) = h(b) \\ \Rightarrow f(a) &\neq f(b) \\ \Rightarrow a &\neq b \end{aligned}$$
The three conditions being satisfied, we conclude that $g \circ f$ is a function.

10.2 Composition of functions viewed as an operation on functions.

Given two functions $f : A \to B$ and $g : B \to C$, we have shown that we can associate with this pair of functions another function $h = g \circ f$ called "the composition of f and g".

This suggests that"composition", denoted by the symbol, $\circ$, can be viewed as an operation on pairs of functions, just like $\times$, $\cup$, $\cap$ and $+$. We wonder:

1) Can we compose any pair of functions?

   The definition of composition of functions makes it quite clear that we can't compose certain pairs of functions. For the composition $g \circ f$ of two functions f and g to be well-defined the image of f must be in the domain of g.

2) Is the composition of functions commutative?

   Again, it is clear from the definition of composition of functions that we can't commute certain pairs of functions with respect to composition.

   - Suppose for example that we can compose the functions $f : A \to B$ and $g : C \to D$ in this order: $f \circ g$.
   - Then, this means that $\operatorname{dom} f \subseteq \operatorname{im} g$. But if $\operatorname{dom} g$ is not contained in the $\operatorname{im} f$, the expression, $g \circ f$, is not meaningful; so we cannot commute the pair f and g.

3) Is the composition of functions associative?

Yes. The following theorem shows that the composition of functions satisfies the associative property.

Theorem 10.3 Let $f : A \to B$, $g : B \to C$ and $h : C \to D$ be three functions. Then $h \circ (g \circ f) = (h \circ g) \circ f$.

Proof: It suffices to show that $h \circ (g \circ f) \subseteq (h \circ g) \circ f$ and $(h \circ g) \circ f \subseteq h \circ (g \circ f)$.

Proof of $h \circ (g \circ f) \subseteq (h \circ g) \circ f$:

$$\begin{aligned}
(x, y) \in h \circ (g \circ f) &\Rightarrow [h \circ (g \circ f)](x) = y \\
&\Rightarrow h(g(f(x)) = y \\
&\Rightarrow h(z) = y \text{ for some } z \in \text{ dom}\, h \subseteq C \\
&\Rightarrow g(f(x)) = z \in C \\
&\Rightarrow g(u) = z \text{ for some } u \in \text{ dom}\, g \subseteq B \\
&\Rightarrow f(x) = u \in B
\end{aligned}$$

$$\begin{aligned}
((h \circ g) \circ f)(x) &= (h \circ g)(f(x)) = (h \circ g)(u) \\
&= h(g(u)) \\
&= h(z) \\
&= y \\
((h \circ g) \circ f)(x) = y &\Rightarrow (x, y) \in (h \circ g) \circ f
\end{aligned}$$

We have shown that $h \circ (g \circ f) \subseteq (h \circ g) \circ f$.
To show $(h \circ g) \circ f \subseteq h \circ (g \circ f)$ we proceed similarly. This is left as an exercise.

An identity element for the operation "$\circ$". Is there a function I such that any function g composed with I will leave that function unchanged, i.e., $g \circ I = I \circ g = g$?

An obvious candidate for an identity element with respect to composition is the identity relation, $I : \mathscr{U} \to \mathscr{U}$.[1] It is defined as, $I(x) = x$, for all $x \in \mathscr{U}$. If A is a set, then I_A will denote the restriction of I to A and so $I_A : A \to A$ is defined as $I_A(x) = x$, for all $x \in A$. The following theorem confirms that this function behaves as expected.

Theorem 10.4 Let $f : A \to B$. Then $I_B \circ f = f$ and $f \circ I_A = f$.

[1] Recall that $\mathscr{U}$ denotes the class of all elements and is called the *Universal class*.

Proof:

Given: $f : A \to B$ and $I_B(x) = x$ for all $x \in B$.

$$\begin{aligned}(x, y) \in I_B \circ f &\Leftrightarrow (x, z) \in f \text{ and } (z, y) \in I_B \text{ for some } z \in B \\ &\Leftrightarrow (x, z) \in f \text{ and } z = y \text{ for some } z \in B \\ &\Leftrightarrow (x, y) \in f\end{aligned}$$

Thus, $I_B \circ f = f$.

The proof of $f \circ I_A = f$ is similar. It is left to the reader.

Inverses with respect to "$\circ$" and the identity I. Once an identity element I has been identified, one naturally wonders whether certain functions f have an inverse g with respect to this identity so that $g \circ f = I$?

We will show that only certain functions have an "inverse" with respect to "composition".

Definition 10.5 Let $f : A \to B$. If $g : B \to A$ is a function satisfying $g \circ f = I_A$, then we will call g an "*inverse of* f"; we represent g as f^{-1}.

Theorem 10.6 Let $f : A \to B$ be a one-to-one onto function.

a) An inverse function, $f^{-1} : B \to A$, of f exists.

b) The function, f^{-1}, is one-to-one and onto.

c) The function, $f : A \to B$, is the inverse of $f^{-1} : B \to A$. That is, $(f^{-1})^{-1} = f$.

d) The inverse function, f^{-1}, of f is unique.

Proof:

What we are given for parts a) to d): That $f : A \to B$ is a one-to-one onto function.

a) What we are required to show: That there exists a function g such that $g(f(x)) = x$. This function g will be f^{-1}.

Define $g : B \to A$ as follows: $g(x) = y$ only if $f(y) = x$. We claim that $g : B \to A$ is a well-defined function:

- Let $x \in B$. Since f is onto B, then there exists $y \in A$ such that $f(y) = x$. Thus, $\operatorname{dom} g = B$. Suppose now that (x, y) and (x, z) are in g. Then y and z are in A such that $f(y) = x$ and $f(z) = x$. Since f is one-to-one, then y and z must be the same element. Thus, $g : B \to A$ is a well-defined function.

Then g satisfies the definition of an inverse of f, $g(f(x)) = x$, and so $g \circ f = I_A$. Thus, $g = f^{-1}$.

b) What we are required to show: That f^{-1} is one-to-one:
Suppose (x, y) and (z, y) both belong to $f^{-1} : B \to A$. Then $f(y) = x$ and $f(y) = z$. Since $f : A \to B$ is a function, then $x = z$. Thus, f^{-1} is one-to-one on its domain as claimed.

c) What we are required to show: If $f^{-1} : B \to A$ and $f : A \to B$, then $f \circ f^{-1} = I_B$:
We are assuming that $A = \operatorname{im} f^{-1}$ and $A = \operatorname{dom} f$. Suppose $f \circ f^{-1}(x) = z$. Then there is some $y \in A$ such that $(x, y) \in f^{-1}$ and $(y, z) \in f$. But since f^{-1} is the inverse of f, $(x, y) \in f^{-1}$ implies $(y, x) \in f$. Since both (y, x) and (y, z) belong to f, then $z = x$. Then $f \circ f^{-1}(x) = x$ for all $x \in B$. We conclude that f is an inverse of f^{-1}.

d) What we are required to show: If $h : B \to A$ is a function satisfying $h \circ f = I_A$, then h can only be f^{-1}.

$$\begin{aligned} h \circ f = I_A = f^{-1} \circ f \quad &\Rightarrow \quad (h \circ f) \circ f^{-1} = (f^{-1} \circ f) \circ f^{-1} \\ &\Rightarrow \quad h \circ (f \circ f^{-1}) = f^{-1} \circ (f \circ f^{-1}) \quad \text{(Associativity)} \\ &\Rightarrow \quad h \circ I_B = f^{-1} \circ I_B \\ &\Rightarrow \quad h = f^{-1} \end{aligned}$$

This theorem confirms that:

- If f is one-to-one on its domain, then f has an inverse f^{-1}
- This function f^{-1} is unique and is one-to-one.

Conversely, if $f : A \to B$ has an inverse f^{-1}, then for $a, b \in A$

$$\begin{aligned} f(a) \neq f(b) \quad &\Rightarrow \quad f^{-1}(f(a)) = a \text{ and } f^{-1}(f(b)) = b \\ &\Rightarrow \quad a \neq b \quad \text{(Otherwise } f \text{ maps } a = b \text{ to distinct points } f(a) \text{ and } f(b)\text{.)} \end{aligned}$$

so the function f must be one-to-one.

We have shown that "f has an inverse if and only if f is one-to-one".

Definition 10.7 *Invertible functions*, $f : A \to B$, on A are precisely the one-to-one functions on A.

10.3 The inverse of the composition of functions.

Suppose we are given the two functions, $f : A \to B$ and $g : B \to C$, both of which are one-to-one and onto functions. The following theorem shows us how to proceed when we wish to find the inverse of their composition, $(g \circ f)^{-1}$.

Theorem 10.8 Let $f : A \to B$ and $g : B \to C$ be two one-to-one and onto functions.

a) Then the function, $(g \circ f) : A \to C$, is also one-to-one and onto C.

b) Then the inverse of $g \circ f$, is $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$

Proof: The proofs of these statements are left as an exercise.

10.4 Comparing the inverse of a function to the inverse of a relation.

Recall that "inverses" were discussed before we defined the notion of a *function* (see definition 5.4). We referred to inverses while studying *relations* and some of their properties. We pause to compare the inverse of a relation to the inverse of a function so that we can better see how they are similar and how they differ.

- *Relations.* Given sets A and B, any subset of $\{(a, b) : (a, b) \in A \times B\}$ is a relation. No other conditions are specified. For any relation $R \subseteq A \times B$ we can construct another relation, R^{-1}, called its inverse. This inverse is defined as: $R^{-1} = \{(y, x) : (x, y) \in R\}$.
- *Functions.* A function $f : A \to B$ is a set of ordered pairs $\{(a, b) : a \in A, b = f(a) \in B\}$ and so is a relation. But those relations we call "functions" must satisfy the condition "$[(a, b) = (a, c)] \Rightarrow [b = c]$". We have declared that a function is an *invertible function* only if f is one-to-one. But when viewing f as a relation (a subset of $A \times B$) we can speak of its "inverse", as a relation, even though it is not one-to-one. That is, if $f = \{(x, y) : y = f(x)\}$ the inverse, f^{-1}, of f is

$$f^{-1} = \{(y, x) : y = f(x), x \in \operatorname{dom} f\} = \{(y, x) : (x, y) \in f\}$$

There is no contradiction here. But we should be more specific by saying: If f is not one-to-one we can speak of its inverse f^{-1}, with the caveat that f^{-1} cannot be referred to as a "function". So, when we say that "f is invertible if and only if it is one-to-one", we actually mean "the inverse f^{-1} of the function, f, is a function if and only if this function f is one-to-one".

We consider the following example. Let A be a set, $\{a, b\} \subseteq A$ where $a \neq b$ and U be a non-empty subset of A. Consider the function $f : A \to A$ defined as follows:

$$f(x) = \begin{cases} a & \text{if} \quad x \in U \\ b & \text{if} \quad x \notin U \end{cases}$$

Then f can be described as

$$f = [U \times \{a\}] \quad \cup \quad [(A - U) \times \{b\}]$$

Its inverse

$$f^{-1} = [\{a\} \times U] \quad \cup \quad [\{b\} \times (A - U)]$$

can only be referred to as a relation on A, unless of course both U and $A - U$ are singleton sets.

Concepts review:

1. Given two functions $f : A \to B$ and $g : B \to C$ what does the expression "$g \circ f$" mean? Under what conditions does this expression make sense?
2. Is the composition of functions commutative? Are there any pairs of functions which always "commute" with each other?
3. Under what condition(s) is the composition of functions associative?
4. Which function plays the role of the identity with respect to "$\circ$"?
5. What does it mean to say that a function f is "invertible"?
6. Under what condition(s) is a function invertible with respect to "$\circ$"?
7. If a function h can be expressed as $h = g \circ f$ where both f and g are invertible, is h invertible? If so, how can we express h^{-1}?
8. If f is not one-to-one on its domain what interpretation can we give to the expression f^{-1}

EXERCISES

A. 1. Suppose $f : A \to B$ and $g : B \to C$ are functions and $D \subseteq A$. Prove that $(g \circ f)|_D = g \circ (f|_D)$.

B. 2. Let $f : A \to B$ and $g : B \to C$ be two functions.
   a) Prove that if $(g \circ f) : A \to C$ is one-to-one, then $f : A \to B$ is one-to-one.
   b) Prove that if $(g \circ f) : A \to C$ is onto C, then g is onto C.

3. Let $g : B \to C$ and $h : B \to C$ be two functions. Suppose $g \circ f = h \circ f$ for every function $f : A \to B$. Prove that $g = h$.

4. Let $g : A \to B$ and $h : A \to B$ be two functions and let C be a set with more than one element. Prove that if $f \circ g = f \circ h$ for every function $f : B \to C$, then $g = h$.

C. 8. Prove the statements a) and b) of theorem 10.8.

---

11 / Images and preimages of sets.

Summary. *Suppose we are given a function $f : A \to B$. In this section we "elevate" this function so that it acts on $\mathscr{P}(A)$, mapping its elements to elements of $\mathscr{P}(B)$ according to the rule determined by f. This provides a mechanism by which we can study the image of sets under the function f. We also define the set-valued inverse of a function, $f^{\leftarrow}$, and provide some examples. We also show how such functions act on unions and intersections of sets.*

11.1 Functions mapping sets to sets.

The domain of all functions in this section are hypothesized to be sets. Given a function, $f : A \to B$, it is sometimes useful to see what effect the function has on subsets of the domain rather than simply on its elements. To study the action of functions on sets we will introduce some special notation.

Definition 11.1 Suppose f is a relation with the set A as domain and the set B as range. If S is a subset of A, then we will represent the *image* of the set S under f as

$$f[S] = \{y \in B : (x, y) \in f \text{ and } x \in S\}$$

If $U \subseteq B$, we will refer to the set

$$f^{\leftarrow}[U] = \{x \in A : (x, y) \in f \text{ and } y \in U\}$$

as the *preimage* of the set U under f.

Remarks. What is new in this definition?

- First observe that the expression $f[S]$ − notice the square brackets − is the image of S under f. So the square brackets are not there just as a matter of style. They have meaning. The function, $f[\ ]$, associates elements of $\mathscr{P}(A)$ to elements of $\mathscr{P}(B)$ where f is predefined either as a relation or, more specifically, as a function.
- The symbol, "$f^{\leftarrow}$", and the words, *preimage of a set*, are new. If f is a function, then $f^{\leftarrow}[U]$ is simply the image of U under the relation f^{-1}. Again notice the square brackets which means we are associating sets to sets, an association governed by $f^{\leftarrow}$.[1] (If f is a function and $x \in \operatorname{im} f$, in some branches of mathematics

[1]Note that use of the notation $f^{\leftarrow}[U]$ is not universal. It is introduced here to avoid confusing the preimage of an element, $f^{-1}(x)$, normally used with one-to-one functions, with the preimage, $f^{-1}[U]$, of a set U. A general topologist might refer to "$B = f^{\leftarrow}[U]$" by saying that "f pulls back the set U to the set B".

$f^{\leftarrow}[\{x\}]$ is referred to as the *fiber of x under the function f.*)

Examples.

1) Let $B = \{a, b, c\}$ and U be a non-empty proper subset of A. Consider the function $f : A \to B$ defined as follows:

$$f(x) = \begin{cases} a & \text{if} \quad x \in U \\ b & \text{if} \quad x \notin U \end{cases}$$

Then

$$\begin{aligned} f[U] = \{a\} \quad &\text{and} \quad f^{\leftarrow}[\{a\}] = U \\ f[A - U] = \{b\} \quad &\text{and} \quad f^{\leftarrow}[\{b\}] = A - U \\ f[\varnothing] = \{c\} \quad &\text{and} \quad f^{\leftarrow}[\{c\}] = \varnothing \end{aligned}$$

We see that the preimage of $\{c\}$ is empty since f maps no elements of A to c. This is another way of saying that c is not in the range of f, or, $c \notin f[A]$.

2) Suppose $f : A \to \mathscr{P}(A)$ is defined as $f(x) = \{x\}$. If $\{a\} \neq \{b\}$, then $a \neq b$, hence, in this case, f is a function. If $B \subseteq A$, then $f[B] = \{\{x\} : x \in B\} \subseteq \mathscr{P}(A)$. In this case f maps the element $B \in \mathscr{P}(A)$ to the element $f[B] \in \mathscr{P}(\mathscr{P}(A))$. For $x, y \in A$, $f^{\leftarrow}[\{\{x\}, \{y\}\}] = \{x, y\}$.

3) Let $A = \{a, b, c, d, e, k, h\}$ and $B = \{u, v, w, z, s\}$ and let $D = \{e, k, h\}$ and $E = \{c, d\}$.
We define $f = \{(a, u), (b, u), (c, u), (d, v), (e, v), (k, z), (h, s)\}$. We describe the function f via images and preimages.

$$\begin{aligned} f^{\leftarrow}[\{u\}] &= \{a, b, c\} \\ f^{\leftarrow}[\{v\}] &= \{d, e\} \\ f^{\leftarrow}[\{z\}] &= \{k\} \\ f|_D[D] = f|_D[\{e, k, h\}] &= \{v, z, s\} \\ (f|_D)^{\leftarrow}[\{v\}] &= \{e\} \\ (f|_E)[E] = (f|_E)[\{c, d\}] &= \{u, v\} \\ (f|_E)^{\leftarrow}[\{u\}] &= \{c\} \end{aligned}$$

11.2 Images and preimages of unions and intersections of sets.

Since we have defined how a function $f : A \to B$ can be elevated to $f : \mathscr{P}(A) \to \mathscr{P}(B)$, mapping sets to sets, we can determine how such functions behave when f acts on unions and intersections of sets (or classes). We present these few properties in the form of a theorem. While reading through these properties, we will

see that $f^{\leftarrow}$ always "respects"[1] unions and intersections. The function f will be seen to "respect" unions; but f will "respect" intersections only in certain circumstances.

Theorem 11.2 Let $f : A \to B$ be a *function* mapping the set A to the set B. Let $\mathscr{A}$ be *a set* of subsets of A and $\mathscr{B}$ be *a set* of subsets of B. Let $D \subseteq A$ and $E \subseteq B$. *Then:*

a) $f\left[\bigcup_{S\in\mathscr{A}} S\right] = \bigcup_{S\in\mathscr{A}} f[S]$

b) $f\left[\bigcap_{S\in\mathscr{A}} S\right] \subseteq \bigcap_{S\in\mathscr{A}} f[S]$ where equality holds true only if f is one-to-one.

c) $f[A - D] \subseteq B - f[D]$. Equality holds true only if f is one-to-one and onto B.[2]

d) $f^{\leftarrow}\left[\bigcup_{S\in\mathscr{B}} S\right] = \bigcup_{S\in\mathscr{B}} f^{\leftarrow}[S]$

e) $f^{\leftarrow}\left[\bigcap_{S\in\mathscr{B}} S\right] = \bigcap_{S\in\mathscr{B}} f^{\leftarrow}[S]$

f) $f^{\leftarrow}[B - E] = A - f^{\leftarrow}[E]$

Proof:

a)
$$\begin{aligned} x \in f\left[\bigcup_{S\in\mathscr{A}} S\right] &\Leftrightarrow x = f(y) \text{ for some } y \in \bigcup_{S\in\mathscr{A}} S \\ &\Leftrightarrow x = f(y) \text{ for some } y \text{ in some } S \in \mathscr{A} \\ &\Leftrightarrow x = f(y) \in f[S] \text{ for some } S \in \mathscr{A} \\ &\Leftrightarrow x \in \bigcup_{S\in\mathscr{A}} f[S] \end{aligned}$$

b) It will be helpful to first prove this statement for the intersection of only two sets U and V. The use of a Venn diagram will also help visualize what is happening.

So we first prove the statement: $f[U \cap V] \subseteq f[U] \cap f[V]$ with equality only if f is one-to-one on $U \cup V$.

Case 1: We consider the case where $U \cap V = \varnothing$.
Then $f[U \cap V] = \varnothing \subseteq f[U] \cap f[V]$. So the statement holds true.

Case 2: We now consider the case where $U \cap V \neq \varnothing$.

$$\begin{aligned} x \in f[U \cap V] &\Leftrightarrow x = f(y) \text{ for some } y \in U \cap V \\ &\Leftrightarrow x = f(y) \text{ for some } y \text{ contained in both } U \text{ and } V \\ &\Rightarrow x = f(y) \in f[U] \text{ and } f[V] \\ &\Leftrightarrow x \in f[U] \cap f[V] \end{aligned}$$

[1] We say that f *respects* unions if it is always true that $f[A \cup B] = f[A] \cup f[B]$. Similarly, f *respects* intersections if it is always true that $f[A \cap B] = f[A] \cap f[B]$.

[2] Remember that $A - B = A \cap B'$ equals A intersection the complement of B.

We now show that if f is one-to-one on $U \cup V$, then $f[U] \cap f[V] \subseteq f[U \cap V]$ and so equality holds true.

- Suppose $x = f(y) \in f[U] \cap f[V]$. Then there exists $u \in U$ and $v \in V$ such that $f(u) = f(v) = f(y)$. Since f is one-to-one, $u = v = y$. This implies $y \in U \cap V$. Hence, $f[U \cap V] = f[U] \cap f[V]$.

The proof of the general statement is left as an exercise.

c) Proof is left as an exercise.

d)
$$\begin{aligned} x \in f^{\leftarrow}\left[\bigcup_{S\in\mathscr{B}} S\right] &\Leftrightarrow x = f(y) \text{ for some } y \in \bigcup_{S\in\mathscr{B}} S \quad \text{(By definition of } f^{\leftarrow}.) \\ &\Leftrightarrow x = f(y) \text{ for some } y \text{ in some } S \in \mathscr{B} \\ &\Leftrightarrow x \in f^{\leftarrow}[\{y\}] \subseteq f^{\leftarrow}[S] \text{ for some } S \in \mathscr{B} \\ &\Leftrightarrow x \in \bigcup_{S\in\mathscr{B}} f^{\leftarrow}[S] \end{aligned}$$

Thus, $f^{\leftarrow}\left(\bigcup_{S\in\mathscr{B}} S\right) = \bigcup_{S\in\mathscr{B}} f^{\leftarrow}(S)$.

e) Proof is left as an exercise.

f) Proof is left as an exercise.

Concepts review:

1. Given a function $f : A \to B$ and $S \subseteq A$, what does the expression $f[S]$ mean?
2. Given a function $f : A \to B$ and $S \subseteq B$ what does the expression $f^{\leftarrow}[S]$ mean?
3. What is the *preimage* of a set S under a function f?
4. Given a function $f : A \to B$ and $x \in B - \operatorname{im} f$, what is $f^{\leftarrow}(\{x\})$?
5. Under what conditions does f respect unions?
6. Under what conditions does f respect intersections?
7. Under what conditions does $f^{\leftarrow}$ respect unions?
8. Under what conditions does $f^{\leftarrow}$ respect intersections?

EXERCISES

A.

1. Give an example of a function $f : A \to B$ where A contains two unequal non-empty subsets D and E satisfying $f[D] = f[E]$.
2. Let $f : A \to B$ be a function where A and B are sets.
   a) Prove that if D and E are equal subsets of A, then $f[D] = f[E]$.
   b) Prove that if U and V are equal subsets of B, then $f^{\leftarrow}[U] = f^{\leftarrow}[V]$.
3. Suppose $f : A \to B$ where A and B are sets.
   a) Show that for any subset D of A, $D \subseteq f^{\leftarrow}(f[D])$.
   b) Give an example where $D \neq f^{\leftarrow}(f[D])$.
   b) Show that for any subset E of B, $f[f^{\leftarrow}[E]] \subseteq E$.
   d) Is it necessarily true that $f[f^{\leftarrow}[E]] = E$?
   e) Prove that if f is one-to-one on A, then, for any subset D of A, $D = f^{\leftarrow}[f[D]]$.
   f) Prove that if f is onto B, then, for any subset E of B, $f[f^{\leftarrow}[E]] = E$.

B.

4. Let S and T be sets and $f : S \times T \to S$ be a function on $S \times T$ defined as: $f((x, y)) = x$ for all $(x, y) \in S \times T$.
   a) If $u \in \text{im } f$ what is $f^{\leftarrow}[\{u\}]$?
   b) If U is a non-empty subset of S what is $f^{\leftarrow}[U]$?
5. Prove the general case of part b) of theorem 11.2.
6. Prove part c) of theorem 11.2.
7. Prove part e) of theorem 11.2.
8. Prove part f) of theorem 11.2.

C.

9. Let $f : A \to B$ be a function mapping the set A to the set B. Prove that if $D \subseteq A$, then
$$f\left[f^{\leftarrow}[f[D]]\right] = f[D]$$
10. Let $f : A \to B$ be a function which maps the set A onto the set B. Prove that
$$f^{\leftarrow}[B] = \bigcup_{x \in B} f^{\leftarrow}[\{x\}]$$
11. Let $f : A \to B$ be a function which maps the set A onto the set B. Prove that if x and y are distinct elements of B, then
$$f^{\leftarrow}[\{x\}] \cap f^{\leftarrow}[\{y\}] = \varnothing$$
12. Let $f : A \to B$ be a function which maps the set A onto the set B. Prove that the set of sets
$$\mathscr{S} = \{f^{\leftarrow}[\{x\}] : x \in B\}$$
forms a partition of the set A.

12 / Equivalence relations induced by functions.

Summary. *In this section we show how a function $f : S \to T$ partitions its domain S. This partition induces an equivalence relation R_f on S which in turn leads to the quotient set S/R_f. We then present a theorem which shows how any function can be expressed as the composition of two functions, neither of which is f itself or the identity function.*

12.1 Partitioning the domain of a function $f : A \to B$.

Suppose $f : A \to B$ is a function which maps a set A into a set B.

We claim that the set $\{f^{\leftarrow}[\{x\}] : x \in f[A]\} \subseteq \mathscr{P}(A)$ forms a *partition* of A.[1]

- Since every $x \in f[A]$ is in the image of f, $f^{\leftarrow}[\{x\}]$ is non-empty for all $x \in f[A]$.
- If $x \neq y$, $f^{\leftarrow}(\{x\}) \cap f^{\leftarrow}(\{y\}) = \varnothing$ otherwise an element $z \in f^{\leftarrow}(\{x\}) \cap f^{\leftarrow}(\{y\})$ would be mapped to distinct points, contradicting the fact that f is a function.
- Finally, the function $f^{\leftarrow}$ sends the image, $f[A]$, of A back to A, i.e., $f^{\leftarrow}[f[A]] = A$. So $A = f^{\leftarrow}[f[A]] = f^{\leftarrow}[\bigcup_{x\in f[A]}\{x\}] = \bigcup_{x\in f[A]} f^{\leftarrow}(\{x\})$. Hence, $\{f^{\leftarrow}[\{x\}] : x \in f[A]\}$ covers all of A.

The set. $\{f^{\leftarrow}(\{x\}) : x \in f[A]\}$, forms a pairwise disjoint set of sets which covers all of A. So this set partitions A, as claimed.

We have seen that the partition of a set is the quotient set of some equivalence relation, R, (see theorem 8.3). The equivalence relation, R, induced by a partition on a set A was defined as follows:

> *Two elements of the set A are related under R if and only if they belong to the same element of the partition induced by R.*

Then, given a function, $f : A \to B$, on A, we can declare that two elements a and b in A are related under a relation R_f if and only if they appear together in $f^{\leftarrow}[\{x\}]$ for some x in the image, $f[A]$, of A. So the set of subsets of A, $\{f^{\leftarrow}[\{x\}] : x \in f[A]\}$, is a quotient set of A induced by R_f. We formalize these thoughts in the following definition.

Definition 12.1 Let $f : A \to B$ be a function which maps a set A into a set B. We define the *equivalence relation, R_f, on A induced (or determined) by f* as follows:

[1] We remind the reader of the definition of a *partition* (also found at Definition 8.1). For a set S we say that a set of subsets $\mathscr{C} \subseteq \mathscr{P}(S)$ forms a *partition of S* if 1) $\bigcup_{A\in\mathscr{C}} A = S$, 2) If A and $B \in \mathscr{C}$ and $A \neq B$, then $A \cap B = \varnothing$, 3) $A \neq \varnothing$ for all $A \in \mathscr{C}$.

Two elements a and b are related under R_f if and only if $\{a, b\} \subseteq f^{\leftarrow}[\{x\}]$ for some x in $\operatorname{im} f$. The quotient set of A induced by R_f is then

$$A/R_f = \mathscr{A}_{R_f} = \{f^{\leftarrow}[\{x\}] : x \in f[A]\}$$

We will refer to A/R_f (or $\mathscr{A}_{R_f}$) as the *quotient set of A induced (or determined) by f.*

We illustrate this in a simple example. Let $U = \{a, \{a\}, \{a, \{a\}\}, \{\{a\}\}\}$ where a is a set. Consider the function $f : U \to U$ defined as follows:

$$f(x) = \begin{cases} a & \text{if} \quad a \in x \\ \{a\} & \text{if} \quad \{a\} \in x \text{ and } a \notin x \\ \{\{a\}\} & \text{if} \quad x = a \end{cases}$$

We see that f is a well-defined function on the set U. So the set

$$\mathscr{C} = \{f^{\leftarrow}[\{a\}], \;\; f^{\leftarrow}[\{\{a\}\}], \;\; f^{\leftarrow}[\{\{\{a\}\}\}] \}$$

partitions U in three pairwise disjoint non-empty sets. From this, we can define the equivalence relation R_f on U where $U/R_f = \mathscr{C}$.
We will list the elements in each set:

$$\begin{aligned} f^{\leftarrow}[\{a\}] &= \{ \;\; \{a\}, \; \{a, \{a\}\}\} \\ f^{\leftarrow}[\{\{a\}\}] &= \{ \; \{\{a\}\} \; \} \\ f^{\leftarrow}[\{\{\{a\}\}\}] &= \{a\} \end{aligned}$$

Observe that all elements of U are represented in the three sets above.

12.2 The canonical decomposition of a function.

Let $f : S \to T$ be a function mapping a set S into a set T. We have seen how this function, f, determines a new set: the *quotient set*, S/R_f, induced by f. For each $x \in S$ we let $S_x = f^{\leftarrow}[\{f(x)\}]$. Then

$$S/R_f = \{S_x : x \in S\}$$

is the quotient set induced by f. Note that the elements of S/R_f are subsets of S. We now show how the function f can be expressed as a composition of two other functions neither of which is the identity function.

The figure below illustrates how we will express the function f as a composition of two functions.

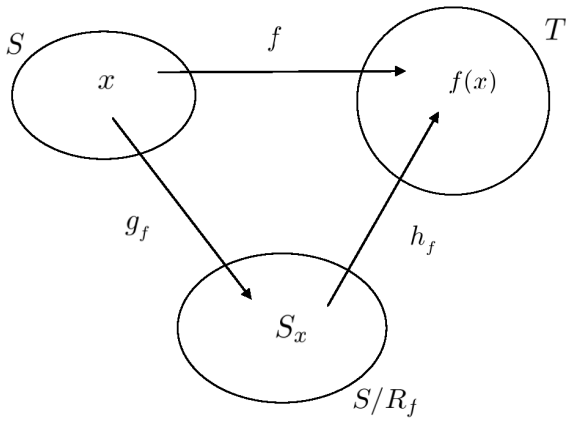


Figure 4: Canonical decomposition of $f : S \to T$

A. We define the function $g_{_f} : S \to S/R_f$ as follows:

$$g_{_f}(x) = S_x$$

We first verify that $g_{_f} : S \to S/R_f$ is a well-defined function on S.

- We first verify that $\operatorname{dom} g_{_f} = S$:

  Let $x \in S$. Then $y = f(x) \in f[A] = \operatorname{im} f$. So $S_x = f^{\leftarrow}[\{y\}]$. Thus, $g_{_f}(x) = S_x = f^{\leftarrow}[\{y\}] \in S/R_f$.
- Next we show that $g_{_f}$ is indeed a function:

  Suppose $S_x \neq S_y$. Since $\{S_x : x \in S\}$ partitions S, then

  $$S_x \cap S_y = f^{\leftarrow}[\{f(x)\}] \cap f^{\leftarrow}[\{f(y)\}] = \varnothing$$

  So $x \neq y$ otherwise we would have $f^{\leftarrow}[\{f(x)\}] = f^{\leftarrow}[\{f(y)\}]$ resulting in a contradiction. Then $g_{_f}$ is a function as claimed.

We now verify that the function g_f is onto S/R_f:

- Since, for any $x \in S$, $S_x = f^{\leftarrow}[\{f(x)\}] = g_{_f}(x)$, then $g_{_f}$ is onto S/R_f

B. We define another function $h_{_f} : S/R_f \to T$ as follows (Remember that T contains the image if f):

$$h_{_f}(S_x) = f(x)$$

We verify that $h_{_f} : S/R_f \to T$ is a well-defined function on S/R_f:

– We first verify that $\operatorname{dom} h_f = S/R_f$:

  Let $S_x \in S/R_f$. Since $x \in S$, then $f(x)$ is defined and so $h_f(S_x) = f(x)$ is defined.

– We now show h_f is indeed a function:
  Suppose $f(x)$, $f(y) \in h_f[S/R_f]$ and $f(x) \neq f(y)$. Then $\varnothing = f^{\leftarrow}[\{f(x)\}] \cap f^{\leftarrow}[\{f(y)\}] = S_x \cap S_y$. So $S_x \neq S_y$. Then h_f is indeed a function.

The function h_f is one-to-one:
The function h_f is one-to-one on S/R_f, since

$$\begin{aligned} S_x \neq S_y &\Rightarrow S_x \cap S_y = \varnothing \\ &\Rightarrow f^{\leftarrow}[\{f(x)\}] \cap f^{\leftarrow}[\{f(y)\}] = \varnothing \\ &\Rightarrow f(x) \neq f(y) \end{aligned}$$

C. Combining the two functions $g_f : S \to S/R_f$ and $h_f : S/R_f \to T$ we obtain the composition $(h_f \circ g_f) : S \to T$ where

$$(h_f \circ g_f)(x) \to f(x)$$

We verify that the function $(h_f \circ g_f) = f$: For $x \in S$,

$$\begin{aligned} f(x) &= h_f(S_x) \\ &= h_f(g_f(x)) \\ &= (h_f \circ g_f)(x) \end{aligned}$$

Thus, the two functions $h_f \circ g_f$ and f agree everywhere on the domain, S, of f. We have just proven the following theorem.

Theorem 12.2 Let $f : S \to T$ be an onto function where S and T are sets. There exists an onto function $g_f : S \to S/R_f$ and a one-to-one function $h_f : S/R_f \to T$ such that

$$h_f \circ g_f = f$$

The function, $h_f \circ g_f = f$, is called the *canonical decomposition of f.*

Example: Let $U = \{a, \{a\}, \{a, \{a\}\}, \{\{a\}\}\}$ where a is a set. Let the function $f : U \to U$ be defined as follows:

$$f(x) = \begin{cases} a & \text{if} \quad a \in x \\ \{a\} & \text{if} \quad \{a\} \in x \text{ and } a \notin x \\ \{\{a\}\} & \text{if} \quad x = a \end{cases}$$

From this we can define the equivalence relation R_f on U where

$$\begin{array}{rcccccc} U/R_f & = & \{ & f^{\leftarrow}[\{a\}], & f^{\leftarrow}[\{\{a\}\}], & f^{\leftarrow}[\{\{\{a\}\}\}] & \} \\ & = & \{ & \{\ \{a\},\ \{a,\{a\}\}\ \}, & \{\ \{\{a\}\}\ \}, & \{a\} & \} \end{array}$$

We can define $g_f : U \mapsto U/R_f$ and $h_f : U/R_f \mapsto U$ as below:

$$\begin{array}{ccccccc} a & \overset{g_f}{\longmapsto} & \{\{\{a\}\}\} & = f^{\leftarrow}[\{\{a\}\}] & \overset{h_f}{\longmapsto} & \{\{a\}\} & = f(a) \\ \{a\} & \overset{g_f}{\longmapsto} & \{\{a\},\{a,\{a\}\}\} & = f^{\leftarrow}[\{a\}] & \overset{h_f}{\longmapsto} & a & = f(\{a\}) \\ \{a,\{a\}\} & \overset{g_f}{\longmapsto} & \{\{a\},\{a,\{a\}\}\} & = f^{\leftarrow}[\{a\}] & \overset{h_f}{\longmapsto} & a & = f(\{a,\{a\}\}) \\ \{\{a\}\} & \overset{g_f}{\longmapsto} & \{a\} & = f^{\leftarrow}[\{\{\{a\}\}\}] & \overset{h_f}{\longmapsto} & \{a\} & = f(\{\{a\}\}) \end{array}$$

We see that

$$\begin{array}{rcc} (h_f \circ g_f)(a) & = & f(a) \\ (h_f \circ g_f)(\{a\}) & = & f(\{a\}) \\ (h_f \circ g_f)(\{a,\{a\}\}) & = & f(\{a,\{a\}\}) \\ (h_f \circ g_f)(\{\{a\}\}) & = & f(\{\{a\}\}) \end{array}$$

Concepts review:

1. If $f : A \to B$ is a function mapping the set A into the set B, describe a partition of the set A induced by the function f.
2. If $f : A \to B$ is a function mapping the set A into the set B, describe and equivalence relation R_f on A induced by f.
3. If $f : A \to B$ is a function mapping the set A into the set B, describe the elements of the quotient set A/R_f induced by f.
4. If $f : A \to B$ is a function mapping the set A into the set B, what does "the canonical decomposition of f" mean?
5. If $f : A \to B$ is a function mapping the set A into the set B, is it always possible to "decompose" f as a composition of two functions? How?

EXERCISES

A. 1. Let $S = \{a, b, c\}$, be a set containing three distinct elements.

a) List all the elements of $\mathscr{P}(S)$.

b) We define a function $f : \mathscr{P}(S) \times S \to \mathscr{P}(S)$ as follows:

$$f((A, x)) = \begin{cases} A & \text{if} \quad x \notin A \\ \{x\} & \text{if} \quad x \in A \end{cases}$$

List all the elements of the function f.

c) Is f onto $\mathscr{P}(S)$? Explain.

d) Is f one-to-one on S? Explain.

e) For every $D \in \mathscr{P}(S)$, give $f^{\leftarrow}(D)$.

f) If the function f determines a partition of $\mathscr{P}(S) \times S$ list the subsets which are members of this partition.

B. 3) Let A be a non-empty subset of the set S and let $T = \{\varnothing, \{\varnothing\}\}$. We define the function $f : S \to T$ as follows: $f(x) = \varnothing$ if $x \notin A$ and $f(x) = \{\varnothing\}$ if $x \in A$. Let R_f denote the equivalence relation determined by f.

a) List the elements of the quotient set S/R_f.

b) If $h_f \circ g_f = f$ is the canonical decomposition of f, list the elements of the functions g_f and of h_f.

Part V

From sets to numbers

13 / The natural numbers.

Summary. *The main objective in this section is to discuss how the natural numbers,* $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$*, are constructed within the Zermelo-Fraenkel axiomatic system. We begin by stating the definitions of "successor set" and "inductive set". The natural numbers,* $\mathbb{N}$*, is then defined as the "smallest" inductive set. A ZF-axiom will guarantee the existence of this "smallest inductive set" called the "natural numbers". We then show how the Principle of mathematical induction is an immediate consequence of this definition of* $\mathbb{N}$*. We define "transitive sets" as sets, A, whose elements are subsets of A. The elements of* $\mathbb{N}$ *are then shown to be "transitive sets". We then prove a few properties possessed by all natural numbers.*

13.1 Preliminary discussion.

We now have enough background material to appropriately define the set commonly known as the *natural numbers*. We have an intuitive understanding of what the numbers $0, 1, 2, 3, \ldots$, mean and so we will let our intuition guide us in our attempt to define $\mathbb{N}$ within our set-theoretic axiomatic system.

When defining $\mathbb{N}$ *what are our options?* If we want $\mathbb{N} = \{0, 1, 2, 3, \ldots, \}$ to be defined within the *ZFC*-axiomatic system then we don't have much choice in the matter: The elements of $\mathbb{N}$ must be sets. But they are sets whose elements have certain characteristics. These characteristics are such that nobody would confuse $\mathbb{N}$ with the real or complex numbers, for example. A way of approaching this question is to ask ourselves what properties of the natural numbers allow us to say with confidence that $1/3$ or $\sqrt{2}$ are not natural numbers. One obvious property of $\mathbb{N}$ is that it is an infinite set. We would necessarily have to define what an "infinite set" is. We then would have to think deeply about the fundamental characteristics of $\mathbb{N}$ and its elements. For example, any set which represents a non-zero natural number must have an immediate predecessor and an immediate successor. We must ask ourselves, "what kind of set can have an *immediate predecessor* and an *immediate successor*?". Furthermore, we will eventually have to determine how arithmetic operations can be performed on such sets. Determining the natural numbers' intrinsic properties which allow us to distinguish them from other types of numbers must thus be our starting point.

13.2 Constructing the natural numbers.

As we reflect on sets which would suitably represent natural numbers we come to realize how very few sets we have actually witnessed up to now in our study of set theory. What kind of sets have we encountered?

1) First, we gave ourselves an axiom (Axiom of class construction) which guarantees

that $\{x : x \neq x\}$ is a well-defined class. We decided to represent this class by "$\varnothing$" and call it the "empty class".

2) Then we gave ourselves an axiom (Axiom of subset) that says that "if S is a set, then any subclass of S must also be a set". We gave ourselves an axiom that guaranteed that there exists at least one set (The Axiom of infinity states that "there exists a non-empty class A called a set such that..."). Once we showed that $\varnothing \subseteq S$, for any set, then $\varnothing$ was our first explicitly constructed set.

3) We then gave ourselves set constructing tools. The *Axiom of pair* allows us to say, for example, that $\{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\{\varnothing, \{\varnothing\}\}\}, \ldots$ are sets. The *Axiom of union* allows to gather together all the elements from a "set of sets" to form a larger set. The *Axiom of power set* allows us to construct a set whose elements are the subsets of a set.

From this we see that nearly all the sets we explicitly constructed up to now have evolved from successive applications of the axioms of pair, union and power set, with the empty set as a starting point. If we explicitly list the elements of these sets we will see repetitive sequences of "curly brackets" "{" and "}" and the symbol "$\varnothing$". We then expect every natural number to be a set of this nature.

If we are asked to define the set of all natural numbers as succinctly as possible we may consider the following definition as a reasonable one:

> The set, $\mathbb{N}$, of all natural numbers is the intersection of all sets S which satisfy the two properties, $0 \in S$ and $[n \in S] \Rightarrow [n + 1 \in S]$.

Given the knowledge and the experience we have with natural numbers, it would be difficult to imagine a natural number which does not belong to such a set. It also seems obvious that numbers such as $\frac{5}{4}$ and $\sqrt{5}$ cannot belong to such a set. This will be our model for formulating a set-theoretic definition of the natural numbers. It seems natural to define, $0 = \varnothing$, as being the smallest of all natural numbers. The challenge is to define the operation "+ 1" using the language of sets. We can view "+ 1" as an "immediate successor constructing mechanism". We begin with the following definition.

Definition 13.1 For any set x, we define the *successor*, x^+, of x as

$$x^+ = x \cup \{x\}$$

We see that this is an operation which adds a single element to a previously known set. For example, if $A = \{a, b, c\}$, then the successor of A is

$$A^+ = \{a, b, c\} \cup \{\{a, b, c\}\} = \{a, b, c, \{a, b, c\}\}^1$$

[1]Note that if x is a set, then the expression $x^+ = x \cup \{x\}$ cannot be simplified since we do not see what the elements of x are.

This is a set constructing mechanism. We need only one set to initiate a non-ending process. Given any set, we can construct a *successor*. By the axiom of union (A7) a successor is always a set, provided the class which initiates the process is a set. Starting with the empty set $\varnothing$, we obtain

$$\begin{aligned} B_0 &= \varnothing \\ B_1 &= \varnothing^+ = \varnothing \cup \{\varnothing\} = \{\varnothing\} \\ B_2 &= (\varnothing^+)^+ = \{\varnothing\}^+ = \{\varnothing\} \cup \{\{\varnothing\}\} = \{\varnothing, \{\varnothing\}\} \end{aligned}$$

Rather than use the symbols, $\{B_0, B_1, B_2, \ldots, \}$, why not use conventional natural number notation:

$$\begin{aligned} 0 &= \varnothing \\ 1 &= 0^+ = \varnothing^+ = \varnothing \cup \{\varnothing\} = \{\varnothing\} = \{0\} \\ 2 &= 1^+ = \{\varnothing\}^+ = \{\varnothing\} \cup \{\{\varnothing\}\} = \{\varnothing, \{\varnothing\}\} = \{0, 1\} \\ 3 &= 2^+ = \{\varnothing, \{\varnothing\}\}^+ = \{\varnothing, \{\varnothing\}\} \cup \{\{\varnothing, \{\varnothing\}\}\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} = \{0, 1, 2\} \end{aligned}$$

We can thus define, one at a time, each symbol $0, 1, 2, 3, \ldots$, as a set. Let's continue for a bit and see what happens. We will define:

$$\begin{aligned} 4 &= 3^+ = \{\ \varnothing,\ \{\varnothing\},\ \{\varnothing, \{\varnothing\}\},\ \{\varnothing,\ \{\varnothing\},\ \{\varnothing, \{\varnothing\}\}\} = \{0, 1, 2, 3\} \\ 5 &= 4^+ = \{0, 1, 2, 3, 4\} \\ 6 &= 5^+ = \{0, 1, 2, 3, 4, 5\} \\ 7 &= 6^+ = \{0, 1, 2, 3, 4, 5, 6\} \end{aligned}$$

This method for constructing natural numbers is worth exploring.

We further examine the properties of a set constructed in this way.

- As mentioned before, construction starts with the set $0 = \varnothing$ moving upwards.
- For each of the elements of the set, $7 = \{0, 1, 2, 3, 4, 5, 6\}$, we see that the property "$n \subset n^+$" is satisfied. And surprisingly enough, $n \in n^+$. That is, each n is both a subset *and* an element of its successor n^+. For convenience, we define the ordering relation "$n < m$" to mean "$n \subset m$" so that $0 < 1 < 2 < 3 < 4 < 5 < 6 < 7$, for example. Witness $3 = 2^+ = \{0, 1, 2\} = \{0, 1, \{0, 1\}\}$, so $2 = \{0, 1\} \subset 3$; hence, $2 < 3$.
- For each element, n, of the set $7 = \{0, 1, 2, 3, 4, 5, 6\}$, we see that n is precisely the set of all "natural numbers" strictly less than itself.
- Also, if we set $0 = \varnothing$

$$\begin{aligned} 1 = \{\varnothing\} = \{0\} \quad &\text{so} \quad 0 \in 1 \text{ and } 0 \subset 1 \\ 2 = \{0, 1\} = \{0, \{0\}\} \quad &\text{so} \quad 1 \in 2 \text{ and } 1 \subset 2 \\ 3 = \{0, 1, 2\} = \{0, \{0\}, \{0, \{0\}\}\} \quad &\text{so} \quad 2 \in 3 \text{ and } 2 \subset 3 \\ &\qquad\quad 1 \in 3 \text{ and } 1 \subset 3 \end{aligned}$$

The set, $7 = \{0, 1, 2, 3, 4, 5, 6\}$, satisfies the property: Every pair of elements contained in the number 7 are comparable with respect to $<$ and they satisfy the property:

$$(n < m) \Rightarrow [(n \in m) \Leftrightarrow (n \subset m)]$$

We want to generalize to all of $\mathbb{N}$ the properties we have just witnessed for the numbers 0, 1, 2, ..., 7. A first draft of a definition of a natural number may look something like this:

> We say that n is a *natural number* if it is either the empty set $\varnothing$ or equal to $m \cup \{m\}$ for some other natural number m.

Then we would have to prove that "for any pair m, n of distinct natural number, $(m \in n) \Leftrightarrow (m \subset n)$". From this we would conclude that both $\in$ and $\subset$ are strict linear orderings of the set of all natural numbers.

Based on such a definition, and the few examples we have seen above, we could conclude that all the elements of $7 = \{0, 1, 2, 3, 4, 5, 6\}$ are natural numbers. Furthermore, we could show that 7 is a natural number. We can then construct $8 = 7^+ = \{0, 1, 2, 3, 4, 5, 6, 7\}$ and see that the elements of 8, as well as 8 itself, are all natural numbers.

Since the expression, $n \cup \{n\}$, is at the core of everything we have seen above, we formally provide some vocabulary to discuss such a concept.

Definition 13.2 If x is a set, then we define

$$x^+ = x \cup \{x\}$$

A set, A, is called an *inductive set* [1] if it satisfies the following two properties:

a) $\varnothing \in A$.

b) $x \in A \Rightarrow x^+ \in A$.

"Inductive set" as defined above nicely describes, in a nutshell, the class $\mathbb{N}$ of natural numbers. But we require an axiom which guarantees that at least one induction set exists. This is done with the *Axiom of infinity* (A8).

> **Axiom** (*The axiom of infinity*): An inductive set exists.

This axiom of infinity provides us with the minimum amount of raw material to do finite mathematics.

Now that we have given ourselves at least one inductive set, we will define the natural numbers as being the smallest one.

[1] The term *successor set* is also used instead of *inductive set*.

Definition 13.3 We define the *set,* $\mathbb{N}$*, of all natural numbers* as the intersection of all inductive sets. That is,

$$\mathbb{N} = \{x : x \in I \text{ for all inductive set } I\}$$

To say that $\mathbb{N}$ is the intersection of all inductive sets is another way of describing the set of all natural numbers. Is the set $\mathbb{N}$ itself inductive? We verify this: By definition, all induction sets contain the element $\varnothing$ and so $\varnothing$ belongs to their intersection, $\mathbb{N}$. Condition one is satisfied. We verify condition two: If $x \in \mathbb{N}$, then x belongs to all inductive sets and so x^+ must belong to all inductive sets; so $x^+ \in \mathbb{N}$. So $\mathbb{N}$ is an inductive set. It immediately follows that if n is any natural number, then so is its successor, n^+. We will now verify that $\mathbb{N}$, thus defined, actually satisfies all the properties that are expected from it.

13.3 Mathematical induction.

The cleverly chosen four words "An inductive set exists" will allow us to prove that the smallest inductive set provides a precise set-theoretic representation of the set of all natural numbers as we know it. This inductive set possesses all the essential properties of the natural numbers, including its linear ordering structure. As we shall soon see, it will allow us to define on it the common arithmetic operations we normally perform on natural numbers. Proving that this inductive set possesses all the essential properties of the natural numbers will require the well known mathematical tool called the *Principle of mathematical induction.* This principle is "hardwired" within the definition of "inductive set".

Theorem 13.4 Let A be a subset of $\mathbb{N}$. If A satisfies the two properties:

a) $0 \in A$

b) $m \in A \Rightarrow m^+ \in A$

then $A = \mathbb{N}$.

Proof:

By hypothesis, A is an inductive set since it satisfies the two required properties. Since $\mathbb{N}$ is the intersection of all inductive sets, then $\mathbb{N} \subseteq A$. By hypothesis, $A \subseteq \mathbb{N}$. Thus, $A = \mathbb{N}$.

Corollary 13.5 (*The Principle of mathematical induction.*) Let P denote a particular set property. Suppose $P(n)$ means "the property P is satisfied depending on the value of the natural number n". If

a) $P(0)$ holds true,

b) $P(n)$ holds true $\Rightarrow P(n^+)$ holds true.

then $P(n)$ holds true for all natural numbers n.

Proof: Let

$$A = \{n \in \mathbb{N} : P(n) \text{ holds true } \}$$

Part a) of the hypothesis states that "$0 \in A$", while part b) states that $[n \in A] \Rightarrow [n^+ \in A]$. Then A is an inductive set and so $A = \mathbb{N}$ (by the theorem). So $P(n)$ is true for all natural numbers n.

A few remarks. The proofs above illustrate how the *Principle of mathematical induction* is intrinsically linked to the definition of the natural numbers. The set of all natural numbers is the only set whose existence is essentially postulated. The other explicitly defined set is the empty class which was shown to be a set (as a consequence of the Axiom of construction followed by the Axiom of subset).

Some readers may not be familiar with "proofs by mathematical induction". These readers will benefit from the many examples of proofs by induction in this section. It is well worth taking a few moments to summarize the main steps to be followed when proving a statement by induction. Induction is used when we are dealing with some property $P(n)$ which is a function of the natural numbers. Let $S = \{n \in \mathbb{N} : P(n) \text{ holds true}\}$. Now this set, S, may possibly be empty, may contain a few elements of $\mathbb{N}$ or may even contain all of its elements. The objective is to show that if two specific conditions are satisfied, then $S = \mathbb{N}$. That is, we want to prove that $P(n)$ holds true for all values of n. For example, suppose $P(n)$ is described as the property

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$$

We want to prove that this holds true no matter what natural number n we use. We highlight the main steps.

Step 1: Write down explicitly the property which is a function of n as illustrated above.

Step 2: Prove the "Base case". This means that we must prove that $P(0)$ is true. In our example, we are required to show that $0 = \frac{0(0+1)}{2} = 0$. We see that the base case holds true. If the property cannot be shown to be true for the base case, then $P(n)$ does not hold true for all n. It sometimes helps us to understand what is going on if we prove that both $P(0)$ and $P(1)$ are true (especially when the base case is

"vacuously true").

Step 3: State the "Inductive hypothesis". In this step we suppose that the $P(n)$ is true for some unspecified natural number n. That this property holds true for a particular n is now considered to be "given".

Step 4: With the help of the assumption that $P(n)$ is true, prove that $P(n^+)$ (equivalently $P(n+1)$) is true. In our example we would write something like this:

$$\begin{aligned} 1+2+\cdots+n=\frac{n(n+1)}{2} \quad &\Rightarrow \quad 1+2+\cdots+n+(n+1)=\frac{n(n+1)}{2}+(n+1) \\ &\Rightarrow \quad 1+2+\cdots+n+(n+1)=\frac{(n+1)(n+2)}{2} \end{aligned}$$

Step 5: Write down the conclusion: Since "$P(n)$ is true" implies that "$P(n+1)$ is true", then, by the principle of mathematical induction, $P(n)$ holds true for all n.

Difficulties encountered when applying this procedure are often due to skipped steps.

13.4 Transitive sets.

The few examples of natural numbers constructed above have illustrated an interesting property: Each element of a natural number, n, is seen to be subset of n. We formally define the words used to express sets which satisfy this property. We then provide in the form of a theorem a useful characterization of such sets.

Definition 13.6 A set, S, which satisfies the property "$x \in S \Rightarrow x \subseteq S$" is called a *transitive set.*

Theorem 13.7 The non-empty set, S, is a transitive set if and only if the property

$$[x \in y \text{ and } y \in S] \Rightarrow [x \in S]$$

holds true.

Proof:

($\Rightarrow$) What we are given: That S is transitive, $x \in y$ and $y \in S$.
What we are required to show: That $x \in S$.
Since S is a transitive set $y \subseteq S$. Then $x \in y \subseteq S$ implies $x \in S$.

($\Leftarrow$) We are given that "$(x \in y \text{ and } y \in S) \Rightarrow (x \in S)$" and $z \in S$.
We are required to show that "$z \subseteq S$".
If $z = \varnothing$, then $z \subseteq S$ and we are done. Suppose that $z \neq \varnothing$. Let $a \in z$. By hypothesis, $a \in S$. Since $a \in z$ implies $a \in S$, then $z \subseteq S$.

The above characterization ($x \in y$ and $y \in S$) $\Rightarrow x \in S$ relates more closely to our idea of transitivity (that is, $a < b$ and $b < c \Rightarrow a < c$).

The next theorem shows that $\mathbb{N}$ is a transitive set.

Theorem 13.8 The set $\mathbb{N}$ of natural numbers is a transitive set.

Proof:

By the characterization of transitive sets stated above, it suffices to show that for each $n \in \mathbb{N}$, $x \in n \Rightarrow x \in \mathbb{N}$. We will prove this by mathematical induction. Let $P(n)$ denote the statement "$(x \in n \in \mathbb{N}) \Rightarrow x \in \mathbb{N}$".

- Base case: The statement, "$x \in 0 = \varnothing$" $\Rightarrow$ "$x \in \mathbb{N}$", is true since there are no elements in $0 = \varnothing$. So $P(0)$ holds true.
- Inductive hypothesis: Suppose the statement $P(n)$ holds true for the natural number n. We are required to show that $P(n^+)$ holds true. Suppose $y \in n^+ = n \cup \{n\}$. Then either $y \in n$ or $y \in \{n\}$. If $y \in n$, then by the inductive hypothesis, $y \in \mathbb{N}$. If $y \in \{n\}$, then $y = n \in \mathbb{N}$.

By mathematical induction, the statement holds true for all elements of $\mathbb{N}$ and so, by definition, $\mathbb{N}$ is a transitive set.

The following theorem first establishes that no natural number is an element of itself. It also shows that for distinct natural numbers one is an element of the other if and only if it is a proper subset of the other. This of course implies that every natural number is a transitive set.

Theorem 13.9

a) For natural numbers n, m, $m \in n \Rightarrow m \subseteq n$. Hence, every natural number is a transitive set.

b) For any natural number n, $n \neq n^+$.

c) For any natural number n, $n \notin n$.

d) For any distinct natural numbers n, m, $m \subset n \Rightarrow m \in n$. [1]

[1] The reader is cautioned not to misread the statement "$m \subset n \Rightarrow m \in n$". It does *not* say that any subset of a natural number n is an element of n. It says that "any *natural number* which is a subset of n is an element of n".

Proof:

a) This is a proof by mathematical induction. Let $P(n)$ be the property "m and n are natural numbers and $m \in n \Rightarrow m \subseteq n$". We are required to prove that the set $\{n \in \mathbb{N} : P(n) \text{ is true } \} = \mathbb{N}$.

- Base case: We claim that $P(0)$ holds true. Recall that $0 = \varnothing$. Suppose $P(\varnothing)$ is false. Then there must be some $x \in \varnothing$ such that $x \notin \varnothing$. This is absurd since $\varnothing$ does not contain any elements. So $P(0)$ holds true.
- Inductive hypothesis: Suppose that for some natural number n, $P(n)$ holds true. We are required to show that $P(n^+)$ holds true.
  * To show that $P(n^+)$ is true, suppose $m \in n^+ = n \cup \{n\}$. We are required to show that $m \subseteq n^+ = n \cup \{n\}$.
    Case 1: If $m \in n$ then by the inductive hypothesis, $P(n)$ is true, and so $m \subseteq n$. Then $m \subseteq n^+ = n \cup \{n\}$ and so $P(n^+)$ is true.
    Case 2: Suppose $m \notin n$.

$$\begin{aligned} m \notin n \text{ and } m \in n^+ = n \cup \{n\} &\Rightarrow m \in \{n\} \\ &\Rightarrow m = n \end{aligned}$$

    Clearly $m = n \subseteq n \cup \{n\}$. Then $P(n^+)$ is true.

  We have shown that if $P(n)$ is true, then $P(n^+)$ is true. By mathematical induction $P(n)$ is true for all $n \in \mathbb{N}$. We conclude that every natural number is a transitive set.

b) We prove that $n \neq n^+$ by induction. Let $P(n)$ denote the statement "$n \neq n \cup \{n\}$".
Base case: Since $\varnothing = \{\ \} \neq \{\varnothing\}$, $P(0)$ holds true.
Inductive hypothesis: Suppose $n \neq n \cup \{n\}$ for some natural number n.
We are required to show that $n \cup \{n\} \neq n \cup \{n\} \cup \{n \cup \{n\}\}$.
Now

$$\begin{aligned} n \cup \{n\} \cup \{n \cup \{n\}\} = n \cup \{n\} &\Rightarrow n \cup \{n\} \in n \cup \{n\} \\ &\Rightarrow n \cup \{n\} \in n \text{ or } n \cup \{n\} = n \end{aligned}$$

The inductive hypothesis does not allow "$n \cup \{n\} = n$". So $n \cup \{n\} \in n$. By part a), $n \cup \{n\} \subseteq n$. Since $n \subseteq n \cup \{n\}$, then $n = n \cup \{n\}$ (Axiom of extent) again contradicting the inductive hypothesis. Then $n \cup \{n\} \cup \{n \cup \{n\}\} \neq n \cup \{n\}$. So "$P(n) \Rightarrow P(n^+)$" holds true.
By mathematical induction, $n \neq n \cup \{n\}$ for all natural numbers.

c) Suppose n is a natural number such that $n \in n$. Then $n \cup \{n\} \subseteq n$. Since $n \subseteq n \cup \{n\}$, $n = n \cup \{n\}$ contradicting the statement of part b). We must conclude that $n \notin n$.

d) We are required to show that for all m, $n \in \mathbb{N}$, $m \subset n \Rightarrow m \in n$. We will prove this by mathematical induction on n. Let $P(n)$ be the property "$[m$ is a natural number and $m \subset n] \Rightarrow [m \in n]$".

– Base cases $n = \varnothing$ or 1: For $n = \varnothing$, the statement $m \subset \varnothing \Rightarrow m \in \varnothing$ is vacuously true. For $n = 1$, $\varnothing \subset 1 = \{\varnothing\}$ and $\varnothing \in 1 = \{\varnothing\}$ hold true. So both base cases $P(0)$ and $P(1)$ hold true. (Actually showing $P(0)$ holds true is sufficient.)

– Inductive hypothesis: Suppose n is a natural number such that $P(n)$ holds true; that is, for any natural number m, "$m \subset n \Rightarrow m \in n$". We are required to show that for any natural number m, "$m \subset n^+ \Rightarrow m \in n^+$".

* Let m be a natural number such that "$m \subset n^+ = n \cup \{n\}$".
Case 1: If $n \notin m$, then $m \subset n$. By the inductive hypothesis, $m \in n \subset n \cup \{n\}$. Hence, $m \in n \cup \{n\}$.
Case 2: Suppose $n \in m$. By part b), $m \neq n$. Since m and n are distinct natural numbers, by part a), $n \subset m$. Then $n \cup \{n\} \subseteq m$. Since $m \subset n \cup \{n\}$, then $n \cup \{n\} \subset n \cup \{n\}$, a contradiction.
So only case 1 applies. So $P(n^+)$ holds true.

By the principle of mathematical induction $m \subset n \Rightarrow m \in n$ for all natural numbers m and n.

To illustrate how the elements of $\mathbb{N}$ satisfy the property "$x \in n \Rightarrow x \subset n$" consider, for example, the natural number 4,

$$4 = \{\ \varnothing,\ \{\varnothing\},\ \{\varnothing, \{\varnothing\}\},\ \{\varnothing,\ \{\varnothing\},\ \{\varnothing, \{\varnothing\}\}\}$$

We see that

$$\begin{aligned}
0 = \varnothing \in 4 \quad &\text{and} \quad 0 = \varnothing \subset 4 \\
1 = \{\varnothing\} \in 4 \quad &\text{and} \quad 1 = \{\varnothing\} \subset 4 \quad \text{Since every element of } \{\varnothing\} \text{ belongs to 4.} \\
2 = \{\varnothing, \{\varnothing\}\} \in 4 \quad &\text{and} \quad 2 = \{\varnothing, \{\varnothing\}\} \subset 4 \quad \text{Since every element of } \{\varnothing, \{\varnothing\}\} \text{ belongs to 4.} \\
3 = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} \in 4 \quad &\text{and} \quad \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} \subset 4 \\
&\qquad \text{Since every element of } \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} \text{ belongs to 4.}
\end{aligned}$$

13.5 Other basic properties of natural numbers.

We prove a few more properties of the set, $\mathbb{N}$, as defined above.

Theorem 13.10 Let m and n be distinct natural numbers.

a) If $m \subset n$, then $m^+ \subseteq n$.

b) Let m and n be any pair of distinct natural numbers. Then either $m \subset n$ or $n \subset m$. Equivalently, $m \in n$ or $n \in m$. Hence, both "$\subset$" and "$\in$" are strict linear orderings of $\mathbb{N}$.

c) There is no natural number m such that $n \subset m \subset n^+$.

Proof:

a) What we are given: That m and n are distinct natural numbers where $m \subset n$. What we are required to show: That $m \cup \{m\} \subseteq n$.

Since $m \subset n$, then $m \in n$ (by theorem 13.9). If $n = m \cup \{m\}$, then we are done since $m \cup \{m\} \subseteq n$. Suppose $n \neq m \cup \{m\}$. Then "$m \subset n$" and "$m \in n$" together imply that $m \cup \{m\} \subset n$. Hence, $m^+ = m \cup \{m\} \subseteq n$, as required.

b) What we are given: That m and n are distinct natural numbers. What we are required to prove: That $m \subset n$ or $n \subset m$.

We will prove this by mathematical induction on n. Let $P(n)$ be the statement "for every natural number $m \neq n$, either $m \subset n$ or $n \subset m$"

- Base cases $n = \varnothing$ or 1: For $n = \varnothing$, the statement $\varnothing \subset m$ holds true for all non-zero natural numbers m. For $n = 1$ and $m = 1$, $\varnothing \subset 1 = \{\varnothing\}$. Suppose m is a natural number other than 0 and 1. Then $\varnothing \subset m \Rightarrow \varnothing^+ = \{\varnothing\} \subseteq m$ (by the theorem 13.9 above). Since $m \neq 1$, then $\{\varnothing\} \subset m$ holds true for any such m (since $\varnothing \in m$ for every non-zero natural number m).
- Inductive hypothesis: Suppose $P(n)$ holds true for some natural number n. That is, suppose n is a natural number such that for any natural number m not equal to n, either $m \subset n$ or $n \subset m$. We are required to show that $P(n^+)$ holds true.

  Let m be a natural number such that $m \neq n^+$. Case 1: If $m = n$, then $m \in n \cup \{n\}$ and so $m \subset n \cup \{n\}$ (by theorem 13.9) and we are done. Case 2: Suppose $m \neq n$. Then, by the inductive hypothesis, either $m \subset n$ or $n \subset m$. If $m \subset n$, then $m \subset n^+ = n \cup \{n\}$. Done. If $n \subset m$, then $n^+ \subseteq m$ (by part a)). Since $m \neq n^+$, $n^+ \subset m$. Then $P(n^+)$ holds true.

By the principle of mathematical induction, for any pair of distinct natural numbers m and n, either $m \subset n$ or $n \subset m$. Since $m \subset n$ if and only if $m \in n$, this property also holds true with respect to the relation "$\in$".

c) What we are given: That n and m are distinct natural numbers. What we are required to prove: That $n \subset m \subset n^+$ is impossible.

Suppose $n \subset m \subset n^+$. Then $m \in n \cup \{n\}$. Since $m \neq n$, then $m \in n$ which means $m \subset n$. This contradicts our hypothesis, $n \subset m$. We have shown that $n \subset m \subset n^+$ is impossible, as required.

13.6 The immediate predecessor of a natural number.

We have seen that the elements, n, of the natural numbers are equipped with an "*immediate successor*" constructing algorithm, $n^+ = n \cup \{n\}$, where $n \subset n^+ = n \cup \{n\}$ and no other natural number m sits between n an n^+. It is normal to ask if every non-zero natural number has an "immediate predecessor". That is, given an arbitrary natural number, n, are we guaranteed that there exists a natural number, k, such that $k \cup \{k\} = k^+ = n$. Part c) of the theorem above guarantees that there can be no

natural number between k and n, and so such a k would be the immediate predecessor of n. If such a k exists, is there a way to construct this predecessor k of n just as we were able to construct an immediate successor of a natural number? The next theorem shows how we can construct the immediate predecessor of a natural number.

Theorem 13.11 If m and n are natural numbers such that $m^+ = n$, then m is called an *immediate predecessor* of n. For any non-zero natural number n, $k = \cup\{m \in \mathbb{N} : m \subset n\}$ is a natural number which is an immediate predecessor of n.

Proof:

What we are given. That n is a non-zero natural number.
What we are required to show: That $k = \cup\{m \in \mathbb{N} : m \subset n\}$ is a natural number and $k^+ = n$.

Proof by induction. Let $P(n)$ be the statement "$k = \cup\{m \in \mathbb{N} : m \subset n\}$ is a natural number and $k^+ = n$".

- Base cases $n = 1$ or 2: If $n = 1$, then $k = \bigcup_{m\in 1} m = \varnothing$ a natural number such that $k^+ = \varnothing\cup\{\varnothing\} = \{\varnothing\} = 1 = n$. If $n = 2$, then $k = \bigcup_{m\in 2} m = \varnothing\cup 1 = \varnothing\cup\{\varnothing\} = \{\varnothing\} = 1$ a natural number such that $k^+ = \{\varnothing\} \cup \{\{\varnothing\}\} = \{\varnothing, \{\varnothing\}\} = 2 = n$.
- Induction hypothesis: Suppose $P(n)$ holds true. That is, suppose n is a natural number for which $\bigcup_{m\in n} m$ is a natural number satisfying $\left(\bigcup_{m\in n} m\right)^+ = n$. To show that $P(n^+)$ holds true, it suffices to show that $\bigcup_{m\in n^+} m$ is a natural number and $\left(\bigcup_{m\in n^+} m\right)^+ = n^+$. Let $k = \bigcup_{m\in n^+} m$. See that

$$\begin{aligned} k &= \cup\{m \in \mathbb{N} : m \in n \cup \{n\}\} \\ &= \cup\{m \in \mathbb{N} : m \in n \text{ or } m = n\} \\ &= \cup\{m \in \mathbb{N} : m \in n\} \cup n \end{aligned}$$

  By the induction hypothesis, $\cup\{m \in \mathbb{N} : m \in n\}$ is a natural number which is an immediate predecessor of n. Then it must be a proper subset of n. It follows that $\cup\{m \in \mathbb{N} : m \in n\} \cup n = n$. Then $k = n$ which implies $k^+ = n^+$. So $P(n^+)$ holds true.

By mathematical induction, for every natural number n, $\cup\{m \in \mathbb{N} : m \subset n\}$ is a natural number and $(\cup\{m \in \mathbb{N} : m \subset n\})^+ = n$.

Theorem 13.12 *Unique immediate predecessors.* Any non-zero natural number has a *unique* immediate predecessor.

Proof:

We prove this by induction. For non-zero natural numbers n let $P(n)$ be the statement " the natural number n has a unique immediate predecessor".

- Base case $n = 1$: By definition, $1 = \{\varnothing\} = \varnothing \cup \{\varnothing\} = \varnothing^+ = 0^+$. So $P(1)$ holds true.
- Induction hypothesis: Suppose n is a non-zero natural number such that $P(n)$ holds true. That is, *there is only one* natural number m such that $m^+ = n$. We are required to show that n^+ has a unique predecessor. Trivially, n is one immediate predecessor of n^+. Suppose k is another natural number such that $k^+ = n^+$. Then $n \cup \{n\} = k \cup \{k\}$. We claim that $n = k$. Suppose not. Then both $n \in k$ and $k \in n$ must hold true. We have shown that every natural number is a transitive set (see theorem 13.9 a)). By theorem 13.7, $n \in k$ and $k \in n$ implies $n \in n$. By 13.9 part c) this cannot be true for any natural number, and so we have a contradiction. Then $n = k$ as claimed. Hence, n^+ has as unique immediate predecessor, n. So $P(n^+)$ holds true.

By mathematical induction, every non-zero natural number has a unique immediate predecessor.

13.7 The second version of the Principle of mathematical induction.

The following theorem is a variation of the Principle of mathematical induction. It may sometimes be more efficient to apply this version when proving certain theorem statements. Although we will not present an application of this version now, we will soon see some proofs in which this version is easier to apply.

Theorem 13.13 (*The Principle of mathematical induction: second version.*) Suppose $P(n)$ is a property whose truth value depends on the natural number n. Suppose that for any natural number n,

$$[P(k) \text{ is true for all } k < n] \Rightarrow [P(n) \text{ is true}]$$

Then $P(n)$ holds true for all natural numbers n.[1]

Proof:

Given hypothesis: $[P(k)$ is true for all $k < n] \Rightarrow [P(n)$ is true$]$
What we are required to show: That $P(n)$ holds true for all natural numbers n.
Let $P^*(n)$ denote the statement "$P(k)$ is true for all $k < n$". We will show by induction (original version) that $P^*(n)$ holds true for all n. From this we will conclude that $P(n)$ holds true for all n.

- Base case: Since $P(k)$ vacuously holds true for all $k \in \varnothing$, then by the given hypothesis, $P(0) = P(\varnothing)$ holds true. So the base case $P^*(0)$ is satisfied.
- Inductive hypothesis: Suppose $P^*(n)$ holds true for some natural number n. This means "$P(k)$ is true for all $k < n$". By the given hypothesis, $P(n)$ must hold true. Then "$P(k)$ is true for all $k < n^+$". That is, $P^*(n^+)$ holds true.

[1] Note that $k < n$ and $k \in n$ are equivalent expressions.

By mathematical induction, $P^*(n)$ holds true for all natural numbers n.

Let m be any natural number. Then, by what we have just shown, $P^*(m^+)$ holds true. That is, "$P(k)$ is true for all $k < m^+$". So $P(m)$ is true. The statement is thus proved.

13.8 A few words about the Peano axioms.

There is a particular set of axioms which serves as a foundation for all mathematical statements related to the natural numbers. These axioms are *not* set-theoretic and slightly predate the *ZFC*-axioms. In 1889, the Italian mathematician Giuseppe Peano proposed a set of 9 mathematical statements from which evolve all mathematical statements relating to the natural numbers. Today these are referred to as the *Peano axioms* (the Italian name "Peano" is pronounced as 'pay-ah-no'). We will see that each of these axioms belongs to *ZFC* set-theoretic universe and so as a group play the role of intermediary − certainly a more easily understandable one − between the Set theory axioms and the body of mathematics we refer to as *number theory*. We will list the 9 Peano axioms below. The symbol "0" is an undefined symbol. The symbol "S" represents a single valued function we refer to as the "*successor function*" on the natural numbers.

P1 The symbol 0 is a natural number.

Peano axioms on equality

P2 Every natural number is equal to itself. That is, equality "=" is a reflexive binary relation on $\mathbb{N}$.

P3 If n and m are natural numbers such that $n = m$ then $m = n$. That is, equality is a symmetric binary relation on $\mathbb{N}$.

P4 If n, m and k are natural numbers such that $n = m$ and $m = k$ then $n = k$. That is, equality is a transitive binary relation on $\mathbb{N}$.

P5 If n is a natural number and "$a = n$" then a is a natural number.

Properties involving the successor function, S.

P6 If n is a natural number then so is the image, $S(n)$, of n under S. We refer to "$S(n)$" as a *successor of* n.

P7 The natural numbers n and m are equal if and only if $S(n) = S(m)$. Hence, every natural number has precisely one successor and a natural number is the successor of, at most, one natural number.

P8 For any natural number n, $S(n) \neq 0$.

Mathematical induction

P9 If M is a set which contains the natural number 0 and whenever n is a natural number then so is $S(n)$ then M contains all natural numbers. That is $\{0,\ S(0),\ S(S(0)),\ S(S(S(0))),\ \ldots,\} \subseteq M$

We verify that each of these 9 statements follows from *ZFC*-axioms. The empty set, $\emptyset$, is easily perceived as being the natural number 0. Equality of sets is reflexive in *ZFC* (this follows just about immediately from axiom A1) and so this must hold true for those sets in *ZFC* we call the natural numbers. Symmetry and transitivity of "=" on sets automatically applies to those sets we call the natural numbers. So P3 and P4 also belong to *ZFC*. Equal sets contain the same elements and so P5 holds true in *ZFC* (See theorem 2.3 c)). By the axiom of pair and union, for any natural number n, $n \cup \{n\} = S(n)$ is a set. So P6, P7 and P8 easily follow from this definition of the "successor of n". The Mathematical induction statement, P9, follows from the Axiom of infinity. Note that the *Axiom of power set*, the *Axiom of replacement*, the *Axiom of regularity* and the *Axiom of choice* are not required to do mathematics with the natural numbers.

Concepts review:

1. If x is a set then what is its *successor*?
2. What is an *inductive set*?
3. What does the *Axiom of infinity* state?
4. How is the set of *natural numbers* defined?
5. List the first four natural numbers using set notation.
6. What is the *Principle of mathematical induction*?
7. What is a *transitive set*?
8. If n is a natural number, is n a transitive set?
9. What is the difference between an *inductive* set and a *transitive set*?
10. Is $\mathbb{N}$ a transitive set?
11. Give a characterization of transitive sets.
12. Is it true that any element of a natural number is a natural number? Why?
13. Can a natural number be an element of itself?

14. Is $\mathbb{N}$ is a natural number? Why?
15. If n is a natural number how many successors can n have?
16. What is a second version of the Principle of mathematical induction?
17. If n is a natural number what does it mean to say that m is its predecessor?
18. Give an expression which describes the predecessor of a natural number n.
19. If m and n are natural numbers such that $m \in n$ can it happen that $m^+ = n$? Can it happen that $n \in m^+$?
20. If m is a subset of the natural number n is it possible that $m \in n$? In which case?
21. Are there any natural numbers which are inductive sets?
22. For three natural numbers m, n and t satisfying $m \in n$ and $n \in t$ does it always follow that $m \in t$?

EXERCISES

A.
1. Let m and n be two natural numbers. Prove that if $m \in n$, then $m \neq n^+$.
2. Show that if n is a natural number, then $n^+ \neq 0$.
3. Write down the natural number 5 using only left and right brackets, commas and the symbol "$\varnothing$".
4. If n is a natural number is n and inductive set? Justify your answer.
5. Is the set of all natural numbers a natural number? Justify your answer.

B.
6. Is $\mathbb{N} \cup \{\mathbb{N}\}$ a natural number? Explain why or why not.
7. Suppose n is a non-zero natural number.
   a) Is $\mathscr{P}(n)$ a natural number? Why?
   b) Does $\mathscr{P}(n)$ contain a natural number? Which one?
   c) List all elements of $\mathscr{P}(3)$.
8. Consider the class $\mathscr{P}(\mathbb{N})$.
   a) Is $\mathscr{P}(\mathbb{N})$ a set? Why?
   b) Does $\mathscr{P}(\mathbb{N})$ contain any natural numbers? Explain.
   c) Does $\mathscr{P}(\mathbb{N})$ contain elements which are not natural numbers? If so, list at least three.
   d) Is $\mathscr{P}(\mathbb{N})$ a natural number? Why?

9. Is $\mathbb{N} \cup \{\mathbb{N}\}$ a transitive set? If so prove it. If not say why.
10. Is $\mathbb{N} \cup \{\mathbb{N}\}$ an inductive set? If so prove it. If not say why.

C.

11. Show that finite unions and finite intersections of transitive sets are transitive sets.
12. Suppose $S \subset \mathbb{N}$. Suppose that the union of all elements of S is S. Prove that S cannot be a natural number.
13. Jo-Anne has defined the natural numbers in the *ZFC*-axiomatic system as follows. She defined an *inductive set* as "S is inductive if, whenever $x \in S$, then $\{x\} \in S$". By first invoking the axiom of infinity she defines the natural numbers $\mathbb{N}$ as the smallest inductive set linearly ordered by "$\in$". She defines $0 = \varnothing$, $1 = \{\varnothing\}$, $2 = \{\{\varnothing\}\}$, $3 = \{\{\{\varnothing\}\}\}$, $4 = \{\{\{\{\varnothing\}\}\}\}$ and so on. We see that $0 \in 1 \in 2 \in 3 \in 4 \cdots$. Will this work as a definition of the natural numbers? If so, say why. If not, explain why.
14. Show that $\mathbb{N} = \cup\{n : n \in \mathbb{N}\}$.

14 / The natural numbers as a well-ordered set.

Summary. *In this section we introduce the notion of "well-ordered set". We show that the set of all natural numbers, when equipped with the membership ordering relation "$\in$", is a well-ordered set. We then define "bounded set" and "the maximal element of a set". Finally we show that bounded subsets of $\mathbb{N}$ must contain a maximal element. We then use $\mathbb{N}$ to construct various other sets, some of which are also well-ordered.*

14.1 Order relations on $\mathbb{N}$.

We have seen that the definition of the natural numbers within the *ZFC*-axiomatic system leads to two equivalent order relations on $\mathbb{N}$. Both "$\subset$" and "$\in$" have been shown to be equivalent strict linear orderings of $\mathbb{N}$, in the sense that $n \subset m$ if and only if $n \in m$.

We can naturally extend the strict order relation "$\subset$" to the non-strict order relation "$\subseteq$" while maintaining the linearity property. That is, $m \subseteq n$ if either $m \subset n$ or $m = n$. We can similarly extend the relation "$\in$" by introducing the following notation.

Notation 14.1 We define the relation "$\in_=$" on $\mathbb{N}$ as follows:

$$m \in_= n \text{ if and only if } m = n \text{ or } m \in n$$

If $m \in_= n$ and we want to state explicitly that $m \neq n$ we write $m \in n$.

14.2 A well-ordering of $\mathbb{N}$.

There is an important property that is not possessed by all linearly ordered classes. It is called the *well-ordering* property. We formally define this property. We will then prove that $(\mathbb{N}, \in)$ is a well-ordered set.

Definition 14.2 Let $(S, \leq)$ be a linearly ordered set. Suppose $T \subseteq S$.

a) We say that the element q is *"a least element of T with respect to $\leq$"* if and only if $q \in T$ and $q \leq m$ for all $m \in T$.

b) If S is equipped with a strict linear ordering "$<$" we say that q is a *least element of T with respect to $<$* if and only if $q \in T$ and $q < m$ for all $m \in T$ where $m \neq q$.

c) The set $(S, \leq)$ is said to be *well-ordered* with respect to "$\leq$" if every non-empty subset T of S contains its least element with respect to $\leq$. Similarly, the set $(S, <)$ is said to be *well-ordered* with respect to "$<$" if every non-empty subset T of S contains its least element with respect to $<$.

We show that $\in$ well-orders the set $\mathbb{N}$.

Theorem 14.3 The set $\mathbb{N}$ of all natural numbers is a strict $\in$-well-ordered set.

Proof:

Given: The relation "$\in$" strictly linearly orders $\mathbb{N}$; the set A is a non-empty subset of $\mathbb{N}$.
Required to show: That A contains a least element with respect to $\in$.

Proof by contradiction: Suppose A does *not* contain a least element. We claim that A must then be empty, thus contradicting our hypothesis.

– Proof of the claim: We invoke the *second version* of the Principle of mathematical induction. For each natural number k, let $P(k)$ denote the statement "$k \notin A$".

Induction hypothesis: Let n be some natural number such that $P(k) =$ "$k \notin A$" holds true for all $k \in n$.
Suppose $n \in A$. Then, for all $a \in A$, $n \in a$ (for if $a \in n$, then, by the induction hypothesis, $P(a) =$ "$a \notin A$" holds true). This means that n is a least element of A with respect to $\in$. This contradicts "A contains no least element with respect to $\in$". Then $n \notin A$. Then, $P(n) =$ "$n \notin A$" holds true. By the second version of the principle of mathematical induction, $P(k) =$ "$k \notin A$" holds true for all $k \in \mathbb{N}$. Then A contains no elements, as claimed.

This contradicts the fact that A is non-empty. The source of this contradiction is our assumption that A does not contain a least element. We must conclude that every subset of $\mathbb{N}$ has a least element with respect to "$\in$".

We have previously shown that the *second version* of the Principle of mathematical induction follows from the *first version* or the Principle of mathematical induction. We can show that if we only assume that $\mathbb{N}$ is $\in$-well-ordered and the second version of the induction principle, then the first version of the induction principle holds true. The proof is as follows.

What we are given: That $\mathbb{N}$ is $\in$-well-ordered and that the second version of the Principle of mathematical induction holds true.

What we are required to show: That the first version of the Principle of mathematical induction must hold true.

Let $P(n)$ be a property whose truth value depends on the natural number n. Suppose $P(0)$ is known to be true. Also suppose that if $P(n)$ holds true, then so does $P(n^+)$. Let $A = \{k \in \mathbb{N} : \ P(k) \text{ is false } \}$. Then $\mathbb{N} - A$ contains 0 and, whenever $n \in \mathbb{N} - A$, then $n^+ \in \mathbb{N} - A$.

We claim that A must be empty (hence, $P(k)$ holds true for all $k \in \mathbb{N}$).

Proof of claim: Suppose A is non-empty. Then, since $\mathbb{N}$ is $\in$-well-ordered, A has a least element, say $s = m^+$. Then $P(k)$ holds true for all $k \in s = m^+$. Then $m \in \mathbb{N} - A$. By hypothesis, $m^+ \in \mathbb{N} - A$. This contradicts the fact that m^+ is the least element of A. The source of the contradiction is the assumption that A is non-empty. Then A must be empty, as claimed.

We conclude that the set A is empty and so $P(k)$ holds true for all natural numbers k. We have thus shown that the first version of the Principle of mathematical induction on $\mathbb{N}$ holds true.

Corollary 14.4 Every natural number n is a $\in$-well-ordered set.

Proof:

Let $n = \{0, 1, 2, \ldots, n-1\}$ be a natural number. We already know that the natural numbers are $\in$-linearly ordered. Let U be a non-empty subset of n. Then U is a non-empty subset of $\mathbb{N}$. When viewed as a subset of the $\in$-well-ordered set $\mathbb{N}$, the set U contains a least natural number, say k. Then $k \in_= m$ for all $m \in U$. So when U is viewed as a subset of n, k is the least element of U. So n is $\in$-well-ordered.

We have thus shown that not only is $\mathbb{N}$ a well-ordered set, but so is every single natural number.

14.3 Bounded subsets of $\mathbb{N}$

The reader may be familiar with the concept of bounded subsets. In the context of a linearly ordered set $(S, <)$, we say that a subset A of S is *bounded above*, or has an *upper bound* if there exists some element $M \in S$ we call an "upper bound of A" such that $x \leq M$ for all $x \in A$. A subset can have many upper bounds. Similarly the subset A is "bounded below" if there exists an element m we call a "lower bound of A" such that $m \leq x$ for all $x \in A$. For example, every non-empty subset of $(\mathbb{N}, \in)$ is bounded below by 0.

Suppose A is a non-empty subset of a linearly ordered set $(S, \leq)$ which contains an upper bound M of A. Then, since A is linearly ordered, for every element $x \in A$, $x \leq M$. Furthermore, M is the only upper bound of A which is contained in A, for if M^* is another upper bound of A which is contained in A, then $M^* \leq M$; if $M^* < M$,

then M^* is not an upper bound of A; so $M^* = M$. In this case, we can refer to M as being *the maximal element of* A or *the maximum element of* A. This corresponds to the definition we have previously provided for the words "maximal element" and "maximum element" of an ordered set. For linearly ordered sets, in this context, the words "maximal" and "maximum" are interchangeable. Similarly, if A contains a lower bound m, then m is the (unique) minimum element of A.

The following theorem shows that any non-empty bounded subset of $\mathbb{N}$ must contain a maximal element with respect to "$\in$".

Theorem 14.5 Any bounded non-empty subset S of $(\mathbb{N}, \in)$ has a maximal element.

Proof:

Suppose S is a non-empty bounded subset of $\mathbb{N}$ with respect to the linear ordering "$\in$". Let Q be the set of all upper bounds of S. Since S is bounded, by definition, it has at least one upper bound and so $Q \neq \varnothing$. Since "$\in$" well-orders $\mathbb{N}$, Q must contain a "least element", say k. We claim that k is a maximal element of S. To prove this claim it suffices to show that k is both an upper bound of S and belongs to S. Since $k \in Q$, then k is an upper bound of S. We now show that $k \in S$: Suppose $k \notin S$. Let t be the unique immediate predecessor of k. That is, $t^+ = k$. Then if $x \in S$,

$$\begin{aligned} k \notin S &\Rightarrow x \neq k, \\ &\Rightarrow x \in k \\ &\Rightarrow x \in t^+ \end{aligned}$$

$$\begin{aligned} x \in t^+ &\Rightarrow x \in (t \cup \{t\}) \\ &\Rightarrow x \in t \text{ or } x \in \{t\} \\ &\Rightarrow x \in t \text{ or } x = t \end{aligned}$$

We have shown that for every $x \in S$, $x \in_= t$; hence, t is an upper bound of S. That is, $t \in Q$. But $t \in t^+ = k$ where k was declared to be the least element in the set Q. This is a contradiction. Then $k \in S$ as claimed. So S has a maximal element.

14.4 Constructing other well-ordered sets from $\mathbb{N}$

We have seen that the order relation "$\in$" is a strict well-ordering of the set of all natural numbers, $\mathbb{N}$. Now that we have given ourselves a large set to work with we will use the set $\mathbb{N}$ as a building block to construct other large well-ordered sets. We will introduce an order relation on various Cartesian products involving $\mathbb{N}$. This particular order relation is defined in terms of the ordering "$\in$" on $\mathbb{N}$.

Lexicographic ordering of the Cartesian product $\{1,2\}\times\mathbb{N}$. Consider the subset $\{1,2\}\times\mathbb{N} = \{(i,n) : i = 1 \text{ or } 2, n \in \mathbb{N}\}$ of $\mathbb{N}\times\mathbb{N}$. We will order the elements of $\{1,2\}\times\mathbb{N}$ by what is called a *lexicographic ordering*[1], denoted by $<_{\text{lex}}$. This means $(a,b) <_{\text{lex}} (c,d)$ if $a \in c$ or, if $a = c$, $b \in d$. For example, $(1,34) <_{\text{lex}} (2,7)$ and $(2,5) <_{\text{lex}} (2,54)$. The elements of $(\{1,2\}\times\mathbb{N}, <_{\text{lex}})$ can then be listed in a strictly increasing order as follows:

$$\{(1,0),(1,1),(1,2),(1,3),\cdots,(2,0),(2,1),(2,2),(2,3),\cdots,(3,0),(3,1),\cdots,\}$$

The lexicographic ordering inherited from "$\in$" is easily seen to be a linear ordering of $\{1,2\}\times\mathbb{N}$. We now investigate specific ordering properties of $(\{1,2\}\times\mathbb{N}, <_{\text{lex}})$.

a) *The order relation* $<_{lex}$ *on* $\{1,2\}\times\mathbb{N}$ *is a well-ordering.* Consider a non-empty subset A of $\{1,2\}\times\mathbb{N}$. If there exists elements of the form $(1,a)$ in A, then the least element of A is $(1,m)$ where m is the least element of $\{n \in \mathbb{N} : (1,n) \in A\}$, guaranteed to exist since $\in$ well-orders $\mathbb{N}$. If all elements of A are of the form $(2,b)$, then the least element of A is $(2,m)$ where m is the least element of $\{n \in \mathbb{N} : (2,n) \in A\}$, again guaranteed to exist since $\in$ well-orders $\mathbb{N}$. Then $(\{1,2\}\times\mathbb{N}, <_{\text{lex}})$ is a well-ordered set.

b) *What are bounded and unbounded subsets of* $\{1,2\}\times\mathbb{N}$ *with respect to* $<_{lex}$ *like?* As examples we consider the following subsets of $\{1,2\}\times\mathbb{N}$.
   - If $S \subseteq \{1\}\times\mathbb{N}$, then any element which has the form $(2,n)$ is an upper bound of S with respect to $<_{\text{lex}}$
   - If T is a bounded subset of $\mathbb{N}$, then there exists a natural number m such that $t \in m$ for all $t \in T$. Then $(2,m)$ would be an upper bound of any subset of $\{1,2\}\times T \subset \{1,2\}\times\mathbb{N}$.
   - There is no natural number n which is an upper bound of $\mathbb{N}$. Then no element of $\{1,2\}\times\mathbb{N}$ is an upper bound of $\{1,2\}\times\mathbb{N}$. In this case we say that the subset $\{2\}\times\mathbb{N}$ is *unbounded* in $\{1,2\}\times\mathbb{N}$.

c) *On maximal elements of subsets of* $(\{1,2\}\times\mathbb{N}, <_{lex})$. We have seen that every bounded subset S of $\mathbb{N}$ has a maximal element. Does every bounded subset of $\{1,2\}\times\mathbb{N}$ contain a maximal element? The subset $T = \{(1,n) : n \in \mathbb{N}\}$ of $\{1,2\}\times\mathbb{N}$ is seen to have an upper bound $(2,0)$. In fact, this element $(2,0)$ is the smallest upper bound of T since any smaller element would be of the form $(1,n)$. But since $(2,0)$ does not belong to T, then T does not contain a maximal element. So the lexicographically ordered set $\{1,2\}\times\mathbb{N}$ possesses at least one ordering properties which are *not* shared by $\mathbb{N}$.

Lexicographic ordering can be used to well-order other sets such as

$$S = \{0,1,2,3\}\times\mathbb{N} \subset \mathbb{N}\times\mathbb{N}$$

[1]Some texts may refer to this as the "dictionary ordering".

This well-ordered set can be visualized as four copies of $\mathbb{N}$ lined up one after the other. The first copy is of the form $\{(0,n) : n \in \mathbb{N}\}$ and the fourth copy is of the form $\{(3,n) : n \in \mathbb{N}\}$. The set $\mathbb{N} \times \mathbb{N} = \{(a_1, a_2) : a_1, a_2 \in \mathbb{N}\}$ can also be ordered lexicographically. When ordered in this way $\mathbb{N} \times \mathbb{N}$ can be viewed as $\{0\} \times \mathbb{N}$ followed by $\{1\} \times \mathbb{N}$ followed by $\{2\} \times \mathbb{N}$, and so on.

14.5 Constructing non-well-ordered sets from $\mathbb{N}$

The sets we will now consider may seem less familiar to many readers who are not yet accustomed to viewing functions as "sets" of ordered pairs. We will in fact go a step further and discuss sets of functions whose domain is $\mathbb{N}$.

Let's consider the set of all functions mapping $\mathbb{N}$ into $\{1,2\}$. This set is normally denoted as $\{1,2\}^{\mathbb{N}}$. An element f of $\{1,2\}^{\mathbb{N}}$ is a set of ordered pairs (n,m) where n takes on any number in $\mathbb{N}$ and m is either the natural number 1 or the natural number 2. Furthermore, only one value m can be the image of a value n in $\mathbb{N}$ under a function f. We can express a function $f : \mathbb{N} \to \{1,2\}$ as the set

$$f = \{(n,m) \in \mathbb{N} \times \mathbb{N} : m = f(n) = 1 \text{ or } 2\}$$

where $[(n,m) \in f \text{ and } (n,k) \in f] \Rightarrow [m = k]$. We see that f is a subset of $\mathbb{N} \times \{1,2\}$ (not an element of $\mathbb{N} \times \{1,2\}$). That is, $f \in \mathscr{P}(\mathbb{N} \times \{1,2\})$. Then any specific function f in $\{1,2\}^{\mathbb{N}}$ can be expressed as an infinite set of ordered pairs

$$\{(0,a_0),(1,a_1),(2,a_2),(3,a_3),\ldots,\}$$

where $a_i = f(i) = 1$ or 2. Notice that we have surreptitiously imposed a lexicographic ordering on the elements of f, since $(n,a_n) < (m,a_m)$ if $n < m$. Actually we could more succinctly express this element f as a sequence

$$\{a_0, a_1, a_2, a_3, \ldots,\}$$

where each $a_i = f(i)$ is the image of i under f.[1] So if $f \in \{1,2\}^{\mathbb{N}}$, f can accurately be described as an infinite sequence of 1's and 2's in an order dictated by the associated elements in the domain of f. This is the particular way we will view the elements of $\{1,2\}^{\mathbb{N}}$. That is,

$$\{1,2\}^{\mathbb{N}} = \{\{a_i\}_{i=0}^{\infty} : \ a_i = 1 \text{ or } 2\}$$

So comparing two functions f and g in $\{1,2\}^{\mathbb{N}}$ is essentially comparing two infinite sequences of 1's and 2's. Given the set $S = \{1,2\}^{\mathbb{N}}$ we want to define on order relation on S. The order relation that we will choose is inspired by the lexicographic ordering defined on subsets of $\mathbb{N} \times \mathbb{N}$ above. We formally define it below.[2]

---

[1] So a_i is in fact shorthand for $(i, f(i))$.

[2] Even though the following definition of the ordering on the set $\{1,2\}^{\mathbb{N}}$ is inspired from the lexicographic ordering of sets of ordered pairs and adopts the notation $<_{\text{lex}}$, it is good to remember that we are not ordering ordered pairs but sets which represent functions.

Definition 14.6 Consider the set $\{1,2\}^{\mathbb{N}}$ of all functions mapping natural numbers to 1 or 2. We define the *lexicographic order* "$<_{\text{lex}}$" on $\{1,2\}^{\mathbb{N}}$ as follows: For any two elements $f = \{a_0, a_1, a_2, a_3, \ldots,\}$ and $g = \{b_0, b_1, b_2, b_3, \ldots,\}$ in $\{1,2\}^{\mathbb{N}}$, $f <_{\text{lex}} g$ if and only if for the first two unequal corresponding terms a_i and b_i, $a_i \in b_i$. Also, $f = g$ if and only if $a_i = b_i$ for all $i \in \mathbb{N}$.[3]

For example, if $f = \{1,2,2,1,1,\cdots,\}$ and $g = \{1,2,2,2,1,2,\cdots\}$, then $f <_{\text{lex}} g$. This ordering is easily seen to be linear. We now investigate basic ordering properties of $(\{1,2\}^{\mathbb{N}}, <_{\text{lex}})$.

a) *The relation $<_{lex}$ is* not *a well-ordering of* $\{1,2\}^{\mathbb{N}}$. At first glance, based on our experience with lexicographic orderings, we may suspect that $<_{\text{lex}}$ well-orders $\{1,2\}^{\mathbb{N}}$. But we should be cautious. Does $\{1,2\}^{\mathbb{N}}$ have a least element? The lexicographic ordering rule shows that no element in $\{1,2\}^{\mathbb{N}}$ can be smaller than $\{1,1,1,1,\cdots\}$. So $\{1,2\}^{\mathbb{N}}$ at least has a smallest element. Let's try to think about what its second smallest element is. We have listed the elements of a subset of $\{1,2\}^{\mathbb{N}}$ in the form of a strictly *decreasing* sequence of elements where each element is larger than $\{1,1,1,1,1,1,1,1,1,1,\cdots\}$:

$$\begin{array}{c}
\{2,1,1,1,1,1,1,1,1,1,\cdots\} \\
\{1,2,1,1,1,1,1,1,1,1,\cdots\} \\
\{1,1,2,1,1,1,1,1,1,1,\cdots\} \\
\{1,1,1,2,1,1,1,1,1,1,\cdots\} \\
\{1,1,1,1,2,1,1,1,1,1,\cdots\} \\
\{1,1,1,1,1,2,1,1,1,1,\cdots\} \\
\{1,1,1,1,1,1,2,1,1,1,\cdots\} \\
\vdots
\end{array}$$

We see that each element is strictly less than its immediate predecessor in this list. We also see that we can never reach $\{1,1,1,1,1,1,1,1,1,1,\cdots\}$ using such a decreasing sequence. This convinces us that the set

$$S = \{f \in \{1,2\}^{\mathbb{N}} : f > \{1,1,1,1,1,1,1,1,\cdots\}\}$$

does not contain its least element since no matter where we insert our first "2" you will be able to insert a "2" further down. Since the non-empty subset S has no least element with respect to the ordering "$<_{\text{lex}}$", then $\{1,2\}^{\mathbb{N}}$ is not a well-ordered set.

b) *The set $\{1,2\}^{\mathbb{N}}$ is bounded.* We easily see that $\{2,2,2,2,\cdots\}$ is a maximal element of $\{1,2\}^{\mathbb{N}}$. So every subset of $\{1,2\}^{\mathbb{N}}$ has at least $\{2,2,2,2,\cdots\}$ as upper bound.

c) *On maximal elements of bounded sets.* Does every bounded subset of $\{1,2\}^{\mathbb{N}}$ contain a maximal element? To help answer this question let's try to find the element

[3] A lexicographic ordering can similarly be defined on $S^{\mathbb{N}}$ where S is any subset of $\mathbb{N}$.

of $\{1,2\}^{\mathbb{N}}$ which immediately precedes $\{2,2,2,2,\cdots\}$. Another way of stating this is: What is the maximal element of $S = \{f \in \{1,2\}^{\mathbb{N}} : f < \{2,2,2,2,\cdots\}\}$? This maximal element must have at least one "1" in it, along with as many 2's as possible. The question is where shall we insert this "1"? No matter where we insert this "1" we will be able to reconsider our choice and reinsert it farther down. So S contains no maximal element.

14.6 The set $\mathbb{N}^{\{1,2\}}$.

Having studied the set $\{1,2\}^{\mathbb{N}}$ of all functions whose domain is $\mathbb{N}$ and codomain is $\{1,2\}$ we now consider the set of all functions with domain $\{1,2\}$ and codomain $\mathbb{N}$. This set is denoted as $\mathbb{N}^{\{1,2\}}$. Of course, if $f \in \mathbb{N}^{\{1,2\}}$, then $f = \{(1,a_1),(2,a_2)\}$ where $a_1 = f(1) \in \mathbb{N}$ and $a_2 = f(2) \in \mathbb{N}$. See that we have purposely lexicographically ordered the two elements $(1,a_1)$ and $(2,a_2)$ of f. Just as for the functions in $\{1,2\}^{\mathbb{N}}$ we can more succinctly represent $f \in \mathbb{N}^{\{1,2\}}$ as $f = \{a_1,a_2\}$, ordered doubletons of natural numbers. That is,

$$\mathbb{N}^{\{1,2\}} = \{\{a_1,a_2\} : a_1,a_2 \in \mathbb{N}\}$$

is a set of ordered doubletons. Since these doubletons are ordered natural numbers we can lexicographically order the elements $\{a_1,a_2\}$ of $\mathbb{N}^{\{1,2\}}$ as if $\mathbb{N}^{\{1,2\}}$ was a set of ordered pairs (a_1,a_2) of natural numbers. That is, a function $\phi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}^{\{1,2\}}$ defined as $\phi((a_1,a_2)) = \{a_1,a_2\}$ can be used to lexicographically order the elements of $\mathbb{N}^{\{1,2\}}$ in such a way that the elements of $\mathbb{N}^{\{1,2\}}$ are ordered in precisely the same way as the elements of $\mathbb{N} \times \mathbb{N}$. This ordering on $\mathbb{N}^{\{1,2\}}$ is also referred to as a lexicographic ordering and is also denoted by the same symbol, "$<_{\text{lex}}$". So $(\mathbb{N}^{\{1,2\}}, <_{\text{lex}})$ has order properties which are identical to those of $(\mathbb{N} \times \mathbb{N}, <_{\text{lex}})$.

Concepts review:

1. What does it mean to say that "$<$" strictly well-orders a set S?
2. Describe two order relations which well-order $\mathbb{N}$?
3. What does it mean to say a subset S of $\mathbb{N}$ ordered by "$\in$" is bounded?
4. What does it mean to say that a non-empty subset S of $\mathbb{N}$ ordered by "$\in$" has a maximal element?
5. Describe the set $\{1,2\} \times \mathbb{N}$ by providing three distinct elements of this set.
6. Define the lexicographic ordering on $\{1,2\} \times \mathbb{N}$.

7. Define the lexicographic ordering on $\{1,2\}^{\mathbb{N}}$.
8. List three elements of $\{1,2\}^{\mathbb{N}}$ in increasing order.
9. Is the lexicographically ordered set $\{1,2\} \times \mathbb{N}$ well-ordered?
10. Is the lexicographically ordered set $\{1,2\}^{\mathbb{N}}$ well-ordered?
11. Does every non-empty subset of the lexicographically ordered set $\{1,2\} \times \mathbb{N}$ have a maximal element?
12. Does every non-empty subset of the lexicographically ordered set $\{1,2\}^{\mathbb{N}}$ have a maximal element?
13. Does $\{1,2\}^{\mathbb{N}}$ have a maximal element?
14. Describe the elements of $\mathbb{N}^{\{1,2\}}$. Propose an ordering for its elements.

EXERCISES

A.

1. Show that $(\mathbb{N}, \in)$ does not contain a maximal element.
2. Can a non-empty bounded subset S of $\mathbb{N}$ ordered by "$\in$" be an inductive set? Why or why not?
3. Construct three non-empty subsets of $\{1,2\}^{\mathbb{N}}$ each of which contains no least element.
4. Consider the lexicographically ordered set $(\mathbb{N} \times \mathbb{N}, <_{\text{lex}})$.
   a) Describe the first few elements of $(\mathbb{N} \times \mathbb{N}, <_{\text{lex}})$.
   b) Does $\mathbb{N} \times \mathbb{N}$ have a maximal element? What is it?
   c) Is $(\mathbb{N} \times \mathbb{N}, <_{\text{lex}})$ a well-ordered set? If so show it. If not produce a non-empty subset which does not contain its least element.
   d) Does every bounded subset of $\mathbb{N} \times \mathbb{N}$ have a maximal element? Why?
5. Consider the lexicographically ordered set $S = \{1,2,3,\cdots,9\}^{\mathbb{N}}$ of all functions mapping $\mathbb{N}$ to the set $\{1,2,3.\cdots,9\}$.
   a) Describe the first few elements of S.
   b) Write the three elements $f = \{5,5,5,5,5,\ldots,\}$, $g = \{4,9,2,2,2,\ldots,\}$ and $h = \{4,9,1,9,9,9,\ldots\}$ in increasing order.
   c) Does S have a maximal element? What is it?
   d) Is S a well-ordered set? If so show it. If not produce a non-empty subset which does not contain its least element.
   e) Does every bounded subset of S have a maximal element? Why?

6. Suppose we represent the set of all functions mapping $\{0, 1, 2\}$ to $\mathbb{N}$ as

$$\mathbb{N}^{\{0,1,2\}} = \{\{a_0, a_1, a_2\} : a_i \in \mathbb{N}\}$$

Suppose we order this set lexicographically.

a) Describe the first few elements of $\mathbb{N}^{\{0,1,2\}}$.

b) Write the three elements $f = \{4, 0, 600\}$, $g = \{600, 9, 8\}$ and $h = \{6, 9, 53\}$ in increasing order.

c) Does $\mathbb{N}^{\{0,1,2\}}$ have a maximal element? What is it?

d) Is $\mathbb{N}^{\{0,1,2\}}$ a well-ordered set? If so show it. If not produce a non-empty subset which does not contain its least element.

e) Does every bounded subset of $\mathbb{N}^{\{0,1,2\}}$ have a maximal element? Why?

C. 7. Consider the lexicographically ordered set $\mathbb{N}^{\mathbb{N}}$ of all functions mapping $\mathbb{N}$ into $\mathbb{N}$.

a) Describe an element of $\mathbb{N}^{\mathbb{N}}$ as a set.

b) Does $\mathbb{N}^{\mathbb{N}}$ have a least element? What is it?

c) Does $\mathbb{N}^{\mathbb{N}}$ have a second element with respect to the lexicographic ordering? What is it?

d) If $f \in \mathbb{N}^{\mathbb{N}}$, can f be viewed as a subset of $\mathbb{N} \times \mathbb{N}$? Explain.

e) If $f \in \mathbb{N}^{\mathbb{N}}$, can f be viewed as an element of $\mathscr{P}(\mathbb{N} \times \mathbb{N})$? Explain.

f) Can $\mathbb{N}^{\mathbb{N}}$ be viewed as a subset of $\mathscr{P}(\mathbb{N} \times \mathbb{N})$? Explain.

15 / Arithmetic of the natural numbers.

Summary. *In this section we define addition, subtraction and multiplication of the natural numbers in a set-theoretic context. We then show that with these definitions, we obtain the expected results. Addition and multiplication on* $\mathbb{N}$ *are defined recursively.*

15.1 Highlights of what we have learned in set theory up to now.

Now that we have defined the *natural numbers* within the framework of set theory it is a good time to look back on what we have learned and provide some insight on what is to come.

In our theory, all objects are *classes* or *sets*. Since we are mainly interested in those objects called "sets", our attention is directed towards these specifically. A few fundamental properties of classes and sets are stated without proof in the form of *axioms*. Axioms are normally expected to be intuitively obvious to users. This does not mean that choosing axioms is done without debate, since what is obvious to one may not be obvious at all to another. We listed ten set-theoretic axioms referred to as the "*ZFC*-axioms". The ZF-axioms refer to the first nine, while the "C" refers to the tenth axiom called "Axiom of choice". The *Axiom of infinity* which postulates the existence of a set which satisfies the essential properties of the natural numbers is not an axiom which is "intuitively obvious". It is however perceived as being essential since it provides a logical basis on which rests most of the mathematics we do today. The most controversial axiom is the *Axiom of choice*. We have not invoked this axiom yet. We will soon see why it is needed if we want our set-theoretic universe to unfold as we think it should.

Once we gave ourselves classes, sets and a few axioms to work with, we gave ourselves the means to construct classes and sets from ones that exist (union, intersection, Cartesian products). This was followed by definitions of relations and functions both of which exist in our set-theoretic universe as sets. Finally, we defined the natural numbers so that they possessed the required well-ordering property. Of course, giving life to the natural numbers is just the beginning. Our next step is to appropriately define addition and multiplication of the natural numbers. Definitions must be such that results we obtain with these operations are what we expect them to be. Definitions of the *integers*, the *rational numbers* and the *real numbers* will then follow. Finally, we will study infinite sets and their various properties.

15.2 Defining addition recursively.

Since natural numbers are sets, we may instinctively attempt to define addition of the natural numbers as follows: $3+5=3\cup 5$. But this doesn't produce the desired result

since we know that $3 \subset 5$ and so $3 + 5 = 3 \cup 5$ would equal 5, not what we want at all. We see that viewing addition of natural numbers by simply joining sets together will not work, particularly if the sets have non-empty intersection. It may be that we are attempting to define too much at the same time. We should maybe try a step by step definition of addition. We will experiment with addition of natural numbers with the natural number "3", specifically. Suppose we define addition of 0 to 3 as follows: $3+0 = 3 \cup \varnothing = 3$. Then we progressively define addition with 3 of successively larger and larger numbers. In what follows the symbol ":=" will serve as a succinct way of saying "is defined as".

$$\begin{aligned}
3+0 &= 3 \\
3+1 &= 3+0^+ := (3+0)^+ = 3^+ = 4 \\
3+2 &= 3+1^+ := (3+1)^+ = 4^+ = 5 \\
&\vdots \\
3+n &= m \\
3+n^+ &:= (3+n)^+ = m^+ \\
&\vdots
\end{aligned}$$

In this sequence of sums only the initial sum $3 + 0 = 3$ is specifically defined while a globally defined rule "$3 + n^+ = (3 + n)^+$" is applied to evaluate the sum of all other natural numbers with 3. If we know the numerical value of $3+n$, then, by the general "rule", $3 + n^+$ has numerical value $(3 + n)^+$. So if we want to determine the value of $3 + 4$ we first have to find the values of $3 + 1$, $3 + 2$ and $3 + 3$.[1]

$$\begin{aligned}
3+4 &= (3+3)^+ \\
&= ((3+2)^+)^+ \\
&= (((3+1)^+)^+)^+ \\
&= ((((3+0)^+)^+)^+)^+ \\
&= (((3^+)^+)^+)^+ \\
&= ((4^+)^+)^+ \\
&= (5^+)^+ \\
&= 6^+ \\
&= 7
\end{aligned}$$

It is a sure (albeit tedious) method of addition that will consistently produce the unique expected values for sums of two natural numbers. For example, once we have computed $3 + 4 = 7$ we can then compute $3+5 = (3+4)^+ = 7^+ = 8$ by applying the algorithm $3+n^+ = (3+n)^+$. Once all values of $3+n$ are obtained we then obtain all values for $4 + n$ as n ranges over $\mathbb{N}$, and so on. This simply shows that it is possible

[1] This is reminiscent of the way we learned addition by using addition tables in elementary school: Before learning that $3+2 = 5$, we learned that $3+0 = 3$, $3+1 = 4$ and deduced from this that $3 + 2$ had to equal 5.

to adequately define addition of natural numbers in a universe of sets in such a way that sums correspond to sums obtained by usual addition algorithms. This does not prevent us from using the various algorithms which allow us to obtain more efficiently the values of sums of natural numbers.

We propose the following formal definition of addition of pairs of natural numbers.

Definition 15.1 Let m be a fixed natural number. Addition of a natural number n with m is defined as the function $r_m : \mathbb{N} \to \mathbb{N}$ satisfying the two conditions

$$\begin{aligned} r_m(0) &= m \\ r_m(n^+) &= [r_m(n)]^+ \end{aligned}$$

The expression $m + n$ is another way of writing $r_m(n)$. Thus,

$$r_m(0) = m \quad \Leftrightarrow \quad m + 0 = m \tag{1}$$

$$r_m(n^+) = [r_m(n)]^+ \quad \Leftrightarrow \quad m + n^+ = (m + n)^+ \tag{2}$$

For example, once the value of $34 + 0 = 34$ is declared, the value of the sum $34 + 123$ is uniquely determined by applying the formula $34 + n^+ = (34 + n)^+$ finitely many times to successively obtain $34 + 1 = 35$, $34 + 2 = 36$, $34 + 3 = 37, \ldots, 34 + 123 = 157$.

Readers may have no doubt noticed that the function $r_m(n)$ is defined by using a mechanism that we have not used or seen before in this text. We are accustomed to defining a function $f : A \to A$ by declaring a rule which associates to each element in A some other element in A without referring to other ordered pairs $(a, f(a))$ in f. For example, the only way we can confirm that the ordered pair $(2, 3 + 2) = (2, 5)$ belongs to the function r_3 is by first determining that $(0, 3 + 0) = (0, 3)$ and $(1, 3 + 1) = (1, 4)$ also belong to r_3. Most readers will intuitively feel that there is no ambiguity in the way we have defined the function r_m on $\mathbb{N}$. We refer to functions which are defined in this way as being *recursively defined functions*. If r_m is indeed a well-defined function, then we must be able to prove that it satisfies the conditions stated in the formal definition of a function. We remind ourselves how we defined a "function" (see Definition 9.1): Given two sets A and B, a function is a subset f of $A \times B$ which satisfies the property "(x, y) and (x, z) belong to f implies $y = z$". There is no reason to deviate from this understanding of functions. We will now formally show that this recursively defined function of addition satisfies the property described in the definition. That is, we will show that if $n = k$, then $r_m(n) = m + n = m + k = r_m(k)$.

Theorem 15.2 Let m be a fixed natural number and let $r_m : \mathbb{N} \to \mathbb{N}$ be a relation satisfying the two properties

$$\begin{cases} r_m(0) &= m \\ r_m(n^+) &= [r_m(n)]^+ \end{cases}$$

Then r_m is a well-defined function on $\mathbb{N}$.

Proof:

Let m be a fixed natural number. Let $\mathscr{S}$ be a class of relations on $\mathbb{N}$ defined as follows:

$$\mathscr{S} = \{R \subseteq \mathbb{N} \times \mathbb{N} : (0, m) \in R \text{ and } (n, y) \in R \Rightarrow (n^+, y^+) \in R\}$$

Now $\mathscr{S}$ is non-empty since it contains $\mathbb{N} \times \mathbb{N}$. Let $r^* = \bigcap_{R \in \mathscr{S}} R$. This means that r^* is the smallest set of ordered pairs satisfying the conditions described for $\mathscr{S}$. The relation r^* looks something like

$$r^* = \{(0, m), (1, m^+), (2, (m^+)^+), \cdots\}$$

We will show two things: 1) That r^* is a function mapping $\mathbb{N}$ into $\mathbb{N}$, and, 2) That $r* = r_m$.

1) *We claim that r^* is a function mapping $\mathbb{N}$ into $\mathbb{N}$.*

Proof of claim:
We first establish (by induction) that $\operatorname{dom} r^* = \mathbb{N}$.

Base case: We first note that since $(0, m) \in r^*$, then $0 \in \operatorname{dom} r^*$.
Inductive hypothesis: Suppose $n \in \operatorname{dom} r^*$. Then $(n, y) \in r^*$ for some $y \in \mathbb{N}$. This implies $(n^+, m^+) \in r^*$ and so $n^+ \in \operatorname{dom} r^*$.
Hence, by induction, the domain of r^* is all of $\mathbb{N}$.

We now proceed to the proof of the claim. The proof of the claim invokes the second version of the principle of mathematical induction. Let $P(n)$ denote the statement "$[(n, x) \in r^* \wedge (n, y) \in r^*] \Rightarrow [x = y]$".

Inductive hypothesis: Suppose $P(m)$ holds true all natural numbers $m < n$. That is, $(m, x) \in r^*$ and $(m, y) \in r^*$ implies $x = y$. We will show that given our hypothesis, $P(n)$ must hold true.

> Suppose not. Suppose $(n, x) \in r^*$ and $(n, y) \in r^*$ where $x \neq y$. Let $U = r^* - \{(n, y)\}$ (the set r^* take away the element (n, y)). Then U is still a relation belonging to $\mathscr{S}$ which is strictly smaller than r^*. But r^* was previously declared to be the smallest of the relations in $\mathscr{S}$. We have a contradiction whose source is our assumption that $x \neq y$. Then x must be equal to y. We conclude that $P(n)$ holds true as required.

By mathematical induction (version two) $P(n)$ holds true for all n.

So r^* is a well-defined function as claimed.

2) *We now claim that the function r^* is the relation r_m as defined above.*

Proof of claim:

The proof is by induction. Let $P(n)$ denote the statement "$r^*(n) = r_m(n)$".

Base case: The statement $P(0)$ holds true since $r^*(0) = m = r_m(0) = m + 0$. *Inductive hypothesis*: Suppose $P(n)$ holds true. That is, suppose $(n, r_m(n)) \in r^*$. Then, by definition of r^*, $(n^+, r_m(n)^+) \in r^*$. Since $r_m(n^+) = r_m(n)^+$, then $(n^+, r_m(n^+)) \in r^*$. That is, $r^*(n^+) = r_m(n^+)$. So $P(n^+)$ holds true.

By mathematical induction $r^*(n) = r_m(n)$ for all n.

Then $r^* = r_m$ as claimed.

We conclude that r_m and r^* are indeed the same relation. Since r^* was shown to be a function, the recursively defined relation $r_m : \mathbb{N} \to \mathbb{N}$ is a well-defined function.

15.3 Basic properties of addition.

We must now be sure that the addition operation we have defined on $\mathbb{N}$ satisfies the basic properties of addition we are accustomed to.

1) *For every natural number n, n^+ and $n+1$ are the same number.* We are required to show that assuming 0^+ is denoted by the symbol 1, then for any non-zero natural number n,

$$n^+ = 1 + n$$

This can be shown by induction:

Proof: By induction. Let $P(n)$ be the property "$n^+ = 1 + n$" (where $1 = 0^+$).
Base case: We see that $P(0)$ holds true since

$$\begin{aligned} r_{0^+}(0) &= 0^+ + 0 = 0^+ \quad \text{(By (1) in the definition of addition above).} \\ 0^+ &= 1 + 0 \quad \text{(By notation).} \end{aligned}$$

Inductive hypothesis: Suppose $P(n)$ holds true for some n. Then

$$(n^+)^+ = (1 + n)^+ = 1 + n^+ \quad \text{(By (2) in the definition of addition above).}$$

So $P(n^+)$ holds true. By mathematical induction $n^+ = 1 + n$ for all natural numbers n.

2) For any natural number n,

$$0 + n = n$$

Proof: By induction. Let $P(n)$ be the property "$0 + n = n$".
Base case: We see that $P(0)$ holds true since $r_0(0) = 0 + 0 = 0$, by (1) in the definition of addition above.
Inductive hypothesis: Suppose $P(n)$ holds true. Then $0 + n^+ = (0 + n)^+ = n^+$, by (2) in the definition of addition. So $P(n^+)$ holds true. By mathematical induction $0 + n = n$ for all natural numbers n.

3) *Addition of the natural numbers is associative.* That is, for any three natural numbers m, n and k

$$(m + k) + n = m + (k + n)$$

Proof: By induction. Let m and k be any two natural numbers. Let $P(n)$ be the property "$(m + k) + n = m + (k + n)$".
Base case: We see that $P(0)$ holds true since $(m + k) + 0 = m + k = m + (k + 0)$, (by (1) in the definition of addition above).
Inductive hypothesis: Suppose $P(n)$ holds true. Then

$$\begin{aligned}
(m + k) + n^+ &= [(m + k) + n]^+ \text{ (By (2) in the definition of addition above.)} \\
&= [m + (k + n)]^+ \text{ (Since } P(n) \text{ holds true.)} \\
&= m + (k + n)^+ \\
&= m + (k + n^+)
\end{aligned}$$

So $P(n^+)$ holds true. By mathematical induction $(m+k)+n = m+(k+n)$ for all natural numbers n. Since m and k were arbitrarily chosen, then $(m+k)+n = m+(k+n)$ holds true for any three natural numbers m, n and k.

4) *Addition of the natural numbers is commutative.* That is, for any two natural numbers m and n

$$m + n = n + m$$

Proof: By induction. Let m be any natural number. Let $P(n)$ be the property "$m + n = n + m$".
Base case: We see that $P(0)$ holds true since

$$\begin{aligned}
m + 0 &= m \text{ (By (1) in the definition above.)} \\
&= 0 + m \text{ (By Property 2).}
\end{aligned}$$

Induction hypothesis: Suppose $P(n)$ holds true. Then

$$\begin{aligned}
m + n^+ &= (m + n)^+ \text{ (By (2) in the definition above.)} \\
&= (n + m)^+ \text{ (Since } P(n) \text{ holds true.)} \\
&= 1 + (n + m) \text{ (By property 1).)} \\
&= (1 + n) + m \text{ (By property 3).)} \\
&= n^+ + m \text{ (By property 1).)}
\end{aligned}$$

So $P(n^+)$ holds true; by mathematical induction $m + n = n + m$ for all natural numbers n. Since m was arbitrarily chosen, then $m+n = n+m$ holds true for any pair of natural numbers m and n.

15.4 Definition of multiplication on $\mathbb{N}$.

As for addition, multiplication will be inductively defined.

Definition 15.3 For any natural number m, multiplication with the natural number m is defined as the function $s_m : \mathbb{N} \to \mathbb{N}$ satisfying the two conditions

$$\begin{aligned} s_m(0) &= 0 \\ s_m(n^+) &= s_m(n) + m \end{aligned}$$

We define the expression mn and $m \times n$ as alternate ways of writing $s_m(n)$. Thus

$$s_m(0) = 0 \quad \Leftrightarrow \quad m0 = m \times 0 = 0 \tag{3}$$

$$s_m(n^+) = s_m(n) + m \quad \Leftrightarrow \quad mn^+ = mn + m = m \times n + m \tag{4}$$

Theorem 15.4 Let m be a fixed natural number and let $s_m : \mathbb{N} \to \mathbb{N}$ be a function satisfying the two properties

$$\begin{cases} s_m(0) &= 0 \\ s_m(n^+) &= s_m(n) + m \end{cases}$$

Then s_m is a well-defined function on $\mathbb{N}$.

Proof: (*Outline*)

Let $\mathscr{S}$ be a class of relations on $\mathbb{N}$ defined as follows:

$$\mathscr{S} = \{R \subseteq \mathbb{N} \times \mathbb{N} : (0,0) \in R \text{ and } (n,y) \in R \Rightarrow (n^+, y+m) \in R\}$$

Now $\mathscr{S}$ is non-empty since it contains $\mathbb{N} \times \mathbb{N}$. Let $s^* = \bigcap_{R \in \mathscr{S}} R$. This means that s^* is the smallest set of ordered pairs satisfying the conditions described for $\mathscr{S}$. The relation s^* looks something like

$$s^* = \{(0,0), (1,m), (2, m+m), (3, m+m+m), \cdots, (n, m \times n), \cdots, \}$$

Claim: *s^* is a function mapping $\mathbb{N}$ into $\mathbb{N}$*. Proof of the claim is left as an exercise.

Claim: *s^* satisfies the properties which characterize s_m*. Proof of the claim is left as an exercise.

We conclude that s_m and s^* are indeed the same relation. Since s^* was shown to be a function, the recursively defined function $s_m : \mathbb{N} \to \mathbb{N}$ is well-defined.

We now verify that the expected properties of multiplication are satisfied.

1) For any natural number
$$0n = 0$$
Proof: By induction. It is left as an exercise.

2) For any natural number n,
$$1n = n$$
Proof: By induction. It is left as an exercise.

3) *Multiplication of the natural numbers is distributive over addition.* That is, for any three natural numbers m, n and k
$$n(m+k) = nm + nk \text{ and } (m+k)n = mn + kn$$
Proof: By induction. Outline of proof for left-hand distribution. Right-hand distribution is left as an exercise. Let m and k be any two natural numbers. Let $P(n)$ be the property "$k(m+n) = km + kn$".
Base case: We see that $P(0)$ holds true since
$$\begin{aligned} k(m+0) &= km \\ &= km + 0 \\ &= km + k0 \end{aligned}$$
Inductive hypothesis: If $P(n)$ holds true, then
$$\begin{aligned} k(m+n^+) &= k(m+n)^+ \\ &= k(m+n) + k \text{ (By 2) in the definition.)} \\ &= km + kn + k \text{ (Since } P(n) \text{ holds true.)} \\ &= km + kn^+ \text{ (By 2) in the definition.)} \end{aligned}$$
So $P(n^+)$ holds true. By mathematical induction $k(m+n) = km+kn$ for all natural numbers n. Since m and k were arbitrarily chosen, then $k(m+n) = km+kn$ holds true for any three natural numbers m, k and n.

4) *Multiplication of the natural numbers is associative.* That is, for any three natural numbers m, n and k,
$$(mn)k = m(nk)$$
Proof: By induction. It is left as an exercise.

5) *Multiplication of the natural numbers is commutative.* That is, for any two natural numbers m and n,
$$mn = nm$$
Proof: By induction. It is left as an exercise.

15.5 Subtraction on the natural numbers.

Subtraction is easily defined in terms of addition. To define subtraction we must first establish the following fact.

Theorem 15.5 For any two natural numbers m and n, $m \in_= n$ if and only if there exists a *unique* natural number k such that $n = m + k$.

Proof:

By induction. Let $P(n)$ be the statement "For any $m \in_= n$ there exists a unique natural number k such that $n = m + k$.

Base case: Suppose $n = 0$. Then, for any $m \in_= 0$, $m = 0$ and so there exists only $k = 0$ such that $0 = n = m + k = 0 + 0$. So $P(0)$ holds true.

Inductive hypothesis: Suppose n is a natural number such that for any natural number $m \in_= n$, there exists a unique natural number k such that $n = m + k$. Suppose m is a natural number such that $m \in_= n^+$. Then $m \in_= n^+ = n \cup \{n\}$ implies $m \in n$ or $m = n$ or $m = n^+$. The equality $m = n^+$ means we can only choose $k = 0$. If $m \in n$, then the existence of a unique natural number k such that $n = m + k$ is guaranteed by our inductive hypothesis. So

$$\begin{aligned} n^+ &= n + 1 \\ &= (m + k) + 1 \\ &= m + (k + 1) \\ &= m + k^+ \end{aligned}$$

In this case the required natural number is k^+. If $m = n$, then $n^+ = k^+ = k+1 = m+1$. The unique required value is $k = 1$. So $P(n^+)$ holds true.

By mathematical induction $P(n)$ holds true for all values of n.

Definition 15.6 For any two natural numbers m and n such that $m \leq n$, the *unique* natural number k satisfying $n = m + k$ is called the *difference between n and m* and is denoted by $n - m$. The operation "$-$" is called *subtraction*.

With respect to the basic operations of addition, subtraction and multiplication on the natural numbers, our work is done. These operations have been shown to be properly defined in our *ZFC*-axiomatic universe. Again the intention is not to adopt the formal definitions as regular methods for doing arithmetic. It is only to ensure that arithmetic is definable in a set-theoretic context.

Concepts review:

1. How is addition on the natural numbers defined?
2. For any natural number n give two ways of describing n^+.
3. How can we prove that $0 + n = n$ for all n from the definition of addition?
4. How can we prove that addition is associative from the definition of addition?
5. How can we prove that addition is commutative from the definition of addition?
6. How is multiplication of natural numbers defined?
7. How is subtraction of natural numbers defined?

EXERCISES

A.

1. Use mathematical induction to prove the following multiplication properties.
   a) For all natural numbers n, $0n = 0$.
   b) For all natural numbers n, $1n = n$.
   c) For all natural numbers m, n and k, $n(m+k) = nm+nk$ and $(m+k)n = mn+kn$.
   d) For all natural numbers m, n and k, $(mn)k = m(nk)$.
   e) For all natural numbers m and n, $mn = nm$.

B.

2. Prove that if $n < k$, then $m + n < m + k$.
3. Prove that if $m + n = m + k$ implies $n = k$.
4. Prove that if $m < n$, then $mk < nk$.
5. Prove that if $mk = nk$ and $k \neq 0$, then $m = n$.

C.

6. Prove that $m + k < n + k$ implies $m < n$.
7. Prove that $mk < nk$ implies $m < n$.
8. Prove that for any two natural numbers m and n, $m \leq n$ if and only if there exists a *unique* natural number k such that $n = m + k$.
9. Prove in detail theorem 15.4.

---

16 / The integers $\mathbb{Z}$ and the rationals $\mathbb{Q}$.

Summary. *In this section we define both the integers, $\mathbb{Z}$, and the rational numbers, $\mathbb{Q}$. The integers are presented as a quotient set of $\mathbb{N} \times \mathbb{N}$ while the rationals are presented as a quotient set of $\mathbb{Z} \times (\mathbb{Z} - \{0\})$. Addition, subtraction and multiplication on each of these are defined within the set-theoretic context. Order relations are defined on each of $\mathbb{Z}$ and $\mathbb{Q}$ so that they are linearly ordered in the way we are accustomed to.*

16.1 Constructing the set of integers $\mathbb{Z}$ from $\mathbb{N}$.

Most people easily recognize an integer when they see one. One might say, given the natural numbers $\mathbb{N} = \{0, 1, 2, 3, \ldots, \}$, if we add to it all the "negative" natural numbers $\{-1, -2, -3, \ldots, \}$ we obtain all integers. We are then only left with the sticky problem of explaining what a "negative" natural number is within our set-theoretic framework.

Remember that the only mathematical objects in our set theoretic universe are *sets*. So the integer -3 must be a set of some sort. But the idea of a "negative set" is not very intuitive. One might say that -3 is the *difference* between the natural numbers 2 and 5. But in the last section, the difference, $n - m$, between natural numbers was only defined for pairs (n, m) where the second natural number is less than or equal to the first one. So the expression $2 - 5$ is, as yet, not defined.

What do the ordered pairs $(5, 2)$, $(10, 7)$ and $(3, 0)$ have in common? We notice that the first entry minus the second entry is 3 for each of the pairs. If we consider, on the other hand, the pairs $(1, 5)$, $(6, 10)$ and $(0, 4)$ we see that in each case, the first entry minus the second entry is -4. This suggests that an equivalence relation of some sort on $\mathbb{N} \times \mathbb{N}$ may provide a useful way of representing negative integers. We explore this avenue to see where it leads us.

Theorem 16.1 Let $Z = \mathbb{N} \times \mathbb{N}$. Let R_z be a relation on Z which is defined as follows: $(a, b)R_z(c, d)$ if and only if $a + d = b + c$. Then R_z is an equivalence relation on Z.

Proof:

Reflexivity: Since $a + b = b + a$, $(a, b)R_z(a, b)$.
Symmetry: $(a, b)R_z(c, d) \Rightarrow a + d = b + c \Rightarrow c + b = d + a \Rightarrow (c, d)R_z(a, b)$.
Transitivity: Suppose $(a, b)R_z(c, d)$ and $(c, d)R_z(e, f)$. Then $a+d = b+c$ and $c+f = d+e$. This implies $a + d + c + f = b + c + d + e$. Subtracting $c + d$ from both sides of the equality gives $a + f = b + e$. Hence, $(a, b)R_z(e, f)$.

Notation: If R is an equivalence relation on S and $x \in S$, then we will use the notation

$$[x]_R$$

to denote the equivalence class of all elements equivalent to x under R. When the context indicates which relation we are referring to and there is no risk of confusion, we will simply write $[x]$ instead of $[x]_R$.

Corollary 16.2 Let $Z = \mathbb{N} \times \mathbb{N}$ be equipped with the equivalence relation R_z defined as:

$$(a, b)R_z(c, d) \Leftrightarrow a + d = b + c$$

for each $n \in \mathbb{N}$, let $[(0, n)]$ and $[(n, 0)]$ denote the R_z-equivalence classes containing the elements $(0, n)$ and $(n, 0)$ respectively. Then the quotient set induced by R_z on Z can be expressed as $Z/R_z = \{[(0, n)] : n \in \mathbb{N}\} \cup \{[(n, 0)] : n \in \mathbb{N}\}$

Proof:

We will start by showing that the equivalence classes in Z/R_z cover all of Z. That is, we will show that

$$Z \subseteq U = \bigcup_{n\in\mathbb{N}} [(0, n)] \quad \cup \quad \bigcup_{n\in\mathbb{N}} [(n, 0)]$$

Let $(a, b) \in Z$. We will show that $(a, b) \in U$. Suppose $(c, d) \in [(a, b,)]$. Then $(a, b)R_z(c, d) \Rightarrow a + d = b + c$. We consider two cases, $d \leq c$ and $c \leq d$

$$\begin{aligned}
d \leq c &\Rightarrow a + d - d = b + c - d \\
&\Rightarrow a + 0 = b + (c - d) \\
&\Rightarrow (a, b)R_z((c - d), 0) \\
c \leq d &\Rightarrow a + d - c = b + c - c \\
&\Rightarrow a + (d - c) = b + 0 \\
&\Rightarrow (a, b)R_z(0, (d - c))
\end{aligned}$$

So if $(a, b) \in Z$, then either $(a, b) \in [((c - d), 0)]$ or $(a, b) \in [(0, (d - c))]$. Thus, $Z \subseteq U = \bigcup_{n\in\mathbb{N}}[(0, n)] \quad \cup \quad \bigcup_{n\in\mathbb{N}}[(n, 0)]$. So every element of $Z = \mathbb{N} \times \mathbb{N}$ is an element of some equivalence class in Z/R_z.

Next we show that if m and n are distinct, then $(0, n)$ and $(0, m)$ are not related with respect to R_z:

$$\begin{aligned}
m \neq n &\Rightarrow 0 + n \neq m + 0 \\
&\Rightarrow (0, m) \notin [(0, n)]
\end{aligned}$$

Since R_z is reflexive, $(m,0) \notin [(n,0)]$. We show that the elements of Z/R_z with distinct representatives do not overlap: For if $m \neq n$ and $(x,y) \in [(0,m)] \cap [(0,n]$, then $(0,m)R_z(x,y)$ and $(x,y)R_z(0,m)$ implies $(0,m)R_z(0,n)$ (by transitivity), a contradiction. So the sets in $\{[(0,n)] : n \in \mathbb{N}\} \cup \{[(n,0)] : n \in \mathbb{N}\}$ represent all equivalences class of Z induced by R_z.

We have set the stage for a set-theoretic definition of the "integers". Some readers may have some insight on where this is leading. It seems that the plan is to have the equivalence class $[(0,n)]$ represent the negative integers $0-n=n$ and $[(n,0)]$ represent the positive integers $n-0=n$. We can equate -5 with $[(0,5)]$ and the integer 5 with $[(5,0)]$. Some may immediately wonder: Why do the positive integers need defining? Aren't positive integers just the natural numbers? How can the natural number $5=\{0,1,2,3,4\}$ be the same set as the integer $5=[(0,5)]$? These two sets are indeed different since they don't contain the same elements. It is true, the "natural number 5" and the "integer 5" have different set representations. The question is: Is this a major problem or is it just a minor annoyance? It *may* be possible to construct the integers with the specific requirement that the sets which represent the positive integers and the sets which represent the natural numbers be the same. But this constraint may present some hurdles around which it may be difficult to maneuver. When we think about it carefully, it is not the sets which represent the natural numbers and the sets which represent the positive integers that are important. What is however crucial is that the arithmetic operations on these sets each produce the expected values. That is, both $5+3$ and $[(5,0)]+[(3,0)]$ produce 8 "the natural number" and $8=[(8,0)]$ "the positive integer" respectively. With this in mind we proceed with a formal definition of the integers.

Definition 16.3 The set of *integers*, $\mathbb{Z}$, is defined as:

$$\mathbb{Z} = Z/R_z = \{[(a,b)] : a,b \in \mathbb{N}\} = \{[(0,n)] : n \in \mathbb{N}\} \cup \{[(n,0)] : n \in \mathbb{N}\}$$

a) *Negative integers*: The set of *negative integers* is defined as being the set

$$\mathbb{Z}^- = \{[(0,n)] : n \in \mathbb{N}\}$$

Positive integers: The set of *positive integers* is defined as being the set

$$\mathbb{Z}^+ = \{[(n,0)] : n \in \mathbb{N}\}$$

If n is not 0, the elements of the form $[(0,n)]$ can be represented by $-n=[(0,n)]$ while the elements of the form $[(n,0)]$ can be represented as $n=[(n,0)]$.

b) *Order relation on* $\mathbb{Z}$: We define a relation $\leq_z$ on $\mathbb{Z}$ as follows: $[(a,b)] \leq_z [(c,d)]$ if and only if $a+d \leq b+c$. It is a routine exercise to show that $\leq_z$ is a linear ordering of $\mathbb{Z}$.

c) *Addition on* $\mathbb{Z}$: We must sometimes distinguish between addition of natural numbers and addition of integers. Where there is a risk of confusion we will use the following notation: "$+_n$" means addition of natural numbers while "$+_z$" means addition of integers.

Addition $+_z$ on $\mathbb{Z}$ is defined as:

$$[(a,b)] +_z [(c,d)] = [(a +_n c,\ b +_n d)]$$

d) *Opposites of integers*: The *opposite* $-[(a,b)]$ of $[(a,b)]$ is defined as

$$-[(a,b)] = [(b,a)]^1$$

e) *Subtraction on integers*: Subtraction "$-_z$" on $\mathbb{Z}$ is defined as:

$$[(a,b)] -_z [(c,d)] = [(a,b)] + (-[(c,d)])^2$$

f) *Multiplication of integers*: Multiplication $\times_z$ on $\mathbb{Z}$ is defined as

$$[(a,b)] \times_z [(c,d)] = [(ac+bd, ad+bc)].^3$$

In particular, $[(0,n)] \times_z [(m,0)] = [(0+0, 0+nm)] = [(0,nm)] = -[(nm,0)]$ and $[(n,0)] \times_z [(m,0)] = [(nm,0)]$.

g) *Absolute value of an integer*: The absolute value, $|n|$, of an integer n is defined as

$$|n| = \begin{cases} n & \text{if} \quad 0 \leq_z n \\ -n & \text{if} \quad n <_z 0 \end{cases}^4$$

h) *Equality of two integers*: If (a,b) and (c,d) are ordered pairs which are equivalent under the relation R_z, then the R_z-equivalence classes $[(a,b)]$ and $[(c,d)]$ are equal sets. To emphasize that they are equal sets under the relation R_z we can write

$$[(a,b)] =_z [(c,d)]$$

i) *Distribution properties*: If $[(a,b)]$, $[(c,d)]$ and $[(e,f)]$ are integers, then

$$[(a,b)] \times_z ([(c,d)] +_z [(e,f)]) =_z [(a,b)] \times_z [(c,d)] +_z [(a,b)] \times_z [(e,f)]$$

and

$$([(c,d)] +_z [(e,f)]) \times_z [(a,b)] =_z [(c,d)] \times_z [(a,b)] +_z [(e,f)] \times_z [(a,b)]$$

---

[1] Note that $-[(n,0)] = [(0,n)] = -n$.

[2] When there is no risk of confusion with subtraction of other types of numbers we will simply use "$-$".

[3] Note that the "center dot" can be used instead of the "$\times_z''$ symbol. When there is no risk of confusion with multiplication of other types of numbers we will simply use "$\times$".

[4] View "absolute value" as a function $|\ \ | : \mathbb{Z} \to \mathbb{Z}$.

It is good to remember that any integer can be written in the form $[(n,0)]$ or $[(0,n)] = -[(n,0)]$. These forms make it easier to add and multiply them without memorizing intricate formulas. For example, the expression $[(2,4)] \times_z [(5,2)]$ can be more easily computed as follows:

$$\begin{aligned} [(2,4)] \times_z [(5,2)] &= [(0,2)] \times_z [(3,0)] \\ &= -[(2,0)] \times_z [(3,0)] \\ &= -[(6,0)] = [(0,6)] \end{aligned}$$

We verify that the product of two non-negative integers produces a positive integer as it should.

$$\begin{aligned} [(2,4)] \times_z [(2,5)] &=_z [(0,2)] \times_z [(0,3)] \\ &=_z [(0+6)+0] \\ &=_z [(6,0)] \end{aligned}$$

Remark: We pause to deconstruct the elements of $\mathbb{Z}$ to better see the nature of the sets that belong to it. Let $u \in \mathbb{Z}$. Then $u = [(a,b)]$ for some $a, b \in \mathbb{N}$. By lemma 4.5, $(a,b) \in \mathbb{N} \times \mathbb{N} \in \mathscr{P}(\mathscr{P}(\mathbb{N}))$. Since $[(a,b)] \subset \mathscr{P}(\mathscr{P}(\mathbb{N}))$ then $[(a,b)] \in \mathscr{P}(\mathscr{P}(\mathscr{P}(\mathbb{N}))) = \mathscr{P}^3(\mathbb{N})$. We conclude that $\mathbb{Z} \subseteq \mathscr{P}^3(\mathbb{N})$.

16.2 Constructing the rational numbers $\mathbb{Q}$ from $\mathbb{Z}$.

We have succeeded in "extracting" the integers $\mathbb{Z}$ from $\mathbb{N} \times \mathbb{N}$ by constructing a quotient set induced by a particular equivalence relation on $\mathbb{N} \times \mathbb{N}$. To construct the rationals we will proceed in a similar way.

When looking at a rational number, $\frac{a}{b}$, it may be useful to view it as an ordered pair of integers (a,b) of integers where the first entry plays the role of the numerator while the (non-zero) second entry plays the role of the denominator. But simply defining a/b as an ordered pair (a,b) in $\mathbb{Z} \times \mathbb{Z}$ would not do, since a rational number, say $-2/3$, can have many equivalent forms: $\frac{2}{-3}$, $\frac{-4}{6}$, $\frac{20}{-30}$. So the associated ordered pairs of integers $(-2,3)$, $(-4,6)$ and $(20,-30)$ should also be equivalent forms of the same number. To overcome this difficulty we will define an equivalence relation on $Q = \mathbb{Z} \times \mathbb{Z}^*$ (where $\mathbb{Z}^* = \mathbb{Z} - \{0\}$) so that all equivalent forms of $(-2,3)$ belong to an equivalence class $[(-2,3)]$ induced by this equivalence relation. We have chosen the Cartesian product $\mathbb{Z} \times \mathbb{Z}^*$ rather than $\mathbb{Z} \times \mathbb{Z}$ since the second entry cannot be zero. The equivalence relation we will use to extract $\mathbb{Q}$ from $Q = \mathbb{Z} \times \mathbb{Z}^*$ will be represented by R_q. To define this equivalence relation R_q we will ask ourselves: What property makes the two rational numbers $-2/3$ and $\frac{8}{-12}$ equivalent? We see that

$$\frac{-2}{3} = \frac{8}{-12} \text{ implies } (-2)(-12) = (3)(8)$$

More generally we see that

$$\frac{a}{b} = \frac{c}{d} \text{ if and only if } ad = bc$$

We want the ordered pairs (a, b) and (c, d) in $\mathbb{Z} \times \mathbb{Z}^*$ to be related under R_q provided they satisfy the property $a \times_z d = b \times_z c$. Proving that this is a valid equivalence relation is routine. It is formally stated as a theorem.

Theorem 16.4 Let $Q = \mathbb{Z} \times \mathbb{Z}^*$ where $\mathbb{Z}^* = \mathbb{Z} - \{0\}$. Let R_q be a relation on Q defined as follows: $(a, b)R_q(c, d)$ if and only if $a \times_z d = b \times_z c$. Then R_q is an equivalence relation on Q.

Proof: The proof is left as an exercise.

For example, consider the two elements $(6, 10)$ and $(15, 25)$ of $\mathbb{Z} \times \mathbb{Z}^*$. Since $6 \times_z 25 = 150 = 10 \times_z 15$, they are equivalent rational numbers. So $[(6, 10)]_q =_q [(15, 25)]_q$. Remember that $[(a, b)]_q$ will represent the set of all elements of $\mathbb{Z} \times \mathbb{Z}^*$ which are R_q-equivalent to the element $(a, b) \in \mathbb{Z} \times \mathbb{Z}^*$. When there is no risk of confusion with the equivalence class of another equivalence relation we will simply use $[(a, b)]$ rather than $[(a, b]_q$.

We now formally define the rational numbers within a set-theoretic context.

Definition 16.5 The set of *rational numbers*, $\mathbb{Q}$, is defined as:

$$\mathbb{Q} = Q/R_q = \{[(a, b)] : a \in \mathbb{Z}, b \in \mathbb{Z}^*\}^1$$

The expression $[(a, b)]$ is normally written in the form $\frac{a}{b}$.

a) We define a relation $\leq_q$ on $\mathbb{Q}$ as follows: If b and d are both positive, $[(a, b)] \leq_q [(c, d)]$ if and only if $a \times_z d \leq_z b \times_z c$.

b) Addition $+_q$ on $\mathbb{Q}$ is defined as:

$$[(a, b)] +_q [(c, d)] = [(ad +_z bc,\ b \times_z d)]$$

c) Subtraction $-_q$ on $\mathbb{Q}$ is defined as:

$$[(a, b)] -_q [(c, d)] = [(a, b)] +_q [(-c, d)]$$

[1] Recall that a and $-b$ is shorthand for expressions of the form $[(a, 0)]$ or $-[(0, b)]$

d) Multiplication $\times_q$ on $\mathbb{Q}$ is defined as

$$[(a,b)] \times_q [(c,d)] = [(a \times_z c, b \times_z d)]$$

e) *Equality of two rational numbers*: If (a,b) and (c,d) are ordered pairs of integers (where neither b nor d is 0) which are equivalent under the relation R_q, then the R_q-equivalence classes $[(a,b)]$ and $[(c,d)]$ are equal sets. To emphasize that they are equal sets under the relation R_q we can write

$$[(a,b)] =_q [(c,d)]$$

f) *Opposites of rational numbers.* If (a,b) is an ordered pair of integers ($b \neq 0$) and $[(a,b)]$ is its R_q-equivalence class, then the opposite of the rational number $[(a,b)]$ is defined as $[(-a,b)]$ and is denoted as $-[(a,b)] =_q [(-a,b)]$.

Proofs that addition, subtraction, multiplication and linear ordering, thus defined, reflect precisely what we normally obtain when performing the usual algorithms on $\mathbb{Q}$ is left as an exercise. Once this is verified, the reader is, of course, free to use the usual algorithms for computation involving rational numbers. When convenient, we will interchangeably represent R_q equivalence classes $[(a,b)]$ as $\frac{a}{b}$ and vice-versa.

Remark: We pause to deconstruct the elements of $\mathbb{Q}$ to better see the nature of the sets that belong to it. Recall that $\mathbb{Z} \subseteq \mathscr{P}^3(\mathbb{N})$. Let $u \in \mathbb{Q}$. Then $u = [(a,b)]$ for some $a, b \in \mathbb{Z}$. By lemma 4.5, $(a,b) \in \mathbb{Z} \times \mathbb{Z}^* \in \mathscr{P}(\mathscr{P}(\mathbb{Z}))$. Since $[(a,b)] \subset \mathscr{P}(\mathscr{P}(\mathbb{Z}))$ then $[(a,b)] \in \mathscr{P}(\mathscr{P}(\mathscr{P}(\mathbb{Z}))) = \mathscr{P}^3(\mathbb{Z})$. We conclude that $\mathbb{Q} \subseteq \mathscr{P}^3(\mathbb{Z}) \subseteq \mathscr{P}^3(\mathscr{P}^3(\mathbb{N})) = \mathscr{P}^6(\mathbb{N})$.

Theorem 16.6 Suppose a and b are positive integers where $b \neq 0$ and $[(a,b)]$ is an R_q equivalence class. Then

a) $[(-a,-b)] = \frac{-a}{-b} = \frac{a}{b} = [(a,b)]$.

b) $-[(a,b)] =_q [(-a,b)] = \frac{-a}{b} = \frac{a}{-b} = [(a,-b)]$

Proof:

a) Since $-a \times_z b = -b \times_z a$, then $(-a,-b) R_q (a,b)$. Then $(-a,-b) \in [(a,b)]$ and so we can write $\frac{-a}{-b} = [(-a,-b)] =_q [(a,b)] = \frac{a}{b}$.

b) Note that if a and b are positive integers, then since $-a \times_z -b = b \times_z a$, then $(-a,b)$ is R_q-equivalent to $(a,-b)$ and so

$$\frac{-a}{b} = [(-a,b)] =_q -[(a,b)] =_q [(a,-b)] = \frac{a}{-b}$$

Hence, moving negative signs from numerators to denominators and vice versa in rational numbers is justified.

Is an integer equal to a rational number in this set-theoretic context? To help answer this question we will compare a particular integer to its equivalent rational number form. In our set-theoretic axiomatic system, the integer -3 looks like $[(0,3)]_z = -[(3,0)]_z \subset \mathbb{N} \times \mathbb{N}$. In the same axiomatic system the rational number $-3/1$ looks like $[(-3,1)]_q \subset \mathbb{Z} \times \mathbb{Z}^*$. So, they are not the same set. However, it can be verified to one's satisfaction that operations made with integers $n = \pm[(n,0)]_z$ when viewed as rational numbers $[(n,1)]_q$ will provide results which consistently match those obtained by performing parallel operations with integer numbers.[1]

Examples.

The following examples illustrate that doing arithmetic by referring to the set-theoretic definitions and the listed properties is a bit awkward and requires some thought. It is of course not an efficient way of doing arithmetic. To see this we compute the following expressions by representing these numbers as equivalence classes of ordered pairs and using the above definitions. When useful we indicate which of the above statements are invoked to justify various steps.

a) Compute $-4(3-7)$ b) Compute $\frac{-2}{5}\left(\frac{6}{7} - \frac{3}{11}\right)$

Solution:

a)

$$\begin{aligned}
-4(3-7) &= [(0,4)] \times_z ([(3,0)] -_z [(7,0)]) \\
&=_z [(0,4)] \times_z ([(3,0)] +_z [(-7,0)]) \quad \text{(By theorem 16.6, b) .)} \\
&=_z [(0,4)] \times_z ([(3,0)] +_z [(0,7)]) \\
&=_z ([(0,4)] \times_z [(3,0)]) +_z ([(0,4)] \times_z [(0,7)]) \\
&=_z [(0,12)] +_z [(28,0)] \\
&=_z [(28,12)] \\
&=_z [(16,0)] \quad \text{(Since } (a,b)R_z(c,d) \Leftrightarrow a+d = b+c.\text{)}
\end{aligned}$$

[1] Just as a matter of interest let's see how our number systems are evolving in our set-theoretic universe: Natural numbers $\mathbb{N}$: Exists by Axiom 8.
The integers $\mathbb{Z}$: $n \in \mathbb{Z} \Rightarrow n = \pm[(n,0]_z \subset \mathbb{N} \times \mathbb{N} \Rightarrow n = \pm[(n,0]_z \in \mathscr{P}(\mathbb{N} \times \mathbb{N})$. Then $\mathbb{Z} \subset \mathscr{P}(\mathbb{N} \times \mathbb{N})$.
The rational numbers $\mathbb{Q}$: $a/b \in \mathbb{Q} \Rightarrow a/b = [(a,b]_q \subset \mathbb{Z} \times \mathbb{Z}^* \Rightarrow a/b = [(a,b]_z \in \mathscr{P}(\mathbb{Z} \times \mathbb{Z}) \Rightarrow a/b \in \mathscr{P}(\mathscr{P}(\mathbb{N} \times \mathbb{N}) \times \mathscr{P}(\mathbb{N} \times \mathbb{N})) \Rightarrow \mathbb{Q} \subset \mathscr{P}(\mathscr{P}(\mathbb{N} \times \mathbb{N}) \times \mathscr{P}(\mathbb{N} \times \mathbb{N}))$.

b) $$\begin{aligned}
\frac{-2}{5}\left(\frac{6}{7}-\frac{3}{11}\right) &= [(-2,5)] \times_q ([(6,7)] -_q [(3,11)]) \\
&=_q [(-2,5)] \times_q ([(6,7)] +_q [(-3,11)]) \quad \text{(By theorem 16.6, b) .)} \\
&=_q [(-2,5)] \times_q [(6 \times_z 11 +_z 7 \times_z -3, \quad 7 \times_z 11)] \\
&=_q [(-2,5)] \times_q [(45,77)] \\
&=_q [(-2 \times_z 45,\ 5 \times_z 77)] \\
&=_q [(-90,385)] \\
&=_q [(-18,77)] \\
&=_q -[(18,77)] \\
&= -\frac{18}{77}
\end{aligned}$$

Concepts review:

1. Describe the equivalence relation R_z on $\mathbb{N} \times \mathbb{N}$ used to define the elements of the integers $\mathbb{Z}$.
2. Describe the equivalence class induced by R_z on $\mathbb{N} \times \mathbb{N}$ which represents the integer -9. What about the integer 3?
3. Do the equivalence classes $\{[(0,n)] : n \in \mathbb{N}\} \cup \{[(n,0)] : n \in \mathbb{N}\}$ account for all the equivalence classes induced by R_z on $\mathbb{N} \times \mathbb{N}$?
4. How is addition $+_z$ defined on $\mathbb{Z}$ in a set-theoretic context?
5. How is multiplication $\times_z$ defined on $\mathbb{Z}$ in a set-theoretic context?
6. Describe the equivalence relation R_q on the Cartesian product $\mathbb{Z} \times \mathbb{Z}^*$ used to define the rational numbers $\mathbb{Q}$.
7. How is addition $+_q$ defined on $\mathbb{Q}$ in a set-theoretic context?
8. How is multiplication $\times_q$ defined on $\mathbb{Q}$ in a set-theoretic context?

EXERCISES

A.

1. Use the definitions in 16.3 to show that the following statements are true:
   a) $-2 \leq 10$.
   b) $0 \leq 3$.
   c) $-5 - 7 = -12$.
   d) $-2 + 6 = 4$.
   e) $7 \times -2 = -14$.
   f) $-1 \times -2 = 2$.
2. Use the definitions in 16.5 to show that the following statements are true:
   a) $\frac{-2}{-3} = \frac{2}{3}$.
   b) $\frac{-2}{3} \leq \frac{3}{2}$.
   c) $\frac{5}{3} + \frac{3}{2} = \frac{19}{6}$.
   d) $-2 - \frac{6}{5} = -\frac{16}{5}$.
   e) $\frac{6}{5} \times \frac{1}{3} = \frac{2}{5}$.
   f) $3 = \frac{6}{2}$.

B.

3. Let $Z = \mathbb{N} \times \mathbb{N}$. Let R_z be a relation on Z that is defined as follows: $((a, b), (c, d)) \in R$ if and only if $a + d = b + c$. Prove that R_z is an equivalence relation on Z.
4. The relation $\leq_z$ on $\mathbb{Z}$ is defined as follows: $[(a, b)] \leq_z [(c, d)]$ if and only if $a+d \leq b+c$. Show that this is a linear ordering.
5. Let $Q = \mathbb{Z} \times \mathbb{Z}^*$ where $\mathbb{Z}^* = \mathbb{Z} - \{0\}$. Let R_q be a relation on Q defined as follows: $((a, b), (c, d)) \in R_q$ if and only if $a \times_z d = b \times_z c$. Show that R_q is an equivalence relation on Q.

C.

6. We define a relation $\leq_q$ on $\mathbb{Q}$ as follows: When both b and d are non-negative, $[(a, b)] \leq_q [(c, d)]$ if and only if $a \times_z d \leq b \times_z c$. Show that this is a linear ordering.

17 / Dedekind cuts: "Real numbers are us!".

Summary. *In this section we show how $\mathbb{R}$ is defined within the confines of the ZFC-set-theoretic axiomatic system. With this objective in mind we begin by defining "initial segments" of rational numbers. These special elements of $\mathscr{P}(\mathbb{Q})$ are referred to as "Dedekind cuts". A linear order relation is defined on these as well as operations of addition and multiplication. We then show that there is a function, f, which maps the real numbers $\mathbb{R}$ to the Dedekind cuts one-to-one and onto while respecting order, addition and multiplication. Finally we show that the set of all Dedekind cuts satisfies the essential Completeness property of the real numbers.*

17.1 Defining the real numbers $\mathbb{R}$.

Numbers such as $\sqrt{2}$ and π were known to be "not rational" long before earnest attempts were made to define the real numbers. Defining a real number by stating that "it is a set of numbers some of which are not rational numbers" is not satisfactory, since today we know of numbers which are neither real numbers nor rational numbers.[1] Making a list of the basic properties which define the real numbers and then, postulating the existence of a set which satisfies these properties is a possibility. But a new axiom is not absolutely necessary since, as we shall soon see, the existence of real numbers can be proven by only invoking the set theory axioms we have seen up to now. At this point, we have invoked, at some point in our study, all of the axioms numbered 1 to 6 and 8. Axiom 8 is the Axiom of replacement, Axiom 9, called the "Axiom of regularity" and the tenth axiom called the "Axiom of choice" have not yet been invoked; we will see that these two axioms are not required to construct the set of all real numbers.

The main challenge we encounter when studying those real numbers which are not rationals is that we cannot "see" them directly, even though there is evidence that they exist as numbers. We can see the rationals and easily order them on a number line. For example, the number $2/5$ can be seen by subdividing the line from 0 to 1 into five equal subintervals; the number $2/5$ is the right endpoint of the second subinterval. Any attempt to "see" $\sqrt{2}$ by subdividing a line interval in this way will fail, even though this number is a solution of the equation $x^2 - 2 = 0$, and we have a approximate idea of where it sits on the number line. Thousands of years ago, the Greeks knew that the length of the diagonal line of a one by one square was not a quotient of natural numbers and so could not be measured in conventional ways.

[1] We are referring here to the "complex numbers".

It is true that the real numbers are, in many respects, similar to the rational numbers. The set $\mathbb{Q}$ is one which is linearly ordered so that we can associate to any pair of distinct rationals a and b satisfying $a < b$ a third rational number c such that $a < c < b$. The set $\mathbb{R}$ also has a linear ordering which satisfies this property. But there is a property which distinguishes the reals $\mathbb{R}$ from the rationals $\mathbb{Q}$ in a fundamental way. This property is called "*the completeness property*"[1]. It states that "Every non-empty bounded subset S of $\mathbb{R}$ has a least upper bound which is in $\mathbb{R}$".[2] The set of all rational numbers $\mathbb{Q}$ does not satisfy this property. For example, the set S of all rational numbers which are strictly less than $\sqrt{2}$ is bounded above; but the non-empty subset, $\mathbb{Q} - S$, of all upper bounds of S has no "least upper bound" (or if one prefers, no minimal element) which is contained in $\mathbb{Q}$. That is, for any rational upper bound q of S which is larger than $\sqrt{2}$, we can find a smaller upper bound rational number t such that $\sqrt{2} < t < q$. We mention this now, since we will be attempting to identify sets which can represent the elements of $\mathbb{R}$ in a universe governed by *ZFC*. Contenders will have to successfully pass the "completeness property" test. We will come back to this soon.

A very elegant and clever set-theoretic definition of the real numbers was put forward by the mathematician Richard Dedekind (1831-1916). Dedekind identified a subset of $\mathscr{P}(\mathbb{Q})$ which successfully represents the set $\mathbb{R}$. We first express a few preliminary ideas that may have led Dedekind to formulate his definition of $\mathbb{R}$.

17.2 Projecting $\mathbb{R}$ into $\mathscr{P}(\mathbb{Q})$.

In what follows, we will construct a one-to-one function which will map $\mathbb{R}$ into $\mathscr{P}(\mathbb{Q})$ while respecting the order and the basic algebraic properties of the real numbers. We will begin with the definition of special subsets of $\mathbb{R}$ and $\mathbb{Q}$.

Definition 17.1 For any real number r, let $(\leftarrow_{\mathbb{R}} r)$ denote the interval $(-\infty, r)$ in $\mathbb{R}$. It is the subset of all real numbers strictly smaller than r. The subset $(\leftarrow_{\mathbb{R}} r)$ is called an *initial segment in* $\mathbb{R}$. For any real number r, the subset $(\leftarrow_{\mathbb{Q}} r)$ of $\mathbb{Q}$ is defined as follows:

$$(\leftarrow_{\mathbb{Q}} r) = (-\infty, r) \cap \mathbb{Q} = (\leftarrow_{\mathbb{R}} r) \cap \mathbb{Q}$$

The subset $(\leftarrow_{\mathbb{Q}} r)$ of $\mathbb{Q}$ is called an *initial segment in* $\mathbb{Q}$. For both $(\leftarrow_{\mathbb{Q}} r)$ and $(\leftarrow_{\mathbb{R}} r)$ the real number r will be referred to as the *leader* of the initial segment. Note that r may be an irrational number even for $(\leftarrow_{\mathbb{Q}} r) \subset \mathbb{Q}$.[3] The number r is also seen to be the *least upper bound* of $(\leftarrow_{\mathbb{Q}} r)$ in $\mathbb{R}$.[4]

---

[1] It is also referred to as the "Least upper bound property".

[2] We can also say: The set U of all upper bounds of the subset S of $\mathbb{R}$ has a minimal element which is in $\mathbb{R}$.

[3] For example, $(\leftarrow_{\mathbb{Q}} \sqrt{2})$ is a subset of $\mathbb{Q}$ whose leader is the irrational number $\sqrt{2}$.

[4] The reader will recall that m is an *upper bound* of a set S if $s \leq m$ for all s in S and t is the least upper bound of S if t is an upper bound of S and $t \leq u$ for all upper bounds of S.

For each real number r, $(\leftarrow_{\mathbb{Q}} r)$ is a subset of $\mathbb{Q}$ and therefore is an element of $\mathscr{P}(\mathbb{Q})$. So the set

$$\mathscr{D} = \{(\leftarrow_{\mathbb{Q}} r) : r \in \mathbb{R}\}$$

is a subset of $\mathscr{P}(\mathbb{Q})$.

We first show how the elements of $\mathscr{D}$ can be linearly ordered; then we show how to add and multiply its elements.

a) *A linear ordering of $\mathscr{D}$*: Since the elements of $\mathscr{D}$ are subsets of $\mathbb{Q}$ we will order the elements of $\mathscr{D}$ by inclusion. That is, we define "$<$" on $\mathscr{D}$ as follows:

$$(\leftarrow_{\mathbb{Q}} r) < (\leftarrow_{\mathbb{Q}} t) \Leftrightarrow (\leftarrow_{\mathbb{Q}} r) \subset (\leftarrow_{\mathbb{Q}} t)$$

Notice that this ordering of $\mathscr{D}$ is one which rigorously respects the order of their leaders in $\mathbb{R}$. That is $(\leftarrow_{\mathbb{Q}} r) < (\leftarrow_{\mathbb{Q}} t) \Leftrightarrow r < t$. For example, since $-2 < \pi$, then

$$(\leftarrow_{\mathbb{Q}} -2) \subset (\leftarrow_{\mathbb{Q}} \pi) \text{ and so } (\leftarrow_{\mathbb{Q}} -2) < (\leftarrow_{\mathbb{Q}} \pi)$$

Furthermore, $(\leftarrow_{\mathbb{Q}} a) = (\leftarrow_{\mathbb{Q}} b)$ if and only if $a = b$. If $(\leftarrow_{\mathbb{Q}} c)$ and $(\leftarrow_{\mathbb{Q}} d)$ are distinct elements of $\mathscr{D}$, then c and d are distinct real numbers and so either $c < d$ or $d < c$. Hence, either $(\leftarrow_{\mathbb{Q}} c) \subset (\leftarrow_{\mathbb{Q}} d)$ or $(\leftarrow_{\mathbb{Q}} d) \subset (\leftarrow_{\mathbb{Q}} c)$. So "$<$" linearly orders the elements of $\mathscr{D}$.

b) *Addition on $\mathscr{D}$*: We define addition in $\mathscr{D}$ as follows:

$$(\leftarrow_{\mathbb{Q}} r) + (\leftarrow_{\mathbb{Q}} t) = \{x + y : x \in (\leftarrow_{\mathbb{Q}} r) \text{ and } y \in (\leftarrow_{\mathbb{Q}} t)\}$$

Note that addition of the elements of $\mathscr{D}$ rigorously respects the addition of its leaders. For example $(\leftarrow_{\mathbb{Q}} -5) + (\leftarrow_{\mathbb{Q}} 7) = (\leftarrow_{\mathbb{Q}} -5 + 7) = (\leftarrow_{\mathbb{Q}} 2)$.

c) *Multiplication on $\mathscr{D}$*: In the case where both r and t are greater than zero we define multiplication as:

$$(\leftarrow_{\mathbb{Q}} r)\,(\leftarrow_{\mathbb{Q}} t) = \{xy : x \in (\leftarrow_{\mathbb{Q}} r) \text{ and } y \in (\leftarrow_{\mathbb{Q}} t), x, y > 0\} \cup [(-\infty, 0) \cap \mathbb{Q}]$$

The more general definition of multiplication on $\mathscr{D}$, which includes products of negative numbers, is a bit more complicated.
In general:

$$(\leftarrow_{\mathbb{Q}} r)\,(\leftarrow_{\mathbb{Q}} t) = \begin{cases} (\leftarrow_{\mathbb{Q}} 0) & \text{for the case where } r \text{ or } t \text{ is 0.} \\ (\leftarrow_{\mathbb{Q}} |r|)\,(\leftarrow_{\mathbb{Q}} |t|) & \text{for the case where } r \text{ and } t \text{ are both positive or both negative.} \\ -(\leftarrow_{\mathbb{Q}} |r|)\,(\leftarrow_{\mathbb{Q}} |t|) & \text{for the case where precisely one of } r \text{ or } t \text{ is negative.} \end{cases}$$

Note that multiplication of the elements of $\mathscr{D}$ rigorously respects the multiplication of its leaders. For example:

$$\begin{array}{lcll} (\leftarrow_{\mathbb{Q}} 5)\,(\leftarrow_{\mathbb{Q}} 7) & = & (\leftarrow_{\mathbb{Q}} 5 \times 7) & = & (\leftarrow_{\mathbb{Q}} 35) \\ (\leftarrow_{\mathbb{Q}} \sqrt{2})\,(\leftarrow_{\mathbb{Q}} -\sqrt{2}) & = & -(\leftarrow_{\mathbb{Q}} \sqrt{2} \times \sqrt{2}) & = & -(\leftarrow_{\mathbb{Q}} 2) \\ (\leftarrow_{\mathbb{Q}} -4)\,(\leftarrow_{\mathbb{Q}} 0) & = & (\leftarrow_{\mathbb{Q}} 0) & & \\ (\leftarrow_{\mathbb{Q}} -2)\,(\leftarrow_{\mathbb{Q}} -10) & = & (\leftarrow_{\mathbb{Q}} 2 \times 10) & = & (\leftarrow_{\mathbb{Q}} 20) \end{array}$$

Basic addition and multiplication properties.

The definitions of addition and multiplication are such that the following fundamental properties are all satisfied:

$$\begin{array}{rcl} (\leftarrow_{\mathbb{Q}} 0) + (\leftarrow_{\mathbb{Q}} a) = (\leftarrow_{\mathbb{Q}} 0 + a) & = & (\leftarrow_{\mathbb{Q}} a) \\ (\leftarrow_{\mathbb{Q}} 0)(\leftarrow_{\mathbb{Q}} a) = (\leftarrow_{\mathbb{Q}} 0 \times a) & = & (\leftarrow_{\mathbb{Q}} 0) \\ (\leftarrow_{\mathbb{Q}} -a) + (\leftarrow_{\mathbb{Q}} a) = (\leftarrow_{\mathbb{Q}} -a + a) & = & (\leftarrow_{\mathbb{Q}} 0) \\ (\leftarrow_{\mathbb{Q}} 1)\,(\leftarrow_{\mathbb{Q}} b) = (\leftarrow_{\mathbb{Q}} 1 \times b) & = & (\leftarrow_{\mathbb{Q}} b) \\ (\leftarrow_{\mathbb{Q}} 1/a)\,(\leftarrow_{\mathbb{Q}} a) = (\leftarrow_{\mathbb{Q}} 1/a \times a) & = & (\leftarrow_{\mathbb{Q}} 1) \end{array}$$

We see that:

- $(\leftarrow_{\mathbb{Q}} 0)$ plays the role of the additive identity in $\mathscr{D}$
- $(\leftarrow_{\mathbb{Q}} 0)$ plays the role of the multiplicative zero element in $\mathscr{D}$
- every element in $\mathscr{D}$ has an additive inverse
- $(\leftarrow_{\mathbb{Q}} 1)$ plays the role of the multiplicative identity in $\mathscr{D}$
- every non-zero element of $\mathscr{D}$ has a multiplicative inverse

We also see that this subset $\mathscr{D} \subset \mathscr{P}(\mathbb{Q})$, when linearly ordered by inclusion, mimics in every way the behaviour of the real numbers.

Defining a one-to-one onto function from $\mathbb{R}$ onto $\mathscr{D}$.

The function $f : \mathbb{R} \to \mathscr{D}$ defined as

$$f(r) = (\leftarrow_{\mathbb{Q}} r)$$

is a natural one-to-one mapping which copies $\mathbb{R}$ into $\mathscr{P}(\mathbb{Q})$ in the form of $\mathscr{D} = \{\, (\leftarrow_{\mathbb{Q}} r) : r \in \mathbb{R}\}$.

This function, f, respects the linear ordering, addition, multiplication in $\mathbb{R}$:

$$\begin{aligned} r < t &\Leftrightarrow f(r) = (\leftarrow_{\mathbb{Q}} r) \subset (\leftarrow_{\mathbb{Q}} t) = f(t) \\ f(r+t) = (\leftarrow_{\mathbb{Q}} r + t) &= (\leftarrow_{\mathbb{Q}} r) + (\leftarrow_{\mathbb{Q}} t) = f(r) + f(t) \\ f(r \times t) = (\leftarrow_{\mathbb{Q}} r \times t) &= (\leftarrow_{\mathbb{Q}} r) \times (\leftarrow_{\mathbb{Q}} t) = f(r) \times f(t) \end{aligned}$$

We first provide an *informal* definition of "Dedekind cuts" based on an understanding of initial segments of rational numbers as described above.

Definition 17.2 *Informal definition.* The elements of the set $\mathscr{D} = \{(\leftarrow_{\mathbb{Q}} r) : r \in \mathbb{R}\}$ where $(\leftarrow_{\mathbb{Q}} r) = (\infty, r) \cap \mathbb{Q}$, are called *Dedekind cuts.*

17.3 Defining the real numbers $\mathbb{R}$ as Dedekind cuts.

We now have the background and the ingredients needed to provide a set-theoretic definition of the real numbers.

Definition 17.3 The set of all Dedekind cuts $\mathscr{D}$, linearly ordered by inclusion with addition $+$ and multiplication $\times$ (as described above) is called the *real numbers.* Those Dedekind cuts which have no least upper bound in $\mathbb{Q}$ are called *irrational numbers.*

"Those Dedekind cuts which have no least upper bound in $\mathbb{Q}$" means those Dedekind cuts whose leader is a real non-rational number. Now defining the set of real numbers while referring to real numbers (in the definition of Dedekind cuts) is of course inappropriate. To make things right, we will go back to our informal definition of *Dedekind cuts* and remove any reference to real numbers. That is, we rewrite this definition (without altering its meaning) so that it assumes no knowledge of the set of real numbers on the part of the reader.[1]

Definition 17.4 A *Dedekind cut* is a subset S of the rational numbers $\mathbb{Q}$ which satisfies the following properties:

1. The set $S \neq \varnothing$ and $S \neq \mathbb{Q}$.

[1] Kronecker, a very influential leading mathematician of his time was quoted as saying to the mathematician Lindmann: "Of what use is your beautiful investigation of π. Why study such problems when irrational numbers do not exist."

2. For any two rational numbers a and b , if $a \in S$ and $b < a$, then $b \in S$.

3. The set S contains no maximal element.

This definition describes precisely those sets which can be expressed in the form $(-\infty, r) \cap \mathbb{Q} = (_{\leftarrow_{\mathbb{Q}}} r)$ for some real number r. To help us see this we make the following comments:

- By condition 1, S is a non-empty *proper* subset of $\mathbb{Q}$ and so there exists a rational element q which does not belong to S.
- If $a \in S$, then $a \leq q$ since, by condition 2, $q < a$ would imply $q \in S$. Also, by condition 2, $(\infty, a] \subseteq S$. Since S is bounded, then by the completeness property of the real numbers, S has a least upper bound, say r.
- By condition 3, r cannot belong to S. Then $S = (-\infty, r) \cap \mathbb{Q} = (_{\leftarrow_{\mathbb{Q}}} r)$.

We conclude that:

> "Any set which satisfies the three conditions in the formal definition of a Dedekind cut is of the form $(_{\leftarrow_{\mathbb{Q}}} r)$ for some real number r."

Proving that

> "Any set of the form $(_{\leftarrow_{\mathbb{Q}}} r)$ for some real number r is a Dedekind cut."

is left as an exercise.

We conclude that the set $\mathscr{D}$ of all Dedekind cuts can conveniently be expressed as

$$\mathscr{D} = \{(_{\leftarrow_{\mathbb{Q}}} r) : r \in \mathbb{R}\} \text{ where } (_{\leftarrow_{\mathbb{Q}}} r) = (\infty, r) \cap \mathbb{Q}$$

as stated in the informal definition 17.2.

So if $(_{\leftarrow_{\mathbb{Q}}} r)$ is a Dedekind cut with a rational number as leader, then this Dedekind cut can be referred to as a rational number. But if those with eyes "for-rational-numbers-only" cannot perceive a leader for $(_{\leftarrow_{\mathbb{Q}}} r)$, only seeing a ray of rationals lining up towards negative infinity with nothing at its head, then this Dedekind cut is an irrational number.

We present a few examples of Dedekind cuts:

$$\begin{aligned}
(_{\leftarrow_{\mathbb{Q}}} 2/3) &= \{x \in \mathbb{Q} : x < 2/3\} && \text{An initial segment of } \mathbb{Q} \text{ with leader } 2/3. \text{ This is a rational number.} \\
(_{\leftarrow_{\mathbb{Q}}} -150) &= \{x \in \mathbb{Q} : x < -150\} && \text{An initial segment of } \mathbb{Q} \text{ with leader } -150. \text{ This is a rational number.} \\
(_{\leftarrow_{\mathbb{Q}}} -\sqrt{2}) &= \{x \in \mathbb{Q} : x < -\sqrt{2}\} && \text{An initial segment of } \mathbb{Q} \text{ with no rational number as a leader.}
\end{aligned}$$

17.4 Completeness property of the real numbers $\mathbb{R}$.

We will now verify if the set of all Dedekind cuts satisfies the *completeness property* (or as it is often called, the *Least upper bound property*). We remind ourselves what this property states:

> "Every bounded subset of $\mathbb{R}$ has a least upper bound which is a real number."[1]

The *completeness property* is one which distinguishes $\mathbb{R}$ from other infinite linearly ordered sets. If the set $\mathscr{D}$ does not satisfy this property it disqualifies it from being called the "real numbers". We will show that the set of all Dedekind cuts, linearly ordered as described above, passes the test for completeness. We first prove a lemma.

Lemma 17.5 The union of a non-empty set of Dedekind cuts is either itself a Dedekind cut or is the set $\mathbb{Q}$.

Proof:

Given: That U is a non-empty set of Dedekind cuts.
What we are required to show: That $\cup\{V : V \in U\}$ is $\mathbb{Q}$ or is of the form $(\leftarrow_{\mathbb{Q}} r)$ for some $r \in \mathbb{R}$.

Case 1: Suppose $U = \{(\leftarrow_{\mathbb{Q}} t) : t \in \mathbb{R}\}$. Since every rational number q belongs to $(\leftarrow_{\mathbb{Q}} t)$ for some real number t, then $\mathbb{Q} = \bigcup_{t\in\mathbb{R}} (\leftarrow_{\mathbb{Q}} t)$, and we are done.

Case 2: Suppose that $\cup\{V : V \in U\} \neq \mathbb{Q}$. Then there is a proper non-empty subset $M \subset \mathbb{R}$ such that $U = \{(\leftarrow_{\mathbb{Q}} t) : t \in M\}$ and $\bigcup_{t\in M} (\leftarrow_{\mathbb{Q}} t) \neq \mathbb{Q}$. It suffices for us to show that $\bigcup_{t\in M} (\leftarrow_{\mathbb{Q}} t)$ is a Dedekind cut.

Then there exists some $u \in \mathbb{Q}$ such that $u \notin (\leftarrow_{\mathbb{Q}} t)$ for all $t \in M$. Then $t < u$ for all $t \in M$. This means that u is an upper bound of $M \subset \mathbb{R}$.
By the completeness principle for the real numbers, since M is bounded in $\mathbb{R}$, M has a least upper bound, say $v \in \mathbb{R}$.

We claim that $\bigcup_{t\in M} (\leftarrow_{\mathbb{Q}} t) = (\leftarrow_{\mathbb{Q}} v)$.

– We first show that $(\leftarrow_{\mathbb{Q}} v) \subseteq \bigcup_{t\in M} (\leftarrow_{\mathbb{Q}} t)$:
Let $z \in (\leftarrow_{\mathbb{Q}} v)$. Then there exists $t \in M$ such that $z < t < v$ (for if $t \leq z$ for all $t \in M$, then z is an upper bound of M, a contradiction of the definition of v). So $z \in (\leftarrow_{\mathbb{Q}} t) \in U$. Then $(\leftarrow_{\mathbb{Q}} v) \subseteq \bigcup_{t\in M} (\leftarrow_{\mathbb{Q}} t)$ must hold true.

[1]Note that this property is often expressed in many different but equivalent forms. The following properties are all equivalent to the Completeness property: 1) The limit of every infinite decimal sequence is a real number, 2) Every bounded monotonic sequence is convergent, 3) A sequence is convergent if and only if it is a Cauchy Sequence. Googling the words "Completeness property" may direct the internet surfer to any one of these.

– We now show that $\bigcup_{t\in M}(\leftarrow_{\mathbb{Q}} t) \subseteq (\leftarrow_{\mathbb{Q}} v)$:
Let $u \in (\leftarrow_{\mathbb{Q}} t)$ for some $t \in M$. Since $t \in M$, $t < v$. Then $u \in (\leftarrow_{\mathbb{Q}} t) \subset (\leftarrow_{\mathbb{Q}} v)$. So $\bigcup_{t\in M}(\leftarrow_{\mathbb{Q}} t) \subseteq (\leftarrow_{\mathbb{Q}} v)$ as claimed.

So $\bigcup_{t\in M}(\leftarrow_{\mathbb{Q}} t) = (\leftarrow_{\mathbb{Q}} v)$, a Dedekind cut as claimed.
So the union $\bigcup_{t\in M}(\leftarrow_{\mathbb{Q}} t)$ of all elements of $U = \{(\leftarrow_{\mathbb{Q}} t) : t \in M\}$ is a Dedekind cut.

Theorem 17.6 Let $\mathscr{D}$ denote the set of all Dedekind cuts linearly ordered by $\subset$. Then if $\mathscr{S}$ is a non-empty bounded subset of $\mathscr{D}$, $\mathscr{S}$ has a least upper bound (with respect to the ordering $\subset$).

Proof:

We are given that $\mathscr{S}$ is a non-empty bounded subset of $\mathscr{D}$. Then $\mathscr{S}$ is of the form

$$\mathscr{S} = \{(\leftarrow_{\mathbb{Q}} t) : t \in M \subset \mathbb{R}\}$$

for some proper subset M of $\mathbb{R}$. Since $\mathscr{S}$ is a bounded subset of $\mathscr{D}$ with respect to $\subset$, there exists $k \in \mathbb{R}$ such that $(\leftarrow_{\mathbb{Q}} t) \subseteq (\leftarrow_{\mathbb{Q}} k)$ for all $t \in M$. Then its union $U = \bigcup_{t\in M}(\leftarrow_{\mathbb{Q}} t)$ cannot be all of $\mathbb{Q}$. By the lemma, its union U is a Dedekind cut, say, $(\leftarrow_{\mathbb{Q}} v)$.
We claim that $U = (\leftarrow_{\mathbb{Q}} v)$ is the least upper bound of $\mathscr{S}$ with respect to $\subset$: Since $\bigcup_{t\in M}(\leftarrow_{\mathbb{Q}} t) = (\leftarrow_{\mathbb{Q}} v)$, then $(\leftarrow_{\mathbb{Q}} t) \subseteq (\leftarrow_{\mathbb{Q}} v)$ for all t in M and so the Dedekind cut $(\leftarrow_{\mathbb{Q}} v)$ is an upper bound of $\mathscr{S}$. Suppose $(\leftarrow_{\mathbb{Q}} u)$ is another upper bound of $\mathscr{S}$. Then $U = \bigcup_{t\in M}(\leftarrow_{\mathbb{Q}} t) \subseteq (\leftarrow_{\mathbb{Q}} u)$. So $(\leftarrow_{\mathbb{Q}} v) \subseteq (\leftarrow_{\mathbb{Q}} u)$. We must conclude that $(\leftarrow_{\mathbb{Q}} v)$ is a least upper bound of $\mathscr{S}$.
So $\mathscr{D}$ satisfies the completeness property.

From this we conclude that the set of all Dedekind cuts, $\mathscr{D}$, represents the set of all real numbers in the *ZFC* set-theoretic universe. Thus, from the primitive concepts "class", "set" and "belongs to" and the axioms one to eight we have successfully defined the sets of natural numbers, integers, rational numbers and real numbers. The elements of these sets are themselves sets. If the existence of the natural numbers $\mathbb{N}$ is almost an immediate consequence of the Axiom of infinity, the other axioms provided the necessary tools to construct from $\mathbb{N}$ the integers, rationals and real numbers (as sets).

Remark: Having defined the real numbers as Dedekind cuts, we see that every real number u is a subset of $\mathbb{Q}$ and so $u \in \mathscr{P}(\mathbb{Q})$. Since $\mathbb{Q} \subseteq \mathscr{P}^6(\mathbb{N})$ (see page 150), $u \in \mathscr{P}(\mathscr{P}^6(\mathbb{N})) = \mathscr{P}^7(\mathbb{N})$ and so, $\mathbb{R} \subseteq \mathscr{P}^7(\mathbb{N})$. We can now see why the *Axiom of power set* plays an essential role in the *ZFC*-universe. Had we not declared that "$\mathscr{P}(S)$ is a set whenever S is a set" what guarantee would we have that the real

numbers exists in our universe.

We have not yet invoked two axioms: Axiom A9, called the *Axiom of regularity* and the *Axiom of choice*. These two axioms will help us handle certain difficulties encountered while dealing with infinite sets, the main subject of our investigation for the rest of this book.

Concepts review:

1. What is an initial segment? What is its leader?
2. How is addition of initial segments defined?
3. If r and t are positive real numbers, how is multiplication of $(_{\leftarrow_{\mathbb{Q}}}r)$ and $(_{\leftarrow_{\mathbb{Q}}}t)$ defined?
4. Provide a definition of Dedekind cuts.
5. Define a function f which maps $\mathbb{R}$ one-to-one onto the set $\mathscr{D}$ of all Dedekind cuts.
6. Give a set-theoretic definition of the real numbers.
7. What can we say about the union of a family of Dedekind cuts?
8. What does the *Completeness property of the reals* (equivalently, *Least upper bound principle of the real numbers*) state?
9. Does the set of all Dedekind cuts satisfy the *Completeness property*?
10. How are the elements of all Dedekind cuts ordered?
11. How does a Dedekind cut representing a rational number differ from one representing an irrational number?

EXERCISES

A. 1. Perform the following operations on the given Dedekind cuts:
   a) $(_{\leftarrow_{\mathbb{Q}}}-3) \times (_{\leftarrow_{\mathbb{Q}}}1/3)$
   b) $(_{\leftarrow_{\mathbb{Q}}}1/3) + (_{\leftarrow_{\mathbb{Q}}}5)$

c) $(\leftarrow_{\mathbb{Q}} 5) \times (\leftarrow_{\mathbb{Q}} 0)$

d) $(\leftarrow_{\mathbb{Q}} -1) + (\leftarrow_{\mathbb{Q}} 1)$

2. Show why $(\leftarrow_{\mathbb{Q}} -1) + (\leftarrow_{\mathbb{Q}} 5) < (\leftarrow_{\mathbb{Q}} 1/3) + (\leftarrow_{\mathbb{Q}} 10)$.
3. What is the least upper bound of $\bigcup_{t<0} (\leftarrow_{\mathbb{Q}} t)$?
4. Find a Dedekind cut strictly in between the two cuts $(\leftarrow_{\mathbb{Q}} 1/2)$ and $(\leftarrow_{\mathbb{Q}} 2/3)$.
5. Is there a Dedekind cut that contains all Dedekind cuts? If so, what is it? If not, why not?
6. How does the set-theoretic definition of $\mathbb{R}$ guarantee that it is a set?
7. What link can we make between the set theoretic definition of $\mathbb{R}$ and the natural numbers $\mathbb{N}$?

B.

8. Show that any set of the form $(\leftarrow_{\mathbb{Q}} r)$, where r is a real number, satisfies the three conditions in the formal definition of a Dedekind cut.

C.

9. Let $f : \mathbb{R} \to \mathscr{P}(\mathbb{Q})$ be a function defined as follows: $f(x)$ is the least upper bound of $\{y \in \mathbb{Q} : y < x\}$. What is the image of $\mathbb{R}$ under the function f?

Part VI

Infinite sets

18 / Infinite sets versus finite sets.

Summary. *In this section we give a definition of "infinite set" as put forward by Richard Dedekind. We then establish a few of the most basic properties of infinite sets and compare them to those of finite sets. These properties allow us to characterize finite sets as those sets which are in one-to-one correspondence with some natural number, n. We also show that for any finite set S, $\mathscr{P}(S)$ is finite. Finally, we prove a version of the "Recursively defined function theorem".*

18.1 Infinite sets.

The notions of "finite set" and "infinite set" are often viewed as being opposites of each other, in the sense that if we define one of these, then the other is its negation. We all have an intuitive idea of what a *finite set* is and how it differs from an *infinite set*. Most would agree with the statement "A finite set is a set whose elements you can count so as to determine how large or how small it is". We would of course first have to explain what it means to "count" and what it means to determine "how many elements there are in a set". Our definition of "finite" would have to be such that we can say "every natural number is finite".

The concept of "infinite set" is abstract, purely an idealization of something we find useful when discussing certain topics, even though it impossible to perceive. Yet, for anyone who studies or uses mathematics in any field, whether it be engineering, physics, or social sciences, doing mathematics without referring to "infinity" or "infinite sets" would feel like trying to walk with both shoelaces tied together. When the word "infinite" is used in a conversation, only a mathematician might possibly ask "What do you mean when you say the word *infinite*?". Most individuals might embarrassingly respond "I don't really know, but I think everybody intuitively understands what we mean when using that word." Appropriately defining "infinite sets" is important, not because it would be an interesting mathematical exercise to do so, but because if we want to make sense of the mathematics we do today, we have no choice. Experience shows that doing mathematics without clearly defining the concepts we are referring to can lead to contradictions or erroneous results. In this text, we have informally used the words "finite" and "infinite" before, but never in a mathematical statement proven to be true or false. If we want to refer to infinite and finite sets in theorem statements or definitions these must be precisely defined. We choose to provide a definition of "*infinite set*" and then define "*finite set*" as being one which is not infinite.

18.2 Dedekind's definition of *infinite set*.

We start with the definition of an infinite set as evoked by the mathematician Richard Dedekind (1831 - 1916).

Definition 18.1 A set S is said to be an *infinite set* if there exists a one-to-one function mapping S onto a proper subset of itself. If a set S is not infinite, then we say that it is a *finite set.*[1]

Surprised? Well, it is better than saying that an infinite set is "a set with lots of things in it". (A set containing a billion pencils has a lot of things in it and yet no one would describe it as being an infinite set.) Dedekind's definition is succinct and without ambiguities since it is only expressed using words that have been previously defined. An infinite set is a set which properly contains a one-to-one image of itself. If we were to define finite sets as being those sets that are not infinite, then we could say that "a set S is finite if and only if S contains no proper subset T which is a one-to-one image of itself". For example, since the natural number $6 = \{0, 1, 2, 3, 4, 5\}$ cannot be in one-to-one correspondence with any of its elements, it cannot be infinite. When a function $f : A \to B$ maps a set A one-to-one into a set B we often say that f *embeds* A inside B in the sense that B contains a "copy" of A. Using this vocabulary we can say that "S is infinite if and only if it is *embedded into a proper subset of itself*".

18.3 Example: $\mathbb{N}$ is an infinite set.

The set, $\mathbb{N}$, of all natural numbers was defined as being the *smallest inductive set.* We test our definition of "infinite set" on $\mathbb{N}$. The set $\mathbb{N}$ is infinite only if we can produce a function $f : \mathbb{N} \to \mathbb{N}$ which embeds $\mathbb{N}$ into a proper subset of itself. Consider the function $f : \mathbb{N} \to \mathbb{N}$ defined as $f(n) = 2n$. We see that f is one-to-one and that the image, $f[\mathbb{N}] = \{0, 2, 4, 6, \ldots, \}$, of $\mathbb{N}$ under f is a proper subset of $\mathbb{N}$. We have proven that:

"*The natural numbers* $\mathbb{N}$ *is an infinite set.*"

18.4 Properties of infinite sets and finite sets.

The following theorems confirm that infinite and finite sets satisfy the properties we expect from them. Ultimately, we want to show that the finite sets are precisely those sets which are in one-to-one correspondence with some natural number. Surprisingly, this does not follow immediately from our definitions of infinite and finite sets.

Theorem 18.2 Basic properties of infinite and finite sets.

[1] Actually records show that Bolzano suggested in 1847, (before Dedekind) that an infinite set was a set that could be mapped one-to-one onto a proper subset of itself, a property was is not satisfied by finite subsets.

a) The empty set is a finite set.

b) Any singleton set is a finite set.

c) A set which has a subset which is infinite must itself be infinite.

d) A subset of a finite set must be finite.

Proof:

a) The empty set $\varnothing$ has no proper subsets and so a function f cannot map $\varnothing$ into a proper subset of $\varnothing$. So $\varnothing$ is finite.

b) The singleton set, $\{x\}$, contains only one element x. Since x is not a proper subset of x, the only proper subset of $\{x\}$ is $\varnothing$.[1] Then, for any well-defined function $f : \{x\} \to \{x\}$, $\varnothing$ cannot be the one-to-one image of $\{x\}$ under f. So singleton sets are finite.

c) What we are given: That X is an infinite subset of a set S.
What we are required to show: That S is infinite.
If $X = S$, then S is infinite and we are done. Suppose $X \neq S$. Since X is infinite, then there exists a one-to-one function f which maps X onto $f[X] \subset X$. Since $f[X] \subset X \subset S$, the set S is the pairwise disjoint union of the three sets $S - X$, $X - f[X]$ and $f[X]$. Define the map $g : S \to (S - X) \cup f[X]$ as follows:

$$g(x) = \begin{cases} x & \text{if} \quad x \in S - X \\ f(x) & \text{if} \quad x \in X \end{cases}$$

Then g maps S one-to-one onto $(S - X) \cup f[X]$, a proper subset of S. So, by definition, S is an infinite set.

d) Suppose F is a finite set and $X \subseteq F$. We are required to show that X is finite. Suppose X is an infinite set. Then by the statement in part c) F must be infinite, contradicting our hypothesis. So X must be finite.

Theorem 18.3 Let $f : X \to Y$ be a one-to-one function mapping X onto Y. The set X is infinite if and only if the set Y is infinite.

Proof:

($\Rightarrow$) What we are given: That X is an infinite set and that $f : X \to Y$ is a *one-to-one* function mapping X onto a set Y.
What we are required to show: That Y is infinite.
Since X is infinite, by definition, there exists a function, $g : X \to g[X]$, mapping X

[1] Even in the odd case where $x \in x$, x can *not* be said to be a proper subset of x since $x = x$.

one-to-one onto a proper subset $g[X]$ of X. Then $f|_{g[X]} : g[X] \to Y$ maps the proper subset $g[X]$ of X one-to-one into Y. Since f maps X one-to-one onto Y, it has an inverse f^{-1} mapping Y one-to-one and onto X. Then the function

$$\left[(f|_{g[X]}) \circ g \circ f^{-1}\right] : Y \to Y$$

maps Y one-to-one onto $f|_{g[X]}[g[X]] = f[g[X]]$, a proper subset of $f[X] = Y$. So, by definition, Y is an infinite set.

($\Leftarrow$) Suppose Y is infinite. Then, since $f^{-1} : Y \to X$ is a one-to-one map from Y onto X, by the first part of this proof, X is infinite.

Corollary 18.4 The one-to-one image of a finite set is finite.

Proof: The proof is left as an exercise.

Lemma 18.5 If S is an infinite set and $a \in S$, then $S - \{a\}$ is an infinite set.

Proof:

What we are given: That S be an infinite set and $a \in S$.
What we are required to show: That $S - \{a\}$ is an infinite set.
Since S is infinite, then there exists a one-to-one function $g : S \to S$ such that $g[S] \subset S$. We will show that $S - \{a\}$ is infinite by exhibiting a one-to-one function h on $S - \{a\}$ such that $h[S - \{a\}]$ is a proper subset of $S - \{a\}$. Choose an arbitrary element $k \in S - g[S]$.

- Case A: Suppose $a \in g[S]$. Then there is some $u \in S$ such that such that $g(u) = a$.
  - Subcase A-1: Suppose $u \neq a$. Define a function h on $S - \{a\}$ as follows:

$$h(x) = \begin{cases} g(x) & \text{if} \quad x \in S - \{a, u\} \\ k & \text{if} \quad x = u \end{cases}$$

    Since g is one-to-one on $S - \{a, u\}$, then so is h. Furthermore, h uniquely maps u to $k \in S - g[S]$. So h is one-to-one on $S - \{a\}$. See that neither of the elements $g(u)$ and a belong to $h[S - \{a\}]$. So $h[S - \{a\}]$ is a proper subset of $S - \{a\}$.

  - Subcase A-2: Suppose $u = a$. Choose an element $v \neq a$ in $g[S]$. Define a function h on $S - \{a\}$ as follows:

$$h(x) = \begin{cases} g(x) & \text{if} \quad x \in S - \{a, v\} \\ k & \text{if} \quad x = v \end{cases}$$

    Since g is one-to-one on $S - \{a, v\}$, then so is h. Furthermore, h uniquely maps v to k in $S - g[S]$. So h is one-to-one on $S - \{a\}$. Also see that neither $g(v)$ nor a belong to $h[S - \{a\}]$. So $h[S - \{a\}]$ is a proper subset of $S - \{a\}$.

– Case B: Suppose $a \in S - g[S]$. Define a function h on $S - \{a\}$ as follows:

$$h(x) = g|_{S-\{a\}}(x) \text{ for all } x \in S - \{a\}$$

Since no other element in $S - \{a\}$ is mapped to $g(a)$ by g, neither $g(a)$ nor a belong to $h[S - \{a\}]$. So $h[S - \{a\}]$ is a proper subset of $S - \{a\}$.

We conclude that $S - \{a\}$ is infinite.

Theorem 18.6 Every natural number n is a finite set.

Proof:

The proof is by induction. Let $P(n)$ be the property "The natural number n is finite". Since $0 = \varnothing$ is finite, then P(0) holds true. Suppose the natural number $n = \{0, 1, 2, 3, \ldots, n-1\}$ is finite. We claim that $n + 1 = n^+ = \{0, 1, 2, 3, \ldots, n\}$ must be finite. Suppose not. That is, suppose n^+ is infinite. By the lemma 18.5, $(n + 1) - \{n\} = n$ must also be infinite contradicting the fact that $P(n)$ holds true. So $P(n + 1)$ must hold true. By the principle of mathematical induction, $P(n)$ holds true for all natural numbers n. Thus, every natural number is a finite set.

In the proof of the following corollary we require the *Axiom of choice* to justify a particular step. This is the first time we invoke this axiom. We will discuss the axiom of choice in length later on. For now we will simply state it and point out the step where it is invoked:

> **Axiom of choice**: For every set $\mathscr{A}$ of non-empty sets there is a function f which associates to every set A in $\mathscr{A}$ an element $a \in A$.

At first, it seems rather harmless enough. It says that if we have a set of non-empty sets, then we can choose from each set one element. If the set of sets has only finitely many sets, then the Axiom of choice is not required. The sticky point is the one encountered when the set contains infinitely many sets. If a theorem statement invokes the *Axiom of choice* in its proof, it is common practice to alert the reader to this fact by posting the acronym [AC].[1]

Corollary 18.7 [AC] A set S is finite if and only if S is empty or it is in one-to-one correspondence with some natural number n.

Proof:

[1]The reader may see the posting of this acronym as a challenge by the author asking: "Is it possible to prove this statement without invoking the Axiom of choice?".

($\Leftarrow$) Suppose S is empty or is the one-to-one image of a natural number n. Since $\varnothing$ is finite and every natural number n is finite, then S must be finite (by Corollary 18.4 and theorem 18.6).

($\Rightarrow$) Conversely, suppose S is a non-empty finite set.
We are required to show that there exists a natural number n which can be mapped one-to-one onto S.
Suppose not. That is, suppose there does not exist a natural number n which maps one-to-one onto S. We claim that S must then be infinite, contradicting our hypothesis.

Proof of claim: We prove the claim by constructing a one-to-one function $f : \mathbb{N} \to S$ which maps $\mathbb{N}$ into S.

- Choose an element s_0 in S to form the subset $S_1 = \{s_0\}$ of S. Define the function $f : \{0\} \to \{s_0\}$ as $f(0) = s_0$. Then $S - S_1$ is non-empty, for if it was empty, then $S = S_1$ would be the one-to-one image of $\{0\}$ under the function f contradicting the fact that S is not the one-to-one image of a natural number. So we can choose an element s_1 from $S - S_1$ to construct the subset $S_2 = \{s_0, s_1\}$. Define the one-to-one function $f : \{0, 1\} \to \{s_0, s_1\}$ as $f(i) = s_i$ for $i = 1, 2$.
- Suppose we have inductively constructed the subset $S_n = \{s_0, s_1, s_2, \ldots, s_{n-1}\}$ of S where $f : \{0, 1, \ldots, n-1\} \to S_n$ is the one-to-one function defined as $f(i) = s_i$. Then to avoid a contradiction, $S - S_n$ must be non-empty. The *Axiom of choice* provides us with the choice function $k : \mathscr{P}(S) \to S$ which allows us to choose from each set $S - S_n$ an element s_n from which we define the one-to-one function $f : \{0, 1, \ldots, n\} \to S_{n+1}$ defined as $f(i) = s_i$ if $i < n$ and $f(n) = k(S - S_n) = s_n$. We can in this way "inductively" construct a one-to-one function $f : \mathbb{N} \to S$ mapping $\mathbb{N}$ into S. Then S contains a one-to-one image $f[\mathbb{N}] = \{s_0, s_1, s_2, s_3, \ldots\}$ of $\mathbb{N}$. By part c) of theorem 18.2, S must be infinite. This contradicts the part of our hypothesis in which S was declared to be finite. The source of this contradiction is the statement "there does not exist a natural number n which maps one-to-one onto S." So any finite subset is the one-to-one image of some natural number.

A few words on the "inductively" constructed function in the above proof. In the proof of the corollary, we have constructed a one-to-one function $f : \mathbb{N} \to S$ by defining $f(i)$ for one number i at a time. For each n, the value of $f(n)$ depends on the values of $f(i)$ for each $i < n$. This is because the choice function k assigns to the set $S - \{s_0, s_1, \ldots, s_{n-1}\}$ an element s_n. The value of $f(n)$ is then set to be equal to s_n.

We conveniently declared that an "inductively constructed function f" is produced in this way, as if it was clear that this process will automatically produce well-defined functions. Even though it seems like a reasonably safe method for constructing functions, we should not be blind to an element of uncertainty involved in this process. It does *not* immediately follow from the definition of a function that this method for

constructing functions will always produce a function. If one asserts that this is obvious, then why not produce a proof that shows us how "obvious" it really is. We will immediately state the theorem which guarantees that functions constructed in this way are valid but defer its proof to the end of this section to avoid digressing from our discussion of finite and infinite sets.

Theorem 18.8 *The recursive function theorem.* Let S be a set. Let $k : \mathscr{P}(S) \to S$ be a function on $\mathscr{P}(S)$ and $f \subseteq \mathbb{N} \times S$ be a relation. We write $f(n) = a$ if and only if $(n, a) \in f$. Let $m \in S$. Suppose the relation, f, satisfies the two properties

$$\begin{cases} f(0) = m & \Rightarrow \quad (0, m) = (0, f(0)) \in f \\ (n, f(n)) \in f & \Rightarrow \quad (n+1, k(S - \{f(0), f(1), \ldots, f(n)\}) = (n+1, f(n+1)) \in f \end{cases}$$

Then f is a well-defined function on $\mathbb{N}$.

Proof: The proof appears at the end of this section.

We have shown that "counting" the elements in a finite set S comes down to determining which natural number n is mapped one-to-one onto S. We are essentially assigning to each of the n elements of S the labels $0, 1, 2, 3, \ldots, n-1$. The corollary above shows that we could have defined finite sets as follows:

> **Definition**: A set S is a *finite set* if and only if it can be mapped one-to-one onto some natural number n. If we say that
>
> "the finite set S *contains* n *elements*"
>
> we mean that S is the one-to-one image of the natural number n. So "S is a finite set" and "S contains n elements for some n" are equivalent expressions. A set is an *infinite set* if it is not the one-to-one image of some natural number.

All the definitions and theorems stated and proved above would logically follow from this definition of finite sets. The following theorem provides another characterization of infinite sets.

Theorem 18.9 [AC] A set S is an infinite set if and only if it contains a one-to-one image of the set of natural numbers $\mathbb{N}$.

Proof:

($\Leftarrow$) If S contains a subset U which is a one-to-one image of $\mathbb{N}$, then U is infinite (by theorem 18.3) and so S is infinite (by theorem 18.2).

($\Rightarrow$) The proof is left as an exercise. (The proof mimics the proof of Corollary 18.7.)

Theorem 18.10 [AC] *If the set S is a finite set and $f : S \to X$ is a function, then $f[S]$ is finite*[1].

Proof:

Suppose the set S is a finite set and $f : S \to X$ is a function mapping S into some set X. We are required to show that $f[S]$ is a finite set. Suppose $f[S]$ is an infinite set. Then there exists a function $g : \mathbb{N} \to f[S]$ mapping $\mathbb{N}$ into $f[S]$ (18.9). Say $g[\mathbb{N}] = \{g(0), g(1), g(2), g(3), \ldots\} \subseteq f[S]$. Then for each $i \in \mathbb{N}$, $f^{\leftarrow}(g(i))$ is a non-empty subset of S. For each $i \in \mathbb{N}$ we can choose an element s_i from $f^{\leftarrow}(g(i))$ (the Axiom of choice allows us to choose an infinite number of elements in this way). So S contains a subset $\{s_1, s_2, s_3, \ldots\}$ of distinct elements. Let $h : \mathbb{N} \to S$ be the function defined as $h(i) = s_i$. This set is infinite since it is a one-to-one image of $\mathbb{N}$ under the function h. Since S contains an infinite subset, then it must be infinite (by 18.2). A contradiction! So $f[S]$ is a finite set.

Theorem 18.11 If a set S contains n elements, then $\mathscr{P}(S)$ contains 2^n elements[2]. Hence, if a set S is a finite set, then the set $\mathscr{P}(S)$ is finite.

Proof:

The proof is by induction.

For a natural number n, let S_n denote a subset of S which contains n elements. That is, $S_0 = \varnothing$, $S_1 = \{s_0\}$, $S_2 = \{s_0, s_1\}, \ldots, S_n = \{s_0, s_1, \ldots, s_{n-1}\}$. For each natural number n let $P(n)$ denote the statement

"The power set $\mathscr{P}(S_n)$ contains 2^n elements"

Base case: If $n = 0$, then $S_0 = \varnothing$ and so $\mathscr{P}(S_0) = \{\varnothing\}$ contains $1 = 2^0$ element so $P(0)$ holds true.

Inductive hypothesis: Suppose $P(n)$ holds true. That is, suppose that for any set of n elements the set $\mathscr{P}(S_n)$ contains 2^n elements. Let $S_{n+1} = \{s_0, s_1, s_2, \ldots, s_n\}$ be a set

[1]Note that f need not be a one-to-one function for this to hold true.

[2]If $n = 0$ we define $2^0 = 1, 2^1 = 2^0 \times 2$. If n is a natural number other than 0 we define $2^n = \underbrace{2 \times 2 \times \cdots \times 2}_{n \text{ times}}$.

containing $n+1$ distinct elements. Then if $S_n = \{s_0, s_1, s_2, \ldots, s_{n-1}\}$, by the inductive hypothesis, we can express $\mathscr{P}(S_n)$

$$\mathscr{P}(S_n) = \{U_0, U_1, U_2, \ldots, U_{2^n-1}\}$$

where the U_i's represent all distinct subsets of S_n; we will suppose that $U_0 = \varnothing$ and $U_{2^n-1} = S_n$. Since the elements of the set S_{n+1} are distinct, then $s_n \notin S_n$, and so $\{s_n\} \notin \mathscr{P}(S_n)$. Then $\mathscr{P}(S_n) \subset \mathscr{P}(S_{n+1})$. For each $i = 0$ to $2^n - 1$ define $V_i = U_i \cup \{s_n\}$. We see that $\mathscr{P}(S_n) \cap \{V_0, V_1, \ldots, V_{2^n-1}\} = \varnothing$ since every V_i contains the element s_n. Furthermore, every element in $\mathscr{P}(S_{n+1})$ is accounted for in $\mathscr{P}(S_n) \cup \{V_0, V_1, \ldots, V_{2^n-1}\}$. Then

$$\mathscr{P}(S_{n+1}) = \{U_0, U_1, U_2, \ldots, U_{2^n-1}, V_0, V_1, V_2, \ldots, V_{2^n-1}\}$$

We see that $\mathscr{P}(S_{n+1})$ contains $2^n \times 2 = 2^{n+1}$ elements. So $P(n+1)$ holds true.
By the principle of mathematical induction, $P(n)$ holds true for all natural numbers n. We conclude that for any set S which contains n elements, the set $\mathscr{P}(S)$ contains 2^n elements. Thus, if S is finite, then so is $\mathscr{P}(S)$.

18.5 Proof of the *recursive function theorem.*

A recursively defined function $f : \mathbb{N} \to S$ (on $\mathbb{N}$) is a function which is defined one term at a time. The process begins by defining $f(0)$ at 0. Then, for each $n \geq 0$, the value of $f(n)$ is determined based on the values previously assigned to each of $f(0)$, $f(1)$, $f(2)$, ..., $f(n-1)$. The theorem explicitly states the conditions under which this method of defining a function is valid.

The recursive function theorem: Let S be a set. Let $k : \mathscr{P}(S) \to S$ be a function mapping subsets of S to elements of S and $f \subseteq \mathbb{N} \times S$ be a relation. We write "$f(n) = a$" if and only if $(n, a) \in f$. Let $m \in S$. Suppose the relation f satisfies the two properties

$$\begin{cases} f(0) = m & \Rightarrow \quad (0, m) = (0, f(0)) \in f \\ (n, f(n)) \in f & \Rightarrow \quad (n+1, k(S - \{f(0), f(1), \ldots, f(n)\}) \\ & \qquad\qquad = (n+1, f(n+1)) \in f \end{cases}$$

Then f is a well-defined function on $\mathbb{N}$.

Proof:

Let $\mathscr{S}$ be a class of relations R in $\mathbb{N} \times S$ which contain $(0, f(0)) = (0, m)$ and satisfy the condition:

$$\{(n, f(i)) : i \leq n\} \subseteq R \quad \Rightarrow \quad (n+1, k(S - \{f(0), f(1), \ldots, f(n)\}) = (n+1, f(n+1)) \in R$$

Now $\mathscr{S}$ is non-empty since it contains $\mathbb{N} \times S$. Let $f^* = \bigcap_{R \in \mathscr{S}} R$. This means that f^* is the smallest set of ordered pairs satisfying the conditions described for $\mathscr{S}$. The relation f^* looks something like

$$f^* = \{(0, f(0)), (1, k(S - \{f(0)\})), (2, k(S - \{f(0), f(1)\})), \cdots\}$$

We claim that f^* is a function mapping $\mathbb{N}$ into S.

Proof of claim: We first establish that $\operatorname{dom} f^* = \mathbb{N}$. We first note that $0 \in \operatorname{dom} f^*$. If $n \in \operatorname{dom} f^*$, then $(n, f(n)) \in f^*$. This implies

$$(n+1, k(S - \{f(0), f(1), \ldots, f(n)\})) = (n+1, f(n+1)) \in f^*$$

and so $n + 1 \in \operatorname{dom} f^*$. Hence, by induction, the domain of f^* is all of $\mathbb{N}$. We now proceed to the proof of the claim.
The proof of the claim is by the second version of mathematical induction. Let $P(n)$ denote the statement "$[(n, x) \in f^*$ and $(n, y) \in f^*] \Rightarrow [x = y]$".

- Inductive hypothesis: Suppose $P(m)$ holds true for all natural numbers $m < n$. That is, $(m, x) \in f^*$ and $(m, y) \in f^*$ implies $x = y$. We will show that given our hypothesis, $P(n)$ must hold true.
  - Suppose not. Suppose $(n, x) \in r$ and $(n, y) \in f^*$ where $x \neq y$. Let $U = f^* - \{(n, y)\}$ (the set f^* take away the element (n, y)). Then we easily see that U is one of the relations in $\mathscr{S}$. The fact that U is strictly smaller than f^*, previously declared to be the smallest of the relations in $\mathscr{S}$, is a contradiction. Then x must be equal to y. We conclude that $P(n)$ holds true as required.

By mathematical induction, $P(n)$ holds true for all n. So f^* is a well-defined function as claimed.

We will now show that f^* is unique. Let g be another function satisfying the conditions $(0, m) = (0, g(0)) \in g$ and

$$\{(n, g(i)) : i \leq n\} \subseteq f^* \Rightarrow (n+1, k(S - \{g(0), g(1), \ldots, g(n)\}) = (n+1, g(n+1)) \in f^*$$

We prove uniqueness by induction: Let $P(n)$ denote the statement "$f^*(n) = g(n)$".

- Inductive hypothesis: Suppose $P(m)$ holds true for all $m < n$. That is, $f^*(m) = g(m)$ for all $m < n$.
  Then by definition of f^*,

$$\begin{aligned} (n, k(S - \{f(0), f(1), \ldots, f(n-1)\})) &= (n, f(n)) \\ &= (n, k(S - \{g(0), g(1), \ldots, g(n-1)\})) \\ &= (n, g(n)) \end{aligned}$$

  Hence, $P(n)$ holds true.
  By mathematical induction $f^*(n) = g(n)$ for all n, and so f^* is unique as claimed.

We conclude that the function $f = f^*$ is a function which is uniquely defined by the given conditions.

Concepts review:

1. What is the definition of *infinite set* as put forward by Dedekind?
2. From Dedekind's definition of *infinite set* how can we show that $\mathbb{N}$ is infinite?
3. How do we define a *finite set*?
4. Is the empty set a finite set?
5. Is a subset of a finite set finite?
6. If S is an infinite set and $u \in S$, must $S - \{u\}$ be infinite?
7. If a set S has a subset which is infinite is S necessarily infinite?
8. If a set S is infinite and $f : S \to Y$ is a one-to-one mapping onto a set Y, what can we say about the set Y?
9. What can we say about one-to-one images of finite sets?
10. If a set S is infinite and we remove two elements from this set is the resulting set necessarily infinite?
11. We know that natural numbers are sets. Are all natural numbers necessarily finite sets?
12. Statement: "Any finite set S is necessarily the one-to-one image of some natural number." Is this statement true or false?
13. Statement: "An infinite set is the one-to-one image of the natural numbers." Is this statement true or false?
14. Statement: "An infinite set necessarily contains a subset which is a one-to-one image of the natural numbers." Is this statement true or false?
15. What can we say about the image (not necessarily one-to-one) of a finite set?
16. Is saying "S is a finite set" equivalent to saying "S contains n elements" where n is a suitable natural number?
17. If S is a finite set, is it necessarily true that $\mathscr{P}(S)$ is a finite set?
18. If S contains 6 elements how many elements does $\mathscr{P}(S)$ contain?

EXERCISES

A.
1. Prove that the one-to-one image of a finite set is finite.
2. Prove that:
   a) $\mathbb{Z}$ is an infinite set.
   b) $\mathbb{Q}$ is an infinite set.
   c) $\mathbb{R}$ is an infinite set.

B.
3. Prove that if S is infinite, then $S \times S$ is infinite.
4. Prove that if S is finite, then $S \times S$ is finite.
5. Prove that if S and T are infinite, then $S \cup T$ is infinite.
6. Prove that the union of two finite sets is finite.
7. Prove that the union of a finite set of sets is finite. (You may use the result in question 6 combined with a proof by induction on the number of finite sets.)
8. Prove that if $S \cup T$ is infinite, then either S or T is infinite.
9. Prove that if the set A is infinite and B is any set, then the set $A \times B$ must be infinite.

C.
10. Prove that if F is a finite subset of an infinite set S, then $S - F$ is infinite.
11. Prove that if a set S is an infinite set, then it contains a one-to-one image of the set of natural numbers $\mathbb{N}$.

---

19 / Countable and uncountable sets.

Summary. *In this section we define the words "equipotent sets", "countable sets" and "uncountable sets". We show that* $\mathbb{Q}$ *and* $\mathbb{Z}$ *are countable since we can produce a one-to-one correspondence between each of these and* $\mathbb{N}$*. We also show that* $\mathbb{N}\times\mathbb{Z}$ *and* $\mathbb{N}\times\mathbb{N}$ *are countable. Countable unions of countable sets are shown to be countable. We also show that no such one-to-one correspondence can exist between* $\mathbb{R}$ *and* $\mathbb{N}$ *and so* $\mathbb{R}$ *is uncountable.*

19.1 Can we compare infinite sets as we do finite sets?

"Finite sets are precisely those sets which are in one-to-one correspondence with some natural number n" is one of the simplest and most intuitive characterizations of finite sets. It essentially allows us to characterize sets according to their size. For example, suppose the two sets S_n and S_m can be mapped one-to-one and onto the natural numbers n and m, respectively. We can declare S_m to be *larger* than S_n if and only if $n \subset m$. We can declare them to be the same *size* if and only if n and m are the same natural number. Note that declaring two sets S and T to be the same size does not mean that they are equal. It simply states that they are both one-to-one images of the same natural number.

Can the described method for comparing finite sets be used to compare infinite sets? One might ask why we would want to compare infinite sets in this way and argue that "infinite sets need no comparing since, intuitively, they are all as big as a set can be". But our intuition is not always reliable, particularly when referring to infinite sets as they are defined in the *ZFC*-universe of sets. Verifying whether the infinite sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ are one-to-one images of each other is a question worth investigating.

19.2 Equipotent sets.

One of the set theory axioms states that two sets are equal provided they contain the same elements. We also say that two finite sets A and B (not necessarily equal) which are both one-to-one images of the same natural number n are the same *size*, or *contain the same number of elements*. If two sets A and B contain n elements it of course follows that these two sets are one-to-one images of each other. Any two *infinite* sets S and T can also be one-to-one images of the other. The set $\mathbb{N}$ and the set of even natural numbers is an example. We introduce a term used to describe this relation between two sets.

Definition 19.1 Two sets, A and B, are said to be *equipotent* sets if there exists a one-to-one function, $f : A \to B$, mapping A onto B. If A and B are *equipotent* we will say that

"A is equipotent to B" or "A is equipotent with B".

So *equipotent* finite sets are precisely those finite sets which are equipotent to the same natural number. We know that for finite sets A and B, "A is equipotent to a proper subset of B" and "A is smaller than B" are equivalent statements; we cannot however say that an infinite set A which is equipotent to to a proper subset of a set B is necessarily "smaller" than B. For example, even if the set $\mathbb{N}$ is easily seen to be equipotent to the proper subset, $\{0, 4, 8, 12, \ldots, \}$, of itself (via the function $f(n) = 4n$), we instinctively hesitate to say that $\mathbb{N}$ is a "smaller" set than $\{0, 2, 4, 6, \ldots, \}$ or that $\{0, 2, 4, 6, \ldots, \}$ is smaller than $\mathbb{N}$. The words "smaller than" seem to have a precise meaning only when discussing finite sets.

19.3 Countable sets.

We will encounter many kinds of infinite sets. It will be helpful if we can categorize infinite sets into subfamilies of equipotent sets. We will be particularly interested in those infinite sets which are equipotent with $\mathbb{N}$. Such sets are said to be *countably infinite sets*.

Definition 19.2 *Countable sets* are those sets that are either finite or equipotent to $\mathbb{N}$. Infinite countable sets are said to be *countably infinite*. Those infinite sets which are not countable are called *uncountable sets*.

So the adjective "countable" means "to be a one-to-one image of some subset of $\mathbb{N}$". Does it make sense to speak of a proper class of elements which is "countable"? The Axiom of replacement (A7) guarantees that all "countable classes" are sets. To see this we recall the Axiom for replacement.

> *Axiom of replacement*: Let A be a set. Let $\phi(x, y)$ be a formula which associates to each element, x, of A an element y in such a way that whenever both $\phi(x, y)$ and $\phi(x, z)$ hold true, $y = z$. Then there exists a set B which contains all elements y such that $\phi(x, y)$ holds true for some $x \in A$.[1]

This axiom dictates that if A is a set and B is a non-empty class of elements and $f \subseteq A \times B$ is a relation which satisfies the property "$(x, y) \in f$ and $(x, z) \in f$, then $y = z$", then there exists a set C which contains the elements, $f(x)$, for all $x \in A$. This axiom guarantees that if A is a set and $f[A]$ contains only elements, then the

[1]This axiom is more often expressed as the *Replacement axiom schema* since it is in fact many axioms each differing only by the formula ϕ it refers to. So to be more precise, given a formula ϕ in set theory language, we would refer to it as axiom A7(ϕ) rather than A7.

functional image, $f[A] = \{x : f(a) = x \text{ for some } a \in A\}$, is a set. So any class which is a one-to-one image of a subset of $\mathbb{N}$ must be a set.

Our definition of countable sets does not guarantee that uncountable sets exist. We simply stated that those sets which are not countable (if any exist) will be called "uncountable sets". We will carefully study the sets we have encountered to this point and determine which ones are countable and which ones are not. Recall (from theorem 18.2 a)) that the empty-set, $\varnothing$, is finite and so, by definition, is countable. Of course, the set, $\mathbb{N}$, of all natural numbers is equipotent to itself and so is countable. Since no element of $\mathbb{N}$ is infinite every natural number is countable. We now investigate the set of all integers.

Example: Show that the set $\mathbb{Z}$ of all integers is countable.

Solution: Define the function $f : \mathbb{N} \to \mathbb{Z}$ as follows:

$$f(n) = \begin{cases} -\frac{n+1}{2} & \text{if } n \text{ is odd} \\ \frac{n}{2} & \text{if } n \text{ is even} \end{cases}$$

It is easily verified that f maps $\mathbb{N}$ one-to-one onto $\mathbb{Z}$. For example,

$$\begin{aligned} f(0) &= 0 \\ f(1) &= -1 \\ f(2) &= 1 \\ f(3) &= -2 \\ f(4) &= 2 \\ &\vdots \end{aligned}$$

So $\mathbb{Z}$ is an infinite countable set[1].

There are quite a few general statements that we can state about countable sets. We will find it very useful to know that subsets of countable sets and images of countable sets are countable.

Theorem 19.3 A subset of a countable set is countable.

Proof:

What we are given: That S is a countable set and T is a subset of S.
What we are required to show: That T is countable.

Case 1: If S is finite, then by theorem 18.2 d), every subset of S is finite and so T is countable; we are done.

[1]The reader should note that this is just one of many ways of proving that $\mathbb{Z}$ is countable.

Case 2: Suppose both S and T are infinite.

Since S is countable it is a one-to-one image, say $h[\mathbb{N}]$, of $\mathbb{N}$. Then the function $h : \mathbb{N} \to S$ can be expressed as $h(i) = x_i \in S$. We can then express S as a sequence $S = \{x_0, x_1, x_2, \ldots, x_n, \ldots\}$.

We are required to show that T is countable. This is done by constructing a function $f : \mathbb{N} \to T$ mapping $\mathbb{N}$ one-to-one onto T as follows:

- Let $g : \mathscr{P}(\mathbb{N}) \to \mathbb{N}$ be the function on $\mathscr{P}(\mathbb{N})$ which maps any subset of $\mathbb{N}$ to its least element. Since $\mathbb{N}$ has been shown to be well-ordered the function g is well-defined.
  - We recursively define the function $m : \mathbb{N} \to \mathbb{N}$ as follows:

$$\begin{aligned} m(0) &= g(\{i \in \mathbb{N} : x_i \in T\}) \\ m(k) &= g(\{i \in \mathbb{N} : x_i \in T - \{x_{m(0)}, x_{m(1)}, \ldots, x_{m(k-1)}\}) \end{aligned}$$

    By theorem 18.8, $m : \mathbb{N} \to \mathbb{N}$ is a well-defined function. Since T is infinite, the domain of m is $\mathbb{N}$. Furthermore m is a one-to-one strictly increasing function.
  - Note that $\{x_{m(i)} : i \in \mathbb{N}\} \subseteq T$. We claim: That $\{x_{m(i)} : i \in \mathbb{N}\} = T$: Suppose $x_k \in T$. Let $U = \{m(i) \in \mathbb{N} : m(i) < k\}$. Then

$$g(\{i \in \mathbb{N} : x_i \in T - \{x_{m(i)} : m(i) \in U\}\}) = k$$

    So $x_k \in \{x_{m(i)} : i \in \mathbb{N}\}$. We conclude that $\{x_{m(i)} : i \in \mathbb{N}\} = T$ as claimed.

- Define the function $f : \mathbb{N} \to T$ as $f(i) = x_{m(i)}$. Since f is one-to-one and onto T, then T is countable.

19.4 More examples of infinite countable sets.

Showing that an infinite set is countable can be a challenge since it requires constructing a function mapping $\mathbb{N}$ one-to-one onto a set. The above result stating that "subsets of countable sets are countable" is a useful tool for showing that some infinite sets are countable without specifically exhibiting a function which maps $\mathbb{N}$ onto it. We provide examples of other sets which are countable.

Example 1. Show that the set $\mathbb{N} \times \mathbb{Z}$ is countable.

Solution: Define the function $f : \mathbb{N} \times \mathbb{Z} \to \mathbb{Z} - \{0\}$ as follows:

$$f(m, n) = 2^m(2n - 1)$$

We claim that the function f is onto $\mathbb{Z} - \{0\}$:

– Let z be any non-zero integer. Then we can factor at most a finite number of 2's, say m 2's for some integer m (m possibly equal to 0), leaving behind a single (either positive or negative) odd factor $2n-1$ for some integer n. So $z = 2^m(2n-1)$. Thus, there exists an ordered pair $(m,n) \in \mathbb{N} \times \mathbb{Z}$ which is mapped to z. So f is onto as claimed.

We claim the function f is one-to-one:

– Suppose $x = y$ in $\mathbb{Z}-\{0\}$. Since x and y are integers there exists natural numbers m and n and integers s and t such that[1]

$$x = f(m,s) = 2^m(2s-1) = 2^n(2t-1) = f(n,t) = y$$

Suppose, without loss of generality, that $m \geq n$, then

$$\begin{aligned} 2^m(2s-1) = 2^n(2t-1) &\Rightarrow 2^{m-n}(2s-1) = 2t-1 \\ &\Rightarrow m = n \text{ and } 2s-1 = 2t-1 \text{ (RHS = odd } \Rightarrow \text{ m -n = 0)} \\ &\Rightarrow s = t \\ &\Rightarrow (m,s) = (n,t) \end{aligned}$$

So f is one-to-one as claimed.

We have shown that f is both one-to-one and onto. Thus, the sets $\mathbb{N} \times \mathbb{Z}$ and $\mathbb{Z}-\{0\}$ are equipotent. We have shown that $\mathbb{Z}$ is countable. Since $\mathbb{Z}-\{0\} \subset \mathbb{Z}$, then, by the previous theorem, $\mathbb{Z}-\{0\}$ is also countable. So there exists a one-to-one function g mapping $\mathbb{Z}-\{0\}$ onto $\mathbb{N}$. So $g^{-1} \circ f^{-1}$ maps $\mathbb{N}$ one-to-one onto $\mathbb{N} \times \mathbb{Z}$. So $\mathbb{N} \times \mathbb{Z}$ is countable, as required.

Example 2. Show that the set of rational numbers, $\mathbb{Q}$, is countable.

Solution: We will represent all rational numbers in the form

$$\mathbb{Q} = \{m/n : m \in \mathbb{N}, n \in \mathbb{Z}-\{0\},\ m/n \text{ is irreducible.}\}$$

The rational numbers in this set are irreducible so $a/b = c/d$ if and only if $a = c$ and $b = d$. We define the function $f : \mathbb{N} \times (\mathbb{Z}-\{0\}) \to \mathbb{Q}$ as follows:

$$f(m,n) = m/n$$

This function is clearly one-to-one and onto and so $\mathbb{N} \times (\mathbb{Z}-\{0\})$ and $\mathbb{Q}$ are equipotent. Since $\mathbb{N} \times (\mathbb{Z}-\{0\}) \subset \mathbb{N} \times \mathbb{Z}$ and $\mathbb{N} \times \mathbb{Z}$ was shown to be countable, then $\mathbb{N} \times (\mathbb{Z}-\{0\})$ is countable. Thus, $\mathbb{Q}$ is countable, as claimed.

[1]Note that any non-zero integer can be expressed as a product $2^m(2n-1)$ for some natural numbers m and n. For example, suppose we are given the integer 1584. If we factor out as many 2's as possible from 1584 we obtain 2^4 and we are left with an odd number $2 \cdot 50 - 1$. See that $1584 = 2^4(2 \cdot 50 - 1)$.

Example 3. Show that the set $\mathbb{N} \times \mathbb{N}$ is countable.

Solution: Since $\mathbb{N} \times \mathbb{Z}$ is countable and $\mathbb{N} \times \mathbb{N} \subset \mathbb{N} \times \mathbb{Z}$, then $\mathbb{N} \times \mathbb{N}$ is countable.

Theorem 19.4 Any finite product, $\mathbb{N} \times \mathbb{N} \times \cdots \times \mathbb{N}$, of $\mathbb{N}$ is a countable set.

Proof:

The proof is by induction. It follows from the statement in example 3. The details are left as an exercise.

19.5 Countable unions of countable sets.

How many countable sets can we join together and still obtain a countable set? The following lemma is the first step towards showing that joining together a countable number of countable sets will always result in a countable set. The lemma guarantees that the image of a countable set is countable.

Lemma 19.5 Suppose f maps an infinite *countable* set A onto a set $B = f[A]$. Then B is *countable*.

Proof:

What we are given: The set A is countable, $f : A \to B$ maps A *onto* the set B. What we are required to show: That B is countable.

Since A is countable we can index the elements of A with the natural numbers. Let $A = \{a_i : i \in \mathbb{N}\}$. For each $b \in f[A] = B$ let a_{b^*} be the element in $f^{\leftarrow}(\{b\})$ such that

$$b^* = \min\{i \in \mathbb{N} : a_i \in f^{\leftarrow}(\{b\})\}$$

This minimum element exists since $\mathbb{N}$ is well-ordered.[1] Since $\{a_{b^*} : b \in B\} \subseteq A$, then $\{a_{b^*} : b \in B\}$ is countable. (Subsets of countable sets are countable.) Then the function $g : \{a_{b^*} : b \in B\} \to B$ defined as $g(a_{b^*}) = b$ is one-to-one and onto B. So B is countable as required.

Theorem 19.6 Let $\{A_i : i \in S \subseteq \mathbb{N}\}$ be a countable set of non-empty countable sets A_i. Then $\bigcup_{i \in S} A_i$ is countable.

[1] Note that the axiom of choice is not required for this.

Proof:

What we are given: That the sets in $\{A_i : i \in S \subseteq \mathbb{N}\}$ are all countable sets.
What we are required to show: That $\bigcup_{i \in S} A_i$ is countable.

Since each set A_i is countable, then we can index the elements of each A_i with an initial segment T_i of natural numbers or with all elements of $\mathbb{N}$. For each i, let $A_i = \{a_{(i,j)} : j \in T_i \subseteq \mathbb{N}\}$.

We will define a function $f : \bigcup_{i \in S} A_i \to S \times \mathbb{N}$ as follows: $f(a_{(i,j)}) = (i, j)$.

We see that f maps $\cup_{i \in S} A_i$ one-to-one into $S \times \mathbb{N}$. Since $f[\cup_{i \in S} A_i] \subseteq S \times \mathbb{N} \subseteq \mathbb{N} \times \mathbb{N}$, it is countable. Then $\cup_{i \in S} A_i$ is the one-to-one image of the countable set $f[\cup_{i \in S} A_i]$ under the inverse map f^{-1}.

We conclude that $\cup_{i \in S} A_i$ is countable.

19.6 The set $\mathbb{R}$ of all real numbers is uncountable.

It is only after numerous attempts to show that the set of all real numbers is countably infinite, that mathematicians turned their attention towards showing that $\mathbb{R}$ is not a countable set. We can of course not say that since our very best mathematicians are unable to prove that $\mathbb{R}$ is countable, then $\mathbb{R}$ must be uncountable and leave it at that. We present a clever proof devised by Georg Cantor (1845-1918). Cantor successfully shows that no one-to-one image of $\mathbb{N}$ in $\mathbb{R}$ can be comprised of all real numbers. That is, for every one-to-one function $f : \mathbb{N} \to \mathbb{R}$, the subset, $\mathbb{R} - f[\mathbb{N}]$, of $\mathbb{R}$ will never be empty. This way of proving that $\mathbb{R}$ is uncountable is referred to as *Cantor's diagonalization method*.

Theorem 19.7 The set $\mathbb{R}$ of all real numbers is uncountable.

Proof:

We will show that the open interval (0, 1) is not countable. As a consequence, it will be impossible for $\mathbb{R}$ to be countable, since subsets of countable sets have been shown to be countable.

Proof by contradiction.

Suppose $f : \mathbb{N} \to (0, 1)$ is a one-to-one function mapping $\mathbb{N}$ onto $(0, 1)$. This means that we can index the elements of $(0, 1)$ with the natural numbers as follows: $(0, 1) = \{x_0, x_1, x_2, x_3, \ldots, \}$ where $x_i = f(i)$. We claim that at least one real number does not belong to $f[\mathbb{N}]$ and so f is not "onto" $(0, 1)$:

– We write out each real number as an infinite decimal expansion:

$$\begin{aligned} x_0 &= 0.a_{11}a_{12}a_{13}a_{14}a_{15}\ldots \\ x_1 &= 0.a_{21}a_{22}a_{23}a_{24}a_{25}\ldots \\ x_2 &= 0.a_{31}a_{32}a_{33}a_{44}a_{45}\ldots \\ &\vdots \quad \ldots\ldots \\ x_{n-1} &= 0.a_{n1}a_{n2}a_{n3}a_{n4}a_{n5}\ldots \\ &\vdots \quad \ldots\ldots \end{aligned}$$

We will construct a non-zero real number $0.b_1b_2b_3b_4\ldots$ between 0 and 1 which is not accounted for in this list.

· For each i, if $a_{ii} \in \{0,1,2,3,4\}$ let $b_i = 7$. If $a_{ii} \in \{5,6,7,8,9\}$ let $b_i = 2$. Note that there is nothing special about the numbers 7 and 2. We could have chosen another pair of integers between 0 and 9. But not 9's since a number containing an infinite string of 9's can produce a real number which has two infinite decimal representations.[1]

· So the real number $x = 0.b_1b_2b_3b_4\ldots$ is a string of 2's and 7's. We claim that x is not in the set $\{x_0, x_1, x_2, x_3, \ldots,\}$ said to contain all real numbers in (0, 1). Verify that $|b_1 - a_{11}| \geq 1$, $|b_2 - a_{22}| \geq 1$, and more generally, $|b_i - a_{ii}| \geq 1$ for all i and so the real number x *cannot* be any one the numbers in the list as claimed.

– So the function f is not onto $(0,1)$ as claimed.

Note that this does not only prove that this particular function is not onto; it also proves that all one-to-one functions $f : \mathbb{N} \to (0,1)$ which claim to be onto cannot be so. Since (0, 1) is not countable, then $\mathbb{R}$ cannot be countable.

Maybe for philosophical reasons, some readers may find it difficult to accept that the infinite set $\mathbb{R}$ is not a one-to-one image of $\mathbb{N}$ even though they can point to no obvious errors in Cantor's proof. Skeptical readers may find some comfort in learning that even very skilled mathematicians, when confronted by results which appear counter-intuitive, may harbour some nagging doubts in spite of being presented with an irrefutable proof. Georg Cantor once wrote to Richard Dedekind "Je le vois, mais je ne le crois pas"[2] (I see it, but I don't believe it) after determining that there is a one-to-one correspondence between all points in the plane and the set of points on a line.

[1] It can be shown, for example, that the rational numbers $0.04999999\ldots$ and $0.05000\ldots$ are different representations of the same rational number 5/100. But the decimal representation of a rational a/b is unique provided we do not allow a tail end of 9's in our representation of this number.

[2] Jean Cavaillès, Philosophie mathématique, p. 211.

Having now convinced ourselves that the set of all real numbers is uncountable, we can subdivide the class of all infinite sets into two categories: the subclass of all countably infinite sets and the subclass of all uncountably infinite sets. We will see in the next section that the class of all uncountable sets can itself be divided into other major subcategories of infinite sets.

Concepts review:

1. What does it mean to say that two sets are *equipotent sets*?
2. What does it mean to say that a set is *countable*?
3. Is the set $\varnothing$ countable?
4. What can we say about the image of a countable set under some function f?
5. Which of the sets $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{Z}$, $\mathbb{N} \times \mathbb{Z}$, $\mathbb{R}$ are countable sets.
6. Is it true that the subset of a countable set must be countable?
7. How is the procedure Cantor used to prove the uncountability of $\mathbb{R}$ referred to?
8. What can we say about the countable union of countable sets?

EXERCISES

A.
1. Prove that $\mathbb{N} \times \mathbb{N}$ is an infinite countable subset.
2. Prove that the sets $\mathbb{N} - \{0\}$ and $\mathbb{N}$ are equipotent.
3. Are there sets which are neither countable nor uncountable? Construct a set other than $\mathbb{R}$ which is uncountable and explain what makes this set uncountable.

B.
4. Prove that $\mathbb{Q}$ does not contain a one-to-one image of $\mathbb{R}$.
5. If S is finite show that $\mathscr{P}(S)$ is countable.

C.
6. Prove that $\mathbb{N}$ contains a subset which is equipotent with $\mathbb{Q} \times \mathbb{Q}$.
7. Show that $(\mathbb{N} \times \mathbb{N}) \times \mathbb{N}$ is countable.
8. Show that $\mathbb{N} \times \mathbb{R}$ is uncountable.
9. Suppose $f : \mathbb{R} \to \mathbb{Z}$ is defined as follows: $f(x)$ is the smallest integer y such that $y > x$. Is the quotient set induced on $\mathbb{R}$ by f a countable or uncountable set. Explain.

10. Suppose we tried to use the *Cantor diagonalization method* to prove that $\mathbb{Q}$ is uncountable. Explain why this would not work.
11. Is the set $\mathscr{P}(\mathbb{Q})$ countable or uncountable. Explain why.
12. Prove that the one-to-one image of an uncountable set must be uncountable.

---

20 / Equipotence as an equivalence relation.

Summary. *In this section we first show that the equipotence relation "$\sim_e$" is an equivalence relation on the class $\mathscr{S}$ of all sets. We use this equivalence relation to partition $\mathscr{S}$ into equivalence classes. We show that for any set S, S cannot be equipotent to its power set $\mathscr{P}(S)$. This fact allows us to construct infinitely many distinct classes of mutually equipotent sets. We also show that for any non-empty set S, the two sets $\mathscr{P}(S)$ and 2^S are equipotent. Finally, we show that $\mathscr{P}(\mathbb{N})$ is embedded in $\mathbb{R}$ and $\mathbb{R}$ is embedded in $\mathscr{P}(\mathbb{N})$.*

20.1 Viewing equipotence as a relation.

Let $\mathscr{S} = \{S : S \text{ is a set}\}$ denote the class of all sets. The *Axiom of class construction* guarantees this to be a well-defined class. Now "equipotence" can be viewed as a relation R_e on $\mathscr{S}$: a pair $(A, B) \in \mathscr{S} \times \mathscr{S}$ belongs to R_e if and only if A and B are equipotent. So, the word *equipotence* conveniently refers to the property possessed by sets which are equipotent.[1]

Notation: If two sets, A and B, are equipotent we will represent this property by, $A \sim_e B$.

For example, we have previously shown that $\mathbb{N} \sim_e \mathbb{Q}$ and $\mathbb{N} \sim_e \mathbb{N} \times \mathbb{Z}$; so the pairs $(\mathbb{N}, \mathbb{Q})$ and $(\mathbb{N}, \mathbb{N} \times \mathbb{Z})$ belong to the relation R_e on $\mathscr{S}$. We can also say, for example, that the class

$$\{S \in \mathscr{S} : S \sim_e \mathbb{N} \text{ or } S \sim_e n, n \in \mathbb{N}\}$$

is the class of all countable sets. Also, for all infinite subsets A of $\mathbb{N}$, $\mathbb{N} \sim_e A$. We have also seen that $\mathbb{N} \not\sim_e \mathbb{R}$; hence, $(\mathbb{N}, \mathbb{R}) \notin R_e$. It is natural to wonder whether R_e is an equivalence relation on $\mathscr{S}$. We immediately verify that this is the case.

Theorem 20.1 Let $\mathscr{S}$ be a class of sets. The equipotence relation R_e on $\mathscr{S}$ is an equivalence relation on $\mathscr{S}$.

Proof:

Reflexivity: For every set S, S is equipotent to itself.
Symmetry: If S is equipotent to T, then T is equipotent to S.
Transitivity: If S is equipotent to T and T is equipotent to H, then S is equipotent to H.

---

[1]The word "equinumerous" is also used to describe two sets which are equipotent. The word "equinumerosity" is also used to describe the property of sets which are equipotent. The word "equipotence" has the advantage of having only four syllables rather than the tongue-twisting seven syllables in "equinumerosity".

The equivalence relation, R_e, on the class $\mathscr{S}$ allows the construction of subclasses of $\mathscr{S}$ we call *equivalence classes induced by* R_e. If S is a set we will represent the equivalence class which contains S by $[S]_e$. We list infinitely many distinct equivalence classes:

$$\{[0]_e, [1]_e, [2]_e, [3]_e, [4]_e, \ldots, [n]_e, \ldots, [\mathbb{N}]_e, [\mathbb{R}]_e\}$$

Just about every set we have discussed up to now belongs to one of these equivalence classes. Of course, this affirmation must be verified case by case. This gives rise to a compelling question: Are there any equipotence-induced equivalence classes other than the ones listed here? To show that there are, is one of the main objectives of this section.

20.2 A few fundamental properties of the equipotence relation.

There are far too many sets to determine, on a case by case basis, which pairs belong to the same equipotence-induced equivalence class and which don't. There are, however, a few fundamental equipotence relation properties we can derive immediately which will help us classify sets by equipotence. We prove some of these now.

Theorem 20.2 Suppose A, B, C and D are sets such that $A \sim_e B$ and $C \sim_e D$ where $A \cap C = \varnothing = B \cap D$. Then $(A \cup C) \sim_e (B \cup D)$.

Proof:

What we are given: $A \sim_e B$ and $C \sim_e D$, $A \cap C = \varnothing$, $B \cap D = \varnothing$
What we are required to show: $(A \cup C) \sim_e (B \cup D)$

Since $A \sim_e B$ and $C \sim_e D$, there exists one-to-one onto functions $f : A \to B$ and $g : C \to D$. Define the function $h : A \cup C \to B \cup D$ as follows: $h|_A = f$ and $h|_C = g$. Since $A \cap C = \varnothing = B \cap D$, then h is well-defined, one-to-one and onto $B \cup D$. So $(A \cup C) \sim_e (B \cup D)$ as required.

Theorem 20.3 Suppose A, B, C and D are sets such that $A \sim_e B$ and $C \sim_e D$. Then $A \times C \sim_e B \times D$.

Proof:

What we are given: $A \sim_e B$ and $C \sim_e D$
What we are required to show: That $A \times C \sim_e B \times D$.

Since $A \sim_e B$ and $C \sim_e D$, there exists one-to-one onto functions $f : A \to B$ and $g : C \to D$. Define the function $h : A \times C \to B \times D$ as follows: $h(a, c) = (f(a), g(c))$. It is left as an exercise to show that h is a one-to-one onto function. So $A \times C \sim_e B \times D$ as required.

Corollary 20.4 Suppose A and B are infinite sets.

1) If $\{A, B\} \subset [\mathbb{N}]_e$, then $A \times B \in [\mathbb{N}]_e$.

2) If $\{A, B\} \subset [\mathbb{R}]_e$, then $A \times B \in [\mathbb{R}]_e$. Hence, $\mathbb{R} \times \mathbb{R} \sim_e \mathbb{R}$.

Proof:

1) Suppose $\{A, B\} \subset [\mathbb{N}]_e$. Then $A \sim_e \mathbb{N}$ and $B \sim_e \mathbb{N}$. By the theorem 20.3, $A \times B \sim_e \mathbb{N} \times \mathbb{N}$ where $\mathbb{N} \times \mathbb{N}$ is known to be countable (theorem 19.4). Hence, $A \times B \in [\mathbb{N}]_e$.

2) Suppose $\{A, B\} \subset [\mathbb{R}]_e$. Then $A \sim_e \mathbb{R}$ and $B \sim_e \mathbb{R}$ and so, by the theorem 20.3, $A \times B \sim_e \mathbb{R} \times \mathbb{R}$. It is easily seen that $\mathbb{R}$ is equipotent with $\{1\} \times \mathbb{R} \subset \mathbb{R} \times \mathbb{R}$. Since $\{1\} \times \mathbb{R}$ is uncountable, $\mathbb{R} \times \mathbb{R}$ is uncountable. Then $A \times B \notin [\mathbb{N}]_e$. We must now show that $A \times B \in [\mathbb{R}]_e$. This will be the case if $\mathbb{R} \times \mathbb{R} \in [\mathbb{R}]_e$.

We claim that $\mathbb{R} \times \mathbb{R} \in [\mathbb{R}]_e$.

– It is easily seen that $(0, 1) \sim_e \mathbb{R}$.[1] Then, by the theorem 20.3, $(0, 1) \times (0, 1) \sim_e \mathbb{R} \times \mathbb{R}$. To show that $\mathbb{R} \times \mathbb{R} \in [\mathbb{R}]_e$ it then suffices to show that $(0, 1) \times (0, 1) \sim_e (0, 1)$.

– For $0.x_1x_2x_3x_4x_5 \ldots \in (0, 1)$ (ignoring those decimal expansions with infinite strings of 9's) define the function $f : (0, 1) \to (0, 1) \times (0, 1)$ as follows:

$$f(0.x_1x_2x_3x_4x_5 \ldots) = (0.x_1x_3x_5x_7x_9 \ldots, 0.x_2x_4x_6x_4x_8 \ldots)$$

The function f is onto:

· Let $(0.a_1a_2a_3a_4a_5 \ldots, 0.b_1b_2b_3b_4b_5 \ldots) \in (0, 1) \times (0, 1)$. Then

$$f(0.a_1b_1a_2b_2a_3b_3 \ldots) = (0.a_1a_2a_3a_4a_5 \ldots, 0.b_1b_2b_3b_4b_5 \ldots)$$

so f is onto $(0, 1) \times (0, 1)$.

The function f is one-to-one:

[1] The function $f(x) = \tan\left(\frac{\pi x}{2}\right)$ maps the open interval (0, 1) one-to-one onto $(0, \infty)$. If $g(x) = \ln x$, $g \circ f$ maps $(0, 1)$ one-to-one onto $(-\infty, \infty) = \mathbb{R}$.

· Let $(0.a_1a_2a_3a_4\ldots, 0.b_1b_2b_3b_4\ldots) \in (0,1)\times(0,1)$. By definition of f, $(0.a_1b_1a_2b_2a_3b_3\ldots)$ is the only element that can be mapped to $(0.a_1a_2a_3a_4\ldots, 0.b_1b_2b_3b_4\ldots)$. Hence, f is one-to-one.

So $(0,1)\sim_e (0,1)\times(0,1)$ as required.

We have shown that not only is $\mathbb{R}\times\mathbb{R}$ uncountable, $\mathbb{R}\times\mathbb{R}\in[\mathbb{R}]_e$. Since $A\times B\sim_e \mathbb{R}\times\mathbb{R}$, $A\times B\in[\mathbb{R}]_e$.

Example: Show that $(-\pi/2,\pi/2)\times\mathbb{Q}\sim_e \mathbb{R}\times\mathbb{N}$.

Solution:

Since the function $f(x)=\tan x$ maps the interval $(-\pi/2,\pi/2)$ one-to-one and onto $\mathbb{R}$ while $\mathbb{Q}$ has been shown to be equipotent with $\mathbb{N}$, then, by the above theorem, $(-\pi/2,\pi/2)\times\mathbb{Q}\sim_e \mathbb{R}\times\mathbb{N}$.

20.3 Products of countable sets.

Given two infinite countable sets S and T, using theorem 20.3, we can write

$$S\times T\sim_e \mathbb{N}\times\mathbb{N}\sim_e \mathbb{N}$$

So the product of two infinite countable sets is countable. If $\{A_i : i=0,1,2,3,\ldots,n\}$ is a finite set of countable sets then the expression $\prod_{i=0}^{n} A_i$ is defined as:

$$\prod_{i=0}^{n} A_i = A_0\times A_1\times A_2\times\cdots\times A_n = \{(x_0,x_1,x_2,\ldots,x_n) : x_i\in A_i\}$$

We can prove the following slightly more general result.

Theorem 20.5 Let $\{A_i : i\in\mathbb{N}\}$ be a (countable) set of non-empty countable sets. Then $\prod_{i=0}^{n} A_i$ is countable for all n.

Proof:

What we are given: $\{A_i : i\in\mathbb{N}\}$ is a set of countable sets.
What we are required to show: $A_0\times A_1\times A_2\times\cdots\times A_n$ is countable.

We know that the product of any two non-empty countable sets is countable:

We prove the statement by induction. Let $P(n)$ be the statement

$$\text{“}\prod_{i=0}^{n} A_i \text{ is countable ”}$$

– Base case: Since A_0 is countable, then $P(0)$ holds true.
– Inductive hypothesis: Suppose $P(n)$ holds true. Then $\prod_{i=0}^{n} A_i$ is countable. Now

$$\prod_{i=0}^{n+1} A_i \sim_e \left(\prod_{i=0}^{n} A_i \right) \times A_{n+1}$$

is a product of two countable sets. We have initially shown that products of pairs of countable sets is countable. Hence, the product $\prod_{i=0}^{n+1} A_i$ is countable. So $P(n+1)$ holds true.

By the principle of mathematical induction $\prod_{i=0}^{n} A_i$ is countable for all n.

The reader should be careful not to generalize the above theorem when it comes to Cartesian products. It does not say that "The Cartesian product of countably many countable sets is countable". This statement does not hold true in general. We will soon witness infinite products of countable sets which are not countable.

20.4 Adding countable sets to infinite sets.

We will now show that if a set, S, is infinite and another set, T, is countable (finite or infinite), then $S \cup T \in [S]_e$. Adding countably many elements to an infinite set S will always result in a set which is equipotent with S.

Theorem 20.6 Suppose S is an infinite set and T is a countable set such that $S \cap T = \varnothing$. Then

$$S \sim_e S \cup T$$

Proof:

What we are given: That T is countable, S is infinite and $S \cap T = \varnothing$.
What we are required to show: That S and $S \cup T$ are equipotent.
It suffices to construct a one-to-one function $f : S \cup T \to S$ mapping $S \cup T$ onto S.

– Since S is infinite, then it must contain a countably infinite subset, say, X. (By theorem 18.9.)
– Since T is countable and X is countably infinite, then, by theorem 19.6, $T \cup X$ is countably infinite.
– So there exists a one-to-one onto function $g : T \cup X \to X$ mapping $T \cup X$ onto $X \subseteq S$ (since both $T \cup X$ and X belong to $[\mathbb{N}]_e$.)
– We will now construct a function $f : S \cup T \to S$ as follows:

$$f(x) = \begin{cases} x & \text{if} \quad x \in S - X \\ g(x) & \text{if} \quad x \in T \cup X \end{cases}$$

We see that f maps $S \cup T$ one-to-one and onto S, as required.

Example: Let $\mathbb{J}$ denote the set of all irrational numbers. Show that $\mathbb{J} \in [\mathbb{R}]_e$.

It was shown that the set of all rational numbers $\mathbb{Q}$ is countably infinite. If $\mathbb{J}$ was countable, then by the theorem above, $\mathbb{R} = \mathbb{J} \cup \mathbb{Q}$ would be countable, a contradiction. So $\mathbb{J}$ is uncountable. We claim that $\mathbb{J} \in [\mathbb{R}]_e$.

– Since $\mathbb{J} \cap \mathbb{Q} = \varnothing$, then, by theorem 20.6, $\mathbb{J} \sim_e \mathbb{J} \cup \mathbb{Q} = \mathbb{R}$. We conclude that $\mathbb{J} \in [\mathbb{R}]_e$.

20.5 Equipotence classes of power sets.

Given any set S, the *Axiom of power set* allows us to construct a new set, $\mathscr{P}(S)$, by gathering together all subsets of S and viewing these subsets as the elements of $\mathscr{P}(S)$. We have also shown that if a set S contains n elements, then its power set, $\mathscr{P}(S)$, contains precisely 2^n elements. Obviously, at least for finite sets S, S and $\mathscr{P}(S)$ are not equipotent. We would like to see whether this rule generalizes to infinite sets.

We begin by showing that pairs of sets which are equipotent produce equipotent power sets.

Theorem 20.7 If the sets A and B are equipotent, then so are their associated power sets $\mathscr{P}(A)$ and $\mathscr{P}(B)$.

Proof:

Given that A and B are equipotent, there exists a one-to-one function $f : A \to B$ mapping A onto B. We define the function $f^* : \mathscr{P}(A) \to \mathscr{P}(B)$ as follows:

$$f^*(T) = M \Leftrightarrow f[T] = M$$

Claim: The function f^* is onto $\mathscr{P}(B)$.

Proof of claim: Let $M \in \mathscr{P}(B)$. If $M = \varnothing$, then $f^*(\varnothing) = f[\varnothing] = \varnothing$. Suppose M is a non-empty subset of B. Since f is onto B, $M \subseteq f[A]$. Then $f[f^{\leftarrow}[M]] = M$. So f^* maps the element $f^{\leftarrow}[M]$ in $\mathscr{P}(A)$ to the element M in $\mathscr{P}(B)$. We conclude that f^* maps $\mathscr{P}(A)$ onto $\mathscr{P}(B)$.

Claim: The function f^* is well-defined.

Suppose U and V are distinct elements of $\mathscr{P}(B)$ such that $x \in U - V \neq \varnothing$. Since f is one-to-one and onto, $f^{\leftarrow}[\{x\}]$ is a singleton set $\{y\}$ contained in $f^{\leftarrow}[U] \subseteq A$. Since $x \notin V$, $y \notin f^{\leftarrow}[V]$. Then $y \in f^{\leftarrow}[U] - f^{\leftarrow}[V]$; this implies that f^* maps the distinct elements $f^{\leftarrow}[U]$ and $f^{\leftarrow}[V]$ to U and V respectively.

Claim: The function f^* is one-to-one.

Suppose S and T are distinct elements of $\mathscr{P}(A)$ where S is non-empty. Suppose that $x \in S - T \neq \varnothing$. Since f is one-to-one, $f(x) \notin f[T]$. Then $f(x) \in f[S] - f[T]$; this implies $f[S] \neq f[T]$. We conclude that $f^*(S) \neq f^*(T)$. So f^* is one-to-one.

So $\mathscr{P}(A)$ and $\mathscr{P}(B)$ are equipotent.

At this point, we know of only two equipotence-induced equivalence classes of infinite sets. They are $[\mathbb{N}]_e$ and $[\mathbb{R}]_e$. We wonder: Is $\mathscr{P}(\mathbb{N})$ a countable set? That is, does $\mathscr{P}(\mathbb{N})$ belong to $[\mathbb{N}]_e$? Let's gather together a few proven facts and possible deductions which can be made from these:

- We have already shown that that $\mathbb{Q} \sim_e \mathbb{N}$.
- It then follows from the theorem above that $\mathscr{P}(\mathbb{Q}) \sim_e \mathscr{P}(\mathbb{N})$.
- Now $\mathbb{R}$ was defined as the set of all Dedekind cuts; Dedekind cuts were seen to be elements of $\mathscr{P}(\mathbb{Q})$.
- Hence, $\mathbb{R}$ is equipotent to a subset of $\mathscr{P}(\mathbb{Q})$.
- Since $\mathbb{R}$ is uncountable, then $\mathscr{P}(\mathbb{Q})$ must also be uncountable.
- It must then follow that $\mathscr{P}(\mathbb{N}) \notin [\mathbb{N}]_e$. That is, $\mathscr{P}(\mathbb{N})$ is uncountable.
- We would now have to verify whether $\mathscr{P}(\mathbb{N}) \in [\mathbb{R}]_e$.

So, just like $\mathscr{P}(n) \not\sim_e n$ for all natural numbers n, $\mathscr{P}(\mathbb{N}) \not\sim_e \mathbb{N}$.

A general follow-up question might be: Is it possible for any infinite set, S, to be equipotent with its power set $\mathscr{P}(S)$?

We will show that the answer to this question is, no! That is, if S is infinite, $S \notin [\mathscr{P}(S)]_e$.

Theorem 20.8 Any non-empty set S is embedded in its power set $\mathscr{P}(S)$. But *no* subset of S is equipotent with $\mathscr{P}(S)$.

Proof:

What we are given: That S is a non-empty set.
What we are required to prove:

1) That S is embedded in $\mathscr{P}(S)$
2) That $\mathscr{P}(S)$ is not equipotent to K for any $K \subseteq S$.

1) For any element $x \in S$, $\{x\} \in \mathscr{P}(S)$, so the function $f : S \to \mathscr{P}(S)$ defined as $f(x) = \{x\}$ maps S one-to-one into $\mathscr{P}(S)$. So S is embedded in $\mathscr{P}(S)$, as required.

2) Proof by contradiction. Suppose $K \subseteq S$ such that $K \sim_e \mathscr{P}(S)$. Note that $K \neq \varnothing$ since $\mathscr{P}(S) \neq \varnothing$. Then there exists a function

$$g : K \to \mathscr{P}(S)$$

mapping K one-to-one onto $\mathscr{P}(S)$. We claim that this leads to a contradiction.

- If $x \in K$, $g(x)$ is seen to be an element of $\mathscr{P}(S)$ and therefore is a subset of S.
- Either $x \in g(x)$ or $x \notin g(x)$. Let T be the subset of K defined as:

  $$T = \{x \in K : x \notin g(x)\}$$

  Note that $T \subseteq K \subseteq S$ and so $T \in \mathscr{P}(S)$.
- Since the function $g : K \to \mathscr{P}(S)$ is onto $\mathscr{P}(S)$ there must be some element in K, say y, such that $g(y) = T$. Let's determine whether y is in T or not:
  - If $y \in T$, then, by definition of T, $y \notin g(y) = T$. This makes no sense, so $y \notin T$.
- On the other hand, if $y \notin T$, then $y \in g(y) = T$. This contradicts the pre-established fact that $y \notin T$.

The source of this contradiction is the supposition that $g : K = \mathscr{P}(S)$ is one-to-one and onto. So $K \not\sim_e \mathscr{P}(S)$ for any $K \subseteq S$.

The above theorem confirms that for any set S, the equipotence-induced equivalence classes, $[S]_e$ and $[\mathscr{P}(S)]_e$, are always distinct. It also suggests that there are many more equipotence-induced equivalence classes than the ones listed previously on page 190. For example,

$$[0]_e \neq [1]_e \neq [2]_e \neq \cdots \neq [\mathbb{N}]_e \neq [\mathbb{R}]_e \neq [\mathscr{P}(\mathbb{R})]_e \neq [\mathscr{P}(\mathscr{P}(\mathbb{R}))]_e \cdots$$

We will show how a never-ending list of equipotence-induced equivalence classes of infinite sets can be constructed. We first define the expression "properly embedded".

Definition 20.9 We will say that the non-empty set, A, is *properly embedded* in the set, B, if A is equipotent to a proper subset of B but *B is not equipotent to A or any of its subsets*. To represent the relationship "A is properly embedded in B" we will write

$$A \hookrightarrow_e B$$

If A and B are sets such that A is equipotent to a subset of B where A may, or may not, be equipotent to B, we will say that A is *embedded* in B. To represent the relationship "A is embedded in B" we will write

$$A \hookrightarrow_{e\sim} B$$

Both relations, $\hookrightarrow_e$ and $\hookrightarrow_{e\sim}$ are easily seen to be a transitive relations on the class, $\mathscr{S}$, of all sets.

If $\mathscr{S} = \{S : S \text{ is a set}\}$, we define the class $\mathscr{E}$ as

$$\mathscr{E} = \{[S]_e : \ S \in \mathscr{S}\}$$

Both $\hookrightarrow_e$ and $\hookrightarrow_{e\sim}$ induce an order relation on $\mathscr{E}$, which we now define.

Notation 20.10 Let $\mathscr{S} = \{S : S \text{ is a set}\}$ and $\mathscr{E} = \{[S]_e : \ S \in \mathscr{S}\ \}$. Let $[A]_e$ and $[B]_e$ be elements of $\mathscr{E}$. We write

$$[A]_e <_e [B]_e$$

if and only if $A \hookrightarrow_e B$. We write

$$[A]_e \leq_e [B]_e$$

if and only if $A \hookrightarrow_{e\sim} B$.

Some may suspect that $\leq_e$ is a non-strict partial ordering on $\mathscr{E}$. The relation $\leq_e$ is easily seen to be reflexive and transitive on $\mathscr{E}$. But it is not clear whether $\leq_e$ is antisymmetric on $\mathscr{E}$. We see that $[A]_e \leq_e [B]_e$ only if $A \hookrightarrow_{e\sim} B$ and $[B]_e \leq_e [A]_e$ only if $B \hookrightarrow_{e\sim} A$. But does $([A]_e \leq_e [B]_e) \wedge ([B]_e \leq_e [A]_e) \Rightarrow ([A]_e = [B]_e)$? Equivalently, does "$A$ is embedded in B and B is embedded in A" imply that A and B are equipotent sets? We suspect that this is the case. But proving this is *not* a trivial matter. So we can not assume this to be a fact at this time. The statement which guarantees that this holds true is called the Schröder-Bernstein theorem. This will be the main topic of the next section.

We provide a few examples. We have previously shown that any infinite set contains a subset which is equipotent with $\mathbb{N}$, and, since $\mathbb{N}$ and $\mathbb{R}$ are known to be non-equipotent, $\mathbb{N} \hookrightarrow_e \mathbb{R}$; it then follows that $[\mathbb{N}]_e <_e [\mathbb{R}]_e$. Similarly, for any natural number n and infinite set A, $n \hookrightarrow_e \mathbb{N} \hookrightarrow_e A$ and so $[n]_e <_e [A]_e$.

Proposition 20.11 Let S be any set. Suppose $\mathscr{P}^0(S) = S$, $\mathscr{P}^1(S) = \mathscr{P}(S)$ and $\mathscr{P}^n(S) = \mathscr{P}(\mathscr{P}^{n-1}(S))$ for all $n \geq 1$. The set

$$\{[S]_e, [\mathscr{P}(S)]_e, [\mathscr{P}^2(S)]_e, [\mathscr{P}^3(S)]_e, \ldots, [\mathscr{P}^n(S)]_e \ldots, \}$$

forms an infinite $<_e$-ordered chain of distinct classes in $\mathscr{E}$.

Proof:

Proof by induction. Let $P(n)$ denote the statement "$\{[\mathscr{P}^i(S)]_e : i = 0, 1, 2, \ldots, n\}$ forms a $<_e$-ordered chain of distinct classes in $\mathscr{E}$".[1]

Base case: $\{[\mathscr{P}^0(S)]_e\} = \{[S]_e\} \in \mathscr{E}$. Since $\{[S]_e\}$ contains only one element $<_e$ linearly orders $\{[S]_e\}$. The base case holds true.

Inductive hypothesis: Suppose $P(n)$ holds true. That is, suppose $\{[\mathscr{P}^i(S)]_e : i = 0, 1, 2, \ldots, n\}$ forms a $<_e$-ordered chain of distinct classes in $\mathscr{E}$. We are required to show that $P(n+1)$ holds true.

By theorem 20.8, $\mathscr{P}^n(S) \hookrightarrow_e \mathscr{P}^{n+1}(S)$, and so $[\mathscr{P}^n(S)]_e <_e [\mathscr{P}^{n+1}(S)]$. Since $<_e$ is irreflexive, antisymmetric and transitive on $\{[\mathscr{P}^i(S)]_e : i = 0, 1, 2, \ldots, n\}$, then it must the case for $\{[\mathscr{P}^i(S)]_e : i = 0, 1, 2, \ldots, n, n+1\}$. So $P(n+1)$ holds true.

By mathematical induction $\{[\mathscr{P}^n(S)]_e : i = 0, 1, 2, \ldots, n\}$ forms a $<_e$-ordered chain of distinct classes in $\mathscr{E}$ for each $n \in \mathbb{N}$.

From this we conclude that the relation $<_e$ is a strict linear ordering of the infinite set $\{[\mathscr{P}^n(S)]_e : n \in \mathbb{N}\}$.

The above proposition allows us to say that both $\{[\mathscr{P}^n(\mathbb{N})]_e : n = 0, 1, 2, \ldots, \}$ and $\{[\mathscr{P}^n(\mathbb{R})]_e : n = 0, 1, 2, \ldots, \}$ form infinite $<_e$-ordered chains of distinct classes. It will be interesting to determine whether these two chains have any elements in common. We will have the tools required to answer this question only in the next section.

20.6 Equipotence of 2^S and $\mathscr{P}(S)$.

In example two on page 129, we introduced the set $\{1, 2\}^{\mathbb{N}}$ equipped with the *lexicographic* linear ordering. This set was defined as being the set of all functions mapping $\mathbb{N}$ into the set $\{1, 2\}$. It can be equivalently described as

$$\{1, 2\}^{\mathbb{N}} = \{(a_0, a_1, a_2, a_3, \ldots,) : a_i \text{ equals 1 or 2}\}$$

which is the set of all possible countably infinite ordered strings of 1's and 2's. Whether $\mathbb{N}$ is mapped to $\{1, 2\}$ or $\{0, 1\}$ is not considered as being a significantly different set since all possible countably infinite strings of 0's and 1's will essentially produce a set which is equipotent to the set of all possible countably infinite strings of 1's and 2's. Using 0's and 1's will allow us to represent the set, $\{0, 1\}^{\mathbb{N}}$, more succinctly as, $2^{\mathbb{N}}$.

We can generalize the expression by replacing $\mathbb{N}$ with any set S. That is, if S is any non-empty set, 2^S represents all functions which map the set S to $\{0, 1\}$. For example, given some finite set, say, $S = \{3, 4\}$ we can actually list the functions in this set as:

$$2^{\{3,4\}} = \{\ \{(3,0),(4,0)\},\ \{(3,1),(4,1)\},\ \{(3,0),(4,1)\},\ \{(3,1),(4,0)\}\ \}$$

a set equivalent to the set of 4 sequences

$$\{\ \{0_3, 0_4\},\ \{1_3, 1_4\},\ \{0_3, 1_4\},\ \{(1_3, 0_4\}\ \}$$

[1] Note that by the Axiom of power set, $\mathscr{P}^n(S)$ is a set for all $n \in \mathbb{N}$; hence $\{[\mathscr{P}^n(S)]_e : n \in \mathbb{N}\} \subseteq \mathscr{E}$.

This set has four, or 2^2, elements. If S has three elements, say, $S = \{7, 8, 9\}$ and we list all elements of 2^S we would see that it contains precisely $2^3 = 8$ elements. Verify this. It can be shown my mathematical induction that if S has n elements, 2^S must contain 2^n functions (see the exercise section). Recall that in theorem 18.11, we showed by induction that the power set, $\mathscr{P}(S)$, of any n-element set, S, contains 2^n elements. From this fact, we deduce that,

$$\text{"If } S \text{ is finite, } 2^S \sim_e \mathscr{P}(S)\text{"}$$

Question: Can we generalize this statement so that it holds true for *all* sets S, including infinite ones?

Answer: We will convince ourselves that we can. But this will require some careful explaining. To help answer this question let's consider a third way of viewing the elements of 2^S (whether S is finite or not). Suppose $f \in 2^S$.

- Then $f^{\leftarrow}[\{1\}]$ and $f^{\leftarrow}[\{0\}]$ form disjoint subsets, say T and $S-T$, of S respectively. Note that T may possibly be empty or possibly be all of S.
- So f is a function which maps x to 1 if and only if $x \in T$ and all other elements to 0.
- That is, $f = \chi_T \in 2^S = \{0, 1\}^S$.[1] In fact, for every $K \subseteq S$, equivalently for every $K \in \mathscr{P}(S)$, $\chi_K \in 2^S$; conversely, for every $f \in 2^S$, there is precisely one $T \subseteq S$, equivalently $T \in \mathscr{P}(S)$, such that $f = \chi_T$.

Consider the function, $g : \mathscr{P}(S) \to 2^S$, defined as: $g(T) = \chi_T$. We have just shown that g maps $\mathscr{P}(S)$ one-to-one onto 2^S. Hence, $2^S \sim_e \mathscr{P}(S)$. It is worth formally stating this important statement as a theorem.

Theorem 20.12 For any non-empty set S,

$$2^S = \{\chi_T : T \in \mathscr{P}(S)\} \sim_e \mathscr{P}(S)$$

If $S = \mathbb{N}$, then $2^{\mathbb{N}} \sim_e \mathscr{P}(\mathbb{N})$ or equivalently $2^{\mathbb{N}} \in [\mathscr{P}(\mathbb{N})]_e$. Similarly, $2^{\mathscr{P}(\mathbb{N})} \in [\mathscr{P}(\mathscr{P}(\mathbb{N}))]_e$ and $2^{\mathbb{R}} \in [\mathscr{P}(\mathbb{R})]_e$.

20.7 Comparing $\mathscr{P}(\mathbb{N})$ and $\mathbb{R}$

[1] Recall that on page 81 we introduced a function called the "characteristic function of T" represented as, χ_T. The function χ_T on S was defined as a function mapping T to $\{1\}$ and the rest of S to $\{0\}$; it is precisely the same function as f.

We eagerly wonder: "Do the sets $\mathscr{P}(\mathbb{N})$ and $\mathbb{R}$ belong to the same equipotence equivalence class or not?" We know that $\mathbb{N}$ is embedded in $\mathbb{R}$. But it is not clear how $\mathscr{P}(\mathbb{N})$ relates to $\mathbb{R}$ with respect to the equipotence relation. The following theorems do not entirely answer this question, but they constitute a first step towards answering it.

Theorem 20.13 The set, $\mathbb{R}$, is embedded in $\mathscr{P}(\mathbb{N})$.

Proof:

Recall that Dedekind defines of the real numbers $\mathbb{R}$ as being a particular set of initial segments of $\mathbb{Q}$ and so $\mathbb{R} \subseteq \mathscr{P}(\mathbb{Q})$. Since $\mathbb{Q} \sim_e \mathbb{N}$, then $\mathscr{P}(\mathbb{Q}) \sim_e \mathscr{P}(\mathbb{N})$ (by theorem 20.7). Since $\mathbb{R} \subseteq \mathscr{P}(\mathbb{Q}) \sim_e \mathscr{P}(\mathbb{N})$, $\mathbb{R}$ is equipotent to a subset $\mathscr{P}(\mathbb{N})$.

We provide some background that will help follow the proof of the next statement. The theorem 20.12 states that $2^{\mathbb{N}} = \{\chi_T : T \in \mathscr{P}(\mathbb{N})\} \sim_e \mathscr{P}(\mathbb{N})$. Recall that the function $\chi_T : \mathbb{N} \to \{0, 1\}$ is defined as

$$\chi_T(n) = \begin{cases} 0 & \text{if } n \notin T \\ 1 & \text{if } n \in T \end{cases}$$

So, for a specific subset T of $\mathbb{N}$, χ_T can be viewed as a sequence, $\{i_0, i_1, i_2, i_3, \ldots, \}$, of 0's and 1's, where $i_k = 0$ only if $k \notin T$, otherwise $i_k = 1$. We see that there is a one-to-one correspondence between the set $\{\chi_T : T \subseteq \mathbb{N}\}$ and the set of sequences $\big\{ \{i_0, i_1, i_2, i_3, \ldots, \} : i_n \in \{0, 1\}\big\}$. Equivalently, there is a one-to-one correspondence between the set, $\{\chi_T : T \subseteq \mathbb{N}\}$, and the set of decimal expansions, $\big\{ \ i_0.i_1i_2i_3i_4 \cdots : \text{ where } i_n \in \{0, 1\}\big\}$. For example, if $E = \{n \in \mathbb{N} : \ n \text{ is even } \}$, then χ_{E} can uniquely be represented as:

$$\begin{aligned} \chi_{\text{E}} &= \{(0,1),(1,0),(2,1),(3,0),(4,1),(5,0),\ldots,\} \\ &\to \{1_0, 0_1, 1_2, 0_3, 1_4, 0_5, \ldots, \} \\ &\to 1_0.0_1 1_2 0_3 1_4 0_5 \cdots \\ &\to 1.01010101\cdots \end{aligned}$$

If $O = \{n \in \mathbb{N} : \ n \text{ is odd } \}$, then we can write

$$\begin{aligned} \chi_{\text{O}} &= \{(0,0),(1,1),(2,0),(3,1),(4,0),(5,1),\ldots,\} \\ &\to \{0_0, 1_1, 0_2, 1_3, 0_4, 1_5, \ldots, \} \\ &\to 0_0.1_1 0_2 1_3 0_4 1_5 \cdots \\ &\to 0.10101010\cdots \end{aligned}$$

Say we define a function, $f^* : \{\chi_T : T \in \mathscr{P}(\mathbb{N})\} \to \big\{ \ i_0.i_1i_2i_3i_4 \cdots : \text{ where } i_n \in \{0, 1\}\big\}$ as

$$f^*(\chi_T) = i_0.i_1i_2i_3i_4\cdots \text{ if and only if } i_n = \chi_T(n) \text{ for } n \in \mathbb{N}$$

For example:

$$\begin{aligned} f^*(\chi_{\mathbb{N}}) &= 1.111111111111111\cdots = 10/9 \\ f^*(\chi_{\mathrm{E}}) &= 1.010101010101010\cdots = 100/99 \\ f^*(\chi_{\mathrm{O}}) &= 0.101010101010101\cdots = 10/99 \\ f^*(\chi_{\varnothing}) &= 0.000000000000000\cdots = 0 \\ f^*(\chi_{\{0,3\}}) &= 1.001000000000000\cdots = 1001/1000 \end{aligned}$$

We see that the maximum value in the image of $\{\chi_T : T \in \mathscr{P}(\mathbb{N})\}$ under f^* is $10/9$ while the minimum value in the image is 0. Also, if $U \neq V$, then $f^*(\chi_U) \neq f^*(\chi_V)$ so f^* is one-to-one on $\{\chi_T : T \in \mathscr{P}(\mathbb{N})\}$.

We are now set to prove the following theorem.

Theorem 20.14 The set $\mathscr{P}(\mathbb{N})$ is embedded in $\mathbb{R}$.

Proof:

We define a function $f^* : \{\chi_T : T \in \mathscr{P}(\mathbb{N})\} \to \left\{ \ i_0.i_1i_2i_3i_4\cdots : \text{ where } i_n \in \{0,1\}\right\}$ as

$$f^*(\chi_T) = i_0.i_1i_2i_3i_4\cdots \text{ if and only if } i_n = \chi_T(n) \text{ for } n \in \mathbb{N}$$

We see that the function f^* maps $\{\chi_T : T \in \mathscr{P}(\mathbb{N})\}$ one-to-one into the interval $[0, 10/9]$.

It follows that

$$\mathscr{P}(\mathbb{N}) \sim_e \{\chi_T : T \in \mathscr{P}(\mathbb{N})\} \hookrightarrow_e [0, 10/9] \subset \mathbb{R}$$

So $\mathscr{P}(\mathbb{N})$ is embedded in $\mathbb{R}$, as required.

Concepts review:

1. Describe the equivalence relation on the class of all sets which was discussed in this section.
2. What can we say about the finite union of disjoint countable sets?
3. What can we say about the Cartesian products of two countable sets?
4. If we add a countable set to an infinite set S what can we say about the set that results from this union?
5. The set of all irrationals is equipotent with which set?
6. If two sets A and B are equipotent what can we say about their respective power sets?
7. What is the meaning given to the expression "A is properly embedded in B"?
8. From any non-empty set S construct a set B such that $S \hookrightarrow_e B$.
9. Name a set which contains a copy of $\mathbb{R}$ but is not equipotent with $\mathbb{R}$.
10. If S is a set, what set of functions is equipotent with $\mathscr{P}(S)$ other than $\mathscr{P}(S)$ itself?
11. If S is a set, what does the set $\{\chi_T : T \subseteq S\}$ represent? To what set is it equipotent with?

EXERCISES

A.
1. Show that the following pair of sets are equipotent.
   a) $A = (0, 1)$ and $B = (-1, 1)$. (These represent open intervals in $\mathbb{R}$.)
   b) $A = (-1, 1)$ and $\mathbb{R}$.
2. If $S = \{0, \{1, 2\}\}$ and $T = \{\{x\}, y\}$ write out explicitly the elements in the set $\mathscr{P}(S)$ and $\mathscr{P}(T)$.
3. Show that $\mathscr{P}(\mathbb{N})$, $\mathscr{P}(\mathbb{Q})$ and $\mathscr{P}(\mathbb{Z})$ are equipotent uncountable sets.
4. Prove that the set of all non-negative irrational numbers and the set $\mathbb{R}$ of all real numbers are equipotent.
5. Show that the sets $\{0, 1\} \times \mathbb{N}$ is countable.
6. Let $S = \{a, b, c\}$. Show that the three sets 2^S, $\mathscr{P}(S)$ and $\{\chi_T : T \in \mathscr{P}(S)\}$ contain the same number of elements by listing all their elements.

B.

7. Is the set $\mathbb{N}^{\mathbb{N}}$ countable?
8. Is $\mathbb{N}^{\mathbb{N}}$ embeddable in $\mathbb{R}$? If so, find a suitable mapping.
9. Is $\mathbb{R}$ embeddable in $\mathbb{N}^{\mathbb{N}}$? If so, find a suitable mapping.
10. Prove that $2^{(2^S)}$ and $\mathscr{P}(\mathscr{P}(S))$ are equipotent.
11. Prove, by mathematical induction, that if a non-empty set S contains n elements, then the set 2^S contains 2^n elements.
12. Prove that there are infinitely many equipotence-induced equivalence classes of uncountable sets.

C.

12. Prove that $\mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$ (n times) is equipotent to $\mathbb{R}$ for all non-zero natural numbers n.
13. Prove in detail theorem 20.3.

21 / The Schröder-Bernstein theorem.

Summary. *In this section we state and prove the Schröder-Bernstein theorem. We then illustrate some of its consequences. In particular, we use it as a tool to prove that* $\mathbb{R}$ *and* $\mathscr{P}(\mathbb{N})$ *are equipotent. We finally show that* $\mathbb{N}^{\mathbb{N}}$ *and* $\mathbb{R}$ *are equipotent.*

21.1 Reviewing some basic properties of infinite sets.

In the last chapter, we discussed certain properties possessed by infinite sets. We will now build on those results to prove more general statements about these. These results tend to be less intuitive since they relate to sets other than those which represent the numbers we are accustomed to. Before we begin, we list the results from the last section which will serve as the main tools to prove the statements which follow.

Recall that $\mathscr{S} = \{x : x \text{ is a set}\}$.

- The countable union of countable sets is countable.
- The finite product of countable sets is countable.
- $(A \sim_e B) \wedge (C \sim_e D) \Rightarrow A \times C \sim_e B \times D$.
- $[(S \text{ is infinite}) \wedge (T \sim_e \mathbb{N})] \Rightarrow (S \sim_e S \cup T)$.
- $(S \in \mathscr{S}) \Rightarrow (S \hookrightarrow_e \mathscr{P}(S))$
- $\{\chi_T : T \in \mathscr{P}(S)\} \sim_e \mathscr{P}(S)$
- $(S \in \mathscr{S}) \Rightarrow 2^S \sim_e \mathscr{P}(S)$
- $\mathbb{R} \hookrightarrow_{e\sim} \mathscr{P}(\mathbb{N})$ and $\mathscr{P}(\mathbb{N}) \hookrightarrow_{e\sim} \mathbb{R}$

21.2 The Schröder-Bernstein theorem.

Even if two sets are known to be equipotent, it can be quite difficult to construct a function which maps one set one-to-one onto the other. For example, we may suspect that $\mathscr{P}(\mathbb{N})$ and $\mathbb{R}$ are equipotent sets, but proving this by actually producing a function which maps $\mathbb{R}$ one-to-one onto $\mathscr{P}(\mathbb{N})$ could be a challenging task, one that will no longer be necessary once we have proved a statement called *The Schröder-Bernstein theorem.*

Theorem 21.1 *(The Schröder-Bernstein theorem)* If S and T are infinite subsets where S is embedded in T and T is embedded in S, then S and T are equipotent.

The proof is presented once we have proved the following lemma.

Lemma 21.2 Let T be a proper subset of the set, S, and $f : S \to T$ be a one-to-one function mapping S *into* T. Then there exists a one-to-one function, $f^* : S \to T$, mapping S **onto** T.

Proof:

What we are given: That $T \subset S$; that $f : S \to T$ maps S one-to-one into T.
What we are required to show: There exists a one-to-one function $f^* : S \to T$ which maps S onto T.

Since T is a proper subset of S, then $S - T$ is non-empty.
We construct a sequence of sets $\{S_i : i \in \mathbb{N}\}$ as follows:

$$\begin{aligned} S_0 &= S - T \\ S_1 &= f[S - T] = f[S_0] \\ S_2 &= f^2[S - T] = f[S_1] \\ S_3 &= f^3[S - T] = f[S_2] \\ S_4 &= f^4[S - T] = f[S_3] \\ &\vdots \\ S_n &= f^n[S - T] = f[S_{n-1}] \\ &\vdots \end{aligned}$$

Let $U = \bigcup_{i \in \mathbb{N}} S_i$. Since f maps all of S in T, for all $i > 0$, $S_i \subseteq T$. Remember that $S_0 = S - T$. The S_i's can be shown to be pairwise disjoint. Verification of this fact is left as an exercise. (This fact is important for the validity of this proof. Try a proof by induction.)

We define the function $f^* : S \to T$ as follows:

$$f^*(x) = \begin{cases} f(x) & \text{if} \quad x \in U \\ x & \text{if} \quad x \notin U \end{cases}$$

– We verify that the image of S under f^* is a subset of T:
  Any element a in S is either in U or is not in U. If $a \in U$, then $f^*(a) = f(a) \in S_i$ for some $i \geq 1$ and so $f^*(a) \in T$. If $a \notin U$, then it is not in $S_0 = S - T$ and so is in T. So the image of S under f^* is in T.
– We verify that f^* is one-to-one on S:
  Since $f^* = f|_U$ on U (on which f is one-to-one) and is the identity map on $S - U$, then f^* is one-to-one on S. (Some details are left as an exercise. The reader should see clearly why this is true.)
– We verify that f^* is onto T:

Let $t \in T$. If $t \in T - U$, then f^* maps t to t. Suppose $t \in U$. Then $t \in S_i$ for some $i > 0$ (t cannot be in S_0 since $S_0 = S - T$). Then t is in the image of S_{i-1} under f and so is in the image of T under f^*. So every element of T is in the image of S under f^*.

So f^* maps S one-to-one onto T as required.

Proof of the Schröder-Bernstein theorem.

Proof:

What we are given: There exists a one-to-one function, $f : S \to T$, mapping S into T and a one-to-one function, $g : T \to S$, mapping T into S.
What we are required to show: There is a one-to-one function which maps T onto S.
Let $h = g \circ f$. Then h is a function mapping S *into* S. Since both f and g are one-to-one on their respective domains, then h is one-to-one on S. Then

$$h[S] = g[f[S]] \subseteq g[T]$$

By the lemma, there exists a one-to-one function $h^* : S \to g[T]$ mapping S *onto* $g[T] \subseteq S$. Then $h^{*-1} : g[T] \to S$ maps $g[T]$ one-to-one onto S. This means that $(h^{*-1} \circ g)[T] = S$. That is, the function $h^{*-1} \circ g$ maps T one-to-one onto S.
Thus, S and T are equipotent, as required.

21.3 Consequences of the Schröder-Bernstein theorem.

The important Schröder-Bernstein theorem will allow us to prove statements we suspected were true but for which we lacked the tools to confirm our suspicions. It shows that the relation, $\leq_e$, on $\mathscr{E} = \{[A]_e : A \in \mathscr{S}\}$ is antisymmetric since $[A]_e \leq_e [B]_e$ and $[B]_e \leq_e [A]_e$ implies $[A]_e = [B]_e$. Hence, $\leq_e$ is a non-strict ordering of $\mathscr{E}$.

Theorem 21.3 The set $\mathbb{R}$ of all real numbers is equipotent to $\mathscr{P}(\mathbb{N})$.

Proof:

We proved in theorem 20.13 and 20.14 that $\mathbb{R}$ is embedded in $\mathscr{P}(\mathbb{N})$ and $\mathscr{P}(\mathbb{N})$ is embedded in $\mathbb{R}$. Then by the *Schröder-Bernstein theorem* $\mathbb{R} \sim_e \mathscr{P}(\mathbb{N})$.

Since $\mathbb{R} \sim_e \mathscr{P}(\mathbb{N})$, then $[\mathbb{R}]_e = [\mathscr{P}(\mathbb{N})]_e$. By theorem 20.7, $\mathscr{P}(\mathbb{R}) \sim_e \mathscr{P}(\mathscr{P}(\mathbb{N}))$, and so $[\mathscr{P}(\mathbb{R})]_e = [\mathscr{P}(\mathscr{P}(\mathbb{N}))]_e = [\mathscr{P}^2(\mathbb{N})]_e$. It easily follows, by mathematical induction, that the two countably infinite $<_e$-ordered sets

$$\{[0]_e, [1]_e, [2]_e, \ldots, [\mathbb{N}]_e, [\mathbb{R}]_e, [\mathscr{P}(\mathbb{R})]_e, [\mathscr{P}^2(\mathbb{R})]_e, [\mathscr{P}^3(\mathbb{R})]_e, \ldots, \}$$

$$\{[0]_e, [1]_e, [2]_e, \ldots, [\mathbb{N}]_e, [\mathscr{P}(\mathbb{N})]_e, [\mathscr{P}^2(\mathbb{N})]_e, [\mathscr{P}^3(\mathbb{N})]_e, [\mathscr{P}^4(\mathbb{N})]_e, \ldots, \}$$

have the same elements and so represent identical $<_e$-chains in the class $\mathscr{E} = \{[S]_e : S \text{ is a set}\}$.

21.4 The set $\mathbb{N}^{\mathbb{N}}$.

In theorem 20.12, we saw that for any set S,

$$2^S \sim_e \{\chi_T : T \subseteq S\} \sim_e \mathscr{P}(S)$$

Hence, $2^{\mathbb{N}} \sim_e \{\chi_T : T \subseteq \mathbb{N}\} \sim_e \mathscr{P}(\mathbb{N})$, where $2^{\mathbb{N}}$ is the set of all functions mapping $\mathbb{N}$ into $\{0, 1\}$. The set $2^{\mathbb{N}}$ was also seen to be equipotent to the set, $\big\{ \{a_0, a_1, a_2, a_3, \ldots, \} : a_i \in \{0, 1\}\big\}$.[1]

We will now investigate the "set of all functions mapping $\mathbb{N}$ into $\mathbb{N}$". That is, we will consider a set whose elements are of the form, $f = \{(i, a_i) : i \in \mathbb{N}, a_i \in \mathbb{N}\}$. To be consistent with our notation, we will express this set as

$$\mathbb{N}^{\mathbb{N}}$$

For example, $g = \{(0, 1), (2, 4), (3, 9), (4, 16), \ldots, \}$ represents a particular element of the set $\mathbb{N}^{\mathbb{N}}$ where n is mapped to n^2. We could also represent this element as, $(a_0, a_1, a_2, a_3, \ldots,)$, where $a_i = i^2$. That is, each a_i is associated to the element (i, i^2).[2]

The sets, $\mathbb{N}^{\mathbb{N}}$ and $2^{\mathbb{N}}$, are both sets of functions with domain, $\mathbb{N}$, except the functions in $2^{\mathbb{N}}$ have range, $\{0, 1\}$, while the functions in $\mathbb{N}^{\mathbb{N}}$ have range, $\mathbb{N}$. Not surprisingly, if $g \in 2^{\mathbb{N}}$, then $g \in \mathbb{N}^{\mathbb{N}}$; hence, $2^{\mathbb{N}} \subset \mathbb{N}^{\mathbb{N}}$. Of course, $\mathbb{N}^{\mathbb{N}}$ contains many elements which do not belong to $2^{\mathbb{N}}$. For example, $\{(i, i^2) : i \in \mathbb{N}\}$ belongs to $\mathbb{N}^{\mathbb{N}}$ but not to $2^{\mathbb{N}}$. But it may still be possible for $\mathbb{N}^{\mathbb{N}}$ to be equipotent to $2^{\mathbb{N}}$. If we can show that $\mathbb{N}^{\mathbb{N}}$ is embedded in $2^{\mathbb{N}}$, it will follow from the *Schröder-Bernstein theorem* that $[\mathbb{N}^{\mathbb{N}}]_e = [2^{\mathbb{N}}]_e$.

Theorem 21.4 The sets, $\mathbb{N}^{\mathbb{N}}$ and $\mathbb{R}$, are equipotent.

[1] If $A_i = \{0, 1\}$ for $i = 0, 1, 2, 3, \ldots$ we define $\prod_{i\in\mathbb{N}} A_i = \{(a_0, a_1, a_2, \ldots,) : a_i \in \{0, 1\}\}$. Then $\prod_{i\in\mathbb{N}} A_i \sim_e 2^{\mathbb{N}}$.

[2] If $A_i = \mathbb{N}$ for $i = 0, 1, 2, 3, \ldots,$ we define $\prod_{i\in\mathbb{N}} A_i = \{(a_0, a_1, a_2, a_3, \ldots,) : a_i \in \mathbb{N}\}$. Or if one prefers, $\prod_{i\in\mathbb{N}} A_i$ can be viewed as the set of all possible countably infinite sequences of natural numbers. The element $g = \{(0, 1), (2, 4), (3, 9), (4, 16), \ldots, \}$ can be viewed as $(0, 1, 4, 9, 16, \ldots,) \in \prod_{i\in\mathbb{N}} A_i = \{(a_0, a_1, a_2, a_3, \ldots,) : a_i \in \mathbb{N}\}$. In fact, $\prod_{i\in\mathbb{N}} A_i \sim_e \mathbb{N}^{\mathbb{N}}$.

Proof:

What we are given: $\mathbb{N}^{\mathbb{N}}$ is the set of all functions mapping $\mathbb{N}$ into $\mathbb{N}$.
What we are required to show: $\mathbb{N}^{\mathbb{N}}$ and $\mathbb{R}$ are equipotent.

Claim: $\mathbb{R}$ is embedded in $\mathbb{N}^{\mathbb{N}}$.

– We have shown that $\mathbb{R} \sim_e 2^{\mathbb{N}} \subset \mathbb{N}^{\mathbb{N}}$; hence, $\mathbb{R}$ is embedded in $\mathbb{N}^{\mathbb{N}}$.

Claim: $\mathbb{N}^{\mathbb{N}}$ is embedded in $\mathbb{R}$.

– Let $f \in \mathbb{N}^{\mathbb{N}}$. Then f can be expressed in the form $f = \{(0, a_0), (1, a_1), (2, a_2), \ldots\}$ a subset of $\mathbb{N} \times \mathbb{N}$. Since $f \subset \mathbb{N} \times \mathbb{N}$, then $f \in \mathscr{P}(\mathbb{N} \times \mathbb{N})$. Then

$$\begin{aligned} \mathbb{N}^{\mathbb{N}} &\subset \mathscr{P}(\mathbb{N} \times \mathbb{N}) \\ &\sim_e \mathscr{P}(\mathbb{N}) \quad \text{(By 20.4, } \mathbb{N} \times \mathbb{N} \sim_e \mathbb{N}\text{, followed by 20.7.)} \\ &\sim_e \mathbb{R} \end{aligned}$$

Then $\mathbb{N}^{\mathbb{N}}$ is embedded in $\mathbb{R}$ as claimed.

By the Schröder-Bernstein theorem $\mathbb{N}^{\mathbb{N}} \sim_e \mathscr{P}(\mathbb{N}) \sim_e \mathbb{R}$.

21.5 The set B^A of all functions mapping A into B.

Since we are discussing sets of functions such as, $\mathbb{N}^{\mathbb{N}}$ and 2^S, we slightly generalize these notions by considering ranges other than $\{0, 1\}$ and $\mathbb{N}$.

Definition 21.5 If A and B are two sets, then the symbol, B^A, refers to the set of all functions mapping A into B.[1]

Examples:

a) The set $\mathbb{Q}^{\mathbb{N}}$ denotes the set of all functions $f : \mathbb{N} \to \mathbb{Q}$. For example,

$$\{x_i : x_i = 1/(i+1),\ i \in \mathbb{N}\} = \{\tfrac{1}{1}, \tfrac{1}{2}, \tfrac{1}{3}, \ldots,\}$$

is such a function. We can of course say that $\mathbb{Q}^{\mathbb{N}}$ is the set of all infinite countable sequences of rational numbers.

b) If S contains 3 elements and T contains 4 elements we can verify that the set S^T will contain 3^4 elements.

[1]The following argument confirms that if A and B are sets, then A^B is a set: Every element $f \in A^B$ is a subset of the set $B \times A$ (finite products of sets are sets). So for every $f \in A^B$, $f \in \mathscr{P}(B \times A)$. Then $A^B \subseteq \mathscr{P}(B \times A)$. Since $\mathscr{P}(B \times A)$ is a set (Axiom of power set), then A^B must be a set (Axiom of subset).

We wonder how the sets, $\mathbb{Q}^{\mathbb{N}}$ and $\mathbb{N}^{\mathbb{N}}$, are related. It is clear that if $\mathbf{x} = \{(0, n_0), (1, n_1), (2, n_2), (3, n_3), \ldots, \} \in \mathbb{N}^{\mathbb{N}}$, then $\mathbf{x} \in \mathbb{Q}^{\mathbb{N}}$; hence, $\mathbb{N}^{\mathbb{N}} \subset \mathbb{Q}^{\mathbb{N}}$. On the other hand, we know that $\mathbb{Q}$ and $\mathbb{N}$ are equipotent; so there exists a one-to-one function $f : \mathbb{Q} \to \mathbb{N}$ mapping $\mathbb{Q}$ onto $\mathbb{N}$. It then follows that for any element $\mathbf{x} = \{(0, q_0), (1, q_1), (2, q_2), (3, q_3), \ldots, \}$ in $\mathbb{Q}^{\mathbb{N}}$, we can associate a unique element $f^*(\mathbf{x}) = \mathbf{y} = \{(0, f(q_0)), (1, f(q_1)), (2, f(q_2)), (3, f(q_3), \ldots, \})$ in $\mathbb{N}^{\mathbb{N}}$. So $\mathbb{Q}^{\mathbb{N}}$ is embedded in $\mathbb{N}^{\mathbb{N}}$. Hence, by the Schröder-Bernstein theorem, $\mathbb{Q}^{\mathbb{N}}$ and $\mathbb{N}^{\mathbb{N}}$ are equipotent.

Concepts review:

1. What does the Schröder-Bernstein theorem say?
2. Name three sets which are equipotent to the power set $\mathscr{P}(\mathbb{N})$.
3. What do the symbols $2^{\mathbb{N}}$ and $\mathbb{N}^{\mathbb{N}}$ mean?
4. Is the set $\mathbb{N}^{\mathbb{N}}$ equipotent with $\mathbb{R}$?
5. What does the expression B^A mean? If B has 3 elements and A has 2 elements how many elements does B^A contain? How many elements does A^B contain?

EXERCISES

A. 1. Show that an infinite set S is countable if and only if S is equipotent with every one of its infinite subsets.

B. 2. Prove that an infinite countable set S can be expressed as the union of two disjoint infinite countable sets.

3. Prove that if S and T are sets and $S - T$ and $T - S$ are equipotent, then S and T are equipotent.
4. Prove that for any $m \in \mathbb{N}$, $\mathbb{N}^m$ is countable.
5. If $S = \{0, 1, 2\}$ and $T = \{x, y\}$ write out explicitly the elements of the following sets:
   a) S^T
   b) T^S
   c) 2^S
   d) $\mathscr{P}(S)$

C. 6. Show that $\mathbb{N}^{\mathbb{N}}$ is equipotent with a subset of $\mathscr{P}(\mathbb{N} \times \mathbb{N})$.

7. Show that $\mathscr{P}(\mathbb{N} \times \mathbb{N})$ is equipotent with a subset of $\mathbb{R}$.
8. Show that the set $\{S_i : i \in \mathbb{N}\}$ of sets constructed in the proof of lemma 21.2 are pairwise disjoint.
9. Show that the function h constructed in the proof of lemma 21.2 is one-to-one on S.
10. Suppose R is an equivalence relation on an infinite countable set S. Show that the set of all equivalence classes on S induced by R is countable.
11. Let S be the set of all infinite sequences of natural numbers. We will say that a sequence $s = \{s_i : i \in \mathbb{N}\} \in S$ has a "constant tail-end" if there exists a number $k \in \mathbb{N}$ such that $i > k \Rightarrow s_i = s_k$. Let $T = \{s \in S : s \text{ has a constant tail-end }\}$. Show that T is countable.
12. In theorem 20.5 it is proven that if the A_i's are countable, then, for any n, $\prod_{i=0}^{n} A_i$ is countable. Give an example that shows that $\prod_{i \in \mathbb{N}} A_i$ need not be countable. Explain.
13. A subset T of $\mathbb{R}$ is said to be *open* if for every element $x \in T$, x is contained in an open interval entirely contained in T. Let $\mathscr{S}$ be a set of pairwise disjoint open subsets of $\mathbb{R}$. Show that $\mathscr{S}$ is countable.[1] (Hint: Let $U = \mathbb{Q} \cap (\bigcup_{S \in \mathscr{S}} S)$. A step invoking the Axiom of choice will follow.)
14. Show that $\mathbb{Q}^{\mathbb{Q}} \in [\mathbb{R}]_e$.

---

[1] The statement "Any infinite linearly ordered set V such that the set $\mathscr{S}$ of pairwise disjoint open subsets is at most countable must be equipotent with $\mathbb{R}$." is referred to as the *Suslin's problem*. It remained an open question until it was proved that it is impossible to prove or disprove this statement from ZF plus the Axiom of choice.

Part VII

Cardinal numbers

22 / An introduction to cardinal numbers.

Summary. *In this section we state the Continuum hypothesis and the Generalized continuum hypothesis; we discuss their meaning and consequences. We "informally" define the class of cardinal numbers, $\mathscr{C}$, introduce the "aleph" notation for cardinal numbers and define addition, multiplication and exponentiation of cardinal numbers. Finally, we show that the class of all cardinal numbers is a proper class.*

22.1 The equipotence-based classification of sets.

In the last few sections, we have used the equipotence relation to subdivide the class, $\mathscr{S}$, of all sets into subclasses, $[S]_e$, of mutually equipotent sets. As was done previously (on page 197), we will continue to represent the class of all $\sim_e$-equivalence classes on $\mathscr{S}$ as

$$\mathscr{E} = \{[S]_e : S \text{ is a set}\}$$

In 20.10, we defined the relation, $<_e$, on $\mathscr{E}$ as,

$$([A]_e <_e [B]_e) \Leftrightarrow (A \hookrightarrow_e B)$$

The relation, $<_e$, was seen to be a strict order relation on $\mathscr{E}$.

We proved two fundamental results concerning the elements of $\mathscr{E}$:

1) For all sets S, $[S]_e <_e [\mathscr{P}(S)]_e$, proven in theorem 20.8,
2) $[\mathbb{R}]_e = [\mathscr{P}(\mathbb{N})]_e$, proven in theorem 21.3.

Combining these two statements, along with the proposition 20.11, allows us to conclude that the set

$$\mathscr{A} = \{[0]_e, [1]_e, \ldots, [n]_e, \ldots, [\mathbb{N}]_e, [\mathscr{P}(\mathbb{N})]_e, [\mathscr{P}^2(\mathbb{N})]_e, \ldots, [\mathscr{P}^n(\mathbb{N})]_e, \ldots, \}$$

of distinct equipotence-induced equivalence classes, linearly ordered by the relation $<_e$, not only contains $[\mathbb{R}]_e$ but also the equivalence classes of all power sets generated by $\mathbb{R}$. That is, $[\mathscr{P}(\mathbb{N})]_e = [\mathbb{R}]_e$, $[\mathscr{P}^2(\mathbb{N})]_e = [\mathscr{P}(\mathbb{R})]_e$; more generally

$$[\mathscr{P}^{n+1}(\mathbb{N})]_e = [\mathscr{P}^n(\mathbb{R})]_e$$

A few words of caution: Even though we have shown that $<_e$ linearly orders the set, $\mathscr{A}$, described above, we have *not* proven that $<_e$ linearly orders the class $\mathscr{E}$, even though we suspect that it does. We will *not* assume this to be the case until we formally prove it to be true.

22.2 The Continuum hypothesis.

A particularly intriguing question concerning the strictly ordered set of equivalence classes described above baffled mathematicians for decades: "Does there exist an uncountable set S (that is, one which is not equipotent with $\mathbb{N}$) which is properly embedded in $\mathbb{R} \sim_e \mathscr{P}(\mathbb{N})$?" Equivalently, "Does there exist a set S such that $[\mathbb{N}]_e <_e [S]_e <_e [\mathbb{R}]_e = [\mathscr{P}(\mathbb{N})]_e$? "

After numerous attempts to construct such a set S in vain, Georg Cantor came to believe that no such set S exists. In 1878, he conjectured that:

> "No uncountable set can be properly embedded[1] in $\mathbb{R}$."

This conjecture is referred to as the *Continuum hypothesis*, abbreviated by CH.[2] In 1900, the mathematician David Hilbert declared that proving, or disproving, the Continuum hypothesis was one of the 23 most important unsolved mathematical questions of that time. In 1940, Kurt Gödel showed it is impossible to disprove the Continuum hypothesis within *ZFC*. In 1963, Paul Cohen showed that it is impossible to prove that the Continuum hypothesis holds true within *ZFC*. This settled the question: Neither assuming "CH is true" nor assuming "CH is false" can lead us to a contradiction. This means that we are free to work in a universe governed by *ZFC*+CH or by *ZFC*+¬CH, as we prefer, without fear of a contradiction evolving from the annexation of the axiom CH or ¬CH to *ZFC*.[3]

There is a more general version of the Continuum hypothesis called the *Generalized Continuum Hypothesis*, abbreviated by GCH, which states that:

> "For any *infinite* set A there does not exist a set S such that
> $A \hookrightarrow_e S \hookrightarrow_e \mathscr{P}(A)$.

Or we can equivalently say, "For every set A there does not exist a set S such that $[A]_e <_e [S]_e <_e [\mathscr{P}(A)]_e$". The GCH only refers to infinite sets, not finite ones. The *Generalized continuum hypothesis* implies the *Continuum hypothesis*. It is, however, known that GCH does not follow from CH. Assuming *ZFC*+GCH, the linearly ordered set

$$\mathscr{A} = \{[0]_e, [1]_e, \ldots, [n]_e, \ldots, [\mathbb{N}]_e, [\mathscr{P}(\mathbb{N})]_e, [\mathscr{P}^2(\mathbb{N})]_e, \ldots, [\mathscr{P}^n(\mathbb{N})]_e, \ldots, \}$$

for example, is an "initial segment" of equivalence classes in the sense that for any n, $\{[S]_e \in \mathscr{E} : [S]_e <_e \mathscr{P}^n(\mathbb{N})\} \subset \mathscr{A}$. This does not say that this set represents all equipotence-induced equivalence classes, far from it. It simply means that assuming

[1] Recall that "A is properly embedded in B" means that A is equipotent with a subset of B but B is not equipotent with A nor any of its subsets.

[2] The word *continuum* is simply another way of referring to the set $\mathbb{R}$.

[3] The negation of CH is represented as "¬CH". The symbol "¬" is often read as "not".

GCH, such a set forms a string of countably many equivalence classes, with none missing. In the *ZFC*+¬GCH universe, for each $n > 0$, $\{[S]_e \in \mathscr{E} : [S]_e <_e \mathscr{P}^n(\mathbb{N})\} \not\subset \mathscr{A}$.

It is also known that neither GCH nor ¬GCH can be proved in *ZFC*.[1]

This leaves us with an important question: When trying to determine whether a mathematical statement holds true or not, should we assume CH or "not CH"? If a statement can be proved without invoking either of these axioms, most readers will prefer the proof which avoids these statements (viewing them as being extraneous). However, some mathematical statements have as only proof one which assumes CH (or GCH). In such cases, the reader should be alerted to this fact, and should be informed at which point in the proof this axiom is invoked.

22.3 The class, $\mathscr{E}$, of equipotence induced equivalence classes.

The reader has no doubt noticed that we have been careful not to refer to the class

$$\mathscr{E} = \{[S]_e : \ S \in \mathscr{S}\}$$

of all $\sim_e$-induced equivalence classes on $\mathscr{S}$ as a "set of sets". It is not difficult to show that $\mathscr{E}$ can not be a set of sets.

Theorem 22.1 The class, $\mathscr{E} = \{[S]_e : \ S \in \mathscr{S}\}$, is not a set of sets.

Proof:

Proof by contradiction. Suppose $\mathscr{E}$ is a set of sets.

- Then, by the Axiom of choice, there exists a choice function, f, which maps every element $[S]_e \in \mathscr{E}$ to some element s in $[S]_e \subset \mathscr{S}$. That is, f chooses from each element $[S]_e$ of $\mathscr{E}$ a set representative, $f([S]_e) = s \sim_e S$.
- Let $\mathscr{B} = f[\mathscr{E}]$ denote the image of $\mathscr{E}$ under f. By the Axiom of replacement, the image of a set under a well-defined function is a set; hence, $\mathscr{B}$ is a set of sets.
- Then $T = \cup\{s : s \in \mathscr{B}\}$ is the union of a set of sets. Hence, by the Axiom of union, T is a set.
- By the Axiom of power set, $\mathscr{P}(T)$ is also a set. Since $\mathscr{P}(T)$ is a set, $\mathscr{P}(T) \in [S]_e$ for some $[S]_e \in \mathscr{E}$. Then $\mathscr{P}(T) \sim_e s$ for some $s \in \mathscr{B}$. But $s \subset T$. By transitivity, $\mathscr{P}(T) \hookrightarrow_e T$, a contradiction.

The source of our contradiction is the assumption that $\mathscr{E}$ is a set of sets.

---

[1] Interestingly, it was shown by Sierpinski that the Axiom of choice follows from *ZF*+GCH. That is, the Axiom of choice exists in a universe governed by *ZF*+GCH.

22.4 Cardinal numbers.

We pause to examine more closely the equivalence class $[2]_e$, the class of all sets which are equipotent to the natural number 2. Examples of a few elements which belong to $[2]_e$ are the sets

$$\{9, 7\}$$

$$\{\mathbb{R}, \mathbb{N}\}$$

$$\Big\{\ \varnothing, \{\{\ \varnothing\}\}\ \Big\}$$

$$\Big\{\ \{\{\{\ \varnothing\}\}\}, \{\{\ \varnothing\}\}\ \Big\}$$

each of which is equipotent to 2. We also see that $[2]_e = [\{9, 7\}]_e$. Of course, being equipotent to itself, $2 = \{\varnothing, \{\varnothing\}\}$ also belongs to $[2]_e$. In fact, it is the only natural number which is an element of $[2]_e$. If we represent the class, $[2]_e$, in this way, rather than representing it as, say $[\{9, 7\}]_e$, it is because we surreptitiously selected the set $2 = \{\varnothing, \{\varnothing\}\}$ as being the "official" representative of this class. In fact, we have chosen the natural numbers as the official representatives of all equivalence classes whose elements are finite sets. On the other hand, possible representatives of $[\mathbb{R}]_e$ and $[\mathscr{P}(\mathbb{R})]_e$ are $\mathbb{R}$ and $\mathscr{P}(\mathbb{R})$, respectively. But we could of course have used $\mathscr{P}(\mathbb{N})$ and $\mathscr{P}^2(\mathbb{N})$, respectively.

It would be convenient to uniquely specify a class representative for each element of $\mathscr{E}$. Determining *how* we can select a set from each and every equivalence class in $\mathscr{E}$ is, however, not obvious. The *Axiom of choice* states that there is a choice function that allows us to select an element from each set in a "set of sets". But $\mathscr{E}$ is not a "set of sets". So the Axiom of choice is not available to us as a tool for selecting an element in each set in $\mathscr{E}$.[1] We need to identify a specific property possessed by a single set in $[S]_e$ which clearly distinguishes it from all other sets in $[S]_e$. Unfortunately, at this time, we have not yet sufficiently explored our universe of sets to be able to identify what this set property could be.

For the time being, we will postulate the existence of a class, $\mathscr{C}$, containing the unique representatives of each and every single $\sim_e$-equivalence class in $\mathscr{E}$. We will refer to these class representatives as cardinal numbers.[2]

Postulate 22.2 There exists a class of sets $\mathscr{C}$ which satisfies the following properties:

1. Every natural number n is an element of $\mathscr{C}$.

[1] We could use each equivalent class in $\mathscr{E}$ as "self-representatives" and call them cardinal numbers. The problem with this is that these equivalence class are not known to be sets. We want a set which represents each equivalence class in $\mathscr{E}$.

[2] For those readers who wish to read ahead, the sets that we will call "cardinal numbers" are the *initial ordinals*.

2. Any set $S \in \mathscr{S}$ is equipotent to precisely one element in $\mathscr{C}$

The sets in $\mathscr{C}$ are called *cardinal numbers*. When we say that a set, S, *has cardinality* κ, we mean that $\kappa \in \mathscr{C}$ and that $S \sim_e \kappa$, or equivalently, $S \in [\kappa]_e$. If the set S has cardinality κ, we will write, $|S| = \kappa$.

We emphasize that we postulate the existence of the cardinal numbers, $\mathscr{C}$, immediately, for convenience only. We will eventually prove that the class $\mathscr{C}$, as defined above, exists.

22.5 Definitions and notation associated to cardinal numbers.

Given any set S, the expression, $|S|$, was said to denote the cardinality of S. When referring to a "generic" uncountably infinite cardinal number (that is, a cardinal number which is not the cardinality of an explicitly specified uncountable set S) we will represent it by a Greek letter such as κ or λ.[2] For example, we would represent a cardinal number as a Greek letter in a phrase such as "Suppose S is a set whose cardinality is κ, ...". However, we don't normally represent a finite cardinal number by a Greek letter. When referring to some finite set, F, containing, say, n elements, it's cardinality is the natural number n, so we write, $|F| = n$. Whenever the natural number n is viewed as the representative of all sets which contain n elements, we refer to it as a cardinal number.

The aleph notation. It is customary to express the cardinality of countably infinite sets by using the "aleph" notation[3]: We write $|\mathbb{N}| = \aleph_0 \in \mathscr{C}$. The symbol "$\aleph_0$" is pronounced "aleph-not". For example, since we have shown that $\mathbb{N} \times \mathbb{N} \sim_e \mathbb{N}$ and that $\mathbb{N} \hookrightarrow_e \mathbb{R}$, we can write $|\mathbb{N} \times \mathbb{N}| = \aleph_0$, while $|\mathbb{R}| \neq \aleph_0$.

The cardinality of $\mathbb{R}$. The cardinality of the set $\mathbb{R}$ is often represented by the symbol c.[1] Since $\mathbb{R}$, $\mathscr{P}(\mathbb{N})$ and $2^{\mathbb{N}}$ were shown to be equipotent, then we can write $|\mathscr{P}(\mathbb{N})| = |2^{\mathbb{N}}| = c$. If we assume the Continuum hypothesis (CH), there can be no uncountable cardinal number κ such that $\mathbb{N} \hookrightarrow_e \kappa \hookrightarrow_e c$. That is, CH implies that there are no cardinal numbers between $\aleph_0$ and c. If we assume $\neg$CH, then we are saying that there exists an uncountable cardinal κ such that $\mathbb{N} \hookrightarrow_e \kappa \hookrightarrow_e c$.

Since we are referring to class representatives as "numbers" we will replace the symbols $\hookrightarrow_e$ with $<$ and $\hookrightarrow_{e\sim}$ with $\leq$. That is, if $\kappa = |S|$ and $\lambda = |T|$

$$\kappa < \lambda \Leftrightarrow S \hookrightarrow_e T$$

$$\kappa \leq \lambda \Leftrightarrow S \hookrightarrow_{e\sim} T$$

[2]The Greek letter κ is read as "kappa". The Greek letter λ is read as "lambda".

[3]The aleph notation was introduced by Georg Cantor.

[1]The symbol c abbreviates the word "continuum", another way of referring to the set $\mathbb{R}$.

Transfinite cardinal numbers are those cardinal numbers which are infinite sets. The natural numbers are those cardinal numbers which are not *transfinite*.

22.6 Three operations on cardinal numbers.

Having explicitly provided symbols, $0, 1, 2, 3, \ldots, \aleph_0, c$, for the first few cardinal numbers, we now develop methods to construct from these, other cardinal numbers not found in this list. Methods used to construct new sets from known sets will by applied to construct new cardinal numbers from known ones. For example, given two sets, A and B, we defined new sets such as $C = A \cup B$, $D = A \times B$ and A^B. (Recall that A^B denotes the set of all functions mapping the set B into A.) We will translate these three set operations to cardinal number operations. For example, it may be useful to know how the cardinality of the sets $A \cup B$, $A \times B$ and A^B compare with the cardinality of two sets, A and B. Once we formally define operations on cardinal numbers we will try to determine if there are general principles that can be used to more efficiently compute the value of the cardinal numbers which result from these operations.

Definition 22.3 If S and T are sets and $\kappa = |S|$ and $\lambda = |T|$, then we define *addition*"+", *multiplication* "$\times$" and *exponentiation* of two cardinal numbers as follows:

a) If $S \cap T = \varnothing$,
$$\kappa + \lambda = |S \cup T|$$

b)
$$\kappa \times \lambda = |S \times T|$$

c)
$$\kappa^\lambda = |S^T|$$

where S^T represents the set of all functions mapping T into S (as previously defined). That is, $|S|^{|T|} = |S^T|$. For convenience, we define $0^\lambda = 0$ and $\kappa^0 = 1$.

If two sets, S and T, have non-empty intersection, it is still possible to determine $|S| + |T|$ by proceeding as follows:

$$\begin{aligned} (\kappa = |S|) &\Rightarrow (\kappa = |S \times \{0\}|) \\ (\lambda = |T|) &\Rightarrow (\lambda = |T \times \{1\}|) \\ (S \times \{0\}) \cap (T \times \{1\}) = \varnothing &\Rightarrow \kappa + \lambda = |(S \times \{0\}) \cup (T \times \{1\})| \end{aligned}$$

One should verify that the definitions of sums, products and exponents of cardinal numbers agree with the operations we perform with finite cardinal numbers (the natural numbers). Suppose, for example, that A contains 4 elements and B contains

2 elements. Then there are $16 = 4^2$ elements in A^B. Verify this by listing all the elements in A^B. Also there are $4 \times 2 = 8$ elements in $A \times B$ and $4 + 2 = 6$ elements in $A \cup B$ (assuming that A and B have no elements in common). Verify this fact.

Examples:

a) What is the cardinality of the set $F = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$?
The cardinal numbers of all finite sets are the natural numbers; hence, $|F| = 3$.

b) What is the cardinality of the set $\{2, \{3\}\}$?
The cardinal number of a set does not depend on the type of elements it contains, so $|\{2, \{3\}\}| = 2$.

c) What is the cardinal number of each of the sets $\mathbb{Q}$, $\mathbb{Z}$, and $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$?
They all have the same cardinal number since they are equipotent:

$$|\mathbb{Q}| = |\mathbb{Z}| = |\mathbb{N} \times \mathbb{N} \times \mathbb{N}| = |\mathbb{N}| = \aleph_0$$

d) If $A = \{1, 2, 3\}$ and $B = \{13, 14\}$ the cardinalities of A and B are 3 and 2 respectively. If we view 3 and 2 as cardinal numbers show, by referring to the definition of addition, multiplication and exponentiation of cardinal numbers, that $3+2 = 5$, $3 \times 2 = 6$ and $3^2 = 9$.

· By definition: $3 + 2 = |A \cup B| = |\{1, 2, 3, 13, 14\}| = 5$.

· By definition:

$$3 \times 2 = |A \times B| = |\{(1, 13), (1, 14), (2, 13), (2, 14), (3, 13), (3, 14)\}| = 6$$

· By definition: $3^2 = |A^B|$ is equal to the cardinality of the following set of functions mapping $\{13, 14\}$ into $\{1, 2, 3\}$:

$$\begin{gathered}\{(13, 1), (14, 1)\} \\ \{(13, 2), (14, 2)\} \\ \{(13, 3), (14, 3)\} \\ \{(13, 1), (14, 2)\} \\ \{(13, 1), (14, 3)\} \\ \{(13, 2), (14, 1)\} \\ \{(13, 2), (14, 3)\} \\ \{(13, 3), (14, 1)\} \\ \{(13, 3), (14, 2)\}\end{gathered}$$

Since there are precisely nine functions in this set, $3^2 = 9$.

e) What is the cardinality of the two sets $2^{\mathbb{N}}$ and $\mathscr{P}(\mathbb{N})$.

$$|2^{\mathbb{N}}| = |\mathscr{P}(\mathbb{N})| = |\mathbb{R}| = c$$

f) Compute $2^{\aleph_0}$.

$$2^{\aleph_0} = |2|^{|\mathbb{N}|} = |2^{\mathbb{N}}| = c$$

We will formally show that the class, $\mathscr{C}$, of all cardinal numbers is not a set. The proof mimics the one used to show that the class $\mathscr{E}$ is not a set of sets.

Theorem 22.4 The class $\mathscr{C}$ of all cardinal numbers is a proper class.

Proof:

Suppose $\mathscr{C}$ is a set. Then $T = \bigcup_{\kappa \in \mathscr{C}} \kappa$ must be a set (by the Axiom of union). This implies $\mathscr{P}(T)$ must be a set (Axiom of power set). Since $\mathscr{P}(T)$ is a set, it has a cardinality, say, λ. Then $\mathscr{P}(T) \sim_e \lambda \subset T$. So $\mathscr{P}(T)$ is equipotent to a subset of T, contradicting the previously established fact, $T \hookrightarrow_e \mathscr{P}(T)$. So $\mathscr{C}$ cannot be a set.

22.7 Previous theorem statements using cardinal number notation.

Many of the statements proven in the last few sections can now be restated using cardinal number notation. The results are from sections 19, 20 and 21. Let S and T be sets. Suppose that $\kappa = |S|$, $\lambda = |T|$.

a) If $|S| = \aleph_0$ and $T \subseteq S$, then either $|T| = n$ for some $n \in \mathbb{N}$ or $|T| = \aleph_0$ (theorem 19.3)

b) If $|S| = \kappa = |T|$ if and only if $S \sim_e T$. (postulate 22.2)

c) If $\aleph_0 \leq \kappa$, then $\kappa + \aleph_0 = \kappa$. (theorem 20.6)

d) If $\kappa = |S|$, then $|\mathscr{P}(S)| = 2^\kappa$. (theorem 20.12)

e) For all cardinal numbers κ, $\kappa < 2^\kappa$ (equivalently, $\kappa \hookrightarrow_e 2^\kappa$). (theorem 20.8)

f) $|\mathbb{R}| = c = 2^{\aleph_0}$ (theorem 21.3)

g) Continuum hypothesis: There does not exist a cardinal number κ such that $\aleph_0 < \kappa < c = |\mathbb{R}|$.

h) Generalized continuum hypothesis: For any cardinal number κ there does not exist a cardinal number λ such that $\kappa < \lambda < 2^\kappa$.

Concepts review:

1. What does the Continuum hypothesis say? What does the negation of the Continuum hypothesis say? Which one holds true in ZFC?

2. State the Generalized continuum hypothesis.

3. Define the class of all cardinal numbers.

4. Describe the finite cardinal numbers.

5. What symbol is used to represent the cardinality of the set $\mathbb{R}$?

6. What symbol is used to represent the cardinality of $\mathbb{N}$?
7. Which cardinal numbers are referred to as being *transfinite cardinal numbers*?
8. How are the operations of addition, multiplication and exponentiation of cardinal numbers defined?
9. Can the class of all cardinal numbers be referred to as a set? Why?

EXERCISES

A.
1. Show that there can be no largest cardinal number.
2. Show that for every finite cardinal number n, $n \in \aleph_0$.

B.
3. Prove that if $S \hookrightarrow_e T$ and $T \hookrightarrow_e M$, then $S \hookrightarrow_e M$.
4. Find $\cup\{\kappa \in \mathscr{C} : \ \kappa \text{ is a finite cardinal number}\}$.
5. Determine the cardinal number equal to each of the following expressions:
   a) $c + \aleph_0$
   b) $\aleph_0 \times \aleph_0$
   c) $2^{\aleph_0}$
   d) $\aleph_0^2$
6. What is the cardinality of the set $2^{\mathbb{Z}}$?
7. What is the cardinality of the set $\mathbb{N}^{\mathbb{N}}$?

C.
8. Prove that if $S \subseteq T \subseteq M$ and $S \sim_e M$, then $S \sim_e T$.
9. Prove that $\aleph_0 + c = c + \aleph_0$.
10. Prove that $\aleph_0 \times c = c \times \aleph_0$.
11. Prove that $\aleph_0 < 2^{\aleph_0}$.
12. Show that the class of all infinite sets is not a set.

---

23 / Addition and multiplication in $\mathscr{C}$.

Summary. *We have defined addition and multiplication of cardinal numbers in such a way that when adding or multiplying finite cardinal numbers, we obtain precisely the same answers as the ones obtained when performing these operations with natural numbers in the conventional way. In this section we verify that these two operations are well-defined even when adding and multiplying infinite cardinals. We then verify that addition and multiplication of cardinals satisfy, in many cases, the same properties as addition and multiplication of natural numbers. But not all of their properties generalize from the natural numbers to infinite cardinal numbers.*

23.1 Reviewing basic facts about $\mathscr{C}$.

We have seen that every element, κ, of the class $\mathscr{C}$ of all cardinal numbers is a set around which is gathered in a single class, $[\kappa]_e$, all sets S such that $S \sim_e \kappa$. Every $\sim_e$-equivalence class in $\mathscr{S}$ contains exactly one cardinal number. The finite cardinal numbers were declared to be the natural numbers. The elements of $\mathscr{C}$ are ordered by the relation "$<$" where $\kappa < \lambda$ if and only if $\kappa \hookrightarrow_e \lambda$, and $\kappa \leq \lambda$ if and only if $\kappa \hookrightarrow_{e\sim} \lambda$. We can not yet declare that $\leq$ linearly orders $\mathscr{C}$ since we have not yet shown that $\hookrightarrow_{e\sim}$ linearly orders $\mathscr{S}$ (although we certainly would like this to be the case). We will now study the properties of the two operations, $+$ and $\times$, previously defined on $\mathscr{C}$.

23.2 Addition of cardinal numbers.

Addition of two cardinal numbers was defined as follows: "If $\kappa = |S|$ and $\lambda = |T|$ where $S \cap T = \varnothing$, then $\kappa + \lambda$ is equal to the cardinal number $|S \cup T|$". We can easily see that if S and T are disjoint finite sets, addition of the cardinal numbers $|S|$ and $|T|$ is simply the number of elements in the set obtained when we merge both sets into one.

Although we are sure that addition of finite cardinals is well-defined, we should also verify that addition of *any* pair of cardinal numbers is well-defined.

Theorem 23.1 *Addition on $\mathscr{C}$ is well-defined.* That is, if S_1, S_2, T_1 and T_2 are sets such that $\kappa = |S_1| = |S_2|$ and $\lambda = |T_1| = |T_2|$ and $S_1 \cap T_1 = \varnothing = S_2 \cap T_2$, then $|S_1 \cup T_1| = \kappa + \lambda = |S_2 \cup T_2|$.

Proof:

What we are given: That $S_1 \cap T_1 = \varnothing = S_2 \cap T_2$, S_1 and S_2 are equipotent and T_1 and T_2 are equipotent.

What we are required to show: That $|S_1 \cup T_1|$ and $|S_2 \cup T_2|$ are the same cardinal number.

Since S_1, S_2 and T_1, T_2 are equipotent pairs, then there exist one-to-one onto functions

$$\begin{aligned} f : S_1 &\rightarrow S_2 \\ g : T_1 &\rightarrow T_2 \end{aligned}$$

By definition of addition, we have

$$\begin{aligned} \kappa + \lambda = |S_1| + |T_1| &= |S_1 \cup T_1| \\ \kappa + \lambda = |S_2| + |T_2| &= |S_2 \cup T_2| \end{aligned}$$

To prove that addition is well-defined, it suffices to show that $|S_1 \cup T_1| = |S_2 \cup T_2|$, i.e., that $S_1 \cup T_1$ and $S_2 \cup T_2$ are equipotent:

- We define the function $h : S_1 \cup T_1 \Rightarrow S_2 \cup T_2$ as follows: $h|_{S_1} = f$ and $h|_{S_2} = g$.
- Since S_1 and S_2 are disjoint and both f and g are one-to-one and onto, then h is a well-defined one-to-one and onto function.
- So $S_1 \cup T_1$ and $S_2 \cup T_2$ are equipotent. Thus, $|S_1 \cup T_1| = |S_2 \cup T_2|$ as required.

We now verify that addition on $\mathscr{C}$, thus defined, satisfies most of the basic addition properties.

Theorem 23.2 Let κ, λ, ϕ and ψ be any four cardinal numbers. Then

a) $\kappa + \lambda = \lambda + \kappa$ (Commutativity of addition)

b) $(\kappa + \lambda) + \phi = \kappa + (\lambda + \phi)$ (Associativity of addition)

c) $\kappa \leq \kappa + \lambda$

d) $\kappa \leq \lambda$ and $\phi \leq \psi \Rightarrow \kappa + \phi \leq \lambda + \psi$.

Proof

a) Let S and T be disjoint sets such that $\kappa = |S|$ and $\lambda = |T|$. To prove that $\kappa + \lambda = \lambda + \kappa$ it suffices to prove that $S \cup T \sim_e T \cup S$. This is left as an exercise.

b) Let S, T and F be disjoint sets such that $\kappa = |S|$, $\lambda = |T|$ and $\phi = |F|$. To prove that $(\kappa + \lambda) + \phi = \kappa + (\lambda + \phi)$ it suffices to show that $(S \cup T) \cup F \sim_e S \cup (T \cup F)$. This is left as an exercise.

c) Let S and T be disjoint sets such that $\kappa = |S|$ and $\lambda = |T|$. Since S and T are disjoint, we see that S can be mapped one-to-one into the subset S of $S \cup T$. Hence, $\kappa \leq \kappa + \lambda$.

d) Let $\{S, F\}$ and $\{T, P\}$ be to pairs of disjoint sets such that $\kappa = |S| \leq \lambda = |T|$ and $\phi = |F| \leq \psi = |P|$. The case where we have equality is straightforward. We will only prove the case involving the strict inequality "<". Assuming $\kappa < \lambda$ and $\phi < \psi$,

$$\left.\begin{array}{lll} S \hookrightarrow_e T & \hookrightarrow_e & T \cup P \\ F \hookrightarrow_e P & \hookrightarrow_e & T \cup P \end{array}\right\} \Rightarrow \left.\begin{array}{lll} S & \hookrightarrow_e & T \cup P \\ F & \hookrightarrow_e & T \cup P \end{array}\right\}$$

Since T and P are disjoint $S \cup F \hookrightarrow_e T \cup P$. Then $\kappa + \phi < \lambda + \psi$.

On cancelling out terms in addition. Not all addition properties which hold true for finite cardinals extend to infinite cardinals. For example, for finite cardinals m, n, k the statement

$$(m + n = m + k) \Rightarrow n = k$$

is always true. But for arbitrary cardinals κ, λ, ψ, if $\kappa + \lambda = \kappa + \psi$, it does **not** necessarily follow that $\lambda = \psi$. Recall (from theorem 20.6) that if

"If $\kappa \geq \aleph_0$ and $\lambda \leq \aleph_0$, then $\kappa = \kappa + \lambda$"

is shown to be true. For example, $c + 0 = c + \aleph_0$ does not imply that $\aleph_0$ is 0. Even though, for a non-zero finite cardinal n, the expression $n = n + n$ doesn't make sense, we will soon show that for any infinite cardinal κ, it is always true that $\kappa = \kappa + \kappa$.

23.3 Multiplication of cardinal numbers.

As previously stated, multiplication of two cardinal numbers is defined as being the cardinal number of the Cartesian product of the sets they represent. Just as we have done for addition we confirm that multiplication of cardinal numbers is well-defined.

Theorem 23.3 *Multiplication on $\mathscr{C}$ is well-defined.* That is, if S_1, S_2, T_1 and T_2 are sets such that $\kappa = |S_1| = |S_2|$ and $\lambda = |T_1| = |T_2|$, then

$$|S_1 \times T_1| = \kappa \times \lambda = |S_2 \times T_2|$$

Proof:

What we are given: That S_1 and S_2 are equipotent, T_1 and T_2 are equipotent.
What we are required to show: That $|S_1 \times T_1|$ and $|S_2 \times T_2|$ are the same cardinal number.

Since S_1, S_2 and T_1, T_2 are equipotent pairs, then there exist one-to-one onto functions

$$\begin{array}{lll} f : S_1 & \rightarrow & S_2 \\ g : T_1 & \rightarrow & T_2 \end{array}$$

By definition of multiplication, we have

$$\begin{aligned} \kappa \times \lambda = |S_1| \times |T_1| &= |S_1 \times T_1| \\ \kappa \times \lambda = |S_2| \times |T_2| &= |S_2 \times T_2| \end{aligned}$$

To prove that multiplication is well-defined it suffices to show that $|S_1 \times T_1| = |S_2 \times T_2|$, i.e., that $S_1 \times T_1$ and $S_2 \times T_2$ are equipotent:

- We define the function $h : S_1 \times T_1 \Rightarrow S_2 \times T_2$ as follows: $h(s,t) = (f(s), g(t))$.
- Since both f and g are one-to-one and onto, then h is a well-defined one-to-one and onto function.[1]
- So $S_1 \times T_1$ and $S_2 \times T_2$ are equipotent.

So multiplication is well-defined.

We now describe and prove a few of the most basic multiplication properties on $\mathscr{C}$. We will see that most (but not all) of the multiplication properties which hold true for the natural numbers extend to infinite cardinal numbers.

Theorem 23.4 Let κ, λ, ϕ and ψ be any three cardinal numbers. Then

a) $\kappa \times \lambda = \lambda \times \kappa$ (Commutativity of multiplication)

b) $(\kappa \times \lambda) \times \phi = \kappa \times (\lambda \times \phi)$ (Associativity of multiplication)

c) $\kappa \times (\lambda + \phi) = (\kappa \times \lambda) + (\kappa \times \phi)$ (Left-hand distributivity)

d) $\lambda > 0 \Rightarrow \kappa \leq (\kappa \times \lambda)$

e) $\kappa \leq \lambda$ and $\phi \leq \psi \Rightarrow \kappa \times \phi \leq \lambda \times \psi$.

f) $\kappa + \kappa = 2 \times \kappa$.

g) $\kappa + \kappa \leq \kappa \times \kappa$ when $\kappa \geq 2$.

Proof

a) Let S and T be sets such that $\kappa = |S|$ and $\lambda = |T|$.
What we are required to show: That $\kappa \times \lambda = \lambda \times \kappa$.
To attain this result it suffices to show that $S \times T \sim_e T \times S$.

[1] To see this, note that $(f(s_1), g(t_1)) = (f(s_2), g(t_2))$ implies that $f(s_1) = f(s_2)$ and $g(t_1) = g(t_2)$ which implies that $s_1 = s_2$ and $t_1 = t_2 \Rightarrow (s_1, t_1) = (s_2, t_2)$.

Let $h : S \times T \to T \times S$ be defined as $h(s,t) = (t,s)$. Now

$$\begin{aligned}(s,t) = (a,b) &\Leftrightarrow s = a \text{ and } t = b \\ &\Leftrightarrow (b,a) = (t,s)\end{aligned}$$

Then h is a one-to-one function.
Also, if $(u,v) \in T \times S$, then $(v,u) \in S \times T$ and $h(v,u) = (u,v)$. So h is onto $T \times S$.
We conclude that $S \times T \sim_e T \times S$, as required.

b) Let S, T and F be sets such that $\kappa = |S|$, $\lambda = |T|$ and $\phi = |F|$.
What we are required to show: That $\kappa \times (\lambda \times \phi) = (\lambda \times \kappa) \times \phi)$.
To attain this result it suffices to show that $(S \times T) \times F \sim_e T \times (S \times F)$. This is proven in theorem 4.9.

c) Let S, T and F be sets such that $\kappa = |S|$, $\lambda = |T|$ and $\phi = |F|$. Without loss of generality, suppose T and F are disjoint.
What we are required to show: That multiplication of cardinal numbers is left-hand distributive, i.e., $\kappa \times (\lambda + \phi) = (\kappa \times \lambda) + (\kappa \times \phi)$.
By definition, $\lambda + \phi = |T \cup F|$, $\kappa \times \lambda = |S \times T|$ and $\kappa \times \phi = |S \times F|$. So

$$\begin{aligned}\kappa \times (\lambda + \phi) &= |S \times (T \cup U)| \\ &= |(S \times T) \cup (S \times F)| \quad \text{(By theorem 4.7 b)).} \\ &= |S \times T| + |S \times F| \quad \text{(Since } T \text{ and } F \text{ are disjoint} \Rightarrow S \times T \text{ and } S \times F \text{ are disjoint).} \\ &= (\kappa \times \lambda) + (\kappa \times \phi)\end{aligned}$$

d) Let S and T be sets such that $\kappa = |S|$ and $\lambda = |T|$. Let $a \in T$.
The property $\kappa \leq (\kappa \times \lambda)$ follows from the fact that the function $h : S \to S \times T$ defined as $h(s) = (s,a)$ maps S one-to-one onto $S \times \{a\} \subseteq S \times T$. The details are left as an exercise.

e) Let S, T, F and P be sets such that $\kappa = |S|$, $\lambda = |T|$, $\phi = |F|$ and $\psi = |P|$. Suppose $f : S \to T$ maps S one-to-one into T and $g : F \to P$ maps F one-to-one into P. That is, suppose that $\kappa \leq \lambda$ and $\phi \leq \psi$
What we are required to show: That $\kappa \times \phi \leq \lambda \times \psi$.
To show this, it suffices to show a function $h : S \times F \to T \times P$ which maps $S \times F$ one-to-one into $T \times P$.
The function $h : S \times F \to T \times P$ defined as $h(s,u) = (f(s), g(u))$ can be shown to be one-to-one. This is left as an exercise.

f) What we are given: That S is a set such that $\kappa = |S|$ and 2 is the cardinal number of the set $\{0,1\}$.
What we are required to show: That $\kappa + \kappa = 2 \times \kappa$.
Note that $\kappa = |S|$ implies $\kappa = |S \times \{0\}| = |S \times \{1\}|$. Then, by definition,

$$\begin{aligned}
\kappa + \kappa &= |(S \times \{0\}) \cup (S \times \{1\})| \quad \text{(since } S \times \{0\} \text{ and } S \times \{1\} \text{ are disjoint).} \\
&= |S \times (\{0\} \cup \{1\})| \quad \text{(By theorem 4.7 b)).} \\
&= |S \times \{0, 1\}| \\
&= |\{0, 1\} \times S| \quad \text{(since } S \times \{0, 1\} \text{ and } \{0, 1\} \times S \text{ are equipotent).} \\
&= 2 \times \kappa
\end{aligned}$$

g) What we are given: That S is a set such that $\kappa = |S| \geq 2$.
What we are required to show: That $\kappa + \kappa \leq \kappa \times \kappa$.
By part f), $\kappa + \kappa = 2 \times \kappa$. So it suffices to show that $2 \times \kappa \leq \kappa \times \kappa$.
Since $2 \leq \kappa = |S|$, then $2 = \{0, 1\}$ is embedded in S. Let $f : \{0, 1\} \to S$ be a one-to-one function mapping $\{0, 1\}$ into S. Then the function $h : \{0, 1\} \times S \to S \times S$ defined as $h(i, s) = (f(i), s)$ can be seen as being one-to-one. Showing this is left as an exercise. It follows that $2 \times \kappa \leq \kappa \times \kappa$. So $\kappa + \kappa \leq \kappa \times \kappa$, as required.

Concepts review:

1. How do we go about showing that addition and multiplication of cardinal numbers are "well-defined"?
2. Is addition of cardinal numbers commutative? Is it associative?
3. Is multiplication of cardinal numbers commutative? Is it associative?
4. Does $\lambda + \kappa = \lambda + \psi$ imply $\kappa = \psi$? If so why? If not, give an example showing why not.

EXERCISES

A. 1. Show that $|S| \leq |S^T|$ for any set S and non-empty set T.

B. 2. Let κ, λ, ϕ and ψ be any three cardinal numbers. Show the details of the proofs of the following statements:
   a) $\kappa + \lambda = \lambda + \kappa$
   b) $(\kappa + \lambda) + \phi = \lambda + (\kappa + \phi)$
   c) $\kappa \leq \lambda$ and $\phi \leq \psi \Rightarrow \kappa + \phi \leq \lambda + \psi$.

3. Let κ, λ, ϕ and ψ be any three cardinal numbers. Show the details of the proofs of the following statements:

a) $\kappa \times \lambda = \lambda \times \kappa$
b) $(\kappa \times \lambda) \times \phi = \lambda \times (\kappa \times \phi)$
c) $\lambda > 0 \Rightarrow \kappa \leq (\kappa \times \lambda)$
d) $\kappa \leq \lambda$ and $\phi \leq \psi \Rightarrow \kappa \times \phi \leq \lambda \times \psi$.
e) $\kappa + \kappa \leq \kappa \times \kappa$ when $\kappa \geq 2$.

4. Show that $2^{\aleph_0} = |\mathbb{R} - \mathbb{N}|$.
5. Show that for any cardinal number κ, $\kappa + \kappa + \kappa + \kappa = 4 \times \kappa$.
6. Let n be a finite cardinal number. Prove that:
   a) $n + \aleph_0 = \aleph_0$.
   b) $n \times \aleph_0 = \aleph_0$.
   c) $n + 2^{\aleph_0} = 2^{\aleph_0}$.
   d) $n \times 2^{\aleph_0} = 2^{\aleph_0}$.
   e) $\aleph_0 + 2^{\aleph_0} = 2^{\aleph_0}$.
   f) $\aleph_0 \times 2^{\aleph_0} = 2^{\aleph_0}$.

C.

7. Prove that if $\kappa \times \lambda = 0$, then either $\kappa = 0$ or $\lambda = 0$.
8. Prove that if $\kappa \times \lambda = 1$, then either $\kappa = 1$ or $\lambda = 1$.
9. Prove that if $\kappa \times \lambda = \aleph_0$, then either $\kappa = \aleph_0$ or $\lambda = \aleph_0$.

24 / Exponentiation of cardinal numbers.

Summary. *In this section we show that cardinal exponentiation is well-defined. We then prove three of the most basic properties of cardinal exponentiation as well as inequalities involving cardinal exponentiation. We also show that* $|\mathbb{R}^{\mathbb{R}}| = c^c = 2^c$.

24.1 Cardinal exponentiation.

Given two sets, A and B, we have seen that the expression A^B represents the set of all functions mapping B into A. This means that every element, $f \in A^B$, is a subset of $B \times A$. Also, for every pair of ordered pairs in f of the form (x, u) and (x, y), $y = u$. We see that $f \in \mathscr{P}(B \times A)$; hence, $A^B \subset \mathscr{P}(B \times A)$. If both A and B are finite, we more easily understand the use of the notation A^B to represent this set. Suppose $A = \{a, b, c\}$, a three element set, and $B = \{0, 1\}$, a two element set. We then list the functions in the set A^B:

$$\begin{aligned} f_1 &: \{(0,a),(1,a)\} \\ f_2 &: \{(0,a),(1,b)\} \\ f_3 &: \{(0,a),(1,c)\} \\ f_4 &: \{(0,b),(1,b)\} \\ f_5 &: \{(0,b),(1,a)\} \\ f_6 &: \{(0,b),(1,c)\} \\ f_7 &: \{(0,c),(1,c)\} \\ f_8 &: \{(0,c),(1,a)\} \\ f_9 &: \{(0,c),(1,b)\} \end{aligned}$$

There are precisely nine elements in A^B. Or, we can say that the cardinality of $|A^B|$ of A^B is $|A|^{|B|} = 3^2 = 9$. So the notation A^B is designed to remind us of the number of elements contained in such sets when the sets A and B are finite. For convenience, this notation is maintained for sets of all cardinalities. In this section we try to develop a few rules that will help simplify expressions involving exponentiation of infinite cardinals. We will soon see that cardinal exponentiation is a considerably more complex operation than the cardinal addition and multiplication operations.

We remind ourselves of the formal definition of cardinal number exponentiation:

> If κ and λ are the cardinal numbers of the non-empty sets A and B we define $\kappa^\lambda = |A|^{|B|} = |A^B|$. For convenience we define $0^\lambda = 0$ and $\kappa^0 = 1$.

We will begin by showing that exponentiation of cardinal numbers is well-defined.

Theorem 24.1 *Exponentiation on $\mathscr{C}$ is well-defined.* That is, if S, S^*, T and T^* are sets such that $|S| = |S^*|$ and $|T| = |T^*|$, then $|S^T| = |S^{*T^*}|$.

Proof:

What we are given: The sets S and S^* are equipotent as well as the pair T and T^*.
What we are required to prove: That S^T and S^{*T^*} are equipotent.
Since $S \sim_e S^*$ and $T \sim_e T^*$ there exist one-to-one onto functions $\alpha : T \to T^*$ and $\beta : S \to S^*$.

If $g \in S^T$ define

$$\phi(g) = \{(\alpha(t), \beta(g(t))) : t \in T\} \in \mathscr{P}(T^* \times S^*)$$

We claim that for any $g \in S^T$, $\phi(g) \in S^{*T^*}$:

First note that $(\alpha(t), \beta(g(t))) \in T^* \times S^*$ for all $t \in T$ and that the domain of $\{(\alpha(t), \beta(g(t))) : t \in T\}$ is $T^* = \alpha[T]$. If $(\alpha(a), \beta(g(a)))$ and $(\alpha(b), \beta(g(b)))$ are elements of $\phi(g)$ such that $\beta(g(a)) \neq \beta(g(b))$, then $g(a) \neq g(b)$ (since β is a one-to-one function on S). Since $g \in S^T$, $a \neq b$. Since $\alpha : T \to T^*$ is one-to-one, $\alpha(a) \neq \alpha(b)$. We have shown that $\beta(g(a)) \neq \beta(g(b))$ implies that $\alpha(a) \neq \alpha(b)$. So $\phi(g)$ is a function whose domain is T^* with range S^*. Then, for any $g \in S^T$, $\phi(g) \in S^{*T^*}$ as claimed.

We claim $\phi : S^T \to S^{*T^*}$ is a one-to-one function on S^T:

Since ϕ associates to any $g \in S^T$ an element $\phi(g)$ in S^*T^*, ϕ has domain S^T and range S^*T^*. Suppose $h, g \in S^T$.

$$\begin{aligned}
h \neq g &\Leftrightarrow \exists\ u \in T \text{ such that } h(u) \neq g(u) \\
&\Leftrightarrow \beta(g(u)) \neq \beta(h(u)) \quad \text{(Since } \beta \text{ is one-to-one.)} \\
&\Leftrightarrow (\alpha(u), \beta(g(u))) \neq (\alpha(u), \beta(h(u))) \\
&\Leftrightarrow \{(\alpha(t), \beta(g(t))) : t \in T\} \neq \{(\alpha(t), \beta(h(t))) : t \in T\} \\
&\Leftrightarrow \phi(g) \neq \phi(h)
\end{aligned}$$

So $\phi : S^T \to S^{*T^*}$ is one-to-one on S^T as claimed.
Since S^T is embedded in S^{*T^*}. Then $|S^T| \leq |S^{*T^*}|$.

Using the same arguments but replacing $\alpha : T \to T^*$ with $\alpha^{-1} : T^* \to T$ and $\beta : S \to S^*$ with $\beta^{-1} : S^* \to S$ we can show that S^{*T^*} is embedded in S^T. Then $|S^{*T^*}| \leq |S^T|$.
By the Schröder-Bernstein theorem, $|S^{*T^*}| \leq |S^T|$ and $|S^T| \leq |S^{*T^*}|$ implies that $|S^T| = |S^{*T^*}|$ as required.

24.2 Three basic identities involving cardinal exponentiation.

We now verify that exponentiation in $\mathscr{C}$ satisfies the usual three basic exponential properties.

Theorem 24.2 Let κ, λ and ϕ be any three cardinal numbers. Then

a) $\kappa^{\lambda+\phi} = \kappa^{\lambda} \times \kappa^{\phi}$

b) $(\kappa^{\lambda})^{\phi} = \kappa^{\lambda \times \phi}$

c) $(\kappa \times \lambda)^{\phi} = \kappa^{\phi} \times \lambda^{\phi}$.

Proof:

What we are given: That κ, λ and ϕ are cardinal numbers where $\kappa = |S|$, $\lambda = |T|$ and $\phi = |U|$ for sets S, T and U.

a) $\kappa^{\lambda+\phi} = \kappa^{\lambda} \times \kappa^{\phi}$ (where it is assumed that $T \cap U = \varnothing$):

What we are required to show: That $|S^T \times S^U| = |S^{T \cup U}| = |S|^{|T \cup U|}$.
It suffices to show that $S^T \times S^U \sim_e S^{T \cup U}$.
Let $(f, g) \in S^T \times S^U$. Let $h_{\{f,g\}} : T \cup U \to S$ be the function defined as:

$$h_{\{f,g\}}(x) = \begin{cases} f(x) & \text{if} \quad x \in T \\ g(x) & \text{if} \quad x \in U \end{cases}$$

We claim that $h_{\{f,g\}} \in S^{T \cup U}$: If $h_{\{f,g\}}(a) \neq h_{\{f,g\}}(b)$, then either $f(a) \neq f(b)$, $g(a) \neq g(b)$ or $f(a) \neq g(b)$. Since f and g are functions and $T \cap U = \varnothing$, then for one of these three cases, $a \neq b$. Then $h_{\{f,g\}} \in S^{T \cup U}$.

Define the function $\phi : S^T \times S^U \to S^{T \cup U}$ as

$$\phi(f, g) = h_{\{f,g\}}$$

We claim that ϕ maps $S^T \times S^U$ one-to-one onto $S^{T \cup U}$:

- The function ϕ is well-defined: If $(f, g) \neq (k, r)$ in $S^T \times S^U$, then either $f(x) \neq k(x)$ for some $x \in T$ or $g(x) \neq r(x)$ for some $x \in U$; hence, $h_{\{f,g\}}(x) \neq h_{\{k,r\}}(x)$ for some $x \in T \cup U$. So $h_{\{f,g\}} \neq h_{\{k,r\}}$.
- The function ϕ is *onto* $S^{T \cup U}$: Suppose $t \in S^{T \cup U}$. Then $\phi(t|_T, t|_U)(x) = h_{\{t|_T, t|_U\}}(x)$ for all $x \in T \cup U$ since T and U are disjoint sets and $t|_T \in S^T$ and $t|_U \in S^U$.

– The function ϕ is *one-to-one*: Suppose (f, g) and (k, t) are distinct elements of $S^T \times S^U$. We are required to show that $\phi(f, g) \neq \phi(k, t)$. Suppose

$$\phi(f, g) = h_{\{f,g\}} = h_{\{k,t\}} = \phi(k, t)$$

Then $f = k$ on T and $g = t$ on U. So $(f, g) = (k, t)$ on $T \cup U$. Since $T \cap U = \varnothing$, $f = k$ in S^T and $g = t$ in S^U. Thus, $(f, g) = (k, t)$ in $S^T \times S^U$, a contradiction. Then $\phi(f, g) \neq \phi(k, t)$ as claimed.

We conclude that $S^T \times S^U$ and $S^{T \cup U}$ are equipotent. Hence,

$$\kappa^\lambda \times \kappa^\phi = |S|^{|T|} \times |S|^{|U|} = |S^T \times S^U| = |S^{T \cup U}| = |S|^{|T \cup U|} = \kappa^{\lambda + \phi}$$

b) $(\kappa^\lambda)^\phi = \kappa^{\lambda \times \phi}$:

What we are required to show: That $S^{T \times U}$ and $(S^T)^U$ are equipotent.

For each $u \in U$ and $f \in S^{T \times U}$ we define the function $f_u : T \to S$ in S^T as

$$f_u(t) = f|_{T \times \{u\}}(t, u) \in S$$

Then for each $u \in U$, f_u maps T into S. That is,

$$\{f_u : f \in S^{T \times U}, u \in U\} \subseteq S^T$$

Given $f \in S^{T \times U}$ define the function $\phi_f : U \to S^T$ as follows:

$$\phi_f(u) = f_u$$

We claim that for each f, ϕ_f is a well-defined function mapping U into S^T:

$$\begin{aligned} \phi_f(u_1) \neq \phi_f(u_2) &\Rightarrow f_{u_1} \neq f_{u_2} \\ &\Rightarrow f|_{T \times \{u_1\}}(t, u_1) \neq f|_{T \times \{u_2\}}(t, u_2), \quad \text{for all } t \in T. \\ &\Rightarrow f(t, u_1) \neq f(t, u_2), \quad \text{for all } t \in T. \\ &\Rightarrow u_1 \neq u_2 \end{aligned}$$

So $\phi_f : U \to S^T$ is well-defined, as claimed.

We define the relation $\psi : S^{T \times U} \to (S^T)^U$ as

$$\psi(f) = \phi_f \in (S^T)^U$$

We claim that ψ is a one-to-one function on $S^{T \times U}$:

$$\begin{aligned} f \neq g &\Leftrightarrow f(t_0, u_0) \neq g(t_0, u_0) \text{ for some pair } (t_0, u_0) \in T \times U. \\ &\Leftrightarrow f|_{T\times\{u_0\}}(t_0, u_0) \neq g|_{T\times\{u_0\}}(t_0, u_0) \\ &\Leftrightarrow f_{u_0}(t_0) \neq g_{u_0}(t_0) \\ &\Leftrightarrow f_{u_0} \neq g_{u_0} \\ &\Leftrightarrow \phi_f \neq \phi_g \\ &\Leftrightarrow \psi(f) \neq \psi(g) \end{aligned}$$

So ψ is a one-to-one function on $S^{T\times U}$ as claimed.

We claim that the function ψ is onto $(S^T)^U$:

Let ϕ be a function in $(S^T)^U$. Then $\phi(u) \in S^T$ for all u in its domain U.

We are required to exhibit a function $f \in S^{T\times U}$ such that $\psi(f) = \phi$. That is, we must find a function function $f \in S^{T\times U}$ such that $\phi_f(u) = f_u = \phi(u)$ for all $u \in U$.

For each $u \in U$, $\phi(u)$ is an element in S^T. Define, for each $u \in U$, the function g_u mapping $T \times \{u\}$ into S as

$$g_u(t, u) = [\phi(u)](t), \ \forall t \in T$$

Note that $T \times U = \cup\{T \times \{u\} : u \in U\}$, the union of a collection of pairwise disjoint sets. Define $f : T \times U \to S$ as follows:

$$f = \cup\{g_u : u \in U\}$$

Since the respective domains of the g_u's are pairwise disjoint and their union is all of $T \times U$, then $f : T \times U \to S$ is well-defined. Then $g_u = f|_{T\times\{u\}} = f_u$ for each $u \in U$. So $\phi_f(u) = f_u = \phi(u)$ for each $u \in U$. That is, $\phi_f = \psi(f)$.
Hence, the function ψ is onto $(S^T)^U$, as claimed.

So the sets $S^{T\times U}$ and $(S^T)^U$ are equipotent. We conclude that $(\kappa^\lambda)^\phi = \kappa^{\lambda\times\phi}$.

c) $(\kappa \times \lambda)^\phi = \kappa^\phi \times \lambda^\phi$:

What we are required to show: That $S^U \times T^U$ and $(S \times T)^U$ are equipotent.
We define the function $\phi : S^U \times T^U \to (S \times T)^U$ as follows:

$$\phi(f, g) = h$$

where $f \in S^U$, $g \in T^U$ and $h : U \to S \times T$ is defined as $h(u) = (f(u), g(u))$. That is, $\phi(f, g)(u) = h(u) = (f(u), g(u))$ for all $u \in U$.
We claim that ϕ maps $S^U \times T^U$ one-to-one onto $(S \times T)^U$:

The function ϕ is onto $(S \times T)^U$: Suppose $h \in (S \times T)^U$. Then $h : U \to S \times T$. For any $u \in U$, $h(u) = (h_1(u), h_2(u))$. Then $\phi(h_1, h_2) = h$; so ϕ maps $S^U \times T^U$ onto $(S \times T)^U$.

The function ϕ is one-to-one: Let (f_1, g_1) and (f_2, g_2) be pairs of functions in $S^U \times T^U$ and q and r be functions mapping U into $S \times T$ defined as $q(u) = (f_1(u), g_1(u))$ and $r(u) = (f_2(u), g_2(u))$.
Then

$$\begin{aligned}
\phi(f_1, g_1) \neq \phi(f_2, g_2) &\Leftrightarrow q \neq r \\
&\Leftrightarrow q(u) \neq r(u), \text{ for some } u \in U \\
&\Leftrightarrow (f_1(u), g_1(u)) \neq (f_2(u), g_2(u)) \\
&\Leftrightarrow f_1(u) \neq f_2(u) \text{ or } g_1(u) \neq g_2(u) \\
&\Leftrightarrow f_1 \neq f_2 \text{ or } g_1 \neq g_2 \\
&\Leftrightarrow (f_1, g_1) \neq (f_2, g_2)
\end{aligned}$$

Then the function ϕ is one-to-one.

We conclude that ϕ maps $S^U \times T^U$ one-to-one onto $(S \times T)^U$ and so these two sets are equipotent.

The following example shows how these identities can help simplify the computation of cardinal exponentials.

Find the cardinality of $\mathbb{R}^{\mathbb{R}}$.

Solution:

$$\begin{aligned}
|\mathbb{R}^{\mathbb{R}}| &= c^c \\
&= (2^{\aleph_0})^c \\
&= 2^{\aleph_0 \times c}
\end{aligned}$$

The function $f(x) = (0, x) \in \{0\} \times \mathbb{R}$ embeds $\mathbb{R}$ in $\mathbb{N} \times \mathbb{R}$; hence, $c \leq \aleph_0 \times c$. Since $\aleph_0 < c$, then, by theorem 23.4 e), $\aleph_0 \times c \leq c \times c$. By corollary 20.4, $c \times c = c$. We then have

$$c \leq \aleph_0 \times c \leq c \times c = c$$

which implies $\aleph_0 \times c = c$. Then

$$|\mathbb{R}^{\mathbb{R}}| = 2^{\aleph_0 \times c} = 2^c$$

24.3 A few basic inequalities involving cardinal exponentiation.

We verify a few more basic cardinal exponentiation properties.

Theorem 24.3 Let κ, λ, and α be infinite cardinal numbers. Then

a) $\kappa \leq \kappa^\lambda$

b) $\alpha \leq \kappa \Rightarrow \alpha^\lambda \leq \kappa^\lambda$

c) $\alpha \leq \lambda \Rightarrow \kappa^\alpha \leq \kappa^\lambda$

Proof:

What we are given: That $\kappa = |K|$, $\alpha = |A|$ and $\lambda = |L|$.

a) $\kappa \leq \kappa^\lambda$:

What we are required to show: That K is embedded in K^L.
Define the function $f : K \to K^L$ as follows: $f(k) = \{k\}^L \subset K^L$. Note that $\{k\}^L$ contains only one function; it maps all elements of L to the single element k. Since "$k \neq t$ implies $\{k\}^L \neq \{t\}^L$", the function f is one-to-one. Since f embeds K in K^L, then $\kappa \leq \kappa^\lambda$.

b) $\alpha \leq \kappa \Rightarrow \alpha^\lambda \leq \kappa^\lambda$:

What we are also given: That A is embedded in K.
What we are required to show: That A^L is embedded in K^L.
Suppose the function $f : A \to K$ embeds A into K. Define $\phi : A^L \to K \times L$

$$\phi(g) = \{(l, f(g(l))) : l \in L\} \subseteq L \times K$$

We claim that $\phi(g) \in K^L$:

If $(a, f(g(a)))$ and $(b, f(g(b)))$ are elements of $\phi(g)$ such that $f(g(a)) \neq f(g(b))$, then $g(a) \neq g(b)$ (since f is a function mapping A to K). Since $(a, g(a))$ and $(b, g(b))$ both belong to $g \in A^L$, then $a \neq b$ and so $\phi(g)$ is a function in K^L as claimed.

We claim $\phi : A^L \to K^L$ is one-to-one:
Suppose $h, g \in A^L$.

$$\begin{aligned}
h \neq g &\Leftrightarrow \exists\ u \in L \text{ such that } h(u) \neq g(u) \\
&\Leftrightarrow f(g(u)) \neq f(h(u)) \text{ since } f : A \to K \text{ is one-to-one.} \\
&\Leftrightarrow (u, f(g(u))) \neq (u, f(h(u))) \\
&\Leftrightarrow \{(l, f(g(l))) : l \in L\} \neq \{(l, f(h(l))) : l \in L\} \\
&\Leftrightarrow \phi(g) \neq \phi(h)
\end{aligned}$$

So $\phi : A^L \to K^L$ is one-to-one as claimed. Since $\phi : A^L \to K^L$ embeds A^L into K^L, then $\alpha^\lambda \leq \kappa^\lambda$ as required.

c) [AC] $\alpha \leq \lambda \Rightarrow \kappa^\alpha \leq \kappa^\lambda$:

What we are also given: That A is embedded in L.
What we are required to show: That K^A is embedded in K^L.
Suppose the function $f : A \to L$ embeds A into L. Define $\phi : K^A \to f[A] \times K$ as:

$$\phi(g) = \{(f(a), g(a)) : a \in A\} \subseteq f[A] \times K \subseteq L \times K$$

We claim that $\phi(g) \in K^{f[A]}$:

If $(f(a), g(a))$ and $(f(b), g(b))$ are elements of $\phi(g)$ such that $g(a) \neq g(b)$, then $a \neq b$ (since $g \in K^A$). Since f is one-to-one mapping A into L, $a \neq b$ implies $f(a) \neq f(b)$. Then $\phi(g) \in K^{f[A]}$ as claimed.

We claim $\phi : K^A \to K^{f[A]}$ is one-to-one:

Suppose $h, g \in K^A$.

$$\begin{aligned} h \neq g &\Leftrightarrow \exists\ u \in A \text{ such that } h(u) \neq g(u) \\ &\Leftrightarrow (f(u), g(u)) \neq (f(u), h(u)) \\ &\Leftrightarrow \{(f(a), g(a)) : a \in A\} \neq \{(f(a), h(a)) : a \in A\} \\ &\Leftrightarrow \phi(g) \neq \phi(h) \end{aligned}$$

So $\phi : K^A \to K^{f[A]}$ is one-to-one as claimed.

For each function $h \in K^{f[A]}$ the axiom of choice allows us to choose a function $h^* \in K^L$ such that $h^*|_{f[A]} = h$. We define the function $\phi^* : K^A \to K^L$ as $\phi^*(g) = \phi(g)^*$. Since $\phi : K^A \to K^{f[A]}$ is one-to-one, then $\phi^* : K^A \to K^L$ is one-to-one.
We conclude that K^A is embedded in K^L. This implies $\kappa^\alpha \leq \kappa^\lambda$ as required.

Concepts review:

1. If κ and λ are two cardinal numbers how is the expression κ^λ defined?
2. What are the three basic identities for cardinal exponentiation stated and proved in this section?
3. What are the three basic inequalities for cardinal exponentiation stated and proved in this section?

EXERCISES

A. 1. Show that for any cardinal number κ:
a) $\kappa^1 = \kappa$.

b) $1^\kappa = 1$.
c) $\kappa^0 = 1$
d) $0^\kappa = 0$, if $\kappa > 0$.

B. 2. Show that for any finite cardinal number n and any cardinal number κ:
a) $(2^{\aleph_0})^n = 2^{\aleph_0}$
b) $\aleph_0^n = \aleph_0$
c) $\aleph_0^{\aleph_0} = 2^{\aleph_0}$
d) $n^{\aleph_0} = 2^{\aleph_0}$.
e) $(2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0}$

3. Let κ be an infinite cardinal number. Suppose $|K| = \kappa$ and that $\{K_i : i \in K\}$ is a set of pairwise disjoint sets K_i each of which has cardinality κ. Show that $|\cup \{K_i : i \in K\}| = \kappa$.

C. 4. Show that if the cardinal number $\kappa \neq 1$, then $\kappa^\kappa \neq \kappa$.
5. Show that $2^{\aleph_0} < 2^{(2^{\aleph_0})}$.

25 / On sets of cardinality c.

Summary. *In this section we examine a few sets whose cardinality is the same as the cardinality of* $\mathbb{R}$. *In particular, we determine the cardinality of the set of all one-to-one functions mapping* $\mathbb{N}$ *to* $\mathbb{R}$ *and the cardinality of the set of all continuous real-valued functions. We also study the well-known Cantor set, discuss its construction and prove that it has cardinality c.*

25.1 Sets related to $\mathbb{R}$.

It can sometimes be a challenge to determine the cardinality of certain sets. Many of the cardinal arithmetic principles presented in the last section will help us determine the cardinality of sets associated to $\mathbb{R}$.

The cardinality of finite products of the reals, of the complex numbers and the irrational numbers are discussed in the following theorem.

Theorem 25.1 Let $\mathbb{C}$ denote the set of all complex numbers and $\mathbb{J}$ denote the set of all irrational numbers. Let n denote the cardinality of a non-empty finite set.

a) The cardinality of $\mathbb{R}^n$ is c.

b) The cardinality of $\mathbb{C}$ is c.

c) The cardinality of $\mathbb{J}$ is c.

Proof:

a) $|\mathbb{R}^n| = c$:

To prove that $|\mathbb{R}^n| = c$ it suffices to show that $c^n = c$.
We will prove this by mathematical induction.
What we are given: That n is a natural number greater than zero.
What we are required to show: That $c^n = c$.
Let $P(n)$ be the statement "$c^n = c$".

- Base case: Trivially, $P(1)$ holds true.
- Inductive hypothesis: Suppose $P(n)$ holds true. That is, suppose $c^n = n$. Then

$$\begin{aligned} c^{n+1} &= c^n \times c^1 \quad \text{(theorem 24.2 a).)} \\ &= c \times c \quad \text{(By the inductive hypothesis.)} \\ &= c \quad \text{(corollary 20.4)} \end{aligned}$$

So by mathematical induction $c^n = c$ for all finite non-zero cardinals n. So $|\mathbb{R}^n| = c$.

b) $|\mathbb{C}| = c$:

Define the function $f : \mathbb{R}^2 \to \mathbb{C}$ as $f(a, b) = a + bi$. The function f is easily shown to be one-to-one. So $\mathbb{R}^2$ and $\mathbb{C}$ are equipotent. It follows that $|\mathbb{R}^2| = |\mathbb{C}| = c$.

c) $|\mathbb{J}| = c$:
Suppose $\kappa = |\mathbb{J}|$. Since $\mathbb{J} \cup \mathbb{Q} = \mathbb{R}$ and $\mathbb{J} \cap \mathbb{Q} = \varnothing$, then $\kappa + \aleph_0 = c$. If $\kappa \leq \aleph_0$, then $\kappa + \aleph_0 = \aleph_0 \neq c$. So $\aleph_0 < \kappa$. That is, κ is an uncountable set. By theorem, 20.6, $\kappa + \aleph_0 = \kappa$. Hence, $\kappa = c$. That is, $|\mathbb{J}| = c$ as required.

25.2 The cardinality of sets of sequences and functions.

Sets of sequences and functions are more abstract in nature. This sometimes makes it more difficult to determine their cardinality. The following theorem illustrates a few strategies that can be used to determine the cardinality of such sets. In the proof of the following theorem we will invoke the statement, $\mathbb{N}\times\mathbb{N} \sim \mathbb{N}$, proven in theorem 19.4.

Theorem 25.2

a) Let $\mathscr{S}_{\mathbb{R}}$ denote the set of all countably infinite sequences of real numbers. Then the cardinality of $\mathscr{S}_{\mathbb{R}}$ is c.

b) Let $\mathscr{S}_{\mathbb{N}}$ denote the set of all countably infinite sequences of natural numbers. Then the cardinality of $\mathscr{S}_{\mathbb{N}}$ is c.

c) Let $\mathbb{N}^{\mathbb{N}}_{(1-1)}$ denote the set of all one-to-one functions mapping $\mathbb{N}$ to $\mathbb{N}$. Then the cardinality of $\mathbb{N}^{\mathbb{N}}_{(1-1)}$ is c.

d) Let $\mathbb{R}^{\mathbb{N}}_{(1-1)}$ denote the set of all one-to-one functions mapping $\mathbb{N}$ to $\mathbb{R}$. Then the cardinality of $\mathbb{R}^{\mathbb{N}}_{(1-1)}$ is c.

Proof:

a) $|\mathscr{S}_{\mathbb{R}}| = c$:
A sequence of real numbers $\{a_0, a_1, a_2, \ldots\}$ is a function $s : \mathbb{N} \to \mathbb{R}$ mapping each natural number $i \in \mathbb{N}$ to $a_i \in \mathbb{R}$. So each infinite sequence $\{a_0, a_1, a_2, \ldots\}$ is associated to a unique function $s : \mathbb{N} \to \mathbb{R}$. So the set of all infinite sequences of real numbers can be represented by $\mathbb{R}^{\mathbb{N}}$. Then

$$\begin{aligned} |\mathbb{R}^{\mathbb{N}}| &= (2^{\aleph_0})^{\aleph_0} \\ &= 2^{\aleph_0 \times \aleph_0} \quad \text{(By theorem 24.2)} \\ &= 2^{|\mathbb{N}\times\mathbb{N}|} \\ &= 2^{|\mathbb{N}|} = 2^{\aleph_0} \quad \text{(By theorem 19.4)} \\ &= |2^{\mathbb{N}}| = c \end{aligned}$$

Then $|\mathscr{S}_{\mathbb{R}}| = |\mathbb{R}^{\mathbb{N}}| = c$.

b) $|\mathscr{S}_{\mathbb{N}}| = c$:

A sequence of natural numbers $\{a_0, a_1, a_2, \ldots\}$ is a function $f : \mathbb{N} \to \mathbb{N}$ mapping each natural number $i \in \mathbb{N}$ to $a_i \in \mathbb{N}$. So the set of all infinite sequences of natural numbers can be represented by $\mathbb{N}^{\mathbb{N}}$. The cardinality of the set of all infinite sequences of natural numbers is then $|\mathbb{N}^{\mathbb{N}}|$. Note that

$$\begin{aligned} f \in \mathbb{N}^{\mathbb{N}} &\Rightarrow f \subseteq \mathbb{N} \times \mathbb{N} \\ &\Rightarrow f \in \mathscr{P}(\mathbb{N} \times \mathbb{N}) \\ &\Rightarrow \mathbb{N}^{\mathbb{N}} \subseteq \mathscr{P}(\mathbb{N} \times \mathbb{N}) \end{aligned}$$

Then

$$\begin{aligned} c &= 2^{\aleph_0} \\ &\leq \aleph_0^{\aleph_0} \quad \text{(By theorem 24.3 b).)} \\ &= |\mathbb{N}^{\mathbb{N}}| \\ &\leq |\mathscr{P}(\mathbb{N} \times \mathbb{N})| \\ &= |\mathscr{P}(\mathbb{N})| \quad (\mathbb{N} \times \mathbb{N} \sim \mathbb{N} \text{ followed by 20.7.}) \\ &= |\mathbb{R}| = c \end{aligned}$$

We conclude that $|\mathbb{N}^{\mathbb{N}}| = |\mathscr{S}_{\mathbb{N}}| = c$.

c) $|\mathbb{N}^{\mathbb{N}}_{(1-1)}| = c$:

Let $f \in \mathbb{N}^{\mathbb{N}}$. Consider the set

$$S_f = \{(n, (n, f(n)) : n \in \mathbb{N}\} \subset \mathbb{N} \times (\mathbb{N} \times \mathbb{N})$$

We claim that S_f is a one-to-one function mapping $\mathbb{N}$ into $\mathbb{N} \times \mathbb{N}$:

If $(n, (a, b))$ and $(n, (c, d))$ belong to S_f, then $a = n = c$ and, since f is a function, $b = f(n) = d$; hence, $(a, b) = (c, d)$. So S_f is a function mapping $\mathbb{N}$ into $\mathbb{N} \times \mathbb{N}$. If $(a, (n, f(n)))$ and $(b, (n, f(n))$ belong to S_f, then $a = b = n$. So $S_f : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ is a one-to-one function, as claimed.

We know that $\mathbb{N} \times \mathbb{N} \sim_e \mathbb{N}$. Then there exists a function h mapping $\mathbb{N} \times \mathbb{N}$ one-to-one onto $\mathbb{N}$. Since both h and S_f are one-to-one, then, for each $f \in \mathbb{N}^{\mathbb{N}}$, $h \circ S_f : \mathbb{N} \to \mathbb{N}$ is a one-to-one function and so belongs to $\mathbb{N}^{\mathbb{N}}_{(1-1)}$.

Let $H : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}_{(1-1)}$ be a function defined as $H(f) = h \circ S_f$. We claim that H is one-to-one:

$$\begin{aligned} f \neq g &\Rightarrow f(m) \neq g(m) \quad \text{for some natural number } m \\ &\Rightarrow S_f(m) = (m, f(m)) \neq (m, g(m)) = S_g(m) \quad \text{for some natural number } m \\ &\Rightarrow h(S_f(m)) \neq h(S_g(m)) \quad \text{for some natural number } m \\ &\Rightarrow h \circ S_f \neq h \circ S_g \\ &\Rightarrow H(f) \neq H(g) \end{aligned}$$

So H is one-to-one as claimed.

Then the function H embeds the set $\mathbb{N}^{\mathbb{N}}$ into $\mathbb{N}^{\mathbb{N}}_{(1-1)}$. We conclude that $|\mathbb{N}^{\mathbb{N}}| \leq |\mathbb{N}^{\mathbb{N}}_{(1-1)}|$. Since $|\mathbb{N}^{\mathbb{N}}_{(1-1)}| \leq |\mathbb{N}^{\mathbb{N}}|$, then by the Schröder-Bernstein theorem $|\mathbb{N}^{\mathbb{N}}_{(1-1)}| = |\mathbb{N}^{\mathbb{N}}| = c$.

d) $|\mathbb{R}^{\mathbb{N}}_{(1-1)}| = c$:

By part c) $|\mathbb{N}^{\mathbb{N}}_{(1-1)}| = c$. Since $\mathbb{N}^{\mathbb{N}}_{(1-1)} \subset \mathbb{R}^{\mathbb{N}}_{(1-1)}$, then $|\mathbb{R}^{\mathbb{N}}_{(1-1)}| \geq c$. Also given that $\mathbb{R}^{\mathbb{N}}_{(1-1)} \subset \mathbb{R}^{\mathbb{N}}$

$$\begin{aligned} |\mathbb{R}^{\mathbb{N}}_{(1-1)}| &\leq |\mathbb{R}^{\mathbb{N}}| \\ &= |\mathbb{R}|^{|\mathbb{N}|} \\ &= (2^{\aleph_0})^{\aleph_0} \\ &= 2^{\aleph_0 \times \aleph_0} \\ &= 2^{\aleph_0} \\ &= c \end{aligned}$$

So $|\mathbb{R}^{\mathbb{N}}_{(1-1)}| = c$.

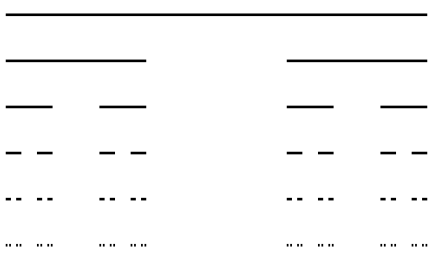
25.3 The Cantor set.

The Cantor set is an inductively constructed subset of the closed interval [0, 1]. Because of its interesting properties it is often discussed in various branches of mathematics. We will present the steps for its construction and discuss its cardinality.

In what follows, the expression (a, b) will mean the open interval in [0, 1], with endpoints a and b. We will construct a countably infinite set of subsets $\{C_n : n \in \mathbb{N}\}$ where each C_n is defined as follows:

$$\begin{aligned} C_0 &= [0, 1] \\ C_1 &= C_0 - (1/3,\, 2/3) \\ C_2 &= C_1 - [(1/9,\, 2/9) \cup (7/9,\, 8/9)] \\ C_3 &= C_2 - [(1/27,\, 2/27) \cup (7/27,\, 8/27) \cup (13/27,\, 14/27) \cup (25/27,\, 26/27)] \\ &\vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots \end{aligned}$$

The construction can be summarized as follows: C_{n+1} is obtained by "punching out" open middle thirds from each closed subinterval in its predecessor C_n. Actually constructing C_0 to C_3 will allow one to see the pattern of construction and develop a mental picture of what each C_n looks like.



Pursuing this process an infinite number of times will yield a countably infinite set $\{C_n : n \in \mathbb{N}\}$ of subsets of the closed interval [0, 1]. For every natural number n, we see that $C_{n+1} \subset C_n$. Furthermore, for each natural number n, C_n will be the union of 2^n closed intervals. We will index these closed intervals with sequences of n zeroes and ones as follows:[2]

$$
\begin{aligned}
C_0 &= [0,1] \\
C_1 &= {}_1I_{\{0\}} \cup {}_1I_{\{1\}} \\
C_2 &= {}_2I_{\{0,0\}} \cup {}_2I_{\{0,1\}} \cup {}_2I_{\{1,0\}} \cup {}_2I_{\{1,1\}} \\
C_3 &= {}_3I_{\{0,0,0\}} \cup {}_3I_{\{0,0,1\}} \cup {}_3I_{\{0,1,0\}} \cup {}_3I_{\{0,1,1\}} \cup {}_3I_{\{1,0,0\}} \cup {}_3I_{\{1,0,1\}} \cup {}_3I_{\{1,1,0\}} \cup {}_3I_{\{1,1,1\}} \\
C_4 &= {}_4I_{\{0,0,0,0\}} \cup {}_4I_{\{0,0,0,1\}} \cup {}_4I_{\{0,0,1,0\}} \cup {}_4I_{\{0,0,1,1\}} \cup {}_4I_{\{0,1,0,0\}} \cup {}_4I_{\{0,1,0,1\}} \cup {}_4I_{\{0,1,1,0\}} \\
&\quad \cup {}_4I_{\{0,1,1,1\}} \cup {}_4I_{\{1,0,0,0\}} \cup {}_4I_{\{1,0,0,1\}} \cup {}_4I_{\{1,0,1,0\}} \cup {}_4I_{\{1,0,1,1\}} \cup {}_4I_{\{1,1,0,0\}} \\
&\quad \cup {}_4I_{\{1,1,0,1\}} \cup {}_4I_{\{1,1,1,0\}} \cup {}_4I_{\{1,1,1,1\}} \\
\vdots &\qquad\qquad \vdots
\end{aligned}
$$

For example,

$$
\begin{aligned}
C_2 &= [0, \tfrac{1}{9}] \cup [\tfrac{2}{9}, \tfrac{3}{9}] \cup [\tfrac{6}{9}, \tfrac{7}{9}] \cup [\tfrac{8}{9}, \tfrac{9}{9}] \\
&= {}_2I_{\{0,0\}} \cup {}_2I_{\{0,1\}} \cup {}_2I_{\{1,0\}} \cup {}_2I_{\{1,1\}} \\
&\subset {}_1I_{\{0\}} \cup {}_1I_{\{1\}} \\
&= [0, \tfrac{1}{3}] \cup [\tfrac{2}{3}, \tfrac{3}{4}] \\
&= C_1
\end{aligned}
$$

More generally, for each $n \in \mathbb{N}$,

$$C_n = \cup\{{}_nI_t : t \in \{0,1\}^n\}$$

[2]Some readers may want to refer to the natural numbers expressed in base 2: 0, 1, 10, 11, 100, 101, 110, 111, ..., to determine the order in which the finite zero-one sequences are ordered.

Since $C_{n+1} \subset C_n$ for each $n \in \mathbb{N}$, the sets $\{C_n : n \in \mathbb{N}\}$ are said to be "nested". We call the intersection

$$C = \bigcap_{n\in\mathbb{N}} C_n$$

the *Cantor set*.

We begin by listing a few facts about the Cantor set, C, and its construction.

– Note that if A is a sequence of n zeroes and ones and ${}_nI_A$ is one of the intervals which forms C_n, then, by the construction rules of C_{n+1},

$${}_{n+1}I_{A\cup\{0\}} \ \cup \ {}_{n+1}I_{A\cup\{1\}} \subset {}_nI_A$$

For example,

$$\begin{aligned}
{}_3I_{\{0,0,0\}} \cup {}_3I_{\{0,0,1\}} &\subset {}_2I_{\{0,0\}}\\
{}_3I_{\{0,1,0\}} \cup {}_3I_{\{0,1,1\}} &\subset {}_2I_{\{0,1\}}\\
{}_3I_{\{1,0,0\}} \cup {}_3I_{\{1,0,1\}} &\subset {}_2I_{\{1,0\}}\\
{}_3I_{\{1,1,0\}} \cup {}_3I_{\{1,1,1\}} &\subset {}_2I_{\{1,1\}}
\end{aligned}$$

Then every level $C_n = C_0 \cap C_1 \cap C_2 \cap \cdots \cap C_n$ can also be obtained by taking the finite union of the intersections of nested sets of closed intervals

$$\bigcap_{i=0}^{n} {}_iI_{A_i}$$

One nested set of closed intervals in C_{100} would be, for example,

$$[0,1] \cap [0,\tfrac{1}{3}] \cap [0,\tfrac{1}{3^2}] \cap [0,\tfrac{1}{3^3}] \cap \cdots \cap [0,\tfrac{1}{3^{100}}]$$

If $\{{}_nI_A : n \in \mathbb{N}\}$ is an infinite set of nested closed intervals, then $\cap_{n\in\mathbb{N}}\{{}_nI_A\} \neq \varnothing$.[1] At any level n there are 2^n closed intervals in C_n; 2^n is precisely the number of different sequences containing n zeroes and ones.

– The length of each closed interval which forms C_n is $\frac{1}{3^n}$. For example, the set C_3 is the union of closed intervals of the form $[\frac{n}{3^3}, \frac{n+1}{3^3}]$. If $\{{}_nI_A : n \in \mathbb{N}\}$ is an infinite set of nested closed intervals, then the length of the interval $\cap_{n\in\mathbb{N}}\{{}_nI_A\}$ must be $\lim_{n\to\infty} \frac{1}{3^n} = 0$. From this we must conclude that the non-empty set $\cap_{n\in\mathbb{N}}\{{}_nI_A\}$ cannot contain any more than a single element x.

– Notice that every endpoint in C_n will still appear in C_{n+1}. Once an endpoint appears in C_n for some n it is never removed in subsequent steps. For example, since $\frac{1}{9} \in C_2$, then $\frac{1}{9} \in C_n$ for all $n \geq 2$. So if E_n denotes the endpoints of the 2^n closed intervals in C_n, $E_n \subseteq \bigcap_{n\in\mathbb{N}} C_n$. So the countably infinite set $\cup_{n\in\mathbb{N}} E_n$ is contained in C. So the Cantor set contains infinitely many points.

[1] This statement is referred to as the *Nested interval lemma*. This lemma is proven in most Calculus texts.

We will show that the Cantor set C is, quite surprisingly, an *uncountably* infinite set, this in spite of the large amount of points removed from [0, 1] to construct it. There are various ways of proving that C is uncountable. We provide a proof that has a set-theoretic flavour to it.

Proposition 25.3 The Cantor set has cardinality c.

Proof:

Since the Cantor set C is a subset of $[0,1]$, then the cardinality of C cannot be larger than c. We will show that $|\{0,1\}^{\mathbb{N}}| \leq |C|$ where the cardinality of $\{0,1\}^{\mathbb{N}}$ is known to be c (see theorem 20.12).

Let s be a specific sequence in $\{0,1\}^{\mathbb{N}}$. For each $n \in \mathbb{N}$, let $s(n)$ denote the finite sequence made of the first n terms (of zeroes and ones) of the infinite sequence, s. Then for each $n \in \mathbb{N}$, ${}_{n+1}I_{s(n+1)} \subset {}_nI_{s(n)}$ and so the set $\{{}_nI_{s(n)} : n \in \mathbb{N}\}$ forms a set of nested closed intervals (uniquely determined by the sequence s) obtained by applying the construction algorithm of C. The Nested interval lemma guarantees that $\cap_{n=0}^{\infty} {}_nI_{s(n)}$ is non-empty for each $s \in \{0,1\}^{\mathbb{N}}$. For each s, we can then choose an element x_s in $\cap_{n=0}^{\infty} {}_nI_{s(n)}$. (To do this we invoke the Axiom of choice.) See that, since $x_s \in {}_nI_{s(n)} \subset C_n$ for all n, then $x_s \in \cap_{n=0}^{\infty} C_n = C$. We define the function $f : \{0,1\}^{\mathbb{N}} \to C$ mapping *into* C, as $f(s) = x_s$. We claim that f is one-to-one: Suppose s and t are distinct elements of $\{0,1\}^{\mathbb{N}}$. Let n be the least natural number such that $s(n) \neq t(n)$. Then the two closed intervals

$$ {}_nI_{s(n)} \text{ and } {}_nI_{t(n)} $$

in C_n have empty intersection. Then the intersection of the two sets of nested closed intervals

$$ \{f(s)\} = \{x_s\} \subseteq \cap\{{}_nI_{s(n)} : n \in \mathbb{N}\} \text{ and } \cap\{{}_nI_{t(n)} : n \in \mathbb{N}\} \supseteq \{x_t\} = \{f(t)\} $$

must be empty. So x_s and x_t cannot the same element. This shows that f is one-to-one, as claimed. Then f embeds $\{0,1\}^{\mathbb{N}}$ into C. Hence, $c = |\{0,1\}^{\mathbb{N}}| \leq |C|$ as claimed. Since $C \subset \mathbb{R}$, $|C| \leq c$. We conclude that $|C| = c$.

It is surprising to see that the cardinality of C is the same as the cardinality of $\mathbb{R}$ since, to obtain C from $[0,1]$ we removed from [0, 1] a total length of open intervals equal to

$$ \frac{1}{3} + \frac{2}{3^2} + \frac{4}{3^3} + \cdots = \frac{1}{3}\left(1 + \frac{2}{3} + \frac{2^2}{3^2} + \cdots\right) = \frac{\frac{1}{3}}{1 - \frac{2}{3}} = 1 $$

There are still uncountably many points that are left behind. One may expect that C is simply the set of all endpoints that appear in all C_n's. But this can't be, since if we

take the union of all endpoints $\bigcup_{n\in\mathbb{N}} E_n$ we obtain only a countably infinite set, while C is uncountable. The Cantor set must then contain uncountably many numbers which are not endpoints! Skeptical readers may want to look at the proof again to see if there is any sleight of hand. Even if one believes the given proof, it does not meant that it will necessarily be what we might call "a satisfying proof". We cannot actually see what is going on at the very high levels of n. The proof doesn't help us understand why the "non-endpoints" in C are not excluded in the construction process.

Identifying numbers in C which are non-endpoints. The following arguments show why some "non-endpoints" of C remain in the infinite intersection of the sets, C_n, which are used to construct the Cantor set C. Consider the sequence of numbers, $\{S_n : n \in \mathbb{N}\}$, where

$$S_n = \sum_{k=0}^{n} \left(\frac{-1}{3}\right)^k$$

By carefully examining this sequence we can deduce the following facts:

- We see that $S_0 = 1$, $S_1 = 1 - \frac{1}{3}$, $S_2 = 1 - \frac{1}{3} + \frac{1}{3^2}$, $S_3 = 1 - \frac{1}{3} + \frac{1}{3^2} - \frac{1}{3^3}$, and so on.
- The subsequence $\{S_{2n} : n \in \mathbb{N}\}$ is a strictly decreasing sequence while the subsequence $\{S_{2n+1} : n \in \mathbb{N}\}$ is a strictly increasing sequence.
- We also see that, for all $n \in \mathbb{N}$, $[S_{2n+1}, S_{2n}] \subset [S_{2n-1}, S_{2n-2}]$, hence the set $\{[S_{2n+1}, S_{2n}] : n \in \mathbb{N}\}$ forms a nested set of closed intervals.
- For each n, $[S_{2n+1}, S_{2n}] \subset C_{2n+1}$.
- Since the elements of the sequence $\{S_n\}$ are partial sums of a geometric series with common ratio $r = \frac{-1}{3}$ then $S_n = \frac{1-r^{n+1}}{1-r} = \frac{1-(-1/3)^{n+1}}{1-(-1/3)}$. The limit of the sequence $\{S_n\}$ is then computed to be $\frac{3}{4}$ where, for all n, $S_{2n+1} < \frac{3}{4} < S_{2n}$.
- We deduce that

$$\{3/4\} = \cap\{[S_{2n+1}, S_{2n}] : n \in \mathbb{N}\} \subseteq \cap\{C_n : n \in \mathbb{N}\} = C$$

So, even though 3/4 is not an endpoint of one of the subintervals which form each level C_n it belongs to the Cantor set. Other such points can be found in C in this way.

25.4 Counting those elements of $\mathscr{P}(\mathbb{R})$ of cardinality c.

In what follows we will let $\mathscr{A}_c$ denote the set

$$\mathscr{A}_c = \{S \subseteq \mathbb{R} : |S| = |\mathbb{R}| = c\}$$

Since $\{\mathbb{R} - \{x\} : x \in \mathbb{R}\} \subset \mathscr{A}_c$ and $\mathscr{A}_c \subset \mathscr{P}(\mathbb{R})$, then $c \leq |\mathscr{A}_c| \leq 2^c$. We know that the cardinality of $\mathscr{P}(\mathbb{R})$ is $2^{|\mathbb{R}|} = 2^c$. We claim that $|\mathscr{A}_c| = 2^c$.

Recall (from corollary 20.4) that $|\mathbb{R} \times \mathbb{R}| = |\mathbb{R}| = c$. Then to each subset $S \subseteq \mathbb{R} \times \mathbb{R}$ of cardinality c we can associate a unique subset $S^* \subset \mathbb{R}$ of cardinality c. Then

$$|\mathscr{A}_c| = |\{S \subseteq \mathbb{R} \times \mathbb{R} : |S| = c\}|$$

For any non-empty subset U of $\mathbb{R}$ and $x \in U$,

$$c = |\{x\} \times \mathbb{R}| \leq |U \times \mathbb{R}| \Rightarrow |U \times \mathbb{R}| \geq c \; \forall \; U \subseteq \mathbb{R}$$

Since

$$|\{U \times \mathbb{R} : U \in \mathscr{P}(\mathbb{R}) - \{\varnothing\}\}| = 2^c$$

then

$$|\mathscr{A}_c| = |\{S \subseteq \mathbb{R} \times \mathbb{R} : |S| = c\}| \geq |\{U \times \mathbb{R} : U \in \mathscr{P}(\mathbb{R}) - \{\varnothing\}\}| = 2^c$$

So $|\mathscr{A}_c| = 2^c$, as required.

25.5 Counting all real-valued continuous functions on $\mathbb{R}$.

Continuous real-valued functions on $\mathbb{R}$ can be characterized as being those functions which satisfy the property

$$\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n) = f(x)$$

for any sequence $\{x_n\}$ of numbers which converges to a number x. Since every irrational number is the limit of a sequence of rational numbers, the value of a function f at an irrational number x is uniquely determined by the value of this function at all rational numbers which surround it. This means that given any continuous real-valued function, f, on $\mathbb{Q}$, the continuous function f^* on $\mathbb{R}$ such that $f^*|_{\mathbb{Q}} = f$ is unique. In set-theoretic language, this means that the sets, $\mathscr{B} = \{f \in \mathbb{R}^{\mathbb{R}} : \ f \text{ is continuous}\}$ and $\mathscr{D} = \{f \in \mathbb{R}^{\mathbb{Q}} : \ f \text{ is continuous}\}$, are equipotent. So to determine the cardinality of the set $\mathscr{B}$ it suffices to determine the cardinality of the set $\mathscr{D}$. It is shown in theorem 25.2 part a) that the cardinality of the set $\mathbb{R}^{\mathbb{N}}$ of all functions mapping $\mathbb{N}$ into $\mathbb{R}$ is c. From the equipotence relation $\mathbb{Q} \sim_e \mathbb{N}$ we deduce that $\mathbb{R}^{\mathbb{Q}} \sim_e \mathbb{R}^{\mathbb{N}}$. Since $\mathscr{D} \subset \mathbb{R}^{\mathbb{Q}}$, $|\mathscr{D}| \leq c$. The uncountable set $\{f \in \mathbb{R}^{\mathbb{Q}} : f = c, c \in \mathbb{R}\}$ of constant functions is a subset of the set $\mathscr{D}$. Then $c = |\{f \in \mathbb{R}^{\mathbb{Q}} : f = c, c \in \mathbb{R}\}| \leq |\mathscr{D}|$. So the cardinality of $\mathscr{D}$ is c. We conclude that the cardinality of the set of all real-valued continuous functions on $\mathbb{R}$ is c.

Concepts review:

1. What is the cardinality of $\mathbb{R}^n$ for any natural number n?

2. Do the real numbers and the complex numbers have the same cardinality?
3. How does the cardinality of the set of all countably infinite sequences of real numbers compare with the cardinality of $\mathbb{R}^{\mathbb{R}}$?
4. What is the cardinality of the set of all irrational numbers?
5. What is the cardinality of the set of all countably infinite sequences of natural numbers?
6. What is the cardinality of the set of all countably infinite sequences of real numbers?
7. Let $\mathbb{N}^{\mathbb{N}}$ denote the set of all functions mapping $\mathbb{N}$ into $\mathbb{N}$ and $\mathbb{N}^{\mathbb{N}}_{1-1}$ denote the set of one-to-one functions mapping $\mathbb{N}$ into $\mathbb{N}$. Are the sets $\mathbb{N}^{\mathbb{N}}$ and $\mathbb{N}^{\mathbb{N}}_{1-1}$ equipotent? What is their cardinality?
8. What is the Cantor set? How is it constructed? What is its cardinality?
9. What is the cardinality of the set of all continuous real-valued functions?

EXERCISES

A. 1. Show that for any finite cardinal n, $n \times 2^{(2^{\aleph_0})} = 2^{(2^{\aleph_0})}$.

B. 2. Let $\mathscr{P}(\mathbb{N})_F$ denote the set of all finite subsets of $\mathbb{N}$.
   a) Show that the set $\mathbb{N}$ is embedded in $\mathscr{P}(\mathbb{N})_F$.
   b) Express $\mathscr{P}(\mathbb{N})_F$ as the union of a countably infinite number of pairwise disjoint subsets of $\mathbb{N}$.
   c) What is the cardinality of the set $\cup_{k\in\mathbb{N}}\mathbb{N}^k$?
   d) Construct a one-to-one function which embeds $\mathscr{P}(\mathbb{N})_F$ in $\cup_{k\in\mathbb{N}}\mathbb{N}^k$.
   e) What is the cardinality of $\mathscr{P}(\mathbb{N})_F$?

   3. Let $\mathscr{P}(\mathbb{R})_F$ denote the set of all finite subsets of $\mathbb{R}$.
   a) Show that the set $\mathbb{R}$ is embedded in $\mathscr{P}(\mathbb{R})_F$.
   b) Express $\mathscr{P}(\mathbb{R})_F$ as the union of a countably infinite number of pairwise disjoint subsets of $\mathbb{R}$.
   c) What is the cardinality of the set $\cup_{k\in\mathbb{N}}\mathbb{R}^k$?
   d) Construct a one-to-one function which embeds $\mathscr{P}(\mathbb{R})_F$ in $\cup_{k\in\mathbb{R}}\mathbb{R}^k$.
   e) What is the cardinality of $\mathscr{P}(\mathbb{R})_F$?

C. 4. Consider the strictly increasing sequence $S = \{2^1, 2^2, 2^3, 2^4, \ldots, 2^n, \ldots\}$ of cardinal numbers.

a) Does the cardinal number $2^{\aleph_0}$ belong to S? Why?

b) Consider the two infinite cardinal numbers $\aleph_0$ and $2^{\aleph_0}$. Is one of these a least upper bound of S?[1] If not say why. If so which one?

---

[1]We remind the reader of the definition of "a least upper bound of an ordered set S": The element m is a *least upper bound* of a set S if m is an upper bound of S and for any other upper bound n, $m \leq n$.

Part VIII

Ordinal numbers

26 / Well-ordered sets.

Summary. *In this section we review a few basic notions about well-ordered sets. We define special subsets of well-ordered sets called "initial segments". "Order isomorphisms" are defined as strictly increasing one-to-one functions between well-ordered sets. Initial segments of well-ordered sets are themselves well-ordered sets. We prove basic properties of order isomorphisms between initial segments and well-ordered sets. We then define the relation, $\leq_{WO}$, on well-ordered sets as follows: "$S \leq_{WO} T$ if S and T are order isomorphic or one is order isomorphic to an initial segment of the other". This section provides the fundamental background for the study of the important set-theoretic topic of "ordinal numbers".*

26.1 Overview.

In the last few sections, we have familiarized ourselves with some of the main properties of infinite sets. We have seen that the *ZFC*-axioms have cleared a path into unfamiliar mathematical territory, populated by uncountably many "infinite sets" in infinite varieties, leading us to reflect on numerous counterintuitive notions. We have discovered, for example, that given any infinite set A we can find another infinite set $B = \mathscr{P}(A)$, not equipotent to A, which properly contains a one-to-one copy of A. To express this relationship we said that A is "properly embedded" in B and wrote $A \hookrightarrow_e B$. We can thus construct infinite chains of sets linearly ordered by the proper embedding $\hookrightarrow_e$-relation. For example:

$$0 \hookrightarrow_e 1 \hookrightarrow_e 2 \hookrightarrow_e \cdots \hookrightarrow_e \mathbb{N} \hookrightarrow_e \mathscr{P}(\mathbb{N}) \hookrightarrow_e \mathscr{P}(\mathscr{P}(\mathbb{N})) \hookrightarrow_e \mathscr{P}(\mathscr{P}(\mathscr{P}(\mathbb{N}))) \hookrightarrow_e \cdots$$

This chain of sets ordered by $\hookrightarrow_e$ begins with the empty set $0 = \{\ \}$. This set is followed by an infinite number of finite sets called the "natural numbers". Once we attain the first infinite set $\mathbb{N}$, an endless sequence of infinite sets can be constructed by successively taking powers of a set. Note that no natural number is constructed by taking the power set of its immediate predecessor. So the method for constructing each natural number from its predecessor is different from the method used to construct each new infinite set. In fact, it is an axiom that allows $\mathbb{N}$ to exist. Another axiom allows us to call the power of a set, a "set". Of course, various chains of sets can be constructed in this way, each depending on the choice of the first set. If we started with the set of all real numbers, $\mathbb{R}$, we then obtained what *initially* appeared to be a different chain of infinite sets, $\mathbb{R} \hookrightarrow_e \mathscr{P}(\mathbb{R}) \hookrightarrow_e \mathscr{P}(\mathscr{P}(\mathbb{R})) \cdots$. It was then determined that $\mathbb{R}$ and $\mathscr{P}(\mathbb{N})$ are in fact equipotent and so the displayed chain containing power sets of $\mathbb{N}$ contains copies of the $\mathbb{R}$-related power sets.

We thought it would be practical to partition the class of all sets into subclasses of mutually equipotent sets. These subclasses were seen to be equivalence classes induced by the equipotence relation $\sim_e$. We defined the notion of $\sim_e$-equivalence class representatives called *cardinal numbers*. A cardinal number was declared to be a set which represents all sets which are equipotent to it. We had to postulate the existence of the cardinal numbers with the promise that once we have developed the required set-theoretic tools, the cardinal numbers would be appropriately defined or constructed.

We have seen that the set of all natural numbers has been extremely useful in determining various properties of countably infinite sets. A critically important tool in our study was the principle of mathematical induction over $\mathbb{N}$. Since any countably infinite set is a one-to-one image of $\mathbb{N}$, this means that the elements of such sets can be indexed by the elements of $\mathbb{N}$. Indexing countable sets in this way allows us to linearly order these sets. We can then apply the principle of mathematical induction to determine some of their properties. When working with uncountable sets, we do not yet have access to uncountable well-ordered sets whose elements can be used to index such sets. We will soon see that ZFC provides the necessary ingredients to construct "universal indexing sets". [1]

26.2 Well-ordered sets revisited.

Recall that "order relations" on a set S are relations which fall into two major categories: *linearly ordered* relations or *partially ordered* relations (also, *non-linearly ordered* relation). Each of these can be strict or non-strict order relations. Non-strict ordered relations are often represented by "$\leq$", while strictly ordered relations are often represented by "$<$" although other symbols may be used. An order relation on a set, S, is linear provided provided any two distinct elements of S are comparable under the given order relation. That is, all elements of S can be lined up on a line, the "larger" elements normally to the right of (or above) the "smaller" ones. Partially ordered classes are often viewed as having many branches where elements on different branches are not comparable by $\leq$ or $<$. It is often said that non-linear order relations have many "chains of elements" (subsets which are linearly ordered) while a linearly ordered class has all its elements lined up in a single chain.

In what follows the hypothesized sets, S, will be linearly ordered by $\leq$ or $<$. We will be investigating those linearly ordered sets which are "well-ordered". We remind ourselves of what "well-ordered" means:

> A *well-ordered* set is a set, S, which is linearly ordered by $\leq$ or $<$ in such a way that every non-empty subset, T, of S contains its least element. If a

[1] These sets will be called *ordinals* (soon to be defined). Cardinal numbers will be defined as being those ordinals which satisfy a specific property. Until we formally define "cardinal numbers" we will refrain from referring to the notion of "cardinality of a set" in the process that leads to this definition.

relation $\leq$ well-orders a class or a set S we will sometimes, more succinctly, say that "S is $\leq$-well-ordered".

Well-ordered classes. Note that in the above definition of "well-ordered set", we can replace the word "set" with the word "class", so that we can speak of an ordered proper class, A, which is well-ordered by some relation, $\leq$. Proper well-ordered classes will be discussed further on in the text.

Before we start we should recall that the set, $\mathbb{N}$, as well as every natural number, n, were shown to be $\in$-well-ordered, (see theorem 14.3 and corollary 14.4). The set, $\mathbb{N}$, of all natural numbers and any natural, n, are also $\subset$-well-ordered. We will invoke this fact to show that given *any* non-empty countable set S, we can define an order relation which well-orders S.

Theorem 26.1 Let $f : T \to S$ be a one-to-one function mapping T onto S. If T is a well-ordered set, then T induces a well-ordering on S. Hence, every countable set can be well-ordered.

Proof:

What we are given: There exists a function $f : T \to S$ which maps the well-ordered set $(T, <_T)$ one-to-one onto the set S.
What we are required to show: There exists an order relation which well-orders the set S.

Since $f : T \to S$ maps T one-to-one onto the set S, we can then index the elements of S as follows: If $s = f(n)$, express s as s_n. Then $S = \{s_n : n \in T\} = f[T]$.
We define the relation "$<_S$" as

$$s_n <_S s_m \text{ if and only if } n <_T m$$

We claim that $<_S$ well-orders S:

– The set S is $<_S$-linearly ordered: It is clear that since f is one-to-one onto, $<_S$ is irreflexive and asymmetric. For transitivity, we see that

$$\begin{aligned} s_n <_S s_m \text{ and } s_m <_S s_r &\Rightarrow n <_T m \text{ and } m <_T r \\ &\Rightarrow n <_T r \\ &\Rightarrow s_n <_S s_r \end{aligned}$$

We now verify that every pair of elements in S are comparable under $<_S$. If $s_n, s_m \in S$, then n and m are the unique corresponding elements in T. Then $n <_T m$ or $m <_T n$. Hence, either $s_n <_S s_m$ or $s_m <_S s_n$. Hence, all pairs of elements of S are $<_S$-comparable and so S is $<_S$-linearly ordered.

– The set S is $<_S$-well ordered: Suppose $A = \{s_i : i \in U \subseteq T\}$ is a non-empty subset of S. Then U is a non-empty subset of T. Since T is well-ordered, U has a least element, say k. Since $k \leq_T i$ for all $i \in U$, then $s_k \leq_S s_i$ for all $s_i \in A$. Thus, A contains a least element.

This proves that the relation, $<_S$, induced on S by T is a well-ordering.

We now show that every non-empty countable set can be well-ordered. Let S be a countably infinite set. Then there exists a function, $f : \mathbb{N} \to S$, mapping $\mathbb{N}$ one-to-one onto S. Since $\mathbb{N}$ is well-ordered, then S has a well-ordering.
If S is finite and non-empty, then it is the one-to-one image of some natural number n (18.7). Since every natural number n is $\in$-well-ordered (14.4), S inherits this well-ordering from n as described above.

We provide a few examples of linearly ordered sets which are well-ordered and some that are not (some of which we have seen before).

a) The set of all even natural numbers, $\mathbb{N}_e$, with the ordering inherited from $(\mathbb{N}, \subset)$ is a well-ordered set since every pair of even numbers are comparable and every subset of even numbers contains a least even number.

b) Every natural number, n, is a well-ordered set. For example, $5 = \{0, 1, 2, 3, 4\}$ is $\in$-linear (or $\subset$-linear) and every subset of 5 contains a least element.

c) The set of all countably infinite sequences of natural numbers, $\mathbb{N}^{\mathbb{N}}$, equipped with the lexicographic ordering[1] has been shown to be a set which is linearly ordered, but not well-ordered, since it contains subsets with no least element. For example, suppose that for each $i \in \mathbb{N}$, $\mathbf{x}_i = \{a_j : j \in \mathbb{N}\}$ where $a_j = 1$ if $j = i$ and $a_j = 0$ otherwise. Then for each $i \in \mathbb{N}$, $\mathbf{x}_i \in \mathbb{N}^{\mathbb{N}}$. The subset $S = \{\mathbf{x}_i : i \in \mathbb{N}\}$ of $\mathbb{N}^{\mathbb{N}}$ does not contain a least element since it does not contain the element $(0, 0, 0, \ldots,)$.

d) The set, $\mathbb{N} \times \mathbb{N}$, can also be equipped with the lexicographic ordering:

$$\{(0,0), (0,1), (0,2), \ldots, (1,0), (1,1), \ldots, (2,0), (2,1), (2,2), (2,3), \ldots\}$$

When ordered in this way, $\mathbb{N} \times \mathbb{N}$ can be seen as being the union of a countably infinite number of copies of $\mathbb{N}$ lined up from end to end. This is easily seen to be a linear ordering. Given any non-empty subset $M = \{(s,t) : s \in S, t \in T\}$ of $\mathbb{N} \times \mathbb{N}$, the least element of M is (u, v) where $u = \text{least}\{S\}$ and $v = \text{least}\{t : (u,t) \in M\}$. The element, (u, v), belongs to M since both u and v are least elements of subsets of a well-ordered set. We conclude that the lexicographically ordered $\mathbb{N} \times \mathbb{N}$ is well-ordered.

e) The set, $\mathbb{R}$, of real numbers equipped with the usual real number ordering is linear but is not well-ordered since the set $\{x \in \mathbb{R} : x > 1\}$ does not have a least

[1] See the definition of lexicographic ordering on page 130.

element. Note that this does not mean that there isn't an order relation which well-orders the real numbers.[2]

Our experience with well-ordered sets is quite limited. Even if we can use the lexicographic ordering tool to construct long chains of copies of $\mathbb{N}$ we have however not been able to exhibit a single uncountably infinite set which is well-ordered. Anyone who has attempted to find a well-ordering relation for $\mathbb{R}$ may wonder if an uncountable well-ordered set exists at all.

26.3 Initial segments revisited.

The "Initial segments of a set" is another concept we will be revisiting at this time. The reader will recall that the notion of an "initial segment" was introduced in 17.1 as special kind of subset of $\mathbb{Q}$. In that section, we referred to it as a *Dedekind cut*. The set of all Dedekind cuts became the "*real numbers*". Initial segments discussed here are the same mathematical objects as the ones discussed in the chapter whose purpose was to define the real numbers $\mathbb{R}$. However, the context is quite different. In this section the initial segments we will study are subsets of abstract well-ordered sets. We will discuss the notion of "initial segments" as though we have never seen these before. We start with the following formal definition.

Definition 26.2 Given a *well-ordered* set $(S, \leq)$, a subset U of S satisfying the property

$$U \neq S \text{ and } \forall u \in U, [x < u] \Rightarrow [x \in U]$$

is called an *initial segment* of S. In this definition, the partial order relation $\leq$ can be used instead of $<$ without altering the meaning of "initial segment".

Formal definitions of abstract concepts are often not expressed in a reader-friendly form. This is because the reader-friendly form is not always the form that is best adapted to the process of proving statements in which carefully formulated definitions are required. The following theorem will allow the reader to more easily visualize what initial segments in well-ordered sets look like.

Theorem 26.3 If $(S, \leq)$ is a well-ordered set, then every initial segment in S is of the form

$$S_a = \{x \in S : x < a\}$$

for some $a \in S$.

[2] We will see later on that the Axiom of choice guarantees that $\mathbb{R}$ can be well-ordered without explicitly stating what such a well-ordering could be.

Proof:

What we are given: That $(S, \leq)$ is a well-ordered set; T is a *proper subset* of S satisfying the property "$\forall t \in T, [x < t] \Rightarrow [x \in T]$".
What we are required to show: That $T = S_a = \{x \in S : x < a\}$ for some $a \in S$.
Since T is a proper subset of S, then $S - T$ is non-empty. So $S - T$ must contain its least element, say a (since S is well-ordered).
Claim $S_a \subseteq T$: Since a is the least element of $S - T$, $x < a \Rightarrow x \notin S - T \Rightarrow x \in T$. So $S_a \subseteq T$, as claimed.
Claim $T \subseteq S_a$: If $x \notin S_a$ then $x \geq a$. Then the element x cannot belong to T, for if $x \in T$, $a \leq x$ would imply that $a \in T$ (by definition of the set T); since $a \in S - T$, we would obtain a contradiction. So $u \in T \Rightarrow u < a$. That is, $T \subseteq S_a$ as claimed.
So the initial segment, T, of the well-ordered set, S, is the set $S_a = \{x \in S : x < a\}$ where a is the least element in $S - T$, as required.

Given the initial segment S_a, we will refer to a as the *leader* of the initial segment. The leader, a, of the initial segment, S_a, is not an element of the initial segment. It is also important to remember that, by definition, *a well-ordered set S is not an initial segment of itself.* We provide a few examples of sets which are initial segments and sets which are not:

a) Note that every natural number n in $\mathbb{N}$ is an initial segment of $\mathbb{N}$. For example, $5 = \{0, 1, 2, 3, 4\} = \{n \in \mathbb{N} : n < 5\} = S_5$ is an initial segment of $\mathbb{N}$.
   – We can view $5 = S_5 = \{0, 1, 2, 3, 4\}$ as being $\subset$-well-ordered. The initial segments of 5 are the following sets only:

$$\begin{aligned} 4 &= S_4 = \{0, 1, 2, 3\} \\ 3 &= S_3 = \{0, 1, 2\} \\ 2 &= S_2 = \{0, 1\} \\ 1 &= S_1 = \{0\} \end{aligned}$$

b) Even though the set $\mathbb{N}_e$ of all even natural numbers is a proper subset of $\mathbb{N}$ it is not an initial segment of $\mathbb{N}$ since $26 \in \mathbb{N}_e$ and $17 < 26$ but $17 \notin \mathbb{N}_e$. However, $S_{26} = \{n \in \mathbb{N}_e : n < 26\}$ is an initial segment of the well-ordered set $(\mathbb{N}_e \subset)$.
c) The subset $S = \{0, 2, 3, 4, 5, \ldots, \}$ in $\mathbb{N}$ is not an initial segment of $\mathbb{N}$ since $3 \in S$ and $1 < 3$ but 1 does not belong to S.
d) Consider the set, $\mathbb{N}^{\{0,1,2\}} = \{\{a_0, a_1, a_2\} : a_i \in \mathbb{N}\}$, of all functions mapping $\{0, 1, 2\}$ into $\mathbb{N}$, ordered lexicographically.[1] This set is easily verified to be well-ordered.[2]

[1] Note that this set is the set of all ordered triples of natural numbers and so is equivalent to the Cartesian product $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$.

[2] If S is a non-empty subset of $\mathbb{N}^{\{0,1,2\}}$ let (b_0, b_1, b_2) be the element in S such that b_0, b_1, and b_2 is the

The set $S_{\{0,1,0\}} = \{\{0,0,i\} : i \in \mathbb{N}\}$ is an initial segment of $\mathbb{N}^{\{0,1,2\}}$. It is the set of all elements in $\mathbb{N}^{\{0,1,2\}}$ which are strictly less than $\{0,1,0\}$.

Initial segments of well-ordered sets are well-ordered. If S_a is an initial segment of a $<$-well-ordered set S, then S_a can inherit the order relation "$<$" from S so that it can itself be viewed as a $<$-well-ordered set.

26.4 "Order isomorphisms" between well-ordered classes.

Given two sets A and B there can be many functions mapping A into B. We may want to classify these functions by "types". For example, we may want to consider only those functions $f : A \to B$ which are one-to-one, or only those functions which are constant, or only those with finite range, and so on. If we are given two linearly ordered sets $(S, \leq_S)$ and $(T, \leq_T)$ we may be interested only in those functions $f : S \to T$ which "respect the order" of these functions. By "respecting the order" we mean that $n \leq_S m \Rightarrow f(n) \leq_T f(m)$. For example, the function $f : \mathbb{N} \to \mathbb{N}$ defined as $f(n) = 5n$ respects the order of the elements of the set $\mathbb{N}$ (for example, $3 < 4$ where $f(3) = 15 < 20 = f(4)$) while the function, $g : (0,1] \to \mathbb{N}$, defined as, $g(x) = 1/x$, does not (since $1/3 < 1/2$ and yet $g(1/3) = 3 \not< 2 = g(1/2)$). One-to-one order-respecting functions between *well-ordered sets* will be called *order isomorphism*. We begin by formally defining this concept.

Definition 26.4 Let $f : (S, \leq_S) \to (T, \leq_T)$ be a function mapping a well-ordered class, $(S, \leq_S)$, onto a well-ordered class, $(T, \leq_T)$. Note that the symbols $\leq_S$ and $\leq_T$ will allow us to distinguish between the order relations applied to the sets S and T, respectively.

a) We will say that the function, f, is *increasing* on $(S, \leq_S)$ if

$$(x \leq_S y) \Rightarrow (f(x) \leq_T f(y))$$

b) We will say that the function f is *strictly increasing* on $(S, \leq_S)$ if

$$(x <_S y) \Rightarrow (f(x) <_T f(y))$$

A strictly increasing function must be one-to-one.

c) If $f : (S, \leq_S) \to (T, \leq_T)$ is strictly increasing, then f is said to be an *order isomorphism* mapping S into T.

---

least element of all first, second and third coordinates of elements in S respectively. Let $(x,y,z) \in S$. If $b_0 < x$, then $(b_0, b_1, b_2) < (x,y,z)$; if $b_0 = x$ and $b_1 < y$, again, $(b_0, b_1, b_2) < (x,y,z)$; if $b_0 = x$ and $b_1 = y$ since $b_2 \leq y$, then $(b_0, b_1, b_2) \leq (x,y,z)$. So (b_0, b_1, b_2) is the least element of S.

If there exists an *onto* order isomorphism between the two well-ordered classes, $(S, \leq_S)$ and $(T, \leq_T)$, we will say that the classes are *order isomorphic*, or that a function maps S *order isomorphically* onto T.

We provide a few examples of order isomorphisms between well-ordered sets introduced in previous chapters.

- Let $(\mathbb{N}_e, \leq)$ denote the even natural numbers equipped with the standard natural number ordering $\leq$. Since the function $f : \mathbb{N} \to \mathbb{N}_e$ defined as $f(n) = 2n$ is one-to-one and strictly increasing, then it maps $\mathbb{N}$ order isomorphically onto $\mathbb{N}_e$.
- On the other hand the function $g : (\mathbb{N}, \leq) \to (\mathbb{N}, \leq)$ defined as $g(n) = n + (-1)^n$ is one-to-one and onto $(\mathbb{N}, \leq)$ but is *not* an order isomorphism. If $g(n) = a_n$, witness $a_0 = 1, a_1 = 0, a_2 = 3, a_3 = 2, \ldots$. We see that g does not respect the order of the elements.
- Consider the set $\mathbb{N} \times \mathbb{N} \times \mathbb{N} = \{(a_0, a_1, a_2) : a_i \in \mathbb{N}\}$ ordered lexicographically. We see that the set
$$S_{(0,1,0)} = \{(0, 0, i) : i \in \mathbb{N}\}$$
is an initial segment of $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$ since $(0, 0, i) < (0, 1, 0)$ for all natural numbers i. Verify that the function $f : \mathbb{N} \to \mathbb{N} \times \mathbb{N} \times \mathbb{N}$ defined as $f(n) = (0, 0, n)$ maps $\mathbb{N}$ order isomorphically onto $S_{(0,1,0)}$.
Does there exist some other order isomorphism which maps $\mathbb{N}$ onto $S_{(0,1,0)}$? (A statement proven in the proposition below will confirm that there can be no other.)
- Suppose $<^*$ orders the elements of $\mathbb{N}$ as follows:
  - If n is even and m is odd, then $n <^* m$.
  - If n and m are both even or both odd, then n and m respect the usual order of natural numbers. That is,
$$(\mathbb{N} <^*) = \{0, 2, 4, 6, \ldots, 1, 3, 5, 7, \ldots\}$$

Then the function $f : \mathbb{N} \to (\mathbb{N}, <^*)$ defined as $f(n) = 2n$, maps $\mathbb{N}$ order-isomorphically onto the initial segment $\{0, 2, 4, 6, \ldots, \}$ of $(\mathbb{N}, <^*)$.

26.5 Basic properties of order isomorphisms.

We now list and prove a few basic properties of order isomorphisms. We will refer to these often in our study of those sets we will call *ordinals*.

Proposition 26.5 Let $(S, \leq_S)$ and $(T, \leq_T)$ be a well-ordered sets.

a) The inverse of an order isomorphism is an order isomorphism.

b) If $f : (S, \leq_S) \to (S, \leq_S)$ is a *strictly increasing* function mapping S into itself, then $x \leq f(x)$, for all $x \in S$.

c) The set S cannot be order isomorphic to an initial segment of itself.

d) If $f : (S, \leq_S) \to (S, \leq_S)$ is an order isomorphism mapping S onto itself, then f is the identity function.[1]

e) If $f : (S, \leq_S) \to (T, \leq_T)$ and $g : (S, \leq_S) \to (T, \leq_T)$ are two order isomorphisms mapping S *onto* T, then $f = g$.

f) Suppose $f : (S, \leq_S) \to (T, \leq_T)$ is an order isomorphism mapping S onto an initial segment of T. Then S and T cannot be order isomorphic.

Proof:

a) Given: That $f : S \to T$ is an order isomorphism mapping the well-ordered set $(S, \leq_S)$ onto $(T, \leq_T)$.
What we are required to show: That $f^{-1} : T \to S$ must also be an order isomorphism:

To see this, let u, v be elements in T such that $u <_T v$. Since f is one-to-one and onto, there exists distinct elements $a = f^{-1}(u)$ and $b = f^{-1}(v)$ in S. Since S is well-ordered, it is linear and so all elements in S are comparable. So either $f^{-1}(u) = a <_S b = f^{-1}(v)$ or $f^{-1}(v) = b <_S a = f^{-1}(u)$. If $b <_S a$, then, since f is order preserving, $f(b) = v <_T u = f(a)$, a contradiction. So $a = f^{-1}(u) <_S f^{-1}(v) = b$. So $f^{-1} : T \to S$ must also be an order isomorphism.

b) What we are given: That $(S, \leq)$ is well-ordered, that $f : S \to S$, and that $x < y$ implies $f(x) < f(y)$ (that is, f is strictly increasing).
What we are required to prove: That $x \leq f(x)$ for all x. That is, f cannot map an element x "below itself".

Suppose there exists an element x of S such that $f(x) < x$. Then, the set
$T = \{x \in S : f(x) < x\}$ is non-empty. We claim that this will lead to a contradiction:

- Since S is well-ordered, T must contain a least element, say a. Since $a \in T$, $f(a) < a$.
- Since f is strictly increasing,

$$f(a) < a \Rightarrow f(f(a)) < f(a)$$

- By definition of T, $f(f(a)) < f(a)$ implies $f(a) \in T$. Since a is the least element of T, $a \leq f(a)$. But $a \leq f(a)$ and $f(a) < a$ are contradictory statements. The source of this contradiction is supposing that $T \neq \varnothing$.

We must conclude that $T = \varnothing$. That is, for all $x \in S$, $x \leq f(x)$.

[1] An order isomorphism from an ordered set onto itself is called an *order automorphism*. Here we are stating that the only automorphism is the identity function.

c) What we are given: That $(S, \leq)$ is a well-ordered set.
What we are required to show: That S cannot be order isomorphic to an initial segment of itself.

An initial segment of S must be of the form $S_a = \{x \in S : x < a\}$ where $a \in S$. If $f : (S, \leq) \to S_a$ is an order isomorphism onto S_a, then f must map a to some element $f(a)$ in S_a; this means $f(a) < a$. By definition of order isomorphism, f is strictly increasing on S and so $f(f(a)) < f(a)$. But, by part b) above, $x \leq f(x)$ for all $x \in S$ and so $f(a) \leq f(f(a))$. The statements $f(f(a)) < f(a)$ and $f(a) \leq f(f(a))$ are contradictory. So no initial segment of S can be the order isomorphic image of S.

d) What we are given: That $f : (S, \leq) \to (S, \leq)$ is an onto order isomorphism.
What we are required to show: That $f(x) = x$ for all $x \in S$.
If f is an order isomorphism from S onto itself, then both f and f^{-1} must be strictly increasing functions. By part b) above, $x \leq f(x)$ and $x \leq f^{-1}(x)$, for all $x \in S$. Suppose $s < f(s)$ for some $s \in S$. Then

$$f^{-1}(s) < f^{-1}(f(s)) = s$$

The statements $f^{-1}(s) < s$ and $s \leq f^{-1}(s)$ are contradictory. So there can be no element $s \in S$ such that $s < f(s)$. Then $x \leq f(x)$ and $x \not< f(x)$ forces $f(x) = x$ for all $x \in S$.

e) Suppose $f : (S, \leq_S) \to (T, \leq_T)$ and $g : (S, \leq_S) \to (T, \leq_T)$ are two order isomorphisms mapping S onto T. Then $f^{-1} : T \to S$ is an order isomorphism and so the function

$$f^{-1} \circ g : (S, \leq_S) \to (S, \leq_S)$$

is an order isomorphism of S onto itself. By part d) $f^{-1} \circ g$ must be the identity map. Then $g = (f^{-1})^{-1} = f$.

f) What we are given: $f : (S, \leq_S) \to (T, \leq_T)$ is an order isomorphism mapping S onto an initial segment T_u of T.
What we are required to show: That S and T cannot be order isomorphic.
If S and T_u are order isomorphic and S and T are order isomorphic, then T is order isomorphic to T_u contradicting the statement in part c). So S and T cannot be order isomorphic.

We highlight some important points that are made in the above statements.

Firstly, if $f : S \to T$ is an order isomorphism between well-ordered sets S and T, then there can be no other one. This is important, since it points to a crucial difference between equipotent sets and order isomorphic sets. There can be many different functions which map a set S one-to-one onto a set T. But if $(S, <_S)$ and $(T, <_T)$ are known to be order isomorphic, then only one order isomorphism can bear witness to

this fact.[1] We might say that an order-isomorphism is "sensitive" to the structure of a well-ordered set, while equipotence is not. For example, the equipotence relation perceives $(\{0,1\} \times \mathbb{N}, <_{\text{lex}})$ simply as a countable set allowing for many ways of mapping $\mathbb{N}$ one-to-one and onto this set, while an order isomorphism is sensitive to the fact that this set is made of two copies of $\mathbb{N}$ lined up one after the other and so cannot view this set as a single copy of $\mathbb{N}$.

Secondly, two initial segments of the same well-ordered set are order isomorphic only if they are equal.

Thirdly, a well-ordered set can never be order isomorphic to an initial segment of itself. This again underlines an important difference with the equipotence relation. By definition, an infinite set is precisely a set which is equipotent with a proper subset of itself.

26.6 Ranking well-ordered sets with order isomorphisms.

We will show how order isomorphisms can be used to "rank" well-ordered sets. We begin by introducing the following notation.

Notation 26.6 Let S and T be two well-ordered sets. Then the expression

$$S \sim_{\text{WO}} T$$

means "S and T are order isomorphic". The expression

$$S <_{\text{WO}} T$$

means "$S \sim_{\text{WO}} T_a$" where T_a is some *initial segment* of T. The expression

$$S \leq_{\text{WO}} T$$

means "$S \sim_{\text{WO}} T$ or $S <_{\text{WO}} T$".

If $\mathscr{W} = \{S \in \mathscr{S} : S \text{ is well-ordered}\}$ denotes the class of all well-ordered sets, see that the relation $\sim_{\text{WO}}$ is reflexive, symmetric and transitive on $\mathscr{W}$ and so is an equivalence relation. See that $\leq_{\text{WO}}$ is also reflexive and transitive on $\mathscr{W}$. The relation $\leq_{\text{WO}}$ is not antisymmetric on $\mathscr{W}$ in the usual sense, since $S \leq_{\text{WO}} T$ and $T \leq_{\text{WO}} S$ implies

[1] Note that even if the order isomorphism between two initial segments is unique, it is still entirely possible for an initial segment to be mapped order-isomorphically onto another subset of a well-ordered set. For example, the initial segment $\{0,1,2,3\}$ can be mapped order-isomorphically to the non-initial-segment $A = \{11,12,13,14\}$. But note that A is not an initial segment of $\mathbb{N}$.

$S \sim_{\text{WO}} T$, not $S = T$.[1] But $\leq_{\text{WO}}$ can always be used as a ranking too for the elements of $\mathscr{W}$. We will show that any two well-ordered sets are $\leq_{\text{WO}}$-comparable. That is, given any two well-ordered sets, S and T, either $S \leq_{\text{WO}} T$ or $T \leq_{\text{WO}} S$. The reader should carefully note how the "well-order properties" are used in various parts of the proof.

Theorem 26.7 Let $(S, \leq_S)$ and $(T, \leq_T)$ be two well-ordered sets. Then either $S \leq_{\text{WO}} T$ or $T \leq_{\text{WO}} S$.

Proof:

What we are given: Two well-ordered sets $(S, \leq_S)$ and $(T, \leq_T)$. The expression S_a represents the initial segment $\{x \in S : x < a\}$ whose leader is a.
We are required to show that: $S \leq_{\text{WO}} T$ or $T \leq_{\text{WO}} S$

The symbol $S_a \sim_{\text{WO}} T_b$ is to be interpreted as "the initial segment S_a of S is order isomorphic to the initial segment T_b of T".

We define the function $f : S \to T$ as follows: $f(a) = b$ if and only if $S_a \sim_{\text{WO}} T_b$. We will carefully examine this function and describe its properties.

- *We verify that f is well-defined*: If $f(a) = b$ and $f(a) = c$, then $S_a \sim_{\text{WO}} T_b$ and $S_a \sim_{\text{WO}} T_c$. This implies $T_b \sim_{\text{WO}} T_c$. Two initial segments of the same well-ordered set are order isomorphic if and only if they are equal. Then $b = c$.
- *We verify that the domain of f is non-empty*: Suppose 0_S and 0_T denote the least elements of S and T respectively. If 1_S and 1_T denote the least element in $S - \{0_S\}$ and $T - \{0_T\}$, respectively, then $S_{1_S} \sim_{\text{WO}} T_{1_T}$. Then $f(1_S) = 1_T$ so the domain of f contains at least the element 1_S. Let D denote the domain of f.
- *We verify that f is strictly increasing*: Suppose a and b are in the domain D of f such that $a <_S b$. If $f(a) = c$ and $f(b) = d$, then $S_a \sim_{\text{WO}} T_c$ and $S_b \sim_{\text{WO}} T_d$. Then $T_c \sim_{\text{WO}} S_a \subset S_b \sim_{\text{WO}} T_d$. So T_c is order isomorphic to an initial segment of T_d. This implies $T_c \subset T_d \Rightarrow c < d$. So f is strictly increasing on D.

If $D = S$, then f maps S order isomorphically into T (since f has been shown to be strictly increasing on D).
It follows that $D = S \leq_{\text{WO}} T$, and we are done. So let us suppose that $D \neq S$.

- *We claim that the domain D is an initial segment of S*: If $u \in D$, then $S_u \sim_{\text{WO}} T_k$ for some $k \in T$. That is, there exist an order isomorphism $g : S_u \to T_k$. If $x <_S u$, then $S_x \subset S_u$ and $g|_{S_x} : S_x \to T_k$ maps S_x onto an initial segment, say T_t, in T_k. Then $S_x \sim_{\text{WO}} T_t$ implies $f(x) = t$. Then $x \in D$. So D is an initial segment of S as claimed.

[1] However, if $\mathscr{W}^* = \{[S]_{\text{WO}} : S \in \mathscr{W}\}$ denotes the class of all equivalence classes induced by the equivalence relation, $\sim_{\text{WO}}$, on $\mathscr{W}$, then the statement in the theorem will allow us to conclude that $\leq_{\text{WO}}$ induces a linear ordering on $\mathscr{W}^*$.

– *We claim that if $f[D] \neq T$, then $f[D]$ is an initial segment of T* : Let $v \in f[D]$. Then there exists an element $a \in D$ such that $f(a) = v$. This implies that $S_a \sim_{\text{WO}} T_v$. Let $u < v$ in T. Then $T_u \subset T_v$. Since $S_a \sim_{\text{WO}} T_v$, then T_u is order isomorphic to an initial segment $S_b \subset S_a$, for some $b \in D$. Then $f(b) = u$. So $u \in f[D]$. Since $f[D] \neq T$, by definition, $f[D]$ is an initial segment of T as claimed.

– *We claim that $f[D] = T$*: Suppose not. Recall that D is an initial segment of S and so there exists $q \in S$ such that $D = S_q$. Then, as we just showed, $f[D] = f[S_q]$ must be equal to an initial segment, say T_r, for some $r \in T$. Then, since f is an order isomorphism, $S_q \sim_{\text{WO}} T_r$. This means that $f(q) = r$. So $q \in D$. But this contradicts the fact that $D = S_q = \{x \in S : x < q\}$. We must conclude that $f[D] = T$ as claimed.

We have thus shown that either f maps S order isomorphically into T or f order isomorphically maps an initial segment D of S onto T. From this we conclude that for any well-ordered sets $(S, \leq_S)$ and $(T, \leq_T)$, either $S \leq_{\text{WO}} T$ or $T \leq_{\text{WO}} S$.

The above theorem states that if we gather all well-ordered sets together to form a class of sets, we can rank them with the relation $\leq_{\text{WO}}$. Note that in this class, distinct well-ordered sets may well be equal or equipotent sets. For example, the set $(\mathbb{N}^*, <_*) = \{0, 2, 4, \ldots, 1, 3, 5, 7, \ldots, \}$ of all natural numbers where the even numbers are first enumerated in the usual order followed by all odd numbers enumerated in the usual way, is simply another way of describing the set $\mathbb{N}$. But, nevertheless, $\mathbb{N} <_{\text{WO}} \mathbb{N}^*$. Even though $\mathbb{N}$ and $\mathbb{N}^*$ are the same set, they are not order isomorphic. On the other hand, we easily see that $\mathbb{N}^*$ and the lexicographically ordered set $\{0, 1\} \times \mathbb{N}$ are order isomorphic.

It is also interesting to note that every well-ordered set is an initial segment of another well-ordered set. Indeed, if S is well-ordered by $\leq$, then S is order isomorphic to the initial segment $\{1\} \times S$ of the lexicographically ordered set $\{1, 2\} \times S$. In relation to such lexicographically ordered sets we present the following more general result.

Proposition 26.8 For every natural number n, the lexicographically ordered set $S = \{1, 2, \ldots, n\} \times \mathbb{N}$ is well-ordered.

Proof:

Let T be a non-empty subset of S. Let u be the least element of the set $\{r \in \{1, 2, \ldots, n\} : (r, t) \in S\}$. Since every natural number is well-ordered (14.4) such a number u exists. Let v be the least number in $\{t \in \mathbb{N} : (u, t) \in S\}$. Since $\mathbb{N}$ is well-ordered (14.3) such a number v exists. Then $(u, v) \leq (i, j)$ for all $(i, j) \in A$. Hence, every non-empty subset A of S has a least element. So S is $<$-well-ordered.

Concepts review:

1. What is a well-ordered set?
2. What is an initial segment of a well-ordered set?
3. Is a well-ordered set an initial segment of itself?
4. Give three examples of well-ordered sets.
5. Is the lexicographic ordering of $\mathbb{N} \times \mathbb{N}$ a well-ordering? Why or why not.
6. List all initial segments of the natural number 7.
7. Give an infinite initial segment of $\mathbb{N}^{\{0,1,2\}}$.
8. What is an order isomorphism between two well-ordered sets?
9. Can a well-ordered set be order isomorphic to one of its initial segments?
10. If $f : S \to T$ where $(S, \leq_S)$ and $(T, \leq_T)$ are linearly ordered sets, what does it mean to say that f is strictly increasing?
11. If a well-ordered set S is order isomorphic to an initial segment of a well-ordered set T, can S and T be order isomorphic?
12. What can we say about two well-ordered sets S and T in reference to order isomorphism?
13. How many order isomorphisms are there between an initial segment and itself?
14. If S and T are order isomorphic sets and f and g are two order isomorphisms mapping S onto T, what can we say about f and g?

EXERCISES

A.
1. Is the set of all prime numbers ordered in the usual way a well-ordered set?
2. List the first three initial segments of the set of all prime numbers ordered in the usual way. Are initial segments of prime numbers initial segments of $\mathbb{N}$?
3. Let $S = \{0\} \cup \{\frac{1}{n+1} : n \in \mathbb{N}\}$ be ordered by $<$ in the usual way. Is the set S a well-ordered set? Justify.

B.
4. Let $(S, \leq)$ be a well-ordered set. We say that an element b is an *immediate successor* of a if there does not exist an element c such that $a < c < b$. Show that if S is a well-ordered set, then every element of S that is not the maximal element of the set must have an immediate successor.

5. Suppose $(S, \leq)$ is a well-ordered set. Is there an order relation we can define on the set $\{T : T \text{ is an initial segment of } S\}$ which will make it a well-ordered set?
6. Let $\mathbb{N}_o$ denote the set of all odd natural numbers ordered in the usual way. Are $\mathbb{N}_o$ and $\mathbb{N}$ order isomorphic? If so say why. If not explain why.
7. Are the sets of all prime numbers and $\mathbb{N}$ both ordered in the usual way order isomorphic? If so say why. If not explain why.
8. Consider the set $\{1, 2\} \times \mathbb{N}$ when ordered lexicographically.
   a) List the first few elements of $\{1, 2\} \times \mathbb{N}$.
   b. Show that $\{1, 2\} \times \mathbb{N}$ ordered lexicographically is a well-ordered set.
   b) List three finite initial segments of $\{1, 2\} \times \mathbb{N}$. List three infinite initial segments of $\{1, 2\} \times \mathbb{N}$.
   c) In how many ways (if any) can we map $\mathbb{N}$ order isomorphically onto an initial segment of $\{1, 2\} \times \mathbb{N}$?
   d) In how many ways (if any) can we map $\mathbb{N}$ order isomorphically onto $\{1, 2\} \times \mathbb{N}$?

C. 9. Suppose S is a countably infinite set. Show that there exists a well-ordering $<_S$ such that $(S, <_S)$ and $(\{0, 1\} \times \mathbb{N}, <_{\text{lex}})$ are order isomorphic.

10. Let $S = \{1 - \frac{1}{n+1} : n \in \mathbb{N}\} \cup \{1\}$ be ordered by $<$ in the usual way.
    a) Is the set S a well-ordered set?
    b) Are the sets S and $\mathbb{N}$ order isomorphic? If so, show why. If not explain why not.

---

27 / Ordinal numbers: Definition and properties.

Summary. *In this section we provide some motivation for constructing what will be called the "ordinal numbers". We then formally define these. We see that* $\mathbb{N}$ *and all of its elements are ordinal numbers. When* $\mathbb{N}$ *is viewed as an ordinal number it is represented by* ω*. We review the notions of "transitive sets" and "*$\in$*-well-ordering". We then define "the immediate successor" of an element of a linearly ordered set, and show that the immediate successor of an ordinal number is an ordinal number. We also exhibit an ordinal number successor formula,* $\alpha^+ = \alpha \cup \{\alpha\}$*, and show how it is used to recursively construct sets of ordinals. We then prove a few basic properties of ordinal numbers from which we deduce that all pairs of distinct ordinals are* $\in$*-comparable. We define "limit ordinals", show how these are constructed and provide methods to recognize them.*

27.1 Introduction.

Our study of infinite sets began with a declaration of what it means for a set to be infinite. We stated that only a set S "which can be mapped one-to-one onto a proper subset of itself" is referred to as being "infinite". All other sets are "finite" sets. So initially, sets were either finite or infinite. Then we discovered that infinite sets could be subdivided into two categories: Those which are one-to-one images of $\mathbb{N}$ − referred to as "countably infinite" sets − and those that are not − referred to as being "uncountably infinite". Then, we discovered that the class of uncountably infinite sets actually has a more complicated structure. We found that not all uncountable sets were pairwise equipotent. We were led to this conclusion when we proved that no infinite set S could be mapped one-to-one onto its power set $\mathscr{P}(S)$. This implied that we could partition the class of all sets into infinitely many subclasses of sets each containing sets which were pairwise equipotent sets. Up to now, our attention has mainly been centered on investigating the properties of those sets which belong to the class of all countably infinite sets and the class of all sets which are equipotent to $\mathbb{R}$ (since the sets $\mathbb{N}$ and $\mathbb{R}$ are the two sets we are the most familiar with).

Within the class of all sets, we investigated the subclass of all well-ordered sets. We saw that order isomorphisms allow us to refine even further our classification of infinite sets. For example, a class of all well-ordered sets can be further partitioned into subclasses of pairwise order isomorphic well-ordered sets. Recall that an "order isomorphism" between two well-ordered sets S and T is a one-to-one function which respects the order of the elements in the domain S and the image T of S. That is, the order of the elements of the domain and the image is preserved by the one-to-one

function. For convenience, we introduced the following notation:

$$\begin{aligned} S \sim_{\text{WO}} T &\Leftrightarrow \text{“}S \text{ and } T \text{ are order isomorphic”} \\ S <_{\text{WO}} T &\Leftrightarrow S \sim_{\text{WO}} W = \text{an initial segment of } T \\ S \leq_{\text{WO}} T &\Leftrightarrow \text{“}S \sim_{\text{WO}} T \text{ or } S <_{\text{WO}} T \end{aligned}$$

We were able to show that all pairs of well-ordered set are $\leq_{\text{WO}}$-comparable. This is in striking contrast with our first attempts at grasping the structure of the class of all sets. The reader will recall that we were not clear on how to prove that “$\hookrightarrow_{e\sim}$” linearly orders the class of all sets, even though we strongly suspect this to be the case. Our ultimate objective in this section and the one that follows will be to construct a “well ordered class of well-ordered sets” which contains an order isomorphic copy of every well-ordered set. We will see that *ZFC* provides us with the tools to construct a class of sets which serves this purpose. The elements of this class of sets will be called *ordinals*.

27.2 Definition of “ordinal number”.

The reader will recall that every natural number is a “transitive set”. *Transitive sets* are those sets S that satisfy the rule:

$$(y \in S) \Rightarrow (y \subset S)$$

This property was shown (in 13.7) to be equivalent to the property

$$x \in y \text{ and } y \in S \Rightarrow x \in S$$

which is more suggestive of the notion of “transitivity” with respect to the membership order relation $\in$. A proper class which satisfies this transitive property will be referred to as a *transitive class*. We showed (in 13.8) that not only is $\mathbb{N}$ a transitive set, but each natural number is also transitive (13.9). This is easy to see if we re-examine how the natural numbers are constructed:

$$\begin{aligned} \varnothing &= 0 \\ \varnothing^+ &= 0 \cup \{0\} = \{0\} = 1 \\ 1^+ &= 1 \cup \{1\} = \{0,1\} = 2 \\ 2^+ &= 2 \cup \{2\} = \{0,1,2\} = 3 \\ 3^+ &= 3 \cup \{3\} = \{0,1,2,3\} = 4 \\ &\vdots \\ n^+ &= n \cup \{n\} = \{0,1,2,\ldots,n\} = n+1 \end{aligned}$$

So $n+1 = \{0,1,2,3,\ldots,n\} \subseteq \mathbb{N}$ is both a subset of $\mathbb{N}$ and an element of $\mathbb{N}$ contained in the subset $n+2 = \{0,1,2,3,\ldots,n+1\} \subseteq \mathbb{N}$.

The "transitive set" property of a set S does not depend on any particular order relation on S. But if every element of a set S is transitive, it makes it possible for $\in$ to take on the role of an order relation on S. It is shown in theorems 14.4 and 14.3 that all natural numbers n, as well as the set $\mathbb{N}$, are strictly $\in$-well-ordered. To say that "$\mathbb{N}$ is strictly $\in$-well-ordered" means that $\mathbb{N}$ is $\in$-irreflexive, $\in$-asymmetric, $\in$-transitive, any two distinct elements are $\in$-comparable, and every non-empty subset S of $\mathbb{N}$ contains a least element x with respect to $\in$. This means that any non-empty set, S, of natural numbers contains an element x such that $x \in y$ for all $y \in S$. We will now discuss $\in$-well-ordered sets which are not necessarily natural numbers.

Definition 27.1 Let S be a set. If S satisfies the two properties,

1) S is a transitive set,
2) S is strictly $\in$-well-ordered

then S is called an *ordinal number*.

The set $\mathbb{N}$ and all its elements are ordinals. Since the set $\mathbb{N}$ of all natural numbers, as well as each natural number, have been shown (in 13.8, 13.9 14.4 and 14.3) to be strictly $\in$-well-ordered transitive sets, the class of all ordinals contains infinitely many finite ordinals and at least one infinite ordinal, namely $\mathbb{N}$. We will continue to represent finite ordinal numbers by the usual lower case letters such as m or n, but infinite ordinal numbers will be represented by lower-case Greek letters, such as ω, α and β.

Notation 27.2 When viewed as an ordinal number, $\mathbb{N}$ will be represented by the lower-case Greek letter ω.[1] We then write,

$$\omega = \{0, 1, 2, 3, \ \ldots, \ \}$$

The reader should note that, by definition, only "sets" can be ordinals. That is, a strictly $\in$-well-ordered proper class is not an ordinal.

27.3 Constructing new ordinals from known ordinals.

Remember that each natural number is constructed using what we referred to as a "successor constructing algorithm" $n + 1 = n^+ = n \cup \{n\}$. We will use the same mechanism to construct numbers beyond ω.

[1]The letter ω is read "omega". So $\mathbb{N}$ has three representations: When simply viewed as a set, we use $\mathbb{N}$, when viewed as a cardinal number we use $\aleph_0$, when viewed as an ordinal number we use ω. Later, the symbol, ω_0, will be used in instead of ω.

We first show that $\omega \notin \omega$: If $\mathbb{N} \in \mathbb{N}$, then, by definition of $\mathbb{N}$, $\mathbb{N}$ is a natural number n. But $\mathbb{N}$ cannot be a natural number n for if it was, then $n \in n$, contradicting $n \notin n$ proven in 13.9. So $\omega \notin \omega$.
Since ω is a set, the expression

$$\omega^+ = \omega \cup \{\omega\}$$

is the union of two sets and so is itself a set. Note that $\omega^+ \neq \omega$ for if it was, then $\omega \in \omega$, a contradiction. So, from ω, we have generated a new set ω^+. This set is represented as, $\omega^+ = \omega + 1$. If we repeat the procedure again starting with $\omega + 1$ we obtain $\omega + 2 = (\omega + 1)^+ = \omega + 1 \cup \{\omega + 1\}$. We confirm immediately that if α is an ordinal number, then $\alpha + 1$ is necessarily an ordinal number.

Theorem 27.3 If α is an ordinal number, then so is its successor, $\alpha^+ = \alpha \cup \{\alpha\}$.

Proof:

What we are given: That α is an ordinal number (i.e., α is a transitive set and strictly $\in$-well-ordered).
What we are required to prove: That α^+ is an ordinal number.

The class α^+ is a set: Note that since α is an ordinal α must be a set. Hence, by Axiom 3 (Axiom of pair), $\{\alpha\}$ is a set. By Axiom 6 (Axiom of union), $a^+ = \alpha \cup \{\alpha\}$ is a set, as claimed.

The set α^+ is transitive: Suppose $x \in \alpha^+ = \alpha \cup \{\alpha\}$. By definition of "transitive" it suffices to show that $x \subset \alpha^+$. If $x = \alpha$, then $x \subset \alpha \cup \{\alpha\} = \alpha^+$ and we are done. Suppose $x \neq \alpha$; then $x \in \alpha$. Since α is transitive $x \subset \alpha$ and so $x \subset \alpha^+$. So α^+ is transitive. It follows that when viewed as a relation on α^+, $\in$ is a transitive relation.

The elements of the set α^+ are $\in$-comparable: Let x and y be distinct elements in $\alpha^+ = \alpha \cup \{\alpha\}$.
Case 1: If $x = \alpha$, then $y \in x$ (since $x \neq y$). Then x and y are $\in$-comparable.
Case 2: If both $x, y \in \alpha$, then, since α is known to be $\in$-linearly ordered, either $x \in y$ or $y \in x$. So all pairs of elements in α^+ are $\in$-comparable.
It follows that the relation "$\in$" linearly orders α^+.

The relation $\in$ is a strict linear ordering of α^+: Since $\in$ strictly orders α, $x \notin x$ for all $x \in \alpha$. Also $\alpha \notin \alpha$, for if $\alpha = x \in \alpha$, then $x \in x$ contradicting the fact that $\in$ strictly orders α.

The set α^+ is $\in$-well-ordered: Let S be a non-empty subset of α^+. Let $T = S \cap \alpha$.
Case 1: If $T = \varnothing$, then $S = \{\alpha\}$. Since $\alpha \notin \alpha$, α must be the least (actually the only) element of S.
Case 2: Suppose $T \neq \varnothing$. Since α is $\in$-well-ordered, there exists an $m \in T$ which is the $\in$-least element of T. If $S = T$, then m is the $\in$-least element of S, as required. If, on the other hand, $S = T \cup \{\alpha\}$, since α is the maximal element in α^+, $m < \alpha$. Again m is the

$\in$-least element of S.
We conclude that α^+ is strictly $\in$-well-ordered. So α^+ is an ordinal number.

Definition 27.4 Suppose the set, S, is $<$-ordered. We say that an element y in S is an *immediate successor* of the element x if $x < y$ and there does not exist any element z in S such that $x < z < y$. We say that x is an *immediate predecessor* of y if y is an immediate successor of x.

If α is an ordinal number, then we naturally expect α^+ to be an immediate successor of α. We verify that this is indeed the case. Suppose there exists an element, β, which is "strictly in between α and α^+" with respect to the $\in$-ordering. That is, suppose $\alpha \in \beta \in \alpha \cup \{\alpha\}$. Since $\beta \neq \alpha$ and $\beta \notin \alpha$, "$\beta \in \alpha \cup \{\alpha\}$" is impossible. So α^+ is an immediate successor with respect to the $\in$-well-ordering.

The ordinal constructing mechanism, $\alpha^+ = \alpha \cup \{\alpha\}$, can now be used to construct infinitely many ordinals beyond ω.

$$\begin{aligned}
\omega &= \{0, 1, 2, 3, \ \ldots, \ \} \\
\omega + 1 &= \omega^+ = \{0, 1, 2, 3, \ \ldots, \ \omega\} \\
\omega + 2 &= (\omega + 1)^+ = \{0, 1, 2, 3, \ \ldots, \omega, \ \omega + 1\} \\
\omega + 3 &= (\omega + 2)^+ = \{0, 1, 2, 3, \ \ldots, \omega, \ \omega + 1, \omega + 2\} \\
&\vdots \\
(\omega + n) + 1 &= (\omega + n)^+ = n \cup \{n\} = \{0, 1, 2, 3, \ \ldots, \ \omega + n\} \\
&\vdots
\end{aligned}$$

We see that this method for constructing ordinals has a limited range. Are there any other transitive "$\in$-well-ordered sets" beyond the set $\{\omega, \omega + 1, \omega + 2, \omega + 3, \ldots\}$ of ordinals? Our experience with ordinals tells us that there can be. Recall that having defined all finite ordinals (natural numbers) $0, 1, 2, 3, \ldots$, we gathered together all natural numbers to form a new set, $\mathbb{N} = \omega = \{0, 1, 2, 3, \ldots\}$. We then explicitly proved that this new infinite set, ω, is itself an ordinal. This illustrates that the "immediate successor constructing algorithm" is *not* the only way to construct ordinals. Consider, for example, the set

$$\omega + \omega = \{0, 1, 2, 3, \ldots, \omega, \omega + 1, \omega + 2, \omega + 3, \ldots\}$$

obtained by gathering together the ordinals $0, 1, 2, 3, \ldots, \omega, \omega + 1, \omega + 2, \omega + 3, \ldots,$. Is $\omega + \omega$ an ordinal? We will soon show that it is. First, we must show how the notions of "ordinal" and "initial segment of ordinals" are different ways of describing the same object.

27.4 Ordinals viewed as initial segments.

If the ordinal numbers look familiar to us it is because their properties are generalizations of those possessed by the natural numbers. Parts of the proofs of the statements that follow mimic the proofs of various properties of the elements of $\mathbb{N}$. Before we begin, we remind ourselves of what those subsets called "initial segments" are: A set, U, is an initial segment of an ordered set, $(S, <)$, if and only if U is a *proper* subset of S, and

$$\forall u \in U, \quad [v < u] \Rightarrow [v \in U]$$

Theorem 27.5 Let α be an ordinal number greater than zero.

a) Every element of α is an initial segment of α.

b) The ordinal α is an initial segment of an ordinal which contains α, namely, $\alpha^+ = \alpha \cup \{\alpha\}$.

c) Every initial segment of an ordinal α is an ordinal number.

d) Every element of the ordinal α is an ordinal number.

Proof:

a) What we are given: That x is an element of the ordinal α.
What we are required to show: That x is an initial segment of α.

Since $x \in \alpha$, and α is strictly $\in$-well-ordered, then $x \neq \alpha$. We confirm that x is a proper subset of α : Since $x \in \alpha$ and α is a transitive set, then $x \subset \alpha$.
Let $u \in x$ and suppose $v \in u$. We are required to show that $v \in x$. Given that $\in$ linearly orders α, then $\in$ is transitive, and so $v \in u \in x \Rightarrow v \in x$.

We have shown that x is a proper subset of α which satisfies the "initial segment" property with respect to $\in$, as required.

b) We have shown that if α is an ordinal, then so is $\alpha^+ = \alpha \cup \{\alpha\}$. Since $\alpha \in \alpha^+$, then by part a) α is an initial segment of α^+.

c) What we are given: x is an initial segment in α where α is an ordinal number.
What we are required to show: x is an ordinal number.

We claim that x is a transitive set: Let $z \in y \in x$. It suffices to show that $z \in x$. Now $y \in x \subset \alpha$ implies $y \in \alpha$. Also $z \in y \in \alpha$ implies $z \in \alpha$ (since α is transitive). So z is $\in$-less than y, with respect to α's order relation $\in$. Since x is an initial segment, z is $\in$-less than y and $y \in x$ implies $z \in x$. So x is a transitive set as claimed.

The relation, $\in$, is a strict linear ordering of x : All elements of x are elements of α (since $x \subset \alpha$) so x inherits from α all $\in$-ordering properties, including $\in$-transitivity and $\in$-linearity. So x is $\in$-linearly ordered. Since $u \notin u$ for all u in α, then this must be the case for all elements in x. So $\in$ is a strict linear ordering of x.

The set x is $\in$-well-ordered : Let T be a non-empty subset of x. We are required to show that T contains an $\in$-least element.
Case 1: If $T = x$, then, since x is an initial segment, T contains the $\in$-least ordinal, 0, and so we are done.
Case 2: Suppose $T \subset x$. Since α is $\in$-well-ordered $T = T \cap x$ contains its $\in$-least element, say y. Since $y \in T \cap x$, $y \in x$. So y is an element of x which is the $\in$-least element of T. So x is $\in$-well-ordered.
So x is a transitive strictly $\in$-well-ordered set. We conclude that x is an ordinal number.

d) If γ is any element of the ordinal α, then by part a) γ is an initial segment of α. Having shown in part b) that initial segments of ordinals are ordinals, then γ is itself an ordinal.

We now show that any infinite ordinal (other than ω itself) contains ω.

Proposition 27.6 Any infinite ordinal not equal to ω contains ω.

Proof:

What we are given: α is an infinite ordinal number.
What we are required to show: $\omega \in \alpha$.

We claim that $\{n : n \in \omega\} \subset \alpha$:
We prove the claim by induction. Let $P(n)$ be the statement "The natural number n belongs to α".

Base case: We are required to show that $P(0)$ holds true. Since α is infinite it is non-empty. Let γ be the $\in$-least ordinal in α. If $\gamma = 0$, then we are done. Suppose $\gamma \neq 0$. That is, suppose γ is non-empty. Then, when viewed as a subset of α, it contains a least element x. We then see that $x \in \gamma \in \alpha$ contradicting the fact that γ is the $\in$-least element of α. The source of the contradiction is our supposition that γ is not zero. Hence, $\gamma = 0 \in \alpha$. So $P(0)$ holds true.

Inductive hypothesis: Suppose $P(n)$ holds true for some natural number n. That is suppose that $n = \{0, 1, 2, 3, \ldots, n-1\} \in \alpha$. Since α is transitive, $n = \{0, 1, 2, 3, \ldots, n-1\} \subset \alpha$. Then $n+1 = \{0, 1, 2, 3, \ldots, n-1, n\} = n \cup \{n\} \subseteq \alpha$. Since α is infinite, we actually have $n+1 \subset \alpha$. Since $n+1 = \{0, 1, 2, 3, \ldots, n-1, n\}$ is an initial segment of α it is an ordinal in α. Then $P(n^+)$ holds true.

By the principle of mathematical induction, α contains every natural number, as claimed. Since $\omega \neq \alpha$, then ω is an initial segment of α. So $\omega \in \alpha$ as required.

Proposition 27.7 Let α and β be distinct ordinal numbers. If $\alpha \subset \beta$, then $\alpha \in \beta$.

Proof:

What we are given: That α and β are distinct ordinal numbers.
What we are required to show: That $(\alpha \subset \beta) \Rightarrow (\alpha \in \beta)$.

Suppose $\alpha \subset \beta$. The set $\beta - \alpha$ is non-empty, and so contains its least element, say γ. We will show that $\gamma = \alpha$. If so, then $\alpha \in \beta$ and we are done. Since elements of ordinals are ordinals, γ is an ordinal number.

We claim that $\alpha \subseteq \gamma$:

– Let $x \in \alpha \subset \beta$. Then x is also an ordinal number. It suffices to show that $x \in \gamma$. Suppose $x \notin \gamma$. Since β is $\in$-linearly ordered, $x \notin \gamma$ implies either $\gamma \in x$ or $\gamma = x$ holds true. But $\gamma \in x \subset \alpha$ or $\gamma = x \subset \alpha$ implies $\gamma \in \alpha$ (since α is transitive). This contradicts $\gamma \in \beta - \alpha$. So $x \in \gamma$. It follows that $\alpha \subseteq \gamma$ as claimed.

We claim that $\alpha = \gamma$:

– Suppose $\alpha \subset \gamma$. Then there exists $x \in \gamma - \alpha \subset \beta - \alpha$. This means x is an element in $\beta - \alpha$ which is strictly $\in$-less than its least element γ. This contradiction is caused by our supposition $x \in \gamma - \alpha$. We conclude that $\alpha = \gamma$ as claimed.

We have shown that if $\alpha \subset \beta$, then α is the the least element of $\beta - \alpha$ and so $\alpha \in \beta$, as required.

The above results have an implication which is worth pointing out immediately. We have shown in theorem 27.5 that for every ordinal α, α is an initial segment of $\beta = \alpha \cup \{\alpha\}$ with respect to the $\in$ order relation. Then we can write

$$\gamma = \{\alpha \in \beta : \alpha \in \gamma\}$$

where the ordinal γ is the leader of its initial segment. This is the case, for any ordinal γ. This is consistent with what we have observed up to now. Witness the ordinal, $3 = \{0, 1, 2\} = \{n : n \in 3\}$, where 3 is the leader of the ordinal 3, and the infinite ordinal, $\omega = \{0, 1, 2, 3, \ldots, \} = \{n : n \in \omega\}$, where ω is the leader of the ordinal ω.[1]

[1]Unfortunately we cannot deduce from this that the set $\{\alpha : \alpha$ is an ordinal, $\alpha \in \omega + \omega\}$" is an ordinal since we have not yet shown that $\omega + \omega$ is an ordinal.

27.5 The membership relation, $\in$, linearly orders the class of all ordinals.

Similarities between the methods of construction of the ordinals and the natural numbers strongly suggests that the relation "$\in$" linearly orders the class of ordinal numbers. This remains to be proved. Before we do this, we must prove the following lemma.

Lemma 27.8 If the ordinals α and β are order isomorphic, then $\alpha = \beta$.

Proof:

What we are given: The sets α and β are ordinals for which there exists an onto order isomorphic map $f : \alpha \to \beta$.
What we are required to show: That $\alpha = \beta$.

Let $S = \{x \in \alpha : f(x) \neq x\}$. Recall that order isomorphisms are strictly increasing. Since 0 is the least ordinal of both α and β, $f(0) = 0$ (if, for example, $f(0) = 1$, then f must map some element $\alpha > 0$ to $0 < 1$, a contradiction). Hence, S is not all of α.

If $S = \varnothing$, then $f(x) = x$ for all x in α; then $\alpha = \beta$ and we are done.

Suppose $S \neq \varnothing$. We claim that this will lead to a contradiction:

– Since α is $\in$-well-ordered, S contains a smallest element, say d. Since $d \in S$, then $f(d) \neq d$.
We claim: $f(d) \subseteq d$.

If $x \in f(d)$ in β, then there exists $z \in \alpha$ such that $f(z) = x$.

Since f respects $\in$-ordering

$$\begin{aligned} x \in f(d) &\Rightarrow f^{-1}(x) \in f^{-1}(f(d)) \\ &\Rightarrow z \in d \end{aligned}$$

Since d is the smallest element such that $f(x) \neq x$, $z \in d \Rightarrow f(z) = z$. But $f(z) = x$. So $z \in d \Rightarrow x \in d$.

This shows $x \in f(d) \Rightarrow x \in d$. So $f(d) \subseteq d$, as claimed.

– But we also see that $d \subseteq f(d)$, since

$$u \in d \Rightarrow f(u) = u \in f(d)$$

So $d \subseteq f(d)$ and $f(d) \subseteq d$ implies $d = f(d)$ which contradicts the fact that d is least ordinal such that $f(d) \neq d$.

So S must be empty. This means that f is the identity map. We can only conclude that $\alpha = \beta$.

Theorem 27.9 The relation "$\in$" linearly orders the class of all ordinals.

Proof:

We have shown in the theorem 26.7 that any two well-ordered sets S and T are either order isomorphic or one is order isomorphic to an initial segment of the other. If the ordinals α and β are not order isomorphic, then one must contain an order isomorphic copy of the other. Suppose, without loss of generality, that α is order isomorphic to an initial segment γ of β. By part b) of 27.5, γ must be an ordinal number. Since α and γ are order isomorphic ordinals, then, by lemma 27.8, they must be the same ordinal number. Hence, $\alpha \in \beta$. We can conclude that any two ordinal numbers are $\in$-comparable; so "$\in$" linearly orders the class of all ordinals.

The immediate successor of an ordinal is unique. Since "$\in$" linearly orders the class of ordinals, then this restricts the number of immediate successors an ordinal can have. For suppose β_1 and β_2 are immediate successors of the ordinal α. If $\beta_1 \neq \beta_2$, then either $\beta_1 \in \beta_2$ or $\beta_2 \in \beta_1$. Suppose without loss of generality that $\beta_1 \in \beta_2$. Then $\alpha \in \beta_1 \in \beta_2$. But this implies that β_2 is not an immediate successor of α, a contradiction. We must conclude that $\beta_1 = \beta_2$. So the immediate successor of an ordinal is unique.

27.6 Limit ordinals.

We have seen that all ordinals are initial segments of ordinals. Now an initial segment of a linearly ordered set may or may not contain a *maximal element*. For example, the ordinal number

$$\omega + 2 = \{0, 1, 2, \ldots, \omega, \omega + 1\} = \{0, 1, 2, \ldots, \omega\} \cup \{\omega + 1\} = \omega + 1 \cup \{\omega + 1\}$$

contains the maximal element, $\omega + 1$, since every element of $\omega + 2$ is either $\omega + 1$ or is contained in $\omega + 1$. On the other hand, the ordinal $\omega = \{0, 1, 2, 3, \ldots, \}$ is an initial segment $\{\gamma : \gamma \in \omega\}$ of the ordinal $\omega + 1$ which has no maximal element. If an ordinal β has a maximal element, say α, then β is the immediate successor of α. (Equivalently, the maximal ordinal α is an immediate predecessor of β.) We can then divide the class of all ordinals into two subclasses:

1) Ordinals β that contain a maximal element with respect to $\in$. These are precisely the ordinals that have an immediate predecessor. That is,

$$\beta^+ = \{0, 1, 2, \ldots, \beta\}$$

2) Ordinals β that do *not* contain a maximal element with respect to $\in$. These are the ordinals that do not have an immediate predecessor. They can be represented as,

$$\beta = \{\alpha : \alpha \in \beta\}$$

For example, $\omega = \{0, 1, 2, \ldots\}$.

Those ordinals which do not have a maximal element (equivalently, do not have an immediate predecessor) are called "limit ordinals". We define this formally.

Definition 27.10 An ordinal α which does not contain a maximal element is called a *limit ordinal*.

27.7 Constructing limit ordinals.

Suppose U is a non-empty set whose elements are ordinals. What can we say about the union, $\cup\{\alpha : \alpha \in U\}$, of all ordinals in the set U? In particular, we ask the question: Is the union of all ordinals in U necessarily an ordinal? We will show that it must be so. Whether this union of ordinals is a limit ordinal, or a non-limit ordinal, will depend on whether the set U contains, or does not contain, a maximal ordinal with respect to $\in$.

Proposition 27.11 If U is a non-empty set of ordinals which contains a maximal element β with respect to $\in$, then the union, $\cup\{\alpha : \alpha \in U\}$, is equal to the maximal ordinal, β, of U.

Proof:

We are given that β is an ordinal which is the maximal element of a set of ordinals U. Then $\alpha \in \beta$, for all α in U which are distinct from β. Since β is an ordinal it is a transitive set and so, for all $\alpha \in U$ such that $\alpha \neq \beta$, $\alpha \subset \beta$. Then $\cup\{\alpha : \alpha \in U, \alpha \neq \beta\} \subseteq \beta$. Since $\beta \in U$, $\beta = \cup\{\alpha : \alpha \in U\}$.

We have shown above that if β is the maximal element of U, $\cup\{\alpha : \alpha \in U\} = \beta$. Hence, in such a case, $\cup\{\alpha : \alpha \in U\}$ cannot be equal to U. For example, if $U = 3 = \{0, 1, 2\}$

$$\begin{aligned} \cup\{\alpha : \alpha \in U\} &= \cup\{0, 1, 2\} \\ &= 0 \cup 1 \cup 2 \\ &= \varnothing \cup \{\varnothing\} \cup \{\varnothing, \{\varnothing\}\} \\ &= \{\varnothing, \{\varnothing\}\} = 2 \neq U \end{aligned}$$

We will now show that if U is a set of ordinals which contains *no* maximal element, then $\cup\{\alpha : \alpha \in U\}$ is a limit ordinal which is not contained in U.

Theorem 27.12 If U is a set of ordinals which does *not* contain a maximal element with respect to "$\in$", then $\gamma = \cup\{\alpha : \alpha \in U\}$ is a limit ordinal which is not contained in U. Furthermore, γ is the $\in$-least ordinal which contains all elements of U.

Proof:

What we are given: That U is a set of ordinals with no maximal element and $\gamma = \cup\{\alpha : \alpha \in U\}$.
What we are required to prove: That γ is a limit ordinal which is not an element of U; that γ is the least ordinal containing all elements of U.

The class γ is a set. Recall that ordinals are sets (by definition) and, since U is declared to be a set, γ is the union of a set of sets. Axiom 6 guarantees that γ is a set.

Every pair of elements in γ are $\in$-comparable. Any two elements of γ are both elements of some ordinal and so are themselves ordinals. By theorem 27.8, they are $\in$-comparable.

The set γ is a transitive set. Let β be an element of γ. Then $\beta \in \alpha$ for some $\alpha \in U$. Since α is transitive, $\beta \subset \alpha$. Since $\alpha \subset \gamma$, $\beta \subset \gamma$. Then γ is a transitive set, as claimed.

The set γ is a well-ordered set. If A is a non-empty subset of γ, $A \cap \alpha \neq \varnothing$, for some $\alpha \in U$. Since α is $\in$-well-ordered, $A \cap \alpha$ contains a least element β. Let ψ be any element (equivalently, ordinal) in A not equal to β. If $\psi \in \beta$, then $\psi \in \alpha$. Then ψ is an element of $A \cap \alpha$ which is strictly $\in$-less than the least element β in $A \cup \alpha$. Since this cannot be, $\beta \in \psi$. So β is the least element of A. Then γ is $\in$-well-ordered, as required.

We have thus shown that $\gamma = \cup\{\alpha : \alpha \in U\}$ is an ordinal.

The ordinal $\gamma = \cup\{\alpha : \alpha \in U\}$ does not belong to U. Note that $\alpha \subseteq \gamma$ for all $\alpha \in U$. Since γ has been shown to be an ordinal, then for any $\alpha \in U$ not equal to γ, $(\alpha \subset \gamma) \Rightarrow (\alpha \in \gamma)$ (by Proposition 27.7). Suppose $\gamma \in U$. Then, for all $\alpha \in U$ such that $\alpha \neq \gamma$, $\alpha \in \gamma$. This implies that γ is a maximal element of U, contradicting our hypothesis stating that U has no maximal element. Then $\gamma \notin U$, as claimed.

The ordinal $\gamma = \cup\{\alpha : \alpha \in U\}$ is a limit ordinal. Suppose not. That is, suppose $\gamma = \beta^+ = \beta \cup \{\beta\}$. Then $\beta \in \gamma$. This means that $\beta \in \phi$ for some ordinal $\phi \in U$. Equivalently, $\beta \subset \phi$ (since ϕ is transitive). Now $(\phi \in U) \Rightarrow (\phi \subset \gamma)$ (we have shown that $\gamma \notin U$, so we can use strict containment "$\phi \subset \gamma$"). It follows that $\phi \subseteq \beta$. But $\beta \subset \phi \subseteq \beta$ implies $\beta \subset \beta$, a contradiction. The source of this contradiction is our supposition that γ has an immediate predecessor β. Then γ has no immediate predecessor and so is, by definition, a limit ordinal.

The ordinal $\gamma = \cup\{\alpha : \alpha \in U\}$ is the least ordinal which contains all elements of U. Suppose $\alpha \in U$. Then $\alpha \subset \gamma$. Since γ has been shown to be an ordinal, then $\alpha \in \gamma$ (by 27.7). Then $U \subseteq \gamma$. We claim that γ is the least ordinal satisfying this property. Suppose ψ is some ordinal such that $\psi \in \gamma$. Then $\psi \in \alpha$ for some $\alpha \in U$. Then $\alpha \notin \psi$. Then $U \not\subseteq \psi$.

So γ is the least ordinal which contains all elements of U, as claimed.

We summarize. The above theorem now provides us an alternate method for constructing new ordinals from known ordinals. Taking the union of a set U of ordinals where U contains no $\in$-maximal element will always produce a limit ordinal which does not belong to U.

Corollary 27.13 Let U be a non-empty set of ordinals which contains no maximal element. If U satisfies the "initial segment property", then U is the limit ordinal $\cup\{\alpha : \alpha \in U\}$.[1]

Proof:

What we are given: That U is a set of ordinals which contains no maximal element and satisfies the initial segment property.
What we are required to show: That $U = \cup\{\alpha : \alpha \in U\}$.

We have shown above that the set, $\cup\{\alpha : \alpha \in U\}$, is the least ordinal which contains all the elements of U. So, certainly, $U \subseteq \cup\{\alpha : \alpha \in U\}$.
Suppose $\beta \in \cup\{\alpha : \alpha \in U\}$. Then $\beta \in \alpha$ for some $\alpha \in U$. Since U satisfies the "initial segment property", $\beta \in U$. Then $\cup\{\alpha : \alpha \in U\} \subseteq U$. We conclude that $U = \cup\{\alpha : \alpha \in U\}$ as required.

Examples.

a) We define

$$\omega + \omega = \{0, 1, 2, 3, \ldots, \omega, \omega + 1, \omega + 2, \omega + 3, \ldots, \} = \omega \cup \{\omega + n\}_{n=0}^{\infty}$$

We conclude that $\omega + \omega$ is a limit ordinal by arguing as follows:

- Since $\omega + \omega$ is the union of two countable sets, it is a set.
- It is easily seen that $\omega + \omega$ contains no maximal element and satisfies the initial segment property.
- By the above corollary, $\omega + \omega = \{\alpha : \alpha \in \omega + \omega\}$ is the least ordinal which contains all finite ordinals and all infinite ordinals of the form $\omega + n$ where $n \in \mathbb{N}$.

Then $\omega + \omega$ is a limit ordinal number.

b) If we define

$$\omega + \omega + \omega = \omega + \omega \cup \{\omega + \omega,\ \omega + \omega + 1,\ \omega + \omega + 2,\ \omega + \omega + 3, \ldots, \}$$

we can similarly conclude that $\omega + \omega + \omega$ is the least ordinal which contains all ordinals in $\omega + \omega$, and ordinals of the form $\omega + \omega + n$, where $n \in \mathbb{N}$.

[1] The set U satisfies the "initial segment property" if $\forall \alpha \in U$, $[\gamma \in \alpha)] \Rightarrow [\gamma \in U]$.

c) Ordinals such as, $\omega+\omega$, $\omega+\omega+\omega$ and $\omega+\omega+\omega+\omega$, are more succinctly written as

$$\omega 2,\ \omega 3,\ \omega 4,\ \omega 5,\ldots$$

and so on. We denote the set of ordinals $\cup\{\omega n : n\in\mathbb{N}\}$ by

$$\omega\omega$$

We continue constructing, in this way, larger and larger ordinals. Following these general principles for representing ordinals we list a few more of these:

$$0,1,2,3,\ldots,\omega,\ \omega 2,\ \omega 3,\ldots,\omega n,\ldots,\omega\omega=\omega^2,\ \ldots,\omega^2+n,\ldots,\omega^2+\omega,\ldots,$$
$$\omega^2+\omega+n,\ldots,\omega^2+\omega 2,\ldots,\omega^2+\omega 3,\ldots,\omega^2+\omega n,\ldots,\omega^2+\omega^2=\omega^2 2,\ldots,\omega^2 3,\ldots,$$
$$\omega\omega^2 n\ldots,\omega\omega\omega=\omega^3,\ldots,\omega^n,\ldots,\omega^\omega,\ldots,\omega^{\omega^2},\ldots\omega^{\omega^\omega},\ldots$$

Since every one of these is the countable union of countably many ordinals, each is countable; so each of these is a set.

27.8 Characterizations of limit ordinals.

We can also describe limit ordinals in terms of "least upper bounds". We first remind the reader of what is meant by "least upper bound" of an ordered set.

Suppose T is a non-empty subset of a strictly ordered set $(S,<)$. If u is an element of S such that $t\le u$ for all $t\in T$, then we say that u is an *upper bound* of the set T.

Definition 27.14 Let T be a non-empty subset of an ordered set $(S,<)$. If u is an upper bound of the set T and, for any other upper bound v of T, $u\le v$, then we say that u is the *least upper bound* of T. We also abbreviate the expression by writing $u=\mathrm{lub}\,T$ or $u=\mathrm{lub}(T)$. [1]

Note that if a set γ is a non-empty ordinal and $\alpha\in_=\beta$ for all $\alpha\in\gamma$, then β is an upper bound of γ.[2] The least upper bound of a "non-limit (non-empty) ordinal" is always its maximal element. For example,

$$\mathrm{lub}(\omega+2)=\mathrm{lub}\{0,1,2,3,\ldots,\omega,\omega+1\}=\omega+1$$

since every element of this set is "$\in$-less than or equal to" $\omega+1$. So

$$\mathrm{lub}(\omega+2)\neq\omega+2$$

[1] Instead of "least upper bound of T" the word *supremum* of T is commonly used.
[2] The symbol $\in_=$ means "equal to or belongs to".

But if β is a limit ordinal (such as $\omega = \{0, 1, 2, 3 \ldots, \}$, for example), then it has no maximal element (immediate predecessor) and so $\text{lub}(\beta) \notin \beta$.

The following theorem provides various ways of recognizing limit ordinals.

Theorem 27.15 Let γ be a non-zero ordinal number. The following are equivalent:

1) The ordinal γ is a limit ordinal.

2) The ordinal γ is such that $\gamma = \cup\{\alpha : \alpha \in \gamma\}$.

3) The ordinal γ is such that $\text{lub}(\gamma) = \gamma$.

Proof:

$(1 \Rightarrow 2)$ Suppose γ is a limit ordinal. Then γ is an initial segment (27.5) which has no maximal element. By corollary 27.13, $\gamma = \cup\{\alpha : \alpha \in \gamma\}$.

$(2 \Rightarrow 1)$ Suppose the ordinal γ is such that $\gamma = \cup\{\alpha : \alpha \in \gamma\}$. If γ contains a maximal ordinal β, then $\alpha \in_= \beta$ for all $\alpha \in \gamma$. Then $\cup\{\alpha : \alpha \in \gamma\} = \beta \neq \gamma$, a contradiction. Then γ contains no maximal element. By definition, γ is a limit ordinal.

$(1 \Rightarrow 3)$ If γ is a limit ordinal, then γ does not contain a maximal element. By theorem 27.12, $\gamma = \cup\{\alpha : \alpha \in \gamma\}$ is the $\in$-least ordinal containing all elements of γ; hence, γ is the $\in$-least upper bound of γ.

$(3 \Rightarrow 1)$ Suppose the ordinal γ is such that $\text{lub}(\gamma) = \gamma$. If γ has a maximal element β, then $\alpha \in_= \beta$ for all $\alpha \in \gamma$. Then $\beta = \text{lub}(\gamma) \in \gamma$, a contradiction. Then γ cannot have a maximal element and so is a limit ordinal.

Here are a few examples of least upper bounds of sets of ordinals.

$$\begin{aligned}
\omega &= \text{lub}\{0, 1, 2, 3, \ldots\} = \{0, 1, 2, 3, \ldots\} = \cup\{n : n \in \omega\} \\
\omega + \omega &= \text{lub}\{0, 1, 2, 3, \ldots, \omega, \omega + 1, \omega + 2, \ldots, \} \\
\omega 3 &= \text{lub}\{0, 1, 2, 3, \ldots, \omega, \omega + 1, \omega + 2 \ldots, \omega 2, \omega 2 + 1, \ldots, \} \\
\omega 4 &= \text{lub}\{\alpha : \alpha \in \omega 4\} \\
4 &= \text{lub}\{0, 1, 2, 3, 4\} = \text{lub}(5)
\end{aligned}$$

Concepts review:

1. What is an "ordinal number"?

2. What is a transitive set?
3. What does it mean to say that a transitive set is strictly $\in$-well-ordered?
4. Given two elements x and y of a well-ordered set what does it mean to say that y is an immediate successor of x?
5. Give an example of an infinite linearly ordered set which contains elements with no immediate successor.
6. Describe a method for constructing immediate successors of ordinals.
7. When viewed as an ordinal number, how do we represent $\mathbb{N}$?
8. Given an ordinal number α, which one of its elements are initial segments of α?
9. Can an ordinal number be an initial segment of itself?
10. Which elements of an ordinal number α are themselves ordinal numbers?
11. Which elements of an ordinal number α are proper subsets of α?
12. Which subsets of an ordinal number α are elements of α?
13. What can be said about two ordinals which are order isomorphic?
14. How are limit ordinals different from non-limit ordinals?
15. How are the ordinals $\omega 2, \omega 3$ and $\omega 4$ described?
16. What kind of ordinals can be represented as $\gamma = \cup\{\alpha : \alpha \in \gamma\}$?
17. What is the least upper bound (supremum) of a limit ordinal α?
18. What is the least upper bound (supremum) of a non-limit ordinal α?
19. What does the expression $\text{lub}(\alpha) = \alpha$ say about the ordinal α?

EXERCISES

1. Find a well-ordered set which is order isomorphic to the ordinal number $\omega + 3 = \{0, 1, 2, \ldots, \omega, \omega + 1, \omega + 2\}$ (other than $\omega + 3$ itself).
2. Let S be a $\in$-well-ordered transitive set. Construct another $\in$-well-ordered transitive set which contains S.
3. What is the smallest ordinal that properly contains the ordinal number ω? Find the smallest ordinal that properly contains all elements of the ordinal number ω.

4. What is the intersection of all non-zero ordinals?
5. Does the ordinal $\omega + 2$ contain an order isomorphic copy of $\mathbb{N}$ which is not an initial segment of $\omega + 2$?
6. True or False: "Every subset of an ordinal is an ordinal".

B.

7. Find an ordinal number which is order isomorphic to the set $\{1, 3, 5, \ldots, 0, 2, 4\}$ of natural numbers ordered in this particular way.
8. Is the union of two distinct ordinal numbers necessarily an ordinal number? Explain.

C.

9. Show that there cannot exist a largest ordinal number.
10. Can there be an ordinal number that is not the immediate successor of another ordinal number? Explain.
11. Find three ordinal numbers which are *equipotent* with the ordinal number $\omega + 3$?
12. Let U be a set of ordinals. Show that $\cap\{\alpha : \alpha \in U\}$ is an ordinal.
13. Let A be a set of ordinals. Show that the intersection of all ordinals in A is the least ordinal of the set A.
14. What is the union of all ordinals in $\omega + 1$?
15. Prove that $\{\omega n : n \in \mathbb{N}\}$ is an ordinal number.
16. What is the smallest limit ordinal in $\omega 3$?
17. What is the union of all ordinal numbers in $\omega 3 + 1$?
18. Show that the ordinal $\omega\omega$ is countable.
19. Construct a well-ordered set which is order isomorphic to the ordinal $\omega\omega$ (other than $\omega\omega$ itself).

28 / Properties of the class of ordinal numbers.

Summary. *In this section we discuss the class, $\mathscr{O}$, of all ordinal numbers. It will be seen to be an $\in$-well-ordered transitive proper class of sets. Initial segments of $\mathscr{O}$ are shown to be ordinals. Once we prove the "Principle of transfinite induction" we show that every well-ordered set is order isomorphic to some ordinal. We then show that for any set, S, there exists an ordinal which cannot be embedded in S. We introduce what is known as the "Hartogs number of a set" and use this concept to construct a strictly increasing sequence of uncountable ordinals indexed with the ordinals.*

28.1 A well-ordering of the class of ordinal numbers.

By definition, each ordinal is an $\in$-well-ordered transitive set. Our investigation of ordinal numbers has revealed that:

1) Each ordinal has an immediate successor ordinal with respect to $\in$ (27.3).
2) Each ordinal is both an element and subset of any ordinal that contains it (27.7).
3) Given any two distinct ordinals, one is an element of the other (27.9).
4) Some ordinals have no immediate predecessor. Examples are

$$\begin{aligned} \omega &= \{0, 1, 2, 3, \ldots, \} \\ \omega 2 &= \{0, 1, 2, \ldots, \omega + 1, \omega + 2, \ldots, \} \\ \omega 3 &= \{0, 1, 2, \ldots, \omega + 1, \omega + 2, \ldots, \omega 2, \omega 2 + 1, \omega 2 + 2, \ldots, \} \end{aligned}$$

These are called *limit ordinals*. Limit ordinals are seen to be those ordinals which do not contain a maximal ordinal. Equivalently, they are those ordinals α, such that $\alpha = \text{lub}(\alpha)$ (27.15).

Such properties are not entirely new to us since the set, $\mathbb{N}$, and its elements are known to satisfy these very same properties. The class of ordinals appears to be a generalization of the natural numbers reaching well-ordered sets which are not order isomorphic to $\mathbb{N}$ or any of its elements. We will now investigate the class which contains all ordinals.

Notation. The class of all ordinal numbers will be denoted by, $\mathscr{O}$.

In the next few pages, we will show that the class, $\mathscr{O}$, itself satisfies most of the properties possessed by its elements. We will begin by showing that $\mathscr{O}$ is an $\in$-well-ordered class.

Theorem 28.1 The class, $\mathscr{O}$, of ordinal numbers is a strict $\in$-linearly ordered class.

Proof:

What we are given: $\mathscr{O}$ is the class of all ordinals.
What we are required to show: That "$\in$" is a strict linear order relation on $\mathscr{O}$.

- Since $\alpha \notin \alpha$ for all $\alpha \in \mathscr{O}$, "$\in$" is irreflexive and asymmetric. We verify transitivity of the order relation $\in$: If $\alpha \in \beta$ and $\beta \in \gamma$, then $\alpha \subset \beta$ and $\beta \subset \gamma$; so $\alpha \subset \gamma$; this implies $\alpha \in \gamma$. So $\in$ is a strict order relation on $\mathscr{O}$.
- That every pair of ordinals are $\in$-comparable has been shown in theorem 27.8.

We conclude that the class, $\mathscr{O}$, of all ordinal numbers is $\in$-linearly ordered.

Theorem 28.2 The class, $\mathscr{O}$, of all ordinal numbers is $\in$-well-ordered.

Proof:

What we are given: That S is a non-empty subset of ordinal numbers in $\mathscr{O}$.
What we are required to show: That S contains an ordinal β such that $\beta \in_= \alpha$ for all $\alpha \in S$.

Suppose $\gamma \in S$. If $\gamma \in \alpha$ for all $\alpha \in S$, then γ is the least ordinal of S and we are done. Suppose there is some ordinal $\alpha \in S$ such that $\alpha \in \gamma$. Then $\gamma \cap S$ is a non-empty subset of the well-ordered set γ. This means $\gamma \cap S$ contains a least element β. We claim that β is the $\in$-least element of S. Suppose ϕ is an ordinal in S such that $\phi \in \beta$. Since β is an element of γ and γ is transitive, $\phi \in \gamma$. Then ϕ is an element of $S \cap \gamma$ which is $\in$-less than the least element, β, of $\gamma \cap S$. This is a contradiction. So β is the least element of S. This shows that $\mathscr{O}$ is $\in$-well-ordered.

We know that each ordinal, α, is an initial segment of any ordinal that contains it and that the initial segment of any ordinal is an ordinal. The following theorem confirms that initial segments of $\mathscr{O}$ (with respect to $\in$) are precisely the ordinal numbers.

Theorem 28.3 A set S is an initial segment of $\mathscr{O}$ if and only if S is an ordinal number.

Proof:

If S is an $\in$-initial segment of $\mathscr{O}$, by theorem 26.3, there exists an ordinal γ such that $S = \{\alpha \in \mathscr{O} : \alpha \in \gamma\}$. But $\{\alpha : \alpha \in \gamma\}$ is an initial segment of $\gamma \cup \{\gamma\}$ and so, by theorem 27.5, is an ordinal.
Conversely, if $S = \gamma$ for some $\gamma \in \mathscr{O}$, then $\gamma = \{\alpha : \alpha \in \gamma\}$, an initial segment of $\mathscr{O}$.

28.2 The $\in$-well-ordered class, $\mathscr{O}$, is not an ordinal number.

Recall that a set or class S satisfies the transitive property if "$x \in S \Rightarrow x \subset S$". We easily verify that $\mathscr{O}$ is a transitive class:

$$\begin{aligned} \beta \in \mathscr{O} \text{ and } \alpha \in \beta &\Rightarrow \alpha \in \mathscr{O} \quad \text{(Since elements of ordinals are ordinals.)} \\ &\Rightarrow \beta \subset \mathscr{O} \\ &\Rightarrow \mathscr{O} \text{ is a transitive class.} \end{aligned}$$

Since $\mathscr{O}$ is a transitive $\in$-well-ordered class of ordinals, it possesses all the essential properties of ordinals. But if $\mathscr{O}$ is to be an ordinal, it must also be a set. The following theorem confirms that $\mathscr{O}$ cannot be a set.

Theorem 28.4 The class, $\mathscr{O}$, of all ordinal numbers is not a set.[1]

Proof:

Suppose the class $\mathscr{O}$ of all ordinal numbers is a set. Then, since $\mathscr{O}$ is transitive and $\in$-well-ordered, it is an ordinal number. Then $\mathscr{O} \in \mathscr{O}$. Since $\mathscr{O}$ is a transitive set, $\mathscr{O} \subset \mathscr{O}$. Since no ordinal number can be order isomorphic to a proper subset of itself (theorem 26.5), this is a contradiction. So $\mathscr{O}$ cannot be a set.

28.3 The *Principle of transfinite induction* over the ordinals.

The principle of mathematical induction over the natural numbers was seen to be an extremely useful tool to prove that certain properties hold true for sets whose elements can be indexed by the natural numbers. We remind ourselves of what it means to prove a statement by *mathematical induction on* $\mathbb{N}$ (or on the ordinal ω as we can now call it). Suppose $P(n)$ is a property which holds true depending on the value of the natural number (the finite ordinal) n. The principle of induction on ω states that if $P(0)$ holds true, and "$P(n)$ holds true" $\Rightarrow$ "$P(n+1)$ holds true", then $P(n)$ holds true for all values of n. We will show that this principle generalizes to mathematical induction over the ordinal numbers. We must, however, keep in mind that there is an important difference between $\mathscr{O}$ and ω. Some of the elements of an ordinal α may be limit ordinals. A limit ordinal β has no immediate predecessor and so applying the induction algorithm "$P(\alpha) \Rightarrow P(\alpha^+)$" cannot be used to prove that $P(\beta)$ holds true. Mathematical induction generalizes to ordinals provided we can verify that $P(\beta)$ holds true for limit ordinals, β, as well as for non-limit ordinals.

[1] Stating that the class of all ordinals is a set leads to what is referred to as the *Burali-Forti paradox*, the contradiction illustrated in this proof. It shows again that some classes are too large to be called "sets".

The principle of mathematical induction can be used to prove statements about *classes or sets whose elements can be indexed by ordinals.* For example, if $\gamma \in \mathscr{O}$ and $S = \{x_\alpha : \alpha \in \gamma\}$, then, since γ is a set and S is a one-to-one image of γ, by the Axiom of replacement, S is a also a set. Furthermore, the ordinal γ induces a well-ordering on the set S. The following theorem shows how this is done. It is followed by a second version of mathematical induction.

Theorem 28.5 *Principle of transfinite induction.* Let $\{x_\alpha : \alpha \in \mathscr{O}\}$ be a class whose elements are indexed by the ordinals. Let P denote a particular element property. Suppose $P(\alpha)$ means "the element x_α satisfies the property P". Suppose that for any $\beta \in \mathscr{O}$,

"$P(\alpha)$ is true $\forall\ \alpha \in \beta$" implies "$P(\beta)$ is true"

Then $P(\alpha)$ holds true for all ordinals $\alpha \in \mathscr{O}$.

Proof:

We are given that for every ordinal β, "$P(\alpha)$ is true for all $\alpha \in \beta$ implies $P(\beta)$ is true". Suppose there exist some ordinal γ such that $P(\gamma)$ is false.

We claim this will lead to a contradiction: By our supposition, the class $A = \{\alpha : P(\alpha) \text{ is false}\}$ is non-empty. Since $\mathscr{O}$ is $\in$-well-ordered, A must have a least element, say λ. That is, λ is the least ordinal such that $P(\lambda)$ is false. Since this is the least element of A, $P(\alpha)$ holds true for all $\alpha \in \lambda$. By hypothesis, $P(\lambda)$ must be true. We obtain a contradiction, as claimed.

Then the set, A, must be empty. So $P(\alpha)$ holds true for all ordinals α.

Corollary 28.6 *Transfinite induction: A second version.* Let $\{x_\alpha : \alpha \in \mathscr{O}\}$ be a class whose elements are indexed by the ordinals. Let P denote a particular element property. Suppose $P(\alpha)$ means "the element x_α satisfies the property P". Suppose that:

1) $P(0)$ holds true,

2) $P(\alpha)$ holds true implies $P(\alpha + 1)$ holds true,

3) If β is a limit ordinal, "$P(\alpha)$ is true for all $\alpha \in \beta$ implies $P(\beta)$ is true".

Then $P(\alpha)$ holds true for all ordinals α.

Proof:

We are given that P is a property for which conditions one, two and three hold true.
We are required to show that $P(\beta)$ holds true for all ordinals β.
Let β be an ordinal.

– If $\beta = 0$, then by condition 1), $P(\beta)$ holds true.

– Suppose β has an immediate predecessor, say $\lambda^+ = \beta$, such that $P(\lambda)$ holds true. By condition 2) $P(\beta) = P(\lambda^+)$ holds true.
– Suppose β is a limit ordinal such that "$P(\alpha)$ is true for all $\alpha \in \beta$". Then by condition 3) $P(\beta)$ holds true.

So $P(\beta)$ holds true for all ordinals β.

28.4 Indexing any well-ordered set with an ordinal.

In theorem 26.7, it was shown that given any two distinct non-order-isomorphic well-ordered sets, one is order isomorphic to an initial segment of the other. Since ordinals are, by definition, well-ordered, given any well-ordered set S and ordinal number α, precisely one of the following two statements must hold true:

1) The set S is order isomorphic to some ordinal $\beta \in \alpha$.
2) The ordinal α is order isomorphic to an initial segment of S.

If the set S is order isomorphic to some ordinal β, then this is equivalent to saying that "the elements of S can be indexed by the elements of the ordinal β". On the other hand, if the ordinal α is order isomorphic to an initial segment of S, clearly the elements of α cannot be used to index the elements of S since α is not "big enough" to be used as an indexing set for S. It is not yet clear whether, given *any* well-ordered set S, there exists some ordinal which can be used to index the elements of S. The following important theorem guarantees that there are sufficiently many ordinals in $\mathscr{O}$ so that *every* single well-ordered set S is order isomorphic to some ordinal. That is, any well-ordered set S can be indexed by the elements of some ordinal number α.

Theorem 28.7 Let S be a $<$-well-ordered set. Then S is order isomorphic to some ordinal number $\alpha \in \mathscr{O}$. Furthermore, the order isomorphism mapping S onto α is unique.

Proof:

What we are given: The set S is a $<$-well-ordered set.
What we are required to show: There exists a unique ordinal α which is order isomorphic to S. The required order isomorphism, $f : S \to \alpha$, is unique.

For each element $k \in S$, let $S_k = \{x \in S : x < k\}$ denote an initial segment of S. Let

$$D = \{u \in S : \ S_u \sim_{\text{WO}} \alpha_u \text{ for some ordinal number } \alpha_u\}$$

Let

$$f = \{(0_S, 0)\} \cup \{(u, \alpha_u) : u \in D\} \subseteq D \times \mathscr{O}^1$$

Now D is non-empty since $0_S \in D$ (where 0_S is the $<$-least element of S).

We claim that f is a function with domain D: Suppose (u, α) and (u, β) both belong to f. Then $S_u \sim_{\text{WO}} \alpha$ and $S_u \sim_{\text{WO}} \beta$. So $\alpha \sim_{\text{WO}} \beta$. By lemma 27.8, $\alpha = \beta$. Since the domain of f is a set, there is a set in the class $\mathscr{O}$ which contains the image of the set D under f (by the Axiom of replacement). So the relation

$$f : D \to \mathscr{O} \text{ defined as } f(u) = \alpha_u$$

is a function with domain D, as claimed.

The function f is strictly increasing and so is an order isomorphism: If $u, v \in D$ such that $u < v$, then $S_u \subset S_v$. Now, $\alpha_u \sim_{\text{WO}} S_u \subset S_v \sim_{\text{WO}} \alpha_v$. Given that α_u and α_v are ordinals, $\alpha_u \in \alpha_v$. Then, $u < v$ implies $f(u) = \alpha_u \in \alpha_v = f(v)$. So f is strictly increasing. We conclude that $f : D \to f[D] \subset \mathscr{O}$ is an order isomorphism.

We claim that D is a subset of S which satisfies the initial segment property: If $u \in D$, then $f(u) = \alpha_u$ for some $\alpha_u \in \mathscr{O}$. That is, there exists an order isomorphism $g : S_u \to \alpha_u$ mapping S_u onto α_u. If $x < u$, then $S_x \subset S_u$. The function $g|_{S_x}$ is an order isomorphism mapping S_x onto an initial segment (equivalently an ordinal), say α_x, in α_u. Then $f(x) = \alpha_x$. We have shown that $\forall u \in D$, $[x < u \in D] \Rightarrow [x \in D]$. Hence, D satisfies the initial segment property, as claimed.

We now claim that $f[D]$ is an ordinal number: It suffices to show that $f[D]$ is an initial segment of $\mathscr{O}$ and invoke theorem 28.3 (which states that any initial segment of $\mathscr{O}$ is an ordinal). Since $f[D]$ has been shown to be a set, then $f[D]$ cannot be equal to the proper class $\mathscr{O}$. Let $\alpha_v \in f[D]$. Then there exists an element $v \in D$ such that S_v is order isomorphic to α_v. Let $\beta \in \alpha_v$. Then $\beta \subset \alpha_v$. Since S_v is order isomorphic to α_v, then β is order isomorphic to an initial segment $S_u \subset S_v$. Then $f(u) = \beta$. So $\beta \in f[D]$. We have shown that if $\beta \in \alpha_v \in f[D]$, then $\beta \in f[D]$, and so, $f[D]$ is an initial segment of $\mathscr{O}$. Then, by theorem 28.3, $f[D]$ is an ordinal number, as claimed.

Finally, we claim that $D = S$: Suppose the domain D of f is *not* all of S. We have shown that $f[D]$ is an ordinal number. Say $f[D] = \gamma \in \mathscr{O}$. Since $D \neq S$, there exists some $q \in S$ such that $S_q = D$ (having shown that D satisfies the initial segment property). Then $f : D \to f[D] = f[S_q]$ is an order isomorphism mapping S_q onto the ordinal γ. So, by definition of f, $(q, \gamma) \in f$. This means that $q \in D$. But $D = S_q = \{x \in S : x < q\}$

[1]To better understand the set f we describe the first few elements. Suppose 1_S, 2_S, 3_S represent the first few elements of the well-ordered set S.

$$\begin{aligned} S_{1_S} = \{u \in S : u < 1_S\} = \{0_S\} \sim_{\text{WO}} \{0\} = 1 = \alpha_{1_S} &\Rightarrow (1_S, 1) \in f \\ S_{2_S} = \{u \in S : u < 2_S\} = \{0_S, 1_S\} \sim_{\text{WO}} \{0, 1\} = 2 = \alpha_{2_S} &\Rightarrow (2_S, 2) \in f \\ S_{3_S} = \{u \in S : u < 3_S\} = \{0_S, 1_S, 2_S\} \sim_{\text{WO}} \{0, 1, 2\} = 3 = \alpha_{3_S} &\Rightarrow (3_S, 3) \in f \end{aligned}$$

implies $q \notin D$. This is a contradiction. The source of the contradiction is the supposition that $D \neq S$. Then $D = S$, as claimed.

Then f maps S order isomorphically onto the ordinal number γ. By theorem 26.5 part e), this order isomorphism is unique, as required.

This means that the elements of any well-ordered set S can be *indexed* by the elements of some ordinal. That is, if $(S, <)$ is a well-ordered set which is order isomorphic to some ordinal β, then S can be expressed as the indexed set $S = \{s_\alpha : \alpha \in \beta\}$. This makes the set S susceptible to proofs by mathematical induction over the ordinal β, an extremely useful tool for proving various mathematical statements.

At this point, it will be useful to introduce some vocabulary that will allow us to state which ordinal is order isomorphic to a given well-ordered set S.

Definition 28.8 Let S be a $<$-well-ordered set. If α is the unique ordinal which is order isomorphic to S, then we will say that S is of *order type* α, or is of *ordinality* α. If S is of ordinality α, we will write

$$^{\text{ord}}S = \alpha$$

The set $\mathbb{N}$, ordered in the usual way, can then be said to have ordinality,

$$^{\text{ord}}\mathbb{N} = \omega$$

On the other hand the set $\mathbb{N}^*$ ordered as

$$\{0, 2, 4, 6, \ldots, 1, 3, 5, \ldots\}$$

has ordinality

$$^{\text{ord}}\mathbb{N}^* = \omega + \omega = \omega 2$$

The lexicographically ordered countably infinite set $S = \{1, 2, 3\} \times \mathbb{N}$ has ordinality,

$$^{\text{ord}}S = \omega + \omega + \omega = \omega 3$$

Viewing ordinals as $\sim_{WO}$-equivalence class representatives. Let $\mathscr{W} = \{S \in \mathscr{S} : S \text{ is well-ordered}\}$ denote the class of all well-ordered sets. We easily see that $\sim_{\text{WO}}$ is reflexive, symmetric and transitive on $\mathscr{W}$, and so is an equivalence relation on this class of sets. Let $\mathscr{W}^* = \{[S]_{\text{WO}} : S \in \mathscr{W}\}$ denote the class of all equivalence classes induced by $\sim_{\text{WO}}$. We have shown that every well-ordered class is order isomorphic to some ordinal. Then if $[T]_{\text{WO}} \in \mathscr{W}^*$, the equivalence class, $[T]_{\text{WO}}$, contains precisely one ordinal which is order isomorphic to every well-ordered set it contains. This means that we can adopt the ordinals in $\mathscr{O}$ as $\sim_{\text{WO}}$-equivalence class representatives of the

elements in $\mathscr{W}^*$. For example, if $\omega_3 \in [S]_{\text{WO}} \in \mathscr{W}^*$, it means that $[S]_{\text{WO}}$ contains precisely all well-ordered sets of ordinality ω_3. If ${}^{\text{ord}}T = \omega_0$, then $[T]_{\text{WO}}$ contains precisely all well-ordered sets which are order isomorphic to $\mathbb{N}$ (when $\mathbb{N}$ is ordered in the usual way). In this case, $[T]_{\text{WO}}$ contains $\mathbb{N}_{\text{even}} = \{0, 2, 4, 6, \ldots, \}$ but not the well ordered set $M = \{0, 2, 4, 6, \ldots, 1, 3, 5, 7, \ldots\}$ where ${}^{\text{ord}}M = \omega + \omega = \omega 2 \neq \omega$.

28.5 Proving the existence of uncountable ordinals with Hartogs' lemma.

Some readers may have noticed that we have not yet exhibited an uncountable ordinal (or even a single uncountable well-ordered set[1]). A potential candidate for the least uncountable ordinal may be the set:

$$\omega_1 = \{\alpha \in \mathscr{O} : \alpha \text{ is a countable ordinal } \}$$

This is a well-defined class of ordinals whose elements are precisely all countable ordinals. We show that it satisfies two fundamental characteristics:

- *The class ω_1 satisfies the "initial segment" property* : Suppose $\alpha \in \omega_1$ and $\beta \in \alpha$. Since α is a countable ordinal, then the ordinal β is countable and so $\beta \in \omega_1$. Then ω_1 satisfies the "initial segment" property.
- *The class ω_1 is uncountable*: Suppose ω_1 is countable. Since it is the one-to-one image of the set $\mathbb{N}$, ω_1 is a set. Then ω_1 is an initial segment of ordinals which is not equal to $\mathscr{O}$ (since $\mathscr{O}$ is not a set). Then, by theorem 28.3, ω_1 is an ordinal. It follows that $\omega_1 \in \omega_1$. Since no ordinal can be an element of itself we have a contradiction. The class ω_1 cannot be a countable set and so must be uncountable.

Can we conclude from this that ω_1 is an ordinal? What if $\omega_1 = \mathscr{O}$? If so, then ω_1 is a proper class. To show that $\omega_1 \neq \mathscr{O}$ it will suffice to show that an uncountable ordinal β exists. We must then first show that at least one uncountable ordinal exists.

The class of all well-ordered subsets of a well-ordered set. Suppose we are given a set $(S, <_S)$ on which we have defined a well-ordering relation $<_S$. To say that R is a well-ordering relation $<_S$ on S is to say that for $a, b \in S$,

$$a <_S b \Leftrightarrow (a, b) \in R$$

So we see in fact that $R = \{(x, y) \in S \times S : x <_S y\}$ is a subset of $S \times S$ and so the relation R is an element of the power set $\mathscr{P}(S \times S)$. Also note that the set S is an element of the power set $\mathscr{P}(S)$. So the well-ordered set $(S, <_S)$ can be precisely expressed as an ordered pair belonging to $\mathscr{P}(S) \times \mathscr{P}(S \times S)$. For example, if $S = \{1, 2, 3\}$ and "$<$" takes on its usual well-ordering "strictly less than", then

$$(S, <) = (\{1, 2, 3\}, \{(1, 2), (1, 3), (2, 3)\})$$

[1] If an uncountable well-ordered set is constructed, then, by theorem 28.7, it must be order isomorphic to some (uncountable) ordinal number.

represents a particular element of $\mathscr{P}(S) \times \mathscr{P}(S \times S)$. Equivalently, it expresses a particular well-ordering of the set S. On the other hand, if the ordering "$<_S$" of $S = \{1, 2, 3\}$ is such that $3 <_S 2 <_S 1$, then

$$(S, <_S) = (\{1, 2, 3\}, \{(3, 2), (3, 1), (2, 1)\})$$

is a different element of $\mathscr{P}(S) \times \mathscr{P}(S \times S)$ testifying to the fact that $(S, <_S)$ and $(S, <)$ are distinct well-orderings of the set S.

For any subset, T of S, on which we have defined a well-ordering relation $<_T$, $(T, <_T)$ is uniquely represented by a specific element of $\mathscr{P}(S) \times \mathscr{P}(S \times S)$. For example, referring to the same set $S = \{1, 2, 3\}$, letting $T = \{1, 2\}$,

$$(T, <) = (\{1, 2\}, \{(1, 2)\}) \in \mathscr{P}(S) \times \mathscr{P}(S \times S)$$

Suppose we let

$$\mathscr{A}_S = \{(T, <_T) \in \mathscr{P}(S) \times \mathscr{P}(S \times S) : <_T \text{ well-orders } T\}$$

The class $\mathscr{A}_S$ is the class of all well-ordered subsets $(T, <_T)$ of S. Since $\mathscr{A}_S$ is a subclass of the set $\mathscr{P}(S) \times \mathscr{P}(S \times S)$, it is also a set[1].

With these facts in mind, we are now ready to state and prove a very clever result known as *Hartogs' lemma*.[2]

Lemma 28.9 *Hartogs' lemma.* Let S be any set. Then there exists an ordinal α which is not equipotent with S or any of its subsets.
Proof:

What we are given: That S is a set.
What we are required to prove: That there exists an ordinal α that cannot be mapped one-to-one into the set S.

Let $\mathscr{A}_S = \{(T, <_T) \in \mathscr{P}(S) \times \mathscr{P}(S \times S) : <_T \text{ well-orders } T\}$ denote the set of all well-ordered subsets of S. By theorem 28.7, every well-ordered set has an ordinality and so there is a well-defined function $f : \mathscr{A}_S \to \mathscr{O}$ defined as $f((T, <_T)) = {}^{\text{ord}}(T, <_T)$ mapping each well-ordered set in $\mathscr{A}_S$ to some ordinal in $\mathscr{O}$. Then $f[\mathscr{A}_S]$ is a subclass of $\mathscr{O}$. The Axiom of replacement guarantees that $f[\mathscr{A}_S]$ is a set (*not* a proper class) and so $f[\mathscr{A}_S] \neq \mathscr{O}$. Since $f[\mathscr{A}_S]$ is not all of $\mathscr{O}$, there exists an ordinal $\beta \in \mathscr{O} - f[\mathscr{A}_S]$.

We claim that the ordinal β cannot be equipotent to any subset of S: For suppose, $h : \beta \to B$ is one-to-one and onto a subset $B \subseteq S$. Then, by theorem 26.1, B inherits a well-ordering $<_B$ from β (that is, $h(u) <_B h(v)$ in B if and only if $u \in v$ in β.) Then $(B, <_B) \in \mathscr{A}_S$. This contradicts the fact that β is not in the image $\mathscr{A}_S$ under f. So β cannot be equipotent with any subset of S, as required.

---

[1] Since S is a set, both $\mathscr{P}(S)$ and $\mathscr{P}(S \times S)$ are sets; hence, $\mathscr{P}(S) \times \mathscr{P}(S \times S)$ is a set.
[2] Notice how the Axiom of replacement plays a fundamental role in the proof of Hartogs' lemma, 28.9.

Theorem 28.10 There exists an uncountable ordinal.

Proof:

Let the set S in the lemma be the set $\mathbb{N}$ of all natural numbers. By Hartogs' lemma, there is an infinite ordinal β which cannot be mapped one-to-one onto any subset of $\mathbb{N}$. Since the ordinal β is not equipotent to any subset of $\mathbb{N}$, then it cannot be countable, So β is an uncountable ordinal. An uncountable ordinal exists.

Corollary 28.11 The class $\omega_1 = \{\alpha \in \mathscr{O} : \alpha \text{ is a countable ordinal } \}$ is the $\in$-least uncountable ordinal.

Proof:

It was shown that ω_1 is a subclass of $\mathscr{O}$ which satisfies the initial segment property. Hartogs' lemma states that there exists an ordinal number γ which is uncountable. Since $\alpha \in \gamma$ for any countable ordinal α, it follows that $\omega_1 \subseteq \gamma$. Since γ is a set, then ω_1 must also be a set; hence, ω_1 cannot be equal to the class $\mathscr{O}$ of all ordinals. Proper subsets of $\mathscr{O}$ which satisfy the initial segment property were shown to be ordinals. Hence, ω_1 is an ordinal. Since any ordinal $\alpha \in \omega_1$ is countable, then ω_1 must be the $\in$-least uncountable ordinal.

We can now lay this problem to rest. Uncountable ordinals exist in the *ZFC*-universe.

A few words of caution. One may be tempted to say that since ω_1 is the least uncountable ordinal number, then the sets ω_1 and $\mathbb{R}$ are equipotent. But nowhere have we shown that a one-to-one function between $\mathbb{R}$ and ω_1 exists. If such a one-to-one function between $\mathbb{R}$ and ω_1 was shown to exist, then the ordinal ω_1 would induce a well-ordering on $\mathbb{R}$, surely an unexpected result.[2]

We pause to review a sampling of (countable and uncountable) ordinal numbers we have seen up to now.

$$0, 1, 2, \ldots, \omega, \omega + 1, \ldots, \omega n, \ldots, \omega^\omega, \ldots, \omega_1, \omega_1 + 1, \ldots, \omega_1 + \omega, \ldots,$$

The ordinals listed here are strictly $\in$-ordered, in the sense that any ordinal is an element of any ordinal which appears on its right. We can also say that the ordinals are strictly $<_{\text{WO}}$-ordered, in the sense that any ordinal is order isomorphic to an initial

[2]One would never expect that $\mathbb{R}$ is well-orderable, but it will turn out to be the case, provided one believes in the Axiom of choice. That $\mathbb{R}$ is well-orderable does not follow from the Schröder-Bernstein theorem as some might suspect.

segment of any ordinal which appears on its right. But how are the elements of this sampling of ordinals ordered with respect to the equipotence relation $\hookrightarrow_e$[3]? In this case we obtain

$$0 \hookrightarrow_e 1 \hookrightarrow_e 2 \hookrightarrow_e 3 \hookrightarrow_e \cdots \hookrightarrow_e n \hookrightarrow_e \omega$$
$$\omega \sim_e \omega + 1 \sim_e \cdots \sim_e \omega n \sim_e \cdots \sim_e \omega^\omega \sim_e \cdots \sim_e \omega^{\omega^\omega}$$
$$\omega^{\omega^\omega} \cdots \hookrightarrow_e \omega_1 \sim_e \omega_1 + 1 \sim_e \cdots \sim_e \omega_1 + \omega \sim_e \cdots \sim_e \omega_1 + \omega^{\omega^\omega} \sim_e \cdots$$

In the first line, each ordinal is properly embedded into any ordinal which appears on its right. In the second line, all ordinals are equipotent to $\mathbb{N}$. In the third line, all ordinals to the right of ω_1 are equipotent to ω_1 (since each of these can be viewed as the union of ω_1 with a countable set). This sampling of ordinals makes us wonder: Are there any uncountable ordinals β such that $\omega_1 \hookrightarrow_e \beta$? Hartogs' lemma guarantees that such ordinals exist. Consider the class $H = \{\alpha \in \mathscr{O} : \alpha \text{ is not equipotent to any subset of the ordinal } \omega_1\}$. By Hartogs' lemma, the set H is non-empty. Since $\mathscr{O}$ is a well-ordered class, the non-empty class, H, must have a least element, which we denote by ω_2. Since $\mathscr{O}$ is linearly ordered, ω_2 is the *unique* least ordinal which is not equipotent to ω_1 or any of its subsets. By successively invoking Hartogs' lemma, we can show, in a similar way, that there exists unique uncountable ordinals $\omega_2, \omega_3, \ldots \omega_n, \ldots$, such that

$$0 \hookrightarrow_e 1 \hookrightarrow_e 2 \hookrightarrow_e 3 \hookrightarrow_e \cdots \hookrightarrow_e n \hookrightarrow_e \omega \hookrightarrow_e \omega_1 \hookrightarrow_e \omega_2 \hookrightarrow_e \omega_3 \hookrightarrow_e \cdots \hookrightarrow_e \omega_n \cdots$$

The $\hookrightarrow_e$-ordered chain of ordinals $\{0, 1, 2, 3, \ldots, n, \ldots, \omega,\ \omega_1,\ \omega_2,\ \ldots,\ \omega_n,\ \ldots,\}$ is seen to be a countably infinite set of ordinals. Furthermore, the ordinals listed are "$\hookrightarrow_e$-complete" in the sense that for any pair, α, β, of successive ordinals in this list, there does not exist an ordinal γ such that $\alpha \hookrightarrow_e \gamma \hookrightarrow_e \beta$. But this list of ordinals is not "$<_{\text{WO}}$-complete" in the sense that for any n, $\omega_n <_{\text{WO}} \omega_n + 1 <_{\text{WO}} \omega_n + 2 <_{\text{WO}} \cdots <_{\text{WO}} \omega_{n+1}$. That is, if we order the ordinals with "$<_{\text{WO}}$", then there are plenty of ordinals between ω_n and ω_{n+1}.

This makes us wonder whether there exists an even larger ordinal, β, such that $\omega_n \hookrightarrow_e \beta$, for all $n \in \mathbb{N}$. We will prove that there is.

28.6 The *Hartogs number* of a set.

To help answer the question above, we introduce the notion of a *Hartogs number* of a set.

---

[3]Recall that $A \hookrightarrow_e B$ if and only if A is equipotent to a proper subset and $A \not\sim_e B$.

Definition 28.12 Let S be any set. Let

$$U_S = \{\alpha \in \mathscr{O} : \ \alpha \text{ not equipotent to any subset of } S\}$$

By Hartogs' lemma the class, U_S, is non-empty. Since $\mathscr{O}$ is $\in$-well-ordered, U_S contains a unique least ordinal, we will denote as, $h(S)$. We will call the ordinal, $h(S)$, the *Hartogs number of the set* S. Then h can be viewed as a class function, $h : \mathscr{S} \to \mathscr{O}$, which associates to each set S in the class of all sets $\mathscr{S}$, a unique ordinal number, α, in the class of all ordinals $\mathscr{O}$.[1]

By *Hartogs' lemma* (28.9), every set S in $\mathscr{S}$ is assigned a unique Hartogs' number, $h(S)$. For example, the Hartogs numbers, $h(10)$, of the ordinal 10 is the ordinal 11, while the Hartogs number, $h(\omega + 7)$, of $\omega + 7$ is the ordinal $\omega_1 = \{\alpha : \alpha \text{ is a countable ordinal}\}$.[3] We will use the above definition of Hartogs numbers to show how we can recursively construct an endless, strictly increasing, $\hookrightarrow_e$-chain of uncountable ordinals.

In the proof of the theorem below, we will use a function constructing procedure called "transfinite recursion". We have encountered recursively defined functions before when defining addition (15.2) and multiplication (15.4) on $\mathbb{N}$. Using transfinite recursion to define a class function over the ordinal numbers is analogous to recursively defining a function on $\mathbb{N}$ (as shown in theorem 18.8). It is in fact a generalization of this process. The theorem which declares that recursively defined functions are well-defined functions is called the *Transfinite recursion theorem*.

The procedure for the construction of a recursively defined function $g : \mathscr{O} \to W$ on $\mathscr{O}$ is as follows:

Suppose W is a well-ordered set.

- The first step is to assign a value $a \in W$ to $g(0)$.
- Then assuming that $g(\alpha)$ is defined for an ordinal α we define $g(\alpha^+)$ in terms of $g(\alpha)$ according to a well-defined function $f : W \to W$. That is, $g(\alpha^+) = f(g(\alpha))$. This is where the similarity with recursively defined definitions on $\mathbb{N}$ ends.
- If β is a limit ordinal, we define $g(\beta) = \text{lub}\{g(\alpha) : \alpha \in \beta\}$.

We will illustrate how the transfinite recursively defined functions are constructed in the proof of the following theorem. We will differ the proof of the Transfinite recursion theorem to the end of this section.

---

[1] See page 82 for the definition of "class function".

[3] The Hartogs' number, $h(\omega + 7)$, of $\omega + 7$ is not $\omega + 8$ since $\omega + 8 \sim_e \omega + 7 \sim_e \omega$.

Theorem 28.13 There exists a strictly $\hookrightarrow_e$-increasing class $\{\omega_\alpha : \alpha \in \mathscr{O}\}$ of (pairwise non-equipotent) infinite ordinals, all of which are uncountable except for $\omega_0 = \omega$.

Proof:

Let h denote the *class* function which associates to each set S the unique least ordinal number which cannot be embedded in S. The ordinal, $h(S)$, is called the Hartogs number of S. By transfinite recursion, we define the class function $g : \mathscr{O} \to \mathscr{O}$ as follows:

$$\begin{array}{lcll} g(0) & = & \omega_0 & = \ {}^{\text{ord}}\mathbb{N} = \omega \\ g(\alpha^+) & = & \omega_{\alpha^+} & = \ h(\omega_\alpha), \text{ for all } \alpha \\ g(\gamma) & = & \omega_\alpha & = \ \text{lub}\{\omega_\gamma : \gamma \in \alpha\}, \text{ for all limit ordinals } \alpha \end{array}$$

Note that the function, $h : \mathscr{O} \to \mathscr{O}$, maps the proper class of all ordinals $\mathscr{O}$ into $\mathscr{O}$. The *Transfinite recursion theorem* guarantees that the sequence

$$\{g(\alpha) : \alpha \in \mathscr{O}\} = \{\omega_\alpha : \alpha \in \mathscr{O}\}$$

indexed by the ordinals is well-defined.

We claim that the sequence $\{\omega_\alpha\}_{\alpha\in\mathscr{O}}$ is (strictly) $\hookrightarrow_e$-increasing:
The proof of the claim is by transfinite induction. Let $P(\alpha)$ denote the statement "$\beta \in \alpha$ implies $\omega_\beta \hookrightarrow_e \omega_\alpha$".

Base case: The statement $P(1)$ holds true since $\omega_0 = \omega$ is properly embedded in all uncountable sets (by 18.9) and so $\omega_0 \hookrightarrow_e \omega_1$.

Inductive hypothesis for the case where α is a non-limit ordinal: Suppose $P(\alpha)$ holds true for the non-limit ordinal α. That is, "$\beta \in \alpha \Rightarrow \omega_\beta \hookrightarrow_e \omega_\alpha$". By definition of the Hartogs number of the ordinal $h(\omega_\alpha) = \omega_{\alpha^+} = \omega_{\alpha+1} \not\hookrightarrow_e B$, for any subset $B \subseteq \omega_\alpha$. Then $\omega_{\alpha+1} \not\in_= \omega_\alpha$ and so $\omega_\alpha \in \omega_{\alpha^+}$. This means that ω_α is order isomorphic to some initial segment of ω_{α^+} (by theorem 26.7). Since $\omega_{\alpha^+} \not\sim_e \omega_\alpha$, then $\omega_\alpha \hookrightarrow_e \omega_{\alpha^+}$. If $\beta \in \alpha^+$ and $\beta \neq \alpha$, then $\omega_\beta \hookrightarrow_e \omega_\alpha \hookrightarrow_e \omega_{\alpha^+}$; hence, $\omega_\beta \hookrightarrow_e \omega_{\alpha^+}$. So $P(\alpha^+)$ holds true.

Inductive hypothesis for the case where α is a limit ordinal: Suppose $P(\gamma)$ holds true for all ordinals $\gamma \in \alpha$ where α is a limit ordinal. We claim that $P(\alpha)$ holds true. Let $\beta \in \alpha$. We are required to show that $\omega_\beta \hookrightarrow_e \omega_\alpha$. Since $\omega_\alpha = \text{lub}\{\omega_\gamma : \gamma \in \alpha\}$ and $\beta \in \alpha$, $\omega_\beta \in_= \omega_\alpha$. Now ω_β is not equal to ω_α for if it was, $\omega_\alpha = \omega_\beta \in \omega_{\beta^+} \in \{\omega_\gamma : \gamma \in \alpha\}$ contradicting the fact that ω_α is an upper bound of $\{\omega_\gamma : \gamma \in \alpha\}$. Then $\omega_\beta \in \omega_\alpha$. Since $\omega_\beta \hookrightarrow_e \omega_{\beta^+} \in \omega_\alpha$, then $\omega_\beta \hookrightarrow_e \omega_\alpha$. So $P(\alpha)$ holds true.

By transfinite mathematical induction, for any ordinal α, $\beta \in \alpha \Rightarrow \omega_\beta \hookrightarrow_e \omega_\alpha$.

Hence, the class, $\{\omega_\alpha : \alpha \in \mathscr{O}\}$, constructed above is a class of (strictly) $\hookrightarrow_e$-increasing ordinals, as claimed.

Notation: From here on, the least infinite ordinal, ${}^{\text{ord}}\mathbb{N}$, previously represented by ω, will be represented as ω_0.

We immediately establish a few facts about the class $\{\omega_\alpha : \alpha \in \mathscr{O}\}$.

Proposition 28.14 Let $\{\omega_\alpha : \alpha \in \mathscr{O}\}$ be the class of ordinals as defined in the previous theorem.

a) Every element of $\{\omega_\alpha : \alpha \in \mathscr{O}\}$ is a limit ordinal.

b) For every ordinal $\alpha \in \mathscr{O}$, either $\alpha \in \omega_\alpha$ or $\alpha = \omega_\alpha$. (Equivalently, $\forall \alpha \in \mathscr{O}$, $\omega_\alpha \notin \alpha$.)

Proof:

a) Let $\omega_\gamma \in \{\omega_\alpha : \alpha \in \mathscr{O}\}$. We consider two cases: 1) γ is a successor ordinal, and 2) γ is a limit ordinal.

Case 1: Suppose $\gamma = \alpha^+$, for some α. We claim that ω_γ must be a limit ordinal. Suppose not. If ω_{α^+} is not a limit ordinal, then $\omega_{\alpha^+} = \beta^+ = \beta \cup \{\beta\}$ for some ordinal β. Since both ω_{α^+} and β are infinite sets, by theorem 20.6, $\omega_{\alpha^+} \sim_e \beta \cup \{\beta\} \sim_e \beta$, so $\omega_{\alpha^+} \sim_e \beta$. Furthermore, $\omega_{\alpha^+} = \beta \cup \{\beta\}$ implies that $\beta \in \omega_{\alpha^+}$. That is, β is $\in$-less than ω_{α^+}. By definition, ω_{α^+} is the least ordinal such that $\omega_{\alpha^+} \not\sim_e \omega_\alpha$. Since $\beta \sim_e \omega_{\alpha^+}$, then β is an ordinal strictly less than ω_{α^+} such that $\beta \not\sim_e \omega_\alpha$, a contradiction. The source of the contradiction is our supposition that ω_{α^+} is not a limit ordinal. So any element in $\{\omega_\alpha : \alpha \in \mathscr{O}\}$ of the form ω_{α^+} must be a limit ordinal.

Case 2: Suppose γ is a limit ordinal. Then, by definition, $\omega_\gamma = \text{lub}\{\omega_\alpha : \alpha \in \gamma\}$. We claim that ω_γ must be a limit ordinal. Suppose not. That is, suppose $\omega_\gamma = \beta \cup \{\beta\} = \beta^+$, for some ordinal β. Then $\beta \in \omega_\gamma$ and $\beta \sim_e \omega_\gamma$.

Claim: That $\beta \in_= \omega_\psi$ for some $\psi \in \gamma$.
Suppose not. That is, suppose $\omega_\alpha \in \beta$ for all $\alpha \in \gamma$. Then $\omega_\gamma = \text{lub}\{\omega_\alpha : \alpha \in \gamma\} \in_= \beta \in \omega_\gamma$. Then β is an upper bound of $\{\omega_\alpha : \alpha \in \gamma\}$ which is strictly less than ω_γ, a contradiction. So $\beta \in_= \omega_\psi$ for some $\psi \in \gamma$, as claimed

Since γ is a limit ordinal, $\psi^+ \in \gamma$; hence $\omega_{\psi^+} \in \{\omega_\alpha : \alpha \in \gamma\}$. Then

$$\omega_\gamma \sim_e \beta = \omega_\psi \in \omega_{\psi^+} \in \{\omega_\alpha : \alpha \in \gamma\}$$

contradicting the fact that ω_γ is an upper bound of $\{\omega_\alpha : \alpha \in \gamma\}$. The source of this contradiction is our assumption that ω_γ is not a limit ordinal. We can only conclude that for case 2, ω_γ is a limit ordinal.

b) We are required to prove that: $\forall \alpha \in \mathscr{O}$, $\alpha \in_= \omega_\alpha$.

The proof is by transfinite induction. Let $P(\alpha)$ denote the statement " $\alpha \in_= \omega_\alpha$".
Base case: The 0-ordinal $\in \omega_0$ since 0 belongs to all ordinals except the ordinal 0. So $P(0)$ holds true.

First inductive hypothesis: Suppose $P(\alpha)$ holds true for some α. That is, suppose $\alpha \in_= \omega_\alpha$. We are required to show that $\alpha^+ \in_= \omega_{\alpha^+}$. Since $\alpha \in_= \omega_\alpha$, then either $\alpha \in \omega_\alpha$ or $\alpha = \omega_\alpha$.

Case 1: Suppose $\alpha \in \omega_\alpha$. Then, since ω_α is a limit ordinal, $\alpha^+ \in \omega_\alpha \in \omega_{\alpha^+}$. So $P(\alpha^+)$ holds true.
Case 2: Suppose $\alpha = \omega_\alpha$. Then again, since ω_{α^+} is a limit ordinal and $\omega_\alpha \in \omega_{\alpha^+}$, $\alpha^+ = \omega_\alpha^+ \in \omega_{\alpha^+}$. So $P(\alpha^+)$ holds true.

Second inductive hypothesis: Suppose γ is a limit ordinal, $\text{lub}\{\omega_\alpha : \alpha \in \gamma\} = \omega_\gamma$ and $P(\alpha)$ holds true for all $\alpha \in \gamma$. That is, "$\alpha \in_= \omega_\alpha$" for all $\alpha \in \gamma$,
We are required to show: That $\gamma \in_= \omega_\gamma$.
If $\psi \in \cup\{\omega_\alpha : \alpha \in \gamma\}$, then for some $\alpha \in \gamma$, $\psi \in \omega_\alpha \in \omega_{\alpha^+} \in \omega_\gamma$; hence,

$$\cup\{\omega_\alpha : \alpha \in \gamma\} \subseteq \omega_\gamma$$

Since γ is a limit ordinal, $\gamma = \cup\{\alpha : \alpha \in \gamma\}$ (by theorem 27.15).
Given that "$\alpha \in_= \omega_\alpha$" for all $\alpha \in \gamma$, then $\alpha \subseteq \omega_\alpha$ for all $\alpha \in \gamma$ and so,

$$\gamma = \cup\{\alpha : \alpha \in \gamma\} \subseteq \cup\{\omega_\alpha : \alpha \in \gamma\} \subseteq \omega_\gamma$$

We conclude that $\gamma \in_= \omega_\gamma$.
By transfinite induction, $\alpha \in_= \omega_\alpha$ for all ordinals α.

A chain of ordinals under two distinct relations. The class $\{\omega_\alpha : \alpha \in \mathscr{O}\}$ can be viewed as a chain of infinite ordinals which is strictly ordered by "$\in$":

$$\omega_0 \in \omega_1 \in \omega_2 \in \cdots \in \omega_{\omega_0} \in \omega_{\omega_0+1} \in \cdots \in \omega_{\omega_0 2} \in \cdots \in \omega_{\omega_0\omega_0} \in \cdots \in \omega_{\omega_1} \in \cdots$$

The class of ordinals $\{\omega_\alpha : \alpha \in \mathscr{O}\}$ is also inductively constructed in a way that ω_α is not equipotent to any of its predecessors. Then it is also strictly ordered by the proper embedding relation "$\hookrightarrow_e$":

$$\omega_0 \hookrightarrow_e \omega_1 \hookrightarrow_e \cdots \hookrightarrow_e \omega_{\omega_0} \hookrightarrow_e \omega_{\omega_0+1} \hookrightarrow_e \cdots \hookrightarrow_e \omega_{\omega_0 2} \hookrightarrow_e \cdots \hookrightarrow_e \omega_{\omega_1} \hookrightarrow_e \cdots$$

Amongst these, only ω_0 is countable. This takes us considerably further down into the realm of uncountable sets. We already knew that the set $\{\mathscr{P}^n(\mathbb{N}) : n \in \mathbb{N}\}$ was a countably infinite set of pairwise non-equipotent sets. What is new here is that the class, $\{\omega_\alpha : \alpha \in \mathscr{O}\}$, is composed of as many pairwise non-equipotent well-ordered uncountable sets as there are ordinals!

28.7 A particular property of the first uncountable ordinal ω_1.

Consider the two sets $\mathbb{R}$ and $\omega_1^+ = \{0, 1, 2, 3, \ldots, \omega_1\}$. These two sets have some properties in common. For example, they are both linearly ordered and are both uncountable. However, because of their structure, they can also be seen as being radically different in nature. For example, $\mathbb{R}$ has infinitely many uncountable initial segments (subsets of the form $(-\infty, a)$), while ω_1^+ has only one uncountable initial

segment, namely $\omega_1 = [0, \omega_1)$. The real numbers $\mathbb{R}$ has another property which is not shared with ω_1^+. It is well known that, for every element $a \in \mathbb{R}$, $\{a\}$ is the intersection of countably many open intervals. For example, $\{a\} = \cap\{(a - \frac{1}{n}, a + \frac{1}{n}) : n = 1, 2, 3, \ldots\}$. We will refer to a set which is the countable intersection of open intervals in a linearly ordered set as a G_δ-set. For every $a \in \mathbb{R}$, $\{a\}$ is a G_δ-set. We cannot say that ω_1^+ shares the same property. We will show that $\{\omega_1\}$ is *not* a G_δ-set in ω_1^+. We will first describe those subsets of ω_1^+ which are open intervals. If α and β are distinct non-zero elements of ω_1^+, the open intervals in ω_1^+ are the sets of the form

$$\begin{aligned} [0, \alpha) &= \{\kappa : 0 \in_= \kappa \in \alpha\} \\ (\beta, \alpha) &= \{\kappa : \beta \in \kappa \in \alpha\} \\ (\beta, \omega_1^+) &= \{\kappa : \beta \in \kappa \in \omega_1^+\} \end{aligned}$$

Suppose $\{(\alpha_n, \omega_1^+) : n = 1, 2, 3, \ldots, \}$ is a countable set of open intervals each of which contains the ordinal ω_1. Then $\alpha_n \in \omega_1$, for all n. Let $\gamma = \text{ lub}\{\alpha_n : n = 1, 2, 3, \ldots, \}$. If $\gamma \in \{\alpha_n : n = 1, 2, 3, \ldots, \}$ then $\alpha_n \in_= \gamma \in \gamma^+ \in \omega_1$, for all n. If $\gamma \notin \{\alpha_n : n = 1, 2, 3, \ldots, \}$ then $\gamma = \cup\{\alpha_n : n = 1, 2, 3, \ldots, \}$. Since γ is the countable union of countable sets, it must itself be countable (theorem 27.12). Then, again, $\alpha_n \in \gamma \in \gamma^+ \in \omega_1$, for all n. We must then conclude that $\gamma^+ \in \cap\{(\alpha_n, \omega_1^+) : n = 1, 2, 3, \ldots, \}$ and so $\{\omega_1\} \neq \cap\{(\alpha_n, \omega_1^+) : n = 1, 2, 3, \ldots, \}$. That is, $\{\omega_1\}$ is not a G_δ-set. However, there may still be some limit ordinals larger than ω_1 which are G_δ-sets. Witness, $\{\omega_1 + \omega_0\} = \cap\{(\omega_1 + n, \omega_1 + \omega_0 + 1) : n = 1, 2, 3, \ldots, \}$.

28.8 The proof of the Transfinite recursion theorem.

We end this section by proving that recursively defined functions over the ordinals $\mathscr{O}$ are indeed well-defined.

Theorem 28.15 *The Transfinite recursion theorem.* Let W be a well-ordered class and $f : W \to W$ be a class function mapping W into W. Let $u \in W$. Then there exists a unique class function $g : \mathscr{O} \to W$ which satisfies the following properties:

a) $g(0) = u$

b) $g(\alpha^+) = f(g(\alpha)),\ \forall \alpha \in \mathscr{O}$

c) $g(\beta) = \text{lub}\{g(\alpha) : \alpha \in \beta\},\ \forall$ limit ordinals β

Proof outline :

Let $\mathscr{H}$ denote the class of all subclasses of $\mathscr{O} \times W$ which satisfy the three properties given in the theorem statement. That is, $U \in \mathscr{H}$ if and only if

a) $(0, u) \in U$
b) $[(\alpha, x) \in U] \Rightarrow [(\alpha^+, f(x)) \in U],\ \forall \alpha \in \mathscr{O}$
c) For any limit ordinal β,

$$[(\alpha, x_\alpha) \in U\ \forall \alpha \in \beta] \Rightarrow (\beta,\ \text{lub}\{x_\alpha : \alpha \in \beta\}) \in U$$

Now $\mathscr{H}$ is non-empty since $\mathscr{O} \times W$ satisfies all three properties and so $\mathscr{O} \times W$ is an element of $\mathscr{H}$. Let $G = \cap_{U \in \mathscr{H}}\{U\}$. Then G is the smallest element of $\mathscr{H}$ (with respect to "$\subseteq$"). The objective is to prove that the class G is the uniquely defined class function $g : \mathscr{O} \to W$ that we seek.
The first step is to show that $G \in \mathscr{H}$. This is straightforward and so is left as an exercise. The second step is to show that G is a class function, while the third step is to show that G is unique.
We will show that G is a class function by transfinite induction. For each ordinal γ let

$$G|_\gamma = \{(\alpha, x_\alpha) \in G : \alpha \in_= \gamma\}$$

For example,

$G\|_0$	contains at least	$(0, u)$
$G\|_1$	contains at least	$(0, u)$ and $(1, f(u))$
$G\|_2$	contains at least	$(0, u), (1, f(u))$ and $(2, f(f(u)))$
$G\|_3$	contains at least	$(0, u), (1, f(u)), (2, f(f(u)))$ and $(3, f(f(f(u))))$
$\vdots$	$\vdots$	$\vdots$

Let $P(\alpha)$ represent the statement "$G|_\alpha$ is a function".

– *Inductive hypothesis*: Case 1. Suppose $P(\alpha)$ holds true for all $\alpha \in \phi^+$ for some non-limit ordinal ϕ^+. This means that $G|_\phi$ is a function.

  We are required to show that $P(\phi^+)$ holds true. That is, we must show that $G|_{\phi^+}$ is also a function. Now $G|_{\phi^+} = G|_\phi \cup \{(\phi^+, x) : (\phi^+, x) \in G\}$. We know that $(\phi, x_\phi) \in G|_\phi$ so $(\phi^+, f(x_\phi)) \in \{(\phi^+, x) : (\phi^+, x) \in G\}$. To show that $G|_{\phi^+}$ is a function it suffices to show that $\{(\phi^+, x) : (\phi^+, x) \in G\}$ is the singleton set $\{(\phi^+, f(x_\phi))\}$. Suppose not. That is, suppose there exists in G an element (ϕ^+, y) such that $y \neq f(x_\phi)$.

  Claim: $G - \{(\gamma, y)\} \in \mathscr{H}$. If so, then this contradicts the fact that G is the smallest element of $\mathscr{H}$. The proof of the claim is left as an exercise.

  Assuming the claim is proved, we conclude that $P(\phi^+)$ holds true.

– *Inductive hypothesis*: Case 2. Suppose $P(\gamma)$ holds true for all ordinals $\alpha \in \gamma$ where γ is a limit ordinal. This means that $G|_\alpha$ is a function for all ordinals $\alpha \in \gamma$. Equivalently, $\{(\alpha, x_\alpha) : \alpha \in \gamma, (\alpha, x_\alpha) \in G\}$ is a function.

  We are required to show that $P(\gamma)$ holds true. That is, we must show that $G|_\gamma$ is also a function. Now

$$G|_\gamma = \{(\alpha, x_\alpha) : \alpha \in \gamma, (\alpha, x_\alpha) \in G\} \cup \{(\gamma, x_\gamma) : (\gamma, x_\gamma) \in G\}$$

Let $s_\gamma = \text{lub}\{x_\alpha : \alpha \in \gamma\}$. We know, by definition of G, that $(\gamma, s_\gamma) \in \{(\gamma, x_\gamma) : (\gamma, x_\gamma) \in G\}$. To show that $G|_\gamma$ is a function it suffices to show that $\{(\gamma, x_\gamma) : (\gamma, x_\gamma) \in G\}$ is the singleton set $\{(\gamma, s_\gamma)\}$. Suppose not. That is, suppose there exists (γ, y) such that $y \neq s_\gamma$.

Claim: $G - \{(\gamma, y)\} \in \mathscr{H}$. If so, then this contradicts the fact that G is the smallest element of $\mathscr{H}$. The proof of the claim is left as an exercise.

Assuming the claim is proved, we conclude that $P(\gamma)$ holds true.

Then by Transfinite induction "$G|_\alpha$ is a function" for all $\alpha \in \mathscr{O}$.

We claim that G must then be a class function. Suppose not. Then there exists α such that $\{(\alpha, x), (\alpha, y)\} \subset G$ where $x \neq y$. Then, $\{(\alpha, x), (\alpha, y)\} \subset G|_\alpha$, where $G|_\alpha$ is a set shown to be a function. Since this is a contradiction, G must then be a class function, as claimed.

The proof that G is unique is left as an exercise.

We then define the class function g in the statement as $g = G$.

Concepts review:

1. Which ordering relation well-orders the class, $\mathscr{O}$, of all ordinals.
2. If S is a subset of ordinals in $\mathscr{O}$ what is one way of describing its least element?
3. What can we say about initial segments of the well-ordered class $\mathscr{O}$?
4. Is $\mathscr{O}$ an ordinal number? Why or why not?
5. How do we define the immediate successor of an element of an ordered set?
6. Give an example of a linearly ordered set where no element has an immediate successor.
7. State the two versions of the principle of induction over the ordinals.
8. What does it mean to say that elements of every well-ordered set can be indexed by the elements of some ordinal?
9. Which ZFC axiom is invoked to prove that every well-ordered set is order isomorphic to a single ordinal.
10. What does "ordinality of a well-ordered set" mean?
11. What does Hartogs' lemma state?
12. How does the existence of an uncountable ordinal follow from Hartogs' lemma?
13. What is the least uncountable ordinal?

14. What is the Hartogs number of a set S?
15. How is the concept of Hartogs number combined with the Transfinite recursion theorem to show that there exists an infinite sequence of uncountable ordinal numbers no two of which are equipotent?

EXERCISES

A.
1. Show that a well-ordered set can only be isomorphic to a single ordinal number.
2. We have seen that $\omega_0 + \omega_0$ is a limit ordinal. Describe the smallest limit ordinal which is larger than $\omega_0 + \omega_0$.
3. Is the non-limit ordinal $\omega_0 + 2$ equipotent with the limit ordinal $\omega 2$?
4. What is the smallest ordinal number which is equipotent with $\omega 3$?

B.
5. Let $\mathscr{P}(\mathbb{Q})$ denote the set of all subsets of the set of rational numbers $\mathbb{Q}$.
   a) Construct a countably infinite subset S of $\mathscr{P}(\mathbb{Q})$ which is well-ordered by the relation $\subseteq$ such that $(S, \subseteq)$ is order isomorphic to the ordinal number ω_0. Prove that $\subseteq$ both linearly orders and well-orders the set S.
   b) Construct a countably infinite subset T of $\mathscr{P}(\mathbb{Q})$ which is well-ordered by the relation $\subseteq$ such that $(T, \subseteq)$ is order isomorphic to the ordinal number $\omega_0 + \omega_0$.

C.
6. Theorem 26.7 states that "any two well-ordered sets S and T are either order isomorphic or one is order isomorphic to an initial segment of the other". Can we replace the word "sets" with the word "classes" in this statement. Justify your answer.
7. Construct a set which is not an ordinal number but whose elements can be indexed by the elements of $\omega 5$.
8. Consider the lexicographically well-ordered set $S = \{1, 2, \ldots, 100\} \times \mathbb{N}$. State the ordinal number which is order isomorphic to the subset
$$\{(1,0), (1,1), (1,2), (1,3), \ldots, (2,0)\}$$
Which ordinal number is order isomorphic to S?
9. Let $S = \{\alpha \in \mathscr{O} : |\alpha| \leq |\mathbb{N}|\}$. Does S have a maximal element? If so what is it? If not state why.
10. Show that there is an ordinal number α which is not equipotent with $\mathbb{R}$.
11. Let $S = \{\alpha \in \mathscr{O} : |\alpha| \leq |\omega_1|\}$. Does S have a maximal element? If so what is it? If not state why.

---

29 / Initial ordinals: "Cardinal numbers are us!"

Summary. *In this section we formally define "initial ordinals". We then prove that the class of all initial ordinals is precisely the class* $\mathscr{I} = \omega_0 \cup \{\omega_\alpha : \alpha \in \mathscr{O}\}$ *defined at the end of the previous section. We state and prove (by invoking the Axiom of choice) the Well-ordering theorem. Finally we define the "cardinal numbers" as being initial ordinals.*

29.1 Initial ordinals.

Hartogs' lemma is one of the first statements which links ordinals to sets S in a way that does not depend on the structure of S. It states that "for every set S, $U_S = \{\alpha \in \mathscr{O} : \alpha \not\hookrightarrow_e S \text{ and } \alpha \not\sim_e S\} \neq \varnothing$". Since the class, U_S, of ordinals is non-empty and $\mathscr{O}$ is $\in$-well-ordered, then it has an $\in$- least element. We called this $\in$-least element of U_S the "Hartogs number", $h(S)$, of the set S. Since ordinals are sets, then every ordinal, α, has a Hartogs number, $h(\alpha)$. The Hartogs number of the ordinal, α, can be viewed as being the unique ordinal, $h(\alpha)$, satisfying the following properties:

1. $\alpha \in h(\alpha)$
2. $h(\alpha) \not\sim_e \gamma$, for every ordinal γ in $h(\alpha)$

For example, the Hartogs number of the ordinal, $\omega_0 + \omega_0$, is

$$h(\omega_0 + \omega_0) = \omega_1 = \{\alpha \in \mathscr{O} : \alpha \text{ is countable } \}$$

since ω_1 is not equipotent to $\omega_0 + \omega_0$ nor to any of its elements, and, $\omega_1 \not\sim_e \beta$, for any $\beta \in \{\alpha \in \mathscr{O} : \alpha \text{ is countable } \} = \omega_1$. The Hartogs number of $\omega_0 + \omega_0$ cannot be $\omega_1 + 3$ since $\omega_1 + 3 \sim_e \omega_1 + 1 \in \omega_1 + 3$, contradicting the fact the Hartogs number, $h(\alpha)$, of an ordinal α, cannot be equipotent to any element of $h(\alpha)$. Also, the Hartogs number of any finite ordinal, n, is $h(n) = n + 1$ since $n + 1$ is not equipotent to n or any of its elements, and $n + 1 \not\sim_e m$ for any $m \in n + 1$.

By definition, the Hartogs number, $h(\alpha)$, of an ordinal, α, is never equipotent to any of its elements $\beta \in h(\alpha)$. An ordinal which is not equipotent with any of its elements is given a particular name.

Definition 29.1 We say that an ordinal, β, is an *initial ordinal* if it is the least ordinal which is equipotent with itself. That is, β is an *initial ordinal* if $\alpha \in \beta \Rightarrow \alpha \not\sim_e \beta$.

We already know of many initial ordinals. Trivially, every finite ordinal, n, is the least ordinal equipotent to n since the only ordinal which is equipotent to the natural number n is n. Also, no ordinal $\alpha \in \omega_0$ is equipotent to ω_0, so ω_0 is seen to be the least countably infinite initial ordinal. Next in line is the ordinal, ω_1, shown to be the least ordinal not equipotent to any countable ordinal. It is then, by definition, the second infinite initial ordinal. Of course, $\omega_0 + \omega_0$ is not an initial ordinal since $\omega_0 + 2 \in \omega_0 + \omega_0$ where $\omega_0 + 2 \sim_e \omega_0 \sim_e \omega_0 + \omega_0$.

We will investigate the elements of the class, $\{\omega_\alpha : \alpha \in \mathscr{O}\}$, of recursively constructed ordinals in theorem 28.13. We recall how these ordinals were defined: For a successor ordinal $\alpha = \phi + 1$, ω_α as defined as being $h(\phi)$, while for a limit ordinal γ, ω_γ was defined as, $\omega_\gamma = \text{lub}\{\omega_\alpha : \alpha \in \gamma\}$. The above examples suggest that ordinals such as ω_α are in fact initial ordinals. We introduce the following notation.

$$\mathscr{I} = \{0, 1, 2, 3, \ldots, \} \cup \{\omega_\alpha : \alpha \in \mathscr{O}\}$$

We will show that the class, $\mathscr{I}$, is precisely the class of all initial ordinals.

Lemma 29.2 Every initial ordinal is an element of the class of ordinals, $\mathscr{I}$.

Proof:

Given: ψ is an initial ordinal and $\mathscr{I} = \{0, 1, 2, 3, \ldots, \} \cup \{\omega_\alpha : \alpha \in \mathscr{O}\}$
Required to show: $\psi \in \mathscr{I}$.

If the initial ordinal, ψ, is a finite ordinal, then $\psi \in \mathscr{I}$, and we are done. We then suppose that ψ is an infinite initial ordinal.
By part b) of theorem 28.14, for any ordinal ψ, $\psi \in_= \omega_\psi$. If $\psi = \omega_\psi$, then $\psi \in \mathscr{I}$ and we are done. Suppose $\psi \in \omega_\psi$.

We claim: That $\psi = \omega_\beta \in \omega_\psi$, for some ordinal β (and so $\psi \in \mathscr{I}$).

The proof of the claim is by transfinite induction.
Let $P(\alpha)$ denote the statement "If ψ is an infinite initial ordinal in ω_α, then $\psi = \omega_\beta \in \omega_\alpha$, for some ordinal β".

Base case: If ψ is an infinite initial ordinal which belongs to ω_1, then $\psi = \omega_0 \in \omega_1$. So $P(1)$ holds true.

First inductive hypothesis: Suppose $P(\alpha)$ holds true for some ordinal α. That is, for any infinite initial ordinal $\psi \in \omega_\alpha$, $\psi = \omega_\beta \in \omega_\alpha$, for some ordinal β". We are required to prove that $P(\alpha^+)$ holds true. Suppose $\psi \in \omega_{\alpha^+} = h(\omega_\alpha)$. Since both ω_α and ψ belong to ω_{α^+} either $\psi \in_= \omega_\alpha$ or $\omega_\alpha \in \psi$.

- Case 1: If $\psi = \omega_\alpha$, then we are done.
- Case 2: If $\psi \in \omega_\alpha$, then by the inductive hypothesis, there exists some ordinal β such that $\psi = \omega_\beta \in \omega_\alpha$.

– Case 3: If $\omega_\alpha \in \psi$, then, since ψ is an initial ordinal, $\omega_\alpha \not\sim_e \psi$. Then $\omega_{\alpha^+} = h(\omega_\alpha) \in_= \psi$. But this contradicts our hypothesis $\psi \in \omega_{\alpha^+}$. So $\omega_\alpha \notin \psi$. This means, $\psi \in_= \omega_\alpha$.

If $\psi = \omega_\alpha$, then we are done. Otherwise, $\psi \in \omega_\alpha$ implies $\psi = \omega_\beta$ for some β, by our inductive hypothesis. In both cases $\psi \in \mathscr{I}$. So $P(\alpha^+)$ holds true.

Second inductive hypothesis: Suppose γ is a limit ordinal and $P(\alpha)$ holds true, for all $\alpha \in \gamma$. We are required to show that $P(\gamma)$ holds true. Let ψ be an infinite initial ordinal such that $\psi \in \omega_\gamma = \text{lub}\{\omega_\alpha : \alpha \in \gamma\}$. Then $\psi \in \omega_\alpha$, for some $\alpha \in \gamma$. By the inductive hypothesis, there exists an ordinal β such that $\psi = \omega_\beta \in \omega_\alpha$. So $\psi \in \mathscr{I}$. Hence, $P(\gamma)$ holds true.

This proves the claim that $(\psi \in \omega_\psi) \Rightarrow (\psi = \omega_\beta \in \omega_\psi)$ for some ordinal β. Hence, $\psi \in \mathscr{I}$, as required.

Theorem 29.3 The class, $\mathscr{I} = \{0, 1, 2, 3, \ldots, \} \cup \{\omega_\alpha : \alpha \in \mathscr{O}\}$, is *precisely* the class of all initial ordinals.

Proof:

We have shown in the lemma that $\{\alpha \in \mathscr{O} :$ α is an initial ordinal $\} \subseteq \mathscr{I}$. To show that $\{\alpha \in \mathscr{O} :$ α is an initial ordinal $\} = \mathscr{I}$, it suffices to show that $\mathscr{I} \subseteq \{\alpha \in \mathscr{O} :$ α is an initial ordinal $\}$. Since we have already shown that the finite ordinals are initial ordinals it suffices to show that $\{\omega_\alpha : \alpha \in \mathscr{O}\} \subseteq \{\alpha \in \mathscr{O} :$ α is an initial ordinal $\}$. We prove the statement by transfinite induction. Let $P(\alpha)$ denote the statement "ω_α is an initial ordinal".

– Trivially, $P(0)$ holds true.
– Suppose $P(\alpha)$ holds true. That is, suppose ω_α is an initial ordinal. We are required to show that $\omega_{\alpha^+} = h(\omega_\alpha)$ is an initial ordinal. Let $\beta \in h(\omega_\alpha)$. It suffices to show that $\beta \not\sim_e h(\omega_\alpha)$. Recall that $h(\omega_\alpha)$ is the least ordinal which is not equipotent to ω_α or any of its elements. If $\beta \in h(\omega_\alpha)$, then the ordinal β must be equipotent to some subset of $\omega_\alpha \in \omega_{\alpha^+}$. Then β cannot be equipotent to $h(\omega_\alpha)$. This means that $\omega_{\alpha^+} = h(\omega_\alpha)$ is an initial ordinal. So $P(\alpha^+)$ holds true.
– Suppose γ is a limit ordinal and $P(\alpha)$ holds true for all $\alpha \in \gamma$. That is, ω_α is an initial ordinal for all $\alpha \in \gamma$. We are required to show that ω_γ is an initial ordinal. Suppose not. Suppose $\omega_\gamma \sim_e \beta$ for some $\beta \in \omega_\gamma = \text{lub}\{\omega_\alpha : \alpha \in \gamma\}$. Then $\beta \in_= \omega_\alpha$ for some $\alpha \in \gamma$. Since $\omega_\gamma \sim_e \beta \in_= \omega_\alpha \in \omega_{\alpha^+} \in \omega_\gamma$, we have a contradiction. So ω_γ is an initial ordinal.

By transfinite induction, every element of $\{\omega_\alpha : \alpha \in \mathscr{O}\}$ is an initial ordinal.

We conclude that $\mathscr{I} = \omega_0 \cup \{\omega_\alpha : \alpha \in \mathscr{O}\}$ precisely represents the class of all initial ordinals.

29.3 The Well-ordering theorem.

We say that a set, S, is *well-orderable* if a well-ordering of S is known to exist. For example, we have shown that every countable set is well-orderable. On the other hand, other than the uncountable ordinals such as $\omega_1, \omega_3 + 4, \omega_5$, we have not witnessed a single uncountable well-orderable set. To show that a set, S, is well-orderable and to actually produce an algorithm that well-orders S are two different things. Of course, producing an algorithm that well-orders S is more useful than simply proving that S is well-orderable. But sometimes, the best we can hope for is to prove that a set S is well-orderable, even though we may be convinced that no algorithm that well-orders S will ever be found.

It may come as a surprise to many readers to learn that in the set-theoretic universe governed by *ZFC*, all sets are well-orderable (including uncountable ones such as $\mathbb{R}$). The statement "All sets are well-orderable" proved below is called the *Well-ordering theorem* or the *Well-ordering principle*. It is a direct consequence of the Axiom of choice.

Since the Axiom of choice plays a fundamental role in the proof of the Well-ordering theorem, it will be useful to remind ourselves of what the Axiom of choice states.

> *Axiom of choice*: Every set of sets has a choice function.

The Axiom of choice states that if $\mathscr{S} = \{X : X \in \mathscr{P}(S), X \neq \varnothing\}$ is an infinite set of non-empty subsets of a set S, then there exists a function $f : \mathscr{S} \to \cup\{X : X \in \mathscr{P}(S)\}$ which maps each non-empty subset X of S to some element $x \in X$. The key point is that there is no globally stated rule which states which element is chosen from each set in an infinite set of sets. A convenient (often used) example is to imagine an infinite set $\mathscr{S}$ whose elements are "pairs of socks". The Axiom of choice states that their exists a function f which associates to each pair one sock... even if, in each pair, one is indistinguishable from the other.

Theorem 29.4 [AC] *The Well-ordering theorem.* Every set can be well-ordered.

Proof:

What we are given: That S is a non-empty set.
What we are required to show: That S is well-orderable.
To do this, it suffices to show that S is the one-to-one image of some ordinal number. Then, by invoking theorem 26.1, we can conclude that S is well-orderable.

Case 1: Suppose S is a countable set. If S is finite, then it is the one-to-one image of some finite ordinal (natural number), and so S is well-orderable. If S is infinite, then it is the one-to-one image of $\mathbb{N}$ (a well-ordered set). Again, we must conclude that S is well-orderable.

Case 2: Suppose that S is uncountable. We will recursively construct a function $g : \alpha \to S$ which maps some ordinal α one-to-one onto S. If such a function g is shown to exist, then S is the one-to-one image of a well-ordered set and so can be declared to be well-orderable.

Let $\mathscr{P}(S)^* = \mathscr{P}(S) - \varnothing$. By the Axiom of choice there exists a function $f : \mathscr{P}(S)^* \to S$ which maps each element $X \in \mathscr{P}(S)^*$ to some element $x \in X \subseteq S$. We recursively construct a function $g : \mathscr{O} \to S$ as follows:

$$\begin{aligned}
g(0) &= s_0 \in S, \text{ for an arbitrarily chosen element } s_0 \text{ in } S \\
g(1) &= s_1 = f(S - g[\{0\}]) = f(S - \{s_0\}) \\
g(2) &= s_2 = f(S - g[\{0,1\}]) = f(S - \{s_0, s_1\}) \\
&\vdots \quad \vdots \quad \vdots \\
g(\alpha^+) &= s_{\alpha^+} = f(S - g[\{0,1,2,\ldots,\alpha\}]) = f(S - \{s_0, s_1, \ldots, s_\alpha\}),\ \forall \alpha \in \mathscr{O} \\
g(\beta) &= s_\beta = f(S - \{g(\alpha) : \alpha \in \beta\}) = f(S - \{s_\alpha : \alpha \in \beta\}),\ \forall \text{ limit ordinals } \beta
\end{aligned}$$

For each γ in the domain, $\operatorname{dom} g$, of g define $g|_\gamma$ as: $g|_\gamma = \{(\alpha, s_\alpha) \in g : \alpha \in \gamma\}$.
Claim: For each ordinal, γ, such that $\gamma \subseteq \operatorname{dom} g$, $g|_\gamma$ is a one-to-one function on γ.

The proof of the claim is by transfinite induction. Let $P(\alpha)$ denote the statement "$g|_\alpha : \alpha \to S$ is a one-to-one function mapping α into S".

Inductive hypothesis: Suppose $\gamma \subseteq \operatorname{dom} g$. Suppose $P(\alpha)$ holds true for all $\alpha \in \gamma$, where γ belongs to the domain of g. We are required to show that $P(\gamma)$ holds true. That is, we are required to show that $g|_\gamma$ is one-to-one on γ.

Suppose (β, s_β) and (μ, s_μ) are two elements in $g|_\gamma$ such that $\beta \in \mu$. Then β and μ are elements of γ. It suffices to show that $s_\beta \neq s_\mu$. Case 1: If γ is a limit ordinal, then (β, s_β) and (μ, s_μ) belong to $g|_{\mu^+} \subset g|_\gamma$. By the inductive hypothesis, $g|_{\mu^+}$ is one-to-one on μ^+; hence, $s_\beta \neq s_\mu$. Case 2: Suppose γ is a successor ordinal. If $\mu^+ \neq \gamma$, then by the inductive hypothesis, $g|_{\mu^+}$ is one-to-one on μ and, since (β, s_β) and (μ, s_μ) belong to $g|_{\mu^+}$, then $s_\beta \neq s_\mu$. Suppose $\mu^+ = \gamma$. By definition of g, $g(\mu) = s_\mu = f(S - \{s_\alpha : \alpha \in \mu\})$. Since $\beta \in \mu$, $s_\beta \notin S - \{s_\alpha : \alpha \in \mu\}$; hence, $s_\beta \neq f(S - \{s_\alpha : \alpha \in \mu\}) = s_\mu$. Then $g|_\gamma$ is one-to-one on γ.
By transfinite induction, $g|_\gamma$ is one-to-one on γ, for all $\gamma \subseteq \operatorname{dom} g$, as claimed. We conclude that g is one-to-one on $\operatorname{dom} g$.

Claim : The function g maps $\operatorname{dom} g$ onto S. That is, for every $s \in S$, $(\alpha, s) \in g$ for some ordinal α.
Let D denote the domain of g. To prove the claim it suffices to show that $S - g[D] = \varnothing$.

- We first show that D satisfies the initial segment property: If $\gamma \subseteq D$, then $g|_\gamma = \{(\alpha, s_\alpha) \in g : \alpha \in \gamma\} \subseteq g$. Hence, $\alpha \in \gamma \Rightarrow (\alpha, s_\alpha) \in g \Rightarrow \alpha \in D$. Then D satisfies the initial segment property in $\mathscr{O}$.
- The domain D of g is a set: Since S is a set and $g : D \to S$ is a one-to-one function on $D \subset \mathscr{O}$, the image $g^{-1}[S]$ of the one-to-one function $g^{-1} : S \to \mathscr{O}$ must be a set (by the Axiom of replacement).
- The domain D of g is an ordinal: We have shown that the domain D of g is a set in $\mathscr{O}$ which satisfies the *initial segment property*. Then $D \neq \mathscr{O}$ and so there exists an ordinal δ such that $D = S_\delta = \{\alpha \in \mathscr{O} : \alpha \in \delta\} = \delta$ (by theorem 26.3, also see page 273). So we can write $D = \delta$.
- It now suffices to show that $S - g[\delta]$ is empty. Suppose the set $S - g[\delta]$ is non-empty. Then the choice function f maps the non-empty set $S - g[\delta]$ to some element s_δ in $S - g[\delta]$. Then $g(\delta) = s_\delta$, which means that $\delta \in D = \delta$, a contradiction. The source of the contradiction is the assumption that $S - g[\delta]$ is non-empty. Then $g[D] = g[\delta] = S$. So every element of S is in the image of the function g, as claimed.

The function $g : \delta \to S$ mapping δ one-to-one onto $S = \{s_\alpha : \alpha \in \delta\}$ then induces a well-ordering on S. We conclude that any non-empty set S is well-orderable.

29.4 Defining the cardinal numbers.

Overview – We are now set to formally define the sets we call "cardinal numbers". First, we review some background material on how we came to discuss the concept of "cardinal numbers".

Given the class, $\mathscr{S}$, of all sets, we defined a relation $\sim_e$ on $\mathscr{S}$ as follows: $S \sim_e T$ if and only if S and T are equipotent. The relation, $\sim_e$, was shown to be an equivalence relation on $\mathscr{S}$ and so allows us to partition $\mathscr{S}$ into a class of equivalence classes. Every set $S \in \mathscr{S}$ then belongs to some equivalence class of pairwise equipotent sets. We represented an equivalence class containing a set S as $[S]_e = \{T \in \mathscr{S} : T \sim_e S\}$. For example, $[\mathbb{R}]_e$ and $[\mathbb{N}]_e$ are equivalence classes containing all sets which are equipotent to $\mathbb{R}$ and $\mathbb{N}$ respectively. Once we had verified that $\mathscr{P}(\mathbb{N}) \sim_e \mathbb{R}$, for example, we could write that $[\mathbb{R}]_e = [\mathscr{P}(\mathbb{N})]_e$ where $\mathbb{R}$ and $\mathscr{P}(\mathbb{N})$ were simply different representatives of the same equivalence class. When we first discussed the concept of "cardinal numbers", the tools available at that time were insufficient to construct a class of sets whose elements could serve as representatives for each of the equipotence-induced equivalence classes.[1] So we postulated the existence of the class of *cardinal numbers* as follows (reproduced from Postulate 22.2):

[1] One may suggest that we could have defined, $[S]_e$, as being the cardinal number of S. The problem with this is that we want cardinal numbers to be sets; the $\sim_e$-equivalence class, $[S]_e$, may be a proper class.

There exists a class $\mathscr{C}$ of sets such that 1) $\mathbb{N} \subset \mathscr{C}$, 2) every set S is equipotent with exactly one element $\kappa \in \mathscr{C}$.

We now have all the ingredients required to prove that a class of sets whose properties characterize the cardinal numbers exists in *ZFC*.

Theorem 29.5 The class of all initial ordinals, $\mathscr{I} = \omega_0 \cup \{\omega_\alpha : \alpha \in \mathscr{O}\}$, satisfies the following properties:

1. Every set S is equipotent to exactly one element in $\mathscr{I}$.
2. Two sets, S and T, are equipotent if and only if they are equipotent to the same element of $\mathscr{I}$.
3. The class, $\mathscr{I}$, is $\in$-linearly ordered.

Proof:

1) Let S be a set. By theorem 29.4, the set S has a well-ordering "$<$". When equipped with this well-ordering, $(S, <)$ is order isomorphic to some ordinal β (since every well-ordered set is order isomorphic to some ordinal number). We know there exists a unique initial ordinal ω_α which is equipotent with β. Then ω_α is the unique initial ordinal which is equipotent to the set S.

2) Suppose S and T are equipotent. Then by part 1) both S and T are equipotent to initial ordinals ω_γ and ω_α respectively. Since distinct initial ordinals cannot be equipotent, $\omega_\gamma = \omega_\alpha$. Conversely, equipotent initial ordinals must be the same ordinal and so the elements of the class of all sets which are equipotent to the same initial ordinal are pairwise equipotent.

3) Every pair of ordinals in $\mathscr{O}$ are $\in$-comparable. So every pair of ordinals in $\mathscr{I} \subset \mathscr{O}$ must be $\in$-comparable. So $\mathscr{I}$ is $\in$-linearly ordered.

Definition 29.6 *Cardinal numbers.* An ordinal is called a *cardinal number* if and only if this ordinal is an initial ordinal. When the elements of $\mathscr{I}$ are viewed as cardinal numbers we represent $\mathscr{I}$ as $\mathscr{C}$.

29.5 Aleph notation.

Although we could use the "ω_α" notation to represent the *cardinal numbers* it is customary to use the aleph notation, $\aleph_\alpha$. That is, we set $\aleph_0 = \omega_0$, $\aleph_1 = \omega_1$, $\aleph_2 = \omega_2$, more generally, for any ordinal α, $\aleph_\alpha = \omega_\alpha$. For example, when we write the expression $\aleph_1$ we are thinking "the cardinal number $\aleph_1$" rather than "the initial ordinal ω_1"

even though they represent the same entity. Since the initial ordinal, $\aleph_1 = \omega_1$, is, by definition, "the least ordinal number which is not countable" it is the first uncountable cardinal (ordinal). If we assume the Continuum hypothesis, then there exists no cardinal number, $\aleph_\alpha$, such that $|\mathbb{N}| = \aleph_0 \hookrightarrow_e \aleph_\alpha \hookrightarrow_e |\mathscr{P}(\mathbb{N})| = |\mathbb{R}| = c$ (theorem 21.3); hence, $\aleph_1 = \omega_1 = c$ is the least ordinal which is not countable.

Similarly, assuming the Generalized continuum hypothesis, the least uncountable ordinal which is not equipotent to ω_1 is the cardinality of the set $\mathscr{P}(\mathbb{R})$, in which case we say "assuming GCH, the cardinality of $\mathscr{P}(\mathbb{R})$ is $\aleph_2$" where $\aleph_2$ is the least ordinal which is not equipotent to $\aleph_1$.

The class of all cardinal numbers $\mathscr{C}$ can now be represented as

$$\mathscr{C} = \aleph_0 \cup \{\aleph_\alpha : \alpha \in \mathscr{O}\}$$

where, for all ordinals α, $\aleph_{\alpha^+}$ is the least ordinal which is not equipotent with $\aleph_\alpha$ and, for limit ordinals γ, $\aleph_\gamma = \text{lub}\{\aleph_\alpha : \alpha \in \gamma\}$. What is also new, is that the class of all infinite cardinal numbers is indexed by the ordinal numbers.

Even though every cardinal number is an initial ordinal, the context usually allows us to determine whether we are referring to a set's "cardinality" or "ordinality". We normally refer to the "ordinality of a set" only if we have a specific (or hypothesized) well-ordering of that set in mind. In the following tables, we describe how the cardinality, and ordinality of a set are perceived when we are assuming the Generalized continuum hypothesis and when we are not.

Without assuming CH nor GCH (in the presence of the Axiom of choice):

Set S	cardinality of S	initial ordinal of S	${}^{\text{ord}}S$
$\{\ \}$	0	0	0
$\{a, b, c\}$	3	3	3
$\mathbb{N}_{\text{standard}}$	$\aleph_0$	ω_0	ω_0
$\omega_0 + 3_{\in\text{-well-ordered}}$	$\aleph_0$	ω_0	$\omega_0 + 3$
$\{1, 2\} \times \mathbb{N}_{\text{lexico}}$	$\aleph_0$	ω_0	$\omega_0 2$
$\mathbb{N} \times \mathbb{N}_{\text{lexico}}$	$\aleph_0$	ω_0	$\omega_0\omega_0$
ω_1	$\aleph_1$	ω_1	ω_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbb{R}$	$c = 2^{\aleph_0} = \aleph_\alpha \geq \aleph_1$	ω_α	
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathscr{P}(\mathbb{R})$	$2^{\aleph_\alpha} = \aleph_\beta \geq \aleph_{\alpha+1}$	ω_β	
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathscr{P}(\mathscr{P}(\mathbb{R}))$	$2^{\aleph_\beta} = \aleph_\gamma \geq \aleph_{\beta+1}$	ω_γ	
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathscr{P}(\mathscr{P}(\mathscr{P}(\mathbb{R})))$	$2^{\aleph_\gamma} = \aleph_\delta \geq \aleph_{\gamma+1}$	ω_δ	

Note that it is the Well-ordering theorem (itself a consequence of the Axiom of choice) which guarantees that for every ordinal, β, $2^{\aleph_\beta} = |2^{\aleph_\beta}| = |\mathscr{P}(\aleph_\beta)|$ is equal to some (initial) ordinal number $\omega_\gamma = \aleph_\gamma \geq \aleph_{\beta+1}$. That is, the Well-ordering theorem states that $\mathscr{P}(\aleph_\beta) \sim_e \aleph_\gamma$, for some cardinal $\aleph_\gamma \geq \aleph_{\beta+1}$. Without the Axiom of choice, the set, $\mathscr{P}(\aleph_\beta)$, may not be well-orderable in which case it would not necessarily be equipotent to some ordinal number.

Assuming GCH, cardinal numbers are more clearly defined:

Set S	cardinality of S	initial ordinal of S	${}^{\text{ord}}S$
$\{a, b, c\}$	3	3	3
$\mathbb{N}_{\text{standard}}$	$\aleph_0$	ω_0	ω_0
$\omega_0 + 3_{\in\text{-well-ordered}}$	$\aleph_0$	ω_0	$\omega_0 + 3$
$\{1, 2\} \times \mathbb{N}_{\text{lexico}}$	$\aleph_0$	ω_0	$\omega_0 2$
$\mathbb{N} \times \mathbb{N}_{\text{lexico}}$	$\aleph_0$	ω_0	$\omega_0\omega_0$
$\mathbb{R}$	$2^{\aleph_0} = \aleph_1$	ω_1	
$\mathbb{R} \times \mathbb{R}$	$\aleph_1$	ω_1	
$\mathscr{P}(\mathbb{R})$	$2^{\aleph_1} = \aleph_2$	ω_2	
$\mathscr{P}(\mathscr{P}(\mathbb{R}))$	$2^{\aleph_2} = \aleph_3$	ω_3	
$\mathscr{P}(\mathscr{P}(\mathscr{P}(\mathbb{R})))$	$2^{\aleph_3} = \aleph_4$	ω_4	
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\vdots$	$\aleph_\alpha$	ω_α	
$\vdots$	$\vdots$	$\vdots$	$\vdots$

What does this say about GCH? The Axiom of choice guarantees that every set can be well-ordered and so all sets can be ranked on an "equipotence based scale $\mathscr{C}$ of sets" called the cardinal numbers. This means that every set is associated to a uniquely specified ordinal number (cardinal number) on this scale of ordinals. We make the following universe comparisons.

In the ZFC$-$ *universe* : For any infinite set S such that $|S| = \aleph_\gamma$, $2^S \sim_e \mathscr{P}(S) \sim_e \aleph_\alpha$ for some $\alpha > \gamma$. The value of α is guaranteed to exist, but cannot be determined. The value of α is simply assumed to be equal to some ordinal greater than or equal to $\gamma + 1$.

In the ZFC + *GCH*$-$ *universe*: If S is any infinite set such that $|S| = \aleph_\gamma$, then $2^S \sim_e \mathscr{P}(S) \sim_e \aleph_{\gamma+1}$. The cardinality of the set $2^S \sim_e \mathscr{P}(S)$ is the least cardinal number (on the equipotence based scale) which is larger than $\aleph_\gamma$. The axiom GCH limits the size of power sets $\mathscr{P}(S)$ relative to the size of S.

In the ZFC + *CH*$-$ *universe*: For any infinite set S such that $|S| = \aleph_\gamma > \aleph_0$, $2^S \sim_e \mathscr{P}(S) \sim_e \aleph_\alpha$ for some $\alpha > \gamma$. The value of α is guaranteed to exist, but

cannot be determined. It is equal to some ordinal greater than or equal to $\gamma + 1$. But if $|S| = \aleph_0$, $2^S \sim_e \mathscr{P}(S)$ is the immediate successor cardinal, $\aleph_1$, of $\aleph_0$. So CH only limits the the size of $2^{\aleph_0}$.

In the ZFC $+ \neg CH -$ *universe* : For any infinite set S such that $|S| = \aleph_\gamma > \aleph_0$, $2^S \sim_e \mathscr{P}(S) \sim_e \aleph_\alpha$ for some $\alpha > \gamma$. The value of α is guaranteed to exist, but cannot be determined. It is equal to some ordinal greater than or equal to $\gamma + 1$. But if $|S| = \aleph_0$, $2^S \sim_e \mathscr{P}(S)$ is *not* the immediate successor cardinal, $\aleph_1$, of $\aleph_0$. That is, $2^{\aleph_0} > \aleph_1$.

Ranking the elements of the class $\mathscr{S}$ of all sets in ZFC with $\hookrightarrow_e$. Suppose S and T are two infinite sets in $\mathscr{S}$. If S and T are equipotent, then they can be viewed as being the "same size" (just like a set of five squirrels and a set of five submarines are viewed as being the same size, in spite of the fact that squirrels and submarines are entirely different types of objects). Suppose now that S and T are *not* equipotent. We wonder whether one of these two infinite sets is necessarily embedded in the other. We show that it must be the case. Say the relation $<_S$ well-orders S and $<_T$ well-orders T. (The Well-ordering theorem guarantees that such well-orderings exist for each of these two sets.) Suppose $\alpha =^{\text{ord}} S$ and $\beta =^{\text{ord}} T$. Now α and β cannot be the same ordinal number for if they were, then S and T would be equipotent. Then, either $\alpha \in \beta$ or $\beta \in \alpha$. Suppose, without loss of generality, that $\alpha \in \beta$. The set S is order isomorphic to an initial segment of T. It follows that that $S \hookrightarrow_e T$. We have shown that any pair of non-equipotent sets S and T are $\hookrightarrow_e$-comparable. Note that comparing sizes of sets in this way would not be possible without the Well-ordering theorem (which follows from the Axiom of choice).

Finally, since the class, $\{\aleph_\alpha : \alpha \in \mathscr{O}\}$, of all infinite cardinal numbers is indexed by the elements of $\mathscr{O}$, no two of which are equipotent, we can then say that our universe of sets contains as many different infinite set sizes as there are ordinals!

29.6 Some consequences of the cardinal number definition, $\mathscr{C} = \mathscr{I}$.

Knowing that the cardinal numbers are in fact initial ordinals (which in turn are known to be limit ordinals) will allow us to prove certain properties of cardinal numbers, a task which was beyond our reach unless we knew more about the nature of the those sets we call "cardinal numbers". We first remind ourselves of the definition of the "product of two cardinal numbers": If κ and λ are cardinal numbers then $\kappa \times \lambda = |K \times L|$ where $\kappa = |K|$ and $\lambda = |L|$.

We have already provided a few properties of cardinal number multiplication in theorem 23.4. In particular, we showed that $\aleph_0 \times \aleph_0 = \aleph_0$. We can now move a step further by showing that for any ordinal α, $\aleph_\alpha \times \aleph_\alpha = \aleph_\alpha$. We begin with the following

lemma.

Lemma 29.7 For any infinite cardinal number, κ, define a relation $<_*$ on $\kappa \times \kappa$ as follows: For pairs, (α, β) and (γ, ψ), of ordered pairs in $\kappa \times \kappa$,

$$(\alpha, \beta) <_* (\gamma, \psi) \text{ whenever } \begin{cases} \alpha \cup \beta \in \gamma \cup \psi \\ \text{or} \\ \beta \in \psi \text{ when } \alpha \cup \beta = \gamma \cup \psi \\ \text{or} \\ \alpha \in \gamma \text{ when } \alpha \cup \beta = \gamma \cup \psi \text{ and } \beta = \psi \end{cases}$$

Then $<_*$ well-orders $\kappa \times \kappa$.

Proof:

The relation $<_*$ is easily verified (on a case by case basis) to be transitive on $\kappa \times \kappa$. It is also easily verified that if $(\alpha, \beta) \not<_* (\gamma, \psi)$ and $(\gamma, \psi) \not<_* (\alpha, \beta)$, $(\alpha, \beta) = (\gamma, \psi)$ and so $<_*$ linearly orders $\kappa \times \kappa$.

Claim: Every non-empty subset of $\kappa \times \kappa$ contains a least ordinal pair element with respect to $<_*$.

Proof of claim: Let U be a non-empty subset of $\kappa \times \kappa$. Consider the set

$$A = \{\delta : \ \delta = \alpha \cup \beta \text{ for some } (\alpha, \beta) \in U\}$$

If $(\mu, \gamma) \in U$, then $\mu \cup \gamma \in A$, so the set A is non-empty. Since A is a subset of the well-ordered class $\mathscr{O}$, there exists an $\in$-least element $\mu_A \in A$. Let $V_A = U \cap \{(\alpha, \beta) : \alpha \cup \beta = \mu_A\}$. Let

$$B = \{\delta : \ (\alpha, \delta) \in V_A\}$$

The set B is non-empty and so there exists an $\in$-least element $\mu_B \in B$. Let $V_B = V_A \cap \{(\alpha, \beta) : \beta = \mu_B\}$. Let

$$C = \{\delta : \ (\delta, \mu_B) \in V_B\}$$

The set C is non-empty and so there exists an $\in$-least element $\mu_C \in C$. See that the element (μ_C, μ_B) is the $<_*$-least element of U.

Then every non-empty subset U of $\kappa \times \kappa$ has an $<_*$-least element, as claimed.

Then $<_*$ well-orders the ordinal pairs in the set $\kappa \times \kappa$.

For example,

$$\begin{array}{rcl}(\aleph_1, \aleph_2+5) \leq_* (\aleph_2, \aleph_3) & \text{since} & \aleph_1 \cup (\aleph_2+5) = (\aleph_2+5) \in \aleph_3 = \aleph_2 \cup \aleph_3 \\ (\aleph_7, \aleph_3+9) \leq_* (\aleph_7, \aleph_4) & \text{since} & \aleph_7 \cup (\aleph_3+9) = \aleph_7 = \aleph_7 \cup \aleph_4 \text{ and } (\aleph_3+9) \in \aleph_4 \\ (\aleph_3, \aleph_7) \leq_* (\aleph_7, \aleph_7) & \text{since} & \aleph_3 \cup \aleph_7 = \aleph_7 = \aleph_7 \cup \aleph_7 \text{ and } \aleph_7 \in \aleph_7 \text{ and } \aleph_3 \in \aleph_7\end{array}$$

The above lemma shows that for any infinite cardinal number κ, $\kappa \times \kappa$ is well-orderable and so any product, $\kappa \times \kappa$, of an infinite cardinal κ with itself is order isomorphic to some ordinal number μ.

Theorem 29.8 [AC] For any ordinal α, $\aleph_\alpha \times \aleph_\alpha = \aleph_\alpha$.

Proof:

Since we have previously shown that the product of infinite countable sets is countable, then $\aleph_0 \times \aleph_0 = \aleph_0$. We will prove the general statement by transfinite induction.

Inductive hypothesis: Suppose $\aleph_\alpha \times \aleph_\alpha = \aleph_\alpha$ for all $\alpha \in \delta$ for some ordinal δ. We are required so show that $\aleph_\delta \times \aleph_\delta = \aleph_\delta$. By the lemma, $<_*$ well-orders $\aleph_\delta \times \aleph_\delta$ and so $\aleph_\delta \times \aleph_\delta$ is order isomorphic to some ordinal μ. It will suffice to show that $\aleph_\delta = \mu$. Since

$$\aleph_\delta \sim_e \aleph_\delta \times \{1\} \hookrightarrow_e \aleph_\delta \times \aleph_\delta \sim_e \mu$$

then $\aleph_\delta \in_= \mu$. To show that $\aleph_\delta = \mu$, it suffices to show that $\aleph_\delta \in \mu$ is impossible.

Claim: $\aleph_\delta \in \mu$ is impossible. Suppose $\aleph_\delta \in \mu$. Suppose $f : \mu \to \aleph_\delta \times \aleph_\delta$ is the order isomorphism mapping μ onto $\aleph_\delta \times \aleph_\delta$. Then, since $\aleph_\delta \in \mu$,

$$f(\aleph_\delta) = (\alpha_0, \beta_0) \in f(\mu) = \aleph_\delta \times \aleph_\delta$$

Since α_0 and β_0 are ordinals in $\aleph_\delta$, then $\alpha_0 \cup \beta_0$ is an ordinal in $\aleph_\delta$, itself a limit ordinal. So $\alpha_0 \cup \beta_0 + 1 \in \aleph_0$. Since $\aleph_\delta$ is an initial ordinal, $|\alpha \cup \beta + 1| \in \aleph_\delta$. By our induction hypothesis,

$$|\alpha \cup \beta + 1| \times |\alpha \cup \beta + 1| = |\alpha \cup \beta + 1| \in \aleph_\delta$$

Since $(\alpha_0, \beta_0) \leq_* (\alpha_0 \cup \beta_0, \alpha_0 \cup \beta_0) <_* (\alpha_0 \cup \beta_0 + 1, \alpha_0 \cup \beta_0 + 1)$, then

$$f[\aleph_\delta] = \{(\alpha, \beta) \in \aleph_\delta \times \aleph_\delta : (\alpha, \beta) <_* (\alpha_0, \beta_0)\} \subset (\alpha_0 \cup \beta_0 + 1) \times (\alpha_0 \cup \beta_0 + 1)$$

We then have:

$$\begin{array}{rcl}\aleph_\delta & \hookrightarrow_e & (\alpha \cup \beta + 1) \times (\alpha \cup \beta + 1) \\ & \sim_e & |(\alpha \cup \beta + 1) \times (\alpha \cup \beta + 1)| \\ & = & |\alpha \cup \beta + 1| \times |\alpha \cup \beta + 1| \quad \text{(22.3)} \\ & = & |\alpha \cup \beta + 1| \in \aleph_\delta \quad \text{(Inductive hypothesis)}\end{array}$$

We have obtained the contradiction $\aleph_\delta \hookrightarrow_e \aleph_\delta$. Then $\aleph_\delta \in \mu$ is impossible, as claimed.
So $\mu = \aleph_\delta$. That is, $\mu \times \mu = \aleph_\delta \times \aleph_\delta = \aleph_\delta = \mu$.
By transfinite induction, for any ordinal α, $\aleph_\alpha \times \aleph_\alpha = \aleph_\alpha$, as required.

Corollary 29.9 Let κ be an infinite cardinal and $\{A_\alpha : \alpha \in \beta\}$ be a set of non-empty sets indexed by the elements of the ordinal $\beta \in_= \kappa$ where $|A_\alpha| \in_= \kappa$ for all $\alpha \in \beta$. Then $|\cup\{A_\alpha : \alpha \in \beta\}| \in_= \kappa$.

Proof:

We are given that κ is an infinite cardinal number, $|\beta| \in_= \kappa$ and for each $\alpha \in \beta$, the cardinality of $|A_\alpha| \in_= \kappa$, A_α is non-empty.
Since, for each $\alpha \in \beta$, $|A_\alpha| \in_= \kappa$, then for each $\alpha \in \beta$, there exists an isomorphism, $f_\alpha : A_\alpha \to \kappa$, mapping A_α one-to-one into the initial ordinal κ.
For each $x \in A_\alpha$, let $\delta(x) =$ least of $\{\alpha \in \beta : x \in A_\alpha\}$.

We define the function $h : \cup\{A_\alpha : \alpha \in \beta\} \to \kappa \times \kappa$ as follows:

$$h(x) = (\delta(x), f_{\delta(x)}(x)) \in \kappa \times |A_\delta| \in_= \kappa \times \kappa$$

Claim: That h is one-to-one on $\cup\{A_\alpha : \alpha \in \beta\}$.

$$\begin{aligned} h(x) = h(y) &\Rightarrow (\delta(x), f_{\delta(x)}(x)) = (\delta(y), f_{\delta(y)}(y)) \\ &\Rightarrow \delta(x) = \delta(y) \\ &\Rightarrow f_{\delta(x)}(x) = f_{\delta(x)}(y) = f_{\delta(y)}(y) \\ &\Rightarrow x = y \quad \text{(Since } f_\alpha : A_\alpha \to \kappa \text{ is an isomorphism for all } \alpha.\text{)} \end{aligned}$$

So h maps $\cup\{A_\alpha : \alpha \in \beta\}$ one-to-one into $\kappa \times \kappa$, as claimed.

Then

$$\begin{aligned} |\cup\{A_\alpha : \alpha \in \beta\}| &= |h[\cup\{A_\alpha : \alpha \in \beta\}]| \\ &\in_= |\kappa \times \kappa| \\ &= |\kappa| \times |\kappa| \\ &= \kappa \times \kappa \\ &= \kappa \quad \text{(By the previous theorem.)} \end{aligned}$$

So $|\cup\{A_\alpha : \alpha \in \beta\}| \in_= \kappa$, as required.

Corollary 29.10 For any infinite cardinal number κ, $\kappa^\kappa = 2^\kappa$.

Proof:

Since $2 < \kappa$, $2^\kappa \leq \kappa^\kappa$ (by theorem 24.3).
It now suffices to show that $\kappa^\kappa \leq 2^\kappa$:
Let $f \in \kappa^\kappa$. Then f is function which maps κ into κ. Then $f \subset \kappa \times \kappa = \kappa$ (by theorem 29.8). Then $f \in \mathscr{P}(\kappa)$ where $|\mathscr{P}(\kappa)| = 2^\kappa$. Then $\kappa^\kappa \leq 2^\kappa$, as claimed.
We conclude that $\kappa^\kappa = 2^\kappa$.

29.7 Properties of some large cardinal numbers.

We have seen that the class of all infinite cardinal numbers, $\{\aleph_\alpha : \alpha \in \mathscr{O}\}$, is the class of all initial ordinals $\{\omega_\alpha : \omega \in \mathscr{O}\}$. Recall that limit ordinals are those ordinals γ which do not contain a maximal element (since the maximal element of an ordinal, if it has one, is its immediate predecessor). Equivalently, limit ordinals are those ordinals, γ, which satisfy the properties, $\text{lub}(\gamma) = \gamma$ and $\gamma = \cup\{\alpha : \alpha \in \gamma\}$ (27.15). It is interesting to reflect on properties possessed by some cardinals perceived to be considerably larger in size then others. We must proceed cautiously, since the elements of the class of all cardinal numbers are indexed by the elements of the class of all ordinal numbers (a few of which are themselves cardinal numbers). We start by providing the following definitions.

Definition 29.11

a) We say that an infinite cardinal number, $\aleph_\gamma$, is a *successor cardinal* if the index, γ, has an immediate predecessor (i.e., $\gamma = \beta + 1$, for some β). The expression, $\aleph_{\alpha^+} = \aleph_{\alpha+1}$, denotes a successor cardinal. We say that an infinite cardinal number, $\aleph_\gamma$, is a *limit cardinal* if γ is a limit ordinal (i.e., $\gamma = \text{lub}\{\alpha : \alpha \in \gamma\}$).

b) We say that a limit cardinal $\aleph_\gamma$ is a *strong limit cardinal* if $\aleph_\gamma$ is uncountable and $\{2^{\aleph_\alpha} : \alpha \in \gamma\} \subseteq \aleph_\gamma$.

Limit cardinals should not be confused with *limit ordinals*. All infinite cardinals, $\aleph_\alpha$, are limit ordinals. But not all infinite ordinals are limit cardinals. See that

$$\begin{aligned} \aleph_1 &= \aleph_{0+1} = \aleph_{0^+} \\ \aleph_2 &= \aleph_{1+1} = \aleph_{1^+} \\ \aleph_{100} &= \aleph_{99+1} = \aleph_{99^+} \end{aligned}$$

are successor cardinals and so are not limit cardinals, even though they are all limit ordinals. A limit cardinal number is a cardinal number which is not attainable from

below by a finite sequence of successor cardinals. Limit cardinals can be easily recognized as being those infinite cardinals, $\aleph_\gamma$, whose index, γ, is a limit ordinal. We see that the first four *limit cardinals* are $\aleph_0$, $\aleph_{\omega_0}$, $\aleph_{\omega_0+\omega_0}$, $\aleph_{\omega_0+\omega_0+\omega_0}$ since

$$\begin{aligned} n \in \aleph_0 &\Rightarrow n+1 \in \aleph_0 \\ \aleph_n \in \aleph_{\omega_0} &\Rightarrow \aleph_{n+1} \in \aleph_{\omega_0} \\ \aleph_{\omega_0+n} \in \aleph_{\omega_0+\omega_0} &\Rightarrow \aleph_{\omega_0+n+1} \in \aleph_{\omega_0+\omega_0} \\ \aleph_{\omega_0+\omega_0+n} \in \aleph_{\omega_0+\omega_0+\omega_0} &\Rightarrow \aleph_{\omega_0+\omega_0+n+1} \in \aleph_{\omega_0+\omega_0+\omega_0} \end{aligned}$$

On strong limit cardinal numbers. The notion of a strong limit cardinal is a bit more difficult to grasp. Since ω_0 is the first limit ordinal, then the first uncountable limit cardinal is $\aleph_{\omega_0}$. Is $\aleph_{\omega_0}$ "strong"? The answer to this question depends on which assumptions are made. Consider, $5 \in \omega_0$. Is $2^{\aleph_5}$ necessarily an element of $\aleph_{\omega_0}$? For $5 \in \omega_0$, $2^{\aleph_5} \geq \aleph_6$ (since $\left|2^{\aleph_5}\right| = |\mathscr{P}(\aleph_5)| > \aleph_5$, while $\aleph_6 = \omega_6$ is the least ordinal which is not order isomorphic to ω_5). But does $2^{\aleph_5}$ belong to $\aleph_{\omega_0}$? It is possible, but the given inequality doesn't guarantee that $2^{\aleph_5} \in \aleph_{\omega_0}$. However, if we assume GCH, then definitely, $2^{\aleph_5} = \aleph_6 \in \aleph_{\omega_0}$.

This example suggests that when assuming GCH, determining when a limit cardinal is "strong" is pretty straightforward. We prove the following result.

Theorem 29.12 [GCH] Every uncountable limit cardinal is a strong limit cardinal.

Proof:

Suppose $\aleph_\gamma$ is an uncountable limit cardinal, and suppose $\alpha \in \gamma$. Since $\aleph_\gamma$ is a limit cardinal, then, by definition, $\aleph_{\alpha+1} \in \aleph_\gamma$. Assuming GCH, we have $2^{\aleph_\alpha} = \aleph_{\alpha+1} \in \aleph_\gamma$. We have shown that $\alpha \in \gamma \Rightarrow 2^{\aleph_\alpha} \in \aleph_\gamma$. So $\aleph_\gamma$ is a strong limit cardinal, as required.

But what if we don't assume GCH? Are there any strong limit cardinals in the *ZFC*-universe? Interestingly enough, we can show that in the presence of the Axiom of choice, there must be.

Theorem 29.13 [AC] There exists a strong limit cardinal number.

Proof:

Let $\mathscr{C}^* = \{\aleph_\alpha : \alpha \in \mathscr{O}, \alpha \neq 0\}$ denote the class of all uncountable infinite cardinals.

Since we are assuming the Axiom of choice every set is well-orderable and so, for each ordinal α, $\mathscr{P}(\aleph_\alpha)$ is well-orderable and so $\mathscr{P}(\aleph_\alpha)$ has a cardinality, $|\mathscr{P}(\aleph_\alpha)|$. By theorem 20.12, $2^{\aleph_\alpha} \sim_e \mathscr{P}(\aleph_\alpha)$; hence, $2^{\aleph_\alpha}$ is well-orderable and so has a cardinality: $|2^{\aleph_\alpha}| = |2|^{|\aleph_\alpha|} = 2^{\aleph_\alpha} = |\mathscr{P}(\aleph_\alpha)| \in \mathscr{C}^*$.

We define the function $g : \mathscr{C}^* \to \mathscr{C}^*$ as follows:

$$g(\aleph_\alpha) = 2^{\aleph_\alpha}$$

Let $U = \{\kappa_n : n \in \omega_0\} \subset \mathscr{C}^*$ be recursively defined as follows:

$$\begin{aligned}
\kappa_0 &= \aleph_0 \\
\kappa_1 &= g(\kappa_0) = 2^{\kappa_0} \\
\kappa_2 &= g(\kappa_1) = 2^{\kappa_1} \\
\kappa_3 &= g(\kappa_2) = 2^{\kappa_2} \\
&\vdots \\
\kappa_{n+1} &= g(\kappa_n) = 2^{\kappa_n} \\
&\vdots
\end{aligned}$$

Then $\{\kappa_n : n \in \omega_0\}$ is a countable (strictly) increasing sequence of cardinals. It follows that $\gamma = \cup\{\kappa_n : n \in \omega_0\}$ is an infinite ordinal which is the least upper bound of the set U (theorem 27.12).

Claim: The ordinal γ is an infinite cardinal number.

It suffices to show that γ is an initial ordinal. Suppose β is an ordinal in γ. It suffices to show that $|\beta| < |\gamma|$. Since $\beta \in \gamma$, then, for some m,

$$\beta \in \kappa_m = 2^{\kappa_{m-1}} \in 2^{\kappa_m} = \kappa_{m+1} \in \gamma$$

Then $|\beta| < \kappa_m < \kappa_{m+1} \leq |\gamma|$. Since the cardinality of β is less than $|\gamma|$, then γ is an (infinite) initial ordinal, and therefore is cardinal number, as claimed.

Since γ is an infinite cardinal number, there is an ordinal κ such that

$$\aleph_\kappa = \gamma$$

Claim: The cardinal $\aleph_\kappa$ is a strong limit cardinal.

Suppose $\alpha \in \kappa$. It suffices to show that $2^{\aleph_\alpha} \in \aleph_\kappa$. Then

$$\aleph_\alpha \in \aleph_\kappa = \gamma = \cup\{\kappa_n : n \in \omega_0\}$$

Then $\aleph_\alpha \in \kappa_m$ for some $m \in \omega_0$. It follows that

$$\aleph_{\alpha+1} \in_= 2^{\aleph_\alpha} \in 2^{\kappa_m} = \kappa_{m+1} \in \aleph_\kappa \quad (2^{\kappa_m} = \kappa_{m+1} \text{ by construction of } \{\kappa_n : n \in \omega_0\})$$

Since $\aleph_{\alpha+1} \in \aleph_\kappa$, $\aleph_\kappa$ is a limit cardinal. By definition, $2^{\aleph_\alpha} \in \aleph_\kappa$ implies, $\aleph_\kappa$ is a strong limit cardinal number.

We have shown that strong limit cardinals exist in a *ZFC*-universe. Unfortunately the proof cannot tell us how to construct one or what it looks like. Strong limit cardinals are completely elusive. But, ethereal as they may be, we can safely make the assumption that they exist. It is important to see that GCH plays no role in the above proof. In a *ZFC*+GCH-universe strong limit cardinals are everywhere and easily seen since they are the limit cardinals.

We now present other categories of cardinal numbers.

Definition 29.14

a) We say that an infinite cardinal, $\aleph_\gamma$, is a *singular cardinal number* if $\aleph_\gamma$ is the least upper bound of a strictly increasing sequence of ordinals, $\{\alpha_\kappa : \kappa \in \beta\}$, indexed by the elements of some ordinal β in $\aleph_\gamma$.

b) An infinite cardinal $\aleph_\gamma$ is said to be a *regular cardinal number* if it is not a singular cardinal number. That is, there does not exist an ordinal, β, in $\aleph_\gamma$ such that $\aleph_\gamma = \text{lub}\{\alpha_\kappa : \kappa \in \beta\}$.

The following limit cardinals are examples of singular cardinals:

$$\begin{aligned}
\aleph_{\omega_0} &= \text{lub}\{\aleph_n : n \in \omega_0\}, \text{ where } \omega_0 \in \aleph_{\omega_0} \\
\aleph_{\omega_0+\omega_0} &= \text{lub}\{\aleph_{\omega_0+n} : n \in \omega_0\}, \text{ where } \omega_0 \in \aleph_{\omega_0+\omega_0} \\
\aleph_{\omega_0 3} &= \text{lub}\{\aleph_{\omega_0 2+n} : n \in \omega_0\}, \text{ where } \omega_0 \in \aleph_{\omega_0 3} \\
\aleph_{\omega_0\omega_0} &= \text{lub}\{\aleph_{\omega_0 \cdot n} : n \in \omega_0\}, \text{ where } \omega_0 \in \aleph_{\omega_0\omega_0} \\
\aleph_{\omega_1} &= \text{lub}\{\aleph_\alpha : \alpha \in \omega_1\}, \text{ where } \omega_1 = \aleph_1 \in \aleph_{\omega_1}
\end{aligned}$$

The above examples suggest that there can be no upper bound to the class of singular cardinals since if γ is any ordinal larger than ω_0,

$$\aleph_{\gamma+\omega_0} = \text{lub}\{\aleph_{\gamma+n} : n \in \omega_0\}, \text{ where } \omega_0 \in \aleph_\gamma$$

How easy is it to find (infinite) regular cardinal numbers? Our first example of a regular cardinal is a simple one.

Example: The countable limit cardinal $\aleph_0$ is a regular cardinal since, for any non-zero $m \in \aleph_0$, $\text{lub}\{n : n \in m\} = m - 1 \neq \aleph_0$.

Determining whether a given cardinal number is regular or singular case by case can be tedious. It will be helpful to determine a few principles that will help us distinguish those limit cardinals which are regular from those that are singular.

Theorem 29.15 Every infinite successor cardinal, $\aleph_{\alpha^+} = \aleph_{\alpha+1}$, is a regular cardinal.

Proof:

Suppose $\aleph_{\alpha^+}$ is a successor cardinal number. To prove that it is regular we will suppose that it is singular and show how this leads to a contradiction. Suppose that $\beta \in \aleph_{\alpha^+}$ such that $\aleph_{\alpha^+} = \text{ lub}\{\alpha_\kappa : \kappa \in \beta\}$ where $\{\alpha_\kappa : \kappa \in \beta\}$ is an increasing sequence of ordinals. Then $|\beta| \in_= \aleph_\alpha$. Since $\aleph_{\alpha^+}$ is a limit ordinal, $\aleph_{\alpha^+} = \cup\{\alpha_\kappa : \kappa \in \beta\}$ (theorem 27.12). Now $\aleph_{\alpha^+} = \aleph_{\alpha+1}$ is the least ordinal whose cardinality is larger than $\aleph_\alpha$. We then have $|\alpha_\kappa| \in_= \aleph_\alpha$ for all $\kappa \in \beta$.

Then

$$\begin{aligned} \aleph_{\alpha^+} &= |\aleph_{\alpha^+}| \\ &= |\cup\{\alpha_\kappa : \kappa \in \beta\}| \\ &\in_= \aleph_\alpha \quad (|\beta| \in_= \aleph_\alpha,\ |\alpha_\kappa| \in_= \aleph_\alpha, \text{ and by corollary 29.9.}) \end{aligned}$$

So $\aleph_{\alpha^+} \in_= \aleph_\alpha$, a contradiction. The source of our contradiction is our assumption that $\aleph_{\alpha^+}$ is singular. Hence, $\aleph_{\alpha^+}$ must be regular, as required.

The above theorem provides us with a tool for constructing infinite regular cardinal numbers. For example, the following are infinite regular cardinals:

$$\begin{aligned} \aleph_1 &= \aleph_{0^+} \\ \aleph_{\aleph_1+1} &= \aleph_{\aleph_1^+} \\ \aleph_{\aleph_{\aleph_0}+1} &= \aleph_{\aleph_{\aleph_0}^+} \end{aligned}$$

Theorem 29.16 Let γ be an infinite cardinal number. If γ is a singular cardinal, then the cardinal number $\aleph_\gamma$ is a singular cardinal.

Proof:

Given: That γ is a singular cardinal. That is, there exists an ordinal $\beta \in \gamma$ such that $\gamma = \text{ lub}\{\alpha_\kappa : \kappa \in \beta\}$ where $\{\alpha_\kappa : \kappa \in \beta\}$ is an increasing sequence of ordinals.
Required to show: That $\aleph_\gamma$ is a singular cardinal.

Since γ is declared to be an infinite cardinal, it is a limit ordinal. Then $\aleph_\gamma$ is a limit cardinal.

Claim: That $\aleph_\gamma = \text{ lub}\{\aleph_{\alpha_\kappa} : \kappa \in \beta\}$:

It suffices to show that $\text{ lub}\{\aleph_{\alpha_\kappa} : \kappa \in \beta\} \in_= \aleph_\gamma$ and $\aleph_\gamma \in_= \text{ lub}\{\aleph_{\alpha_\kappa} : \kappa \in \beta\}$.

– First note that since $\{\alpha_\kappa : \kappa \in \beta\}$ is increasing,

$$[\alpha_\kappa \in \gamma] \Rightarrow [\aleph_{\alpha_\kappa} \in \aleph_\gamma] \text{ for all } \kappa \in \beta$$

Then $\text{lub}\{\aleph_{\alpha_\kappa} : \kappa \in \beta\} \in_= \aleph_\gamma$.

– Let $\mu \in \aleph_\gamma$. Then $|\mu| \in_= \mu \in \aleph_\gamma$. Let δ be such that $\aleph_\delta = |\mu| \in \aleph_\gamma$. Since $\aleph_\gamma$ is a limit cardinal, $\mu \in \aleph_{\delta+1} \in \aleph_\gamma$; hence, $\delta + 1 \in \gamma$. Since γ is a limit ordinal $\gamma = \cup\{\alpha_\kappa : \kappa \in \beta\}$ and so $\delta + 1 \in \alpha_{\kappa_0}$, for some $\kappa_0 \in \beta$. Then $\mu \in \aleph_{\delta+1} \in \aleph_{\alpha_{\kappa_0}} \in_= \text{lub}\{\aleph_{\alpha_\kappa} : \kappa \in \beta\}$. It follows that $\mu \in \text{lub}\{\aleph_{\alpha_\kappa} : \kappa \in \beta\}$; hence, $\aleph_\gamma \in_= \text{lub}\{\aleph_{\alpha_\kappa} : \kappa \in \beta\}$.

We conclude that $\aleph_\gamma = \text{lub}\{\aleph_{\alpha_\kappa} : \kappa \in \beta\}$, as claimed.

Since $\beta \in_= \aleph_\beta$ (by theorem 28.14) and $\aleph_\beta \in \aleph_\gamma$, $\beta \in \aleph_\gamma$. So $\aleph_\gamma$ is singular, as required.

The above theorem allows us to recursively construct infinite sequences of singular cardinals. For example, since we have shown that the cardinal $\aleph_{\omega_0}$ is singular, then the set

$$\aleph_{\omega_0},\ \aleph_{\aleph_{\omega_0}},\ \aleph_{\aleph_{\aleph_{\omega_0}}},\ \aleph_{\aleph_{\aleph_{\aleph_{\omega_0}}}},\ \ldots,$$

contains only singular cardinals.

Inaccessible cardinal numbers.

The examples we provided of singular cardinals were all limit cardinals. We then showed that $\aleph_0$ and every successor cardinal is a regular cardinal. However, uncountable regular *limit* cardinal appear to be elusive. One may be tempted to conclude that no uncountable limit cardinal can be regular. But none of the ideas expressed above allows us to arrive at such a conclusion. If a regular limit cardinal exists what would it look like? Suppose $\aleph_\beta$ is an uncountable regular limit cardinal. Since $\{\aleph_1, \aleph_2, \aleph_3, \aleph_4, , \ldots, \}$ are all successor cardinals, $\aleph_{\omega_0} \in_= \aleph_\beta$. Now β must be a limit ordinal; hence, we can argue (as in the proof of the theorem immediately above) that $\aleph_\beta = \text{lub}\{\aleph_\alpha : \alpha \in \beta\}$. If $\beta \in \aleph_\beta$, then the fact that $\aleph_\beta = \text{lub}\{\aleph_\alpha : \alpha \in \beta\}$ would imply, by definition, that $\aleph_\beta$ is singular. Since $\aleph_\beta$ is regular, it must be the case that $\beta \notin \aleph_\beta$. Hence, if $\aleph_\beta$ is an uncountable regular limit cardinal, it must be the case that $\beta = \aleph_\beta$ (part b) of theorem 28.14).

Uncountable regular limit cardinals (if they exist) are referred to as "inaccessible" cardinals, possibly because, if some exist, they are difficult to find.

Definition 29.17 A regular cardinal number which is a limit cardinal is called an *inaccessible cardinal*. A regular cardinal number which is a strong limit cardinal is called a *strongly inaccessible cardinal*.

Recall that a strong limit cardinal $\aleph_\alpha$ is a limit cardinal which satisfies the property "$\kappa \in \aleph_\alpha \Rightarrow 2^\kappa \in \aleph_\alpha$". We have shown that strong limit cardinals exist in a *ZFC*-universe. We have also seen that in a universe governed by ZFC + GCH, strong limit cardinals are the same as infinite limit cardinals. So if we assume GCH, all weakly inaccessible cardinals are strongly inaccessible.

Interestingly enough, it has been shown that the existence of inaccessible cardinals cannot be proven from *ZFC*. This means that if we assume that the existence of inaccessible cardinals, this will not lead to a contradiction that is not already there. Those mathematicians whose mathematical statements depend on the existence of inaccessible cardinals are roaming in the "$ZFC + \exists$ inaccessible cardinals"−universe, one which is larger than the one which is governed by *ZFC*.

29.8 Cofinality of a cardinal number

Suppose A and B are two subsets of a partially ordered set $(P, \leq)$. We say that the subset B is *cofinal* in A if for any $a \in A$ there exists $b \in B$ such $a \leq b$. A set is always cofinal in itself.

Examples:

- The set $B = \{0, 2, 4, 6, \ldots\}$ is a cofinal subset of $\aleph_0$, since for any natural number $n \in \aleph_0$ there exist $2n \in B$ such that $n \leq 2n \in B$. In this case we see that $\aleph_0$ has a cofinal subset B such that $|B| = |\aleph_0| = \aleph_0$.
- The set of ordinals $\omega_2 + 3$ is not cofinal in $\omega_2 + 4$ since $\omega_2 + 3 \in \omega_2 + 4$ but there is no $\alpha \in \omega_2 + 3$ such that $\omega_2 + 3 \leq \alpha$. In fact, the only subset of $\omega_0 + 3$ which is cofinal in $\omega_0 + 3$ is $\omega_0 + 3$ itself. We generalize: Every successor ordinal α has only α as cofinal subset.
- We can show, on the other hand, that the set $\aleph_7$ is not cofinal in $\aleph_8$: We know that $\aleph_8$ is the least ordinal whose cardinality is larger than $\aleph_7$. Then $\aleph_7 \in \aleph_7 \cup \{\aleph_7\} \in \aleph_8$ (since the cardinality of $\aleph_8$ is strictly larger than the cardinality of $\aleph_7 \cup \{\aleph_7\}$). Since $\kappa \in \aleph_7 \cup \{\aleph_7\}$ for all $\kappa \in \aleph_7$, $\aleph_7$ cannot be cofinal in $\aleph_8$.
- In fact, it seems that all cofinal subsets of $\aleph_8$ must have cardinality $\aleph_8$. We verify this hypothesis. Let us suppose that B is a cofinal subset of $\aleph_8$ such that $|B| \leq \aleph_7$. Then we can index the elements of $B = \{b_\alpha : \alpha \leq \beta \in \aleph_8\}$ where $|\beta| \leq \aleph_7$ (noting that any proper subset B of $\aleph_8$ has cardinality $\aleph_7$ or less). Now if $\text{lub}(B) = \gamma < \aleph_8$ we can choose $\gamma + 1 \in \aleph_8$ where $b_\alpha < \gamma + 1$ for all $b_\alpha \in B$. Since this would contradict the fact that B is cofinal in $\aleph_8$, we must have that $\text{lub}(B) = \aleph_8$. By corollary 29.9, $|\text{lub}(B)| = |\cup \{b_\alpha : \alpha \leq \beta\}| \leq \aleph_7$. We have a contradiction. Hence cofinal subsets of $\aleph_8$ must have cardinality $\aleph_8$.
- If γ is an ordinal then a subset B of γ is cofinal in γ only if $\text{lub}(B) = \cup\{b : b \in B\} = \gamma$ (for if $\text{lub}(B) = \kappa < \gamma$ then $b < \kappa + 1 < \gamma$ for all $b \in B$).

We see that studying the cardinality of cofinal subsets of cardinals might allow us to distinguish even further certain types of cardinals from others. This leads us the definition of the *cofinality* of cardinal numbers.

Definition 29.18 Let $\aleph_\gamma$ be an infinite cardinal. We say the *cofinality of* $\aleph_\gamma$ is β and write $\text{cf}(\aleph_\gamma) = \beta$ if β is the least ordinal in $\aleph_\gamma$ which indexes an increasing set of ordinals $\{\theta_\alpha : \alpha < \beta\}$ such that $\aleph_\gamma = \text{lub}\{\theta_\alpha : \alpha < \beta\}$. If no such β exists in $\aleph_\gamma$ then we say the cofinality, $\text{cf}(\aleph_\gamma)$, is $\aleph_\gamma$ and write $\text{cf}(\aleph_\gamma) = \aleph_\gamma$.

We see that, for an infinite cardinal number $\aleph_\gamma$, $\text{cf}(\aleph_\gamma) = \beta < \aleph_\gamma$ if and only if there is a smallest set of increasing ordinals $\{\theta_\alpha : \alpha < \beta\}$ which is cofinal in $\aleph_\gamma$ (in the sense of the definition of "cofinal subset" described above). Given this definition of "cofinality of a cardinal number" we can now restate our definition of "singular" and "regular" cardinal numbers as follows:

> An infinite cardinal number $\aleph_\gamma$ is said to be a *singular cardinal* if only if $\text{cf}(\aleph_\gamma) < \aleph_\gamma$. If $\text{cf}(\aleph_\gamma) = \aleph_\gamma$ then $\aleph_\gamma$ is said to be a *regular cardinal*.

So singular cardinals are those infinite cardinals that have a *proper* cofinal subset. We showed above that every successor cardinal $\aleph_\gamma$ is such that $\text{cf}(\aleph_\gamma) = \aleph_\gamma$ and so are regular cardinals. We then defined the *inaccessible cardinals* as being those cardinals $\aleph_\gamma$ which are both limit cardinals and satisfy $\text{cf}(\aleph_\gamma) = \aleph_\gamma$.

Whenever $\aleph_\gamma$ is singular the definition of cofinality $\text{cf}(\aleph_\gamma)$ of $\aleph_\gamma$ is declared to be the smallest ordinal β which satisfies a particular property. We will show that the cofinality, $\text{cf}(\aleph_\gamma)$, of a any cardinal $\aleph_\gamma$ must also be a cardinal.

Theorem 29.19 The cofinality $\text{cf}(\aleph_\gamma)$ of an infinite cardinal number $\aleph_\gamma$ is a cardinal number. Hence $\text{cf}(\aleph_\gamma)$ is the smallest cardinality of all sets which are cofinal in $\aleph_\gamma$.

Proof:

The statement is obviously true if $\aleph_\gamma$ is a regular cardinal (that is, if $\text{cf}(\aleph_\gamma) = \aleph_\gamma$). Suppose that $\aleph_\gamma$ is a singular cardinal. Then $\text{cf}(\aleph_\gamma) = \beta < \aleph_\gamma$. Then $\aleph_\gamma = \text{lub}\{\theta_\alpha : \alpha < \beta\}$ for some strictly increasing function $h : \beta \to \aleph_\gamma$ such that $h(\alpha) = \theta_\alpha$. We are required to show that β is a cardinal number. Suppose not. Suppose $\aleph_\lambda = |\beta| = |\text{cf}(\aleph_\gamma)| < \beta$. Then the ordinal $\aleph_\lambda$ is equipotent to β. Let $f : \aleph_\lambda \to \beta$ be a one-to-one function mapping $\aleph_\lambda$

onto β. Let $g : \aleph_\lambda \to \beta$ be defined as:

$$\begin{aligned}
\mu_0 = g(0) &= f(0) \\
\mu_1 = g(1) &= \text{glb}\{f(1), f(2), \ldots, \text{ all of which} > g(0)^+\} > g(0) \\
\mu_2 = g(2) &= \text{glb}\{f(2), f(3), \ldots, \text{ all of which} > g(1)^+\} > g(1) \\
\mu_3 = g(3) &= \text{glb}\{f(3), f(4), \ldots, \text{ all of which} > g(2)^+\} > g(2) \\
&\vdots \\
\mu_n = g(n) &= \text{glb}\{f(n), f(n+1), \ldots, \text{ all of which} > g(n-1)^+\} > g(n) \\
&\vdots \\
\mu_\kappa = g(\kappa) &= \text{glb}\{f(\kappa), f(\kappa+1), \ldots, \text{ all of which} > g(\delta)^+ \text{ when } \delta < \kappa\} \\
&\vdots
\end{aligned}$$

Then $\mu_{\kappa+1} = g(\kappa+1) > g(\kappa) = \mu_\kappa$ for all $\kappa \in \aleph_\lambda$; hence $g : \aleph_\lambda \to \beta$ is strictly increasing on its domain. Since f is onto β, for each $\alpha \in \beta$, there exists $\kappa_\alpha \in \aleph_\lambda$ such that $g(\kappa_\alpha) = \alpha \in \beta$. Then $\aleph_\gamma = \text{lub}\{\theta_\alpha : \alpha < \beta\} = \text{lub}\{\mu_\kappa : \kappa < \aleph_\lambda\}$ where $\mu_\kappa = t(\kappa) = h(g(\kappa))$. This contradicts the fact that β is the least ordinal such that $\aleph_\gamma = \text{lub}\{\theta_\alpha : \alpha < \beta\}$. We must conclude that, for any cardinal $\aleph_\gamma$, $\text{cf}(\aleph_\gamma)$ is a cardinal number.

Hence cofinalities of cardinal numbers are cardinal numbers. We can now show that the cofinality of a cardinal number must be a *regular* cardinal.

Theorem 29.20 The cofinality $\text{cf}(\varphi)$ of an infinite cardinal number φ is a regular cardinal.

Proof:

Let φ be a cardinal and $\delta = \text{cf}(\varphi)$. To show that δ is regular it suffices to show that $\text{cf}(\delta) = \delta$. We know that $\text{cf}(\delta) \leq \delta$. It will then suffice to show that $\text{cf}(\delta) \not< \delta$. Let $\beta = \text{cf}(\delta)$ and suppose $\beta < \delta$. We will show this will lead to a contradiction. Then

$$\delta = \text{ lub}\{\mu_\alpha : \alpha < \beta\}$$

where $\mu_\alpha = f(\alpha) \in \delta$ for some strictly increasing function $f : \beta \to \delta$. Since $\delta = \text{cf}(\varphi)$ then

$$\varphi = \text{ lub}\{\theta_\kappa : \kappa < \delta\}$$

where $\theta_\kappa = g(\kappa) \in \varphi$ for some strictly increasing function $g : \delta \to \varphi$. Then the function $g \circ f : \beta \to \varphi$ is a strictly increasing function which maps β into φ.
So $\varphi = \text{ lub}\{\theta_\kappa : \kappa < \delta\} = \text{ lub}\{\phi_\alpha : \alpha < \beta\}$ where $\phi_\alpha = t(\alpha) = g(f(\alpha))$. Since $\beta < \delta$ this contradicts the fact that δ is the least ordinal which allows the cofinality condition to hold true. Then $\text{cf}(\delta) \not< \delta$. Hence $\text{cf}(\delta) = \text{cf}(\text{cf}(\varphi)) = \text{cf}(\varphi) = \delta$. So $\text{cf}(\varphi)$ is a regular cardinal.

We have seen that for infinite cardinals κ, it occurs that κ^λ is sometimes equal to κ (for example $(2^\kappa)^\kappa = 2^{\kappa\times\kappa} = 2^\kappa$). The following statement shows a relationship between κ and λ which guarantees that $\kappa < \kappa^\lambda$.

Theorem 29.21 If κ is an infinite cardinal and $\text{cf}(\kappa) \leq \lambda$, then $\kappa < \kappa^\lambda$.

Proof:

We are given that κ is infinite and $\text{cf}(\kappa) \leq \lambda$. We are required to show that $\kappa < \kappa^\lambda$.

Case 1: We first consider the case where κ is a regular cardinal. In this case we have $\kappa = \text{cf}(\kappa)$. From the hypothesis we obtain $\kappa \leq \lambda$. By 24.3, $\kappa^\kappa \leq \kappa^\lambda$. Since $\kappa < 2^\kappa = \kappa^\kappa$ (by 29.10), then $\kappa < \kappa^\lambda$, as required.

Case 2: We now consider the case where κ is a singular cardinal. That is, suppose $\text{cf}(\kappa) = \varphi < \kappa$ where $\varphi \leq \lambda$. Then $\kappa = \text{lub}\{\mu_\alpha : \alpha \in \varphi\}$ where $\{\mu_\alpha : \alpha \in \varphi\}$ is a strictly increasing set. Then $\kappa = \cup\{\mu_\alpha : \alpha \in \varphi\}$ (since κ is a limit ordinal).

We claim $\cup\{\mu_\alpha : \alpha \in \varphi\} < \kappa^\varphi$:

Suppose not. That is, suppose $\kappa^\varphi \leq \cup\{\mu_\alpha : \alpha \in \varphi\}$. Then there exists a function $f : \kappa^\varphi \to \cup\{\mu_\alpha : \alpha \in \varphi\}$ mapping κ^φ one-to-one into $\cup\{\mu_\alpha : \alpha \in \varphi\}$. Then f^{-1} is well-defined and onto κ^φ. Notice that if $k \in \kappa^\varphi$ then k can be expressed as $k = (\kappa_0, \kappa_1, \kappa_2, \ldots, \kappa_\alpha, \ldots,)$ where $\kappa_\alpha \in \kappa$ when $\alpha \in \varphi$. Then κ^φ can be expressed in the form of the product $\Pi_{\alpha\in\varphi}\kappa$. We define the map $\pi_q : \Pi_{\alpha\in\varphi}\kappa \to \kappa$ as $\pi_q(\kappa_0, \kappa_1, \kappa_2, \ldots, \kappa_q, \ldots) = \kappa_q$.

For $q \in \varphi$, $f^{\leftarrow}[\mu_q] = \{k \in \kappa^\varphi : f(k) \in \mu_q < \kappa\}$. For each $q \in \varphi$, let

$$K_q = \pi_q[f^{\leftarrow}[\mu_q]] \subset \kappa$$

Next, for each $q \in \varphi$, choose $x_q \in \kappa - K_q$ (Choice!). Then use the selected x_q's to construct the point

$$x = (x_0, x_1, x_2, \ldots, x_q, \ldots,) \in \Pi_{\alpha\in\varphi}\kappa$$

There must exist some point $y \in \cup\{\mu_\alpha : \alpha \in \varphi\}$ such that $f^{-1}(y) = x$. Then there exists some $q \in \varphi$ such that $y \in \mu_q$, hence $f^{-1}(y) \in f^{\leftarrow}[\mu_q]$. Thus $\pi_q(f^{-1}(y)) = \pi_q(x) = x_q \in \pi_q[f^{\leftarrow}[\mu_q]] = K_q$, contradicting $x_q \in \kappa - K_q$ for all $q \in \varphi$. The source of the contradiction is our assumption that $\kappa^\varphi \leq \cup\{\mu_\alpha : \alpha \in \varphi\}$. Then $\cup\{\mu_\alpha : \alpha \in \varphi\} < \kappa^\varphi$, as claimed.

We then have $\kappa = \cup\{\mu_\alpha : \alpha \in \varphi\} < \kappa^\varphi \leq \kappa^\lambda$ (since $\varphi \leq \lambda$ and by 24.3). Then $\kappa < \kappa^\lambda$, as required.

Concepts review:

1. What is a well-orderable set? How is it different from a well-ordered set?

2. What can be said about those sets that are the one-to-one image of an ordinal number.
3. Is it true that any well-ordered set, no matter how large, is order isomorphic to some ordinal number?
4. What does it mean to say that the well-ordered set S has order type (ordinality) α?
5. What does the Well-ordering theorem say?
6. The Well-ordering theorem is a consequence of which fundamental ZFC-axiom?
7. What is an initial ordinal? Are natural numbers initial ordinals? Why?
8. What is the least uncountable initial ordinal? How is it obtained?
9. Are all ordinals in $\omega_0 \cup \{\omega_\alpha : \alpha \in \mathscr{O}\}$ initial ordinals?
10. Define the "cardinal numbers".
11. If we assume the Continuum hypothesis, what is the cardinality of $\mathbb{R}$?
12. If we do not assume the Continuum hypothesis, what is the cardinality of $\mathbb{R}$?
13. What is the cofinality of a cardinal number?
14. Define singular and regular cardinal number.
15. What kind of cardinal is said to be inaccessible?
16. Is the cofinality of a cardinal number necessarily a cardinal number?
17. Give examples of regular cardinals.
18. Give examples of singular cardinals.
19. What is a limit cardinal?
20. What is a successor cardinal?

EXERCISES

A. 1. Consider the set of ordinals defined as $S = \cup\{\omega n : n \in \mathbb{N}, n > 2\}$.
   a) Is $S = \mathscr{O}$? Why?
   b) Is S is an ordinal number?
   c) If S is an ordinal number, is it a limit ordinal number?
   d) What is the least upper bound of S?

2. List a few possible order types of a set of cardinality $\aleph_0$.
3. What is the cardinality of a set whose ordinality is $\omega_3 + \omega^2$?
4. Use the appropriate symbols to represent the ordinality, the associated initial ordinal and the cardinality of the set $S = \mathbb{N} \cup \{\mathbb{N}\}$ (ordered in the usual way).
5. Does the Cantor set have a well-ordering?
6. If we assume CH, what is the least ordinal that is not equipotent to $\mathbb{R}$?
7. What is the least ordinal which does not belong to $\mathscr{I}$?
8. What is the least limit ordinal that does not belong to $\mathscr{I}$?

B.

9. Let $S = \{1, 2, 4, 8, 32, 64\}$. Define the order relation $\leq$ on S as follows: $a \leq b$ if and only if a divides b.
   a) Is $\leq$ a well-defined order relation on S?
   b) Is $\leq$ a well-ordering of S?
   c) Express $(S, \leq)$ as a an element of $\mathscr{P}(S) \times \mathscr{P}(S \times S)$ by explicitly exhibiting all of its elements.

10) Let $S = \{a_\alpha : \alpha \in \omega_0 + \omega\}$ be a set whose elements are defined as follows:

$$a_\alpha = \begin{cases} 2 - \frac{1}{\alpha+1} & \text{when} \quad \alpha \in \omega_0 \\ 2 & \text{when} \quad \alpha = \omega_0 \\ 2 + \frac{1}{f(\alpha)+1} & \text{when} \quad \omega_0 \in \alpha \end{cases}$$

where $f(\omega_0 + n) = n$.

   a) Are the elements of the set S well-defined?
   b) Is the ordering induced on the elements of S by the index set $\omega_0 + \omega$ a well-ordering?
   c) If the elements of S are assumed to be well-ordered by the index set, what is the least element of the set S with respect to this ordering?
   d) What is the least upper bound (supremum) of the set $\{a_\alpha : \alpha \in \omega\}$ with respect to the ordering defined by the index set.

C.

11. Show that if S is a class of ordinals, then the least upper bound of S is $\cup\{\alpha \in \mathscr{O} : \alpha \in S\}$.
12. Is it true that given any two sets S and T, either S is embeddable in T or T is embeddable in S? Why?
13. Is the class $\mathscr{I}$ of all initial ordinals an initial segment of $\mathscr{O}$? Why?
14. Is the class $\mathscr{I}$ of all initial ordinals a transitive class?

Part IX

More on axioms: Choice, regularity and Martin's axiom

30 / Axiom of choice

Summary. *In this section we prove that the Axiom of choice is equivalent to the Well-ordering principle. We provide a few mathematical statements whose proof requires the Axiom of choice. Finally we present Zorn's lemma and show that it is equivalent to the Axiom of choice. A proof of the fact that every vector space has a basis is given by invoking Zorn's lemma.*

30.1 Introduction.

We have seen that there is a subtle difference between proving that a mathematical object "exists" and actually "constructing" or "exhibiting" this mathematical object. If we can construct a mathematical object, then this can serve as a valid proof that this object exists. But some mathematical objects can be proven to exist without ever being able to provide a concrete example of this object or providing a method to construct it. For example, we have proved in the Well-ordering theorem (a consequence of the Axiom of choice) that there exists, in a *ZFC*-universe, a function which maps ω_1 one-to-one onto $\mathbb{R}$. But we have no way of determining what such a function is. "Most" mathematicians are comfortable with the idea of manipulating a mathematical object whose existence has been proven even if they believe that this same mathematical object can never actually be constructed or witnessed. For a few, however, to put a non-constructible mathematical object whose existence has been proven, on the same level as a mathematical object which can actually be constructed is nonsense. This, in a nutshell, describes the controversy surrounding the use of the Axiom of choice. We will try to develop a deeper understanding of what this axiom is about. The Axiom of choice will be seen to be equivalent to other mathematical statements which many find more palatable.

30.2 The Axiom of choice

The Axiom of choice plays a fundamental role in the study of mathematics. This principle was invoked in the proofs of numerous mathematical statements long before mathematicians began acknowledging it as an axiom. It is worth restating it here:

> *The Axiom of choice*: Let $\mathscr{S}$ be a set of nonempty sets whose union is the set, T, of elements. Then we can choose some element s from each set S in $\mathscr{S}$. That is, there exists a function, $f : \mathscr{S} \to T$, with domain, $\mathscr{S}$, such that for each set, $S \in \mathscr{S}$, $f(S) \in S$.

The function $f : \mathscr{S} \to T$ described in this paragraph is referred to as a "choice function".

Given a set of non-empty sets, $\mathscr{U} = \{S_\alpha : \alpha \in \gamma \in \mathscr{O}\}$, we do not always have to invoke the Axiom of choice if we wish to select an element s from each set $S \in \mathscr{U}$.

We consider the following example in which we wish to select a single element from a single non-empty set.

Existence of a choice function for a single non-empty set. Suppose we are given a non-empty set S for which no specific element of S can be distinguished from the others. Suppose we wish to choose some element in the set S. If we argue as follows, it is not necessary to invoke the Axiom of choice:

$$\begin{aligned} S \neq \varnothing \quad &\Rightarrow \quad \mathscr{P}(S) \neq \varnothing \\ &\Rightarrow \quad \mathscr{P}(S) \times S \neq \varnothing \quad \text{(By definition of Cartesian product.)} \\ &\Rightarrow \quad \{S\} \times S \neq \varnothing \quad \text{(Since } S \in \mathscr{P}(S)\text{.)} \\ &\Rightarrow \quad \text{there exists an element } (S, x) \in \{S\} \times S \end{aligned}$$

Observe that $\{(S, x)\}$ is a function with domain $\{S\}$ and range $\{x\}$ which associates to S some element x of S. We did not invoke the Axiom of choice to postulate the existence of the function $f(S) = x$.

Selecting an element from each set of non-empty sets from $\mathscr{P}(\mathbb{N})$. Consider the set $\mathscr{P}(\mathbb{N})^* = \mathscr{P}(\mathbb{N}) - \{\varnothing\}$ of all non-empty subsets of $\mathbb{N}$. Suppose we want to form a new set, $S = \{n_A : A \in \mathscr{P}(\mathbb{N})^*\}$, by selecting from each set, A, in $\mathscr{P}(\mathbb{N})^*$ a single element, n_A. We can argue as follows: Since the set $\mathbb{N}$ has been shown to be well-ordered, we can choose from each non-empty set $A \in \mathscr{P}(\mathbb{N})^*$ the unique smallest number, n_A, in A. The Axiom of choice is not required, since each element A in the sets of $\mathscr{P}(\mathbb{N})^*$ is specifically and unambiguously identified as being the unique least element in A. We can express our choice function, $f : \mathscr{P}(\mathbb{N})^* \to \mathbb{N}$ as follows:

$$f = \{(A, n_A) \in \mathscr{P}(\mathbb{N})^* \times \mathbb{N} : n_A = \text{unique least element of } A\ \}$$

In this case, the Axiom of choice is not required.

Suppose, on the other hand, that we are given an infinite set of identical golf balls distributed in infinitely many boxes, $\mathscr{A} = \{A_\alpha : \alpha \in \gamma\}$, indexed with the ordinal numbers in such a way that no box, A_α, is empty. We are asked to construct a function, $f : \mathscr{A} \to \cup\{A_\alpha : \alpha \in \gamma\}$, which chooses a single ball from each box. Note that the golf balls are neither labeled nor indexed, in the sense that they are all identical in all respects. So there can be no formula for a function f which globally states which specific ball is to be chosen in each box. In this case, the best we can do is to invoke the Axiom of choice which guarantees that at least one choice function, $f : \mathscr{A} \to \cup\{A_\alpha : \alpha \in \gamma\}$, exists.

It was shown by Kurt Gödel in 1938, that no contradictions can result from invoking the Axiom of choice. In 1963, Paul Cohen showed that the Axiom of choice cannot be proved from *ZF*. So the Axiom of choice adds new sets (since functions are sets) to a *ZF*-universe of sets.

The following more general theorem guarantees that the existence of a choice function for *finitely* many sets does not require the intervention of the Axiom of choice.

Theorem 30.1 Suppose $\mathscr{S}$ is a finite set of non-empty sets whose union is the set M. Then there exists a function $f : \mathscr{S} \to M$ which maps each set to one of its elements.

Proof: We prove this by induction on $n \in \mathbb{N}$.

Let $P(n)$ be the statement: "A system of n sets has a choice function."
Base case: The statement $P(0)$ is vacuously true.
Inductive hypothesis: Suppose the statement $P(n)$ holds true.
Suppose $\mathscr{S}$ is a system of $n+1$ non-empty sets whose union is the set M. Let $U \in \mathscr{S}$. Then $\mathscr{S} - \{U\}$ is a system of n sets. By the inductive hypothesis, a choice function $g : \mathscr{S} - \{U\} \to M$ exists.
Since U is non-empty, there exists an element, say $(U, u) \in \{U\} \times U$. Define the choice function $f : \mathscr{S} \to M$ as follows:

$$f(T) = \begin{cases} g(T) & \text{if} \quad T \in \mathscr{S} - \{U\} \\ u & \text{if} \quad T = U \end{cases}$$

So $P(n+1)$ holds true. By mathematical induction, every finite system of sets has a choice function.

A word of caution. The proof above does not show that any countably infinite set of sets has a choice function. The statement $P(n)$ states that "a set of n sets has a choice function" no matter what the value of n is. It only proves that all finite sets have a choice function, nothing more.

30.3 The Axiom of countable choice.

The *Axiom of countable choice* is a weaker form of the Axiom of choice. We state it formally:

> *The Axiom of countable choice.* Let $\mathscr{S}$ be a *countable* set of nonempty sets whose union is the set, T, of elements. Then we can choose some element s from each set S in $\mathscr{S}$. That is, there exists a function $f : \mathscr{S} \to T$ with domain $\mathscr{S}$ such that for each set $S \in \mathscr{S}$, $f(S) \in S$.

Since the Axiom of countable choice is a special case of the Axiom of choice, it obviously follows from it. If T is countable, then it is well-orderable, and so the Axiom of countable choice is not required to justify the existence of a choice function. If T is uncountable and we do not assume the Axiom of choice, then T may not be well-orderable. In this case the Axiom of countable choice is required to justify the

existence of a choice function on $\mathscr{S} = \{S_n : n \in \mathbb{N}\}$.[1]

The reader may wonder why we cannot generalize the arguments used to prove the existence of a choice function on finite sets of sets to justify the existence of a choice function on an infinite number of subsets of an infinite set S. The main reason is that if there are are infinitely many non-empty subsets $\mathscr{F} = \{S_\alpha : \alpha \in \gamma\}$ of an infinite set S, then, for each set $S_\alpha \in \mathscr{F}$, we would have to invoke the existence principle,

$$\{S_\alpha\} \times S_\alpha \neq \varnothing \Rightarrow \text{ there exists } (S_\alpha, s) \in \{S_\alpha\} \times S_\alpha$$

This means that this existence principle would have to be invoked infinitely many times. The ZF-axioms do not define a formula with an infinite chain of existence symbols.

30.4 Equivalent forms of the Axiom of choice.

There are many mathematical statements which are equivalent to the Axiom of choice. One of the simplest statements is the following one.

Theorem 30.2 Let [AC*] denote the statement:

"For any set, $\mathscr{S} = \{S_\alpha : \alpha \in \gamma\}$, of non-empty sets, $\Pi_{\alpha\in\gamma} S_\alpha$ is non-empty."

The Axiom of choice holds true if and only if [AC*] holds true.

Proof:

($\Rightarrow$) Suppose the Axiom of choice holds true. Let $\mathscr{S} = \{S_\alpha : \alpha \in \gamma\}$ be a set of non-empty sets. Then there exists a choice function $f : \mathscr{S} \to \cup_{\alpha\in\gamma} S_\alpha$ such that $f(S_\alpha) = s_\alpha$ where s_α is some element in S_α. Then $(s_\alpha : \alpha \in \gamma) \in \Pi_{\alpha\in\gamma} S_\alpha$. Hence, $\Pi_{\alpha\in\gamma} S_\alpha$ is non-empty. Then [AC*] holds true.

($\Leftarrow$) Suppose the statement [AC*] holds true. Let $\mathscr{S} = \{S_\alpha : \alpha \in \gamma\}$ be a set of non-empty sets. Since $\Pi_{\alpha\in\gamma} S_\alpha$ is non-empty, it contains some element $(s_\alpha : \alpha \in \gamma) \in \Pi_{\alpha\in\gamma} S_\alpha$. Then, for each $\alpha \in \gamma$, $s_\alpha \in S_\alpha$. The function, $f : \mathscr{S} \to \cup_{\alpha\in\gamma} S_\alpha$, defined as, $f(S_\alpha) = s_\alpha$, is well-defined. Then the Axiom of choice holds true.

We will now prove that the statement "Every set is well-orderable" and the Axiom of choice are equivalent statements.

---

[1] Paul Cohen has also proven that the Axiom of countable choice cannot be proven from the ZF-axioms.

Theorem 30.3 The statement "Every set is well-orderable" holds true if and only if the Axiom of choice holds true.

Proof:

($\Leftarrow$) That the Axiom of choice implies "Every set is well-orderable" is proven in Well-ordering theorem 29.4.

($\Rightarrow$) What we are given: That T is a well-orderable set and $\mathscr{S}$ is a class of non-empty subsets of T.
What we are required to show: There exists a function $f : \mathscr{S} \to T$ which maps each set S in $\mathscr{S}$ to some element $s \in S$.
Since T is well-orderable there exists a function f mapping some ordinal $\alpha \in \mathscr{O}$ one-to-one onto T. For any $S \in \mathscr{S}$, $f^{-1}[S]$ is a non-empty subset of α and so must have a least element α^* (since α is well-ordered). Then $f(\alpha^*)$ is the least element of S with respect to the well-ordering induced on T by α. The function $g : \mathscr{S} \to T$ defined as $g(S) = f(\alpha^*)$ is a well-defined function mapping each set S in $\mathscr{S}$ to one of its elements, as required.

30.5 Some consequences of the Axiom of choice.

The Axiom of choice has already been invoked a few times in the earlier sections to prove important statements which are unprovable without it. There are many such statements, intuitively felt to be true, which cannot be proven from the ZF-axioms. That is, the Axiom of choice is the only key we can use to unlock certain "doors" which would remain closed otherwise. A simple example is the following statement:

> If $f : S \to T$ is a function mapping S onto T, then there exists a function $g : T \to S$ mapping T into S such that $f \circ g = I_T$, the identity map on T.

This statement seems obviously true. It is trivially true if the function f is one-to-one and onto. If f is not one-to-one, at first glance the proof appears to be straightforward. It would go something like this:

- Since f is onto T, for each $y \in T$, $f^{\leftarrow}[\{y\}]$ is non-empty.
- For each $y \in T$, choose $u_y \in f^{\leftarrow}[\{y\}]$.
- Define the function $g : T \to S$ as follows: $g(y) = u_y$.
- Then $(f \circ g)(y) = f(g(y)) = f(u_y) = y$. So $f \circ g$ is the identity function on T.

A minor flaw in this proof is that the statement "For each $y \in T$, choose $u_y \in f^{\leftarrow}[\{y\}]$" is not appropriately justified. The Axiom of choice must be invoked to justify the choice of an element in each set of an infinite number of sets.

We state a few other well-known statements whose proofs depend on the Axiom of choice. That is, the following results hold true only if we accept the Axiom of choice as an axiom along with the other ZF-axioms.

> "*Every infinite set contains a one-to-one image of the natural numbers.*" The proof is given in theorem 18.9.
>
> "*Any infinite set can be expressed as the union of a pairwise disjoint set of infinite countable sets.*"

The proof is provided in the theorem below.

Theorem 30.4 [AC] Any infinite set can be expressed as the union of a pairwise disjoint set of infinite countable sets.

Proof:

What we are given: That S is an infinite set.
What we are required to show: That there exists a set $\mathscr{F} = \{U_\alpha : \alpha \in \phi\}$ of countably infinite pairwise disjoint sets such that $S = \cup\{U_\alpha : \alpha \in \phi\}$.

We will recursively construct the set $\mathscr{F}$:

- Let $\mathscr{T} = \{F \in \mathscr{P}(S) : \ F \text{ is countably infinite}\}$. Since S is a set, then $\mathscr{P}(S)$ is a set and so $\mathscr{T}$ is a set. Since S is infinite, $\mathscr{T}$ must contain at least one countably infinite subset U_0 of S (By theorem 18.9).
- If $S - U_0$ is finite, then S is countably infinite and we can let $\mathscr{F} = \{S\}$; we are done.
- More generally: Let $K_\gamma = \{U_\alpha : \alpha \in \gamma\}$ be a set of countably infinite pairwise disjoint subsets of S indexed with the elements of the ordinal γ. Either $S - \cup\{U_\alpha : \alpha \in \gamma\}$ is finite or it is infinite.
  - If $S - \cup\{U_\alpha : \alpha \in \gamma\}$ is infinite, then choose an arbitrary element U_γ of $\mathscr{T}$ which is entirely contained in $S - \cup\{U_\alpha : \alpha \in \gamma\}$. We then obtain the set $K_{\gamma+1} = \{U_\alpha : \alpha \in \gamma + 1\}$ of countably infinite pairwise disjoint subsets of S. The Axiom of choice will allow us to make the selection of U_γ for all such sets K_γ of countably infinite sets.
- Let $\mathscr{F} = \{U_\alpha : \alpha \in \phi\}$ be the set of all countably infinite sets obtained in this way. Then $\mathscr{F}$ is a set of pairwise disjoint countably infinite subsets of S. Now either $\cup\{U_\alpha : \alpha \in \phi\}$ is equal to S or it is not. If it is equal to S, then we are done. If it is not equal to S, then $S - \cup\{U_\alpha : \alpha \in \phi\}$ is finite. In such a case we can throw those last few elements in U_α for some $\alpha \in \phi$. We then obtain $S = \cup\{U_\alpha : \alpha \in \phi\}$ as required.

30.6 Zorn's lemma

Zorn's lemma is one of the most commonly used statements which is equivalent to the Axiom of choice. It refers to a specific property of a partially ordered set. Recall that a chain of a partially ordered set X is a linearly ordered subset of X. A maximal element of the partially ordered set is an element m of X for which there does not exist an element x of X such that $m < x$. Zorn's lemma states:

> "*If every chain of a partially ordered set* $(X, <)$ *has an upper bound, then* X *has a maximal element.*"

We will first prove that the Axiom of choice implies Zorn's lemma holds true. This will be followed by the proof of its converse.

Theorem 30.5 [AC] Let $(X, <)$ be a partially ordered set. If every chain of X has an upper bound, then X has a maximal element.

Proof:

What we are given: That X is a partially ordered set. For every chain C in X, there is an element k_C in X such that $c \leq k_C$ for all $c \in C$.
What we are required to show: That X contains an element m such that no element x of X satisfies the property $m < x$.

Let ϕ be the Hartogs number of X. That is, ϕ is the least ordinal number which is not equipotent with any subset of X.

Proof by contradiction. Suppose X has no maximal element. Then, for every element $s \in X$, the set $s^* = \{x : x > s\}$ is non-empty. The Axiom of choice guarantees the existence of a choice function $f : \mathscr{P}(X) \to X$ which maps the set s^* to some element $f(s^*) \in s^*$.

We recursively define the function $g : \phi \to X$ as follows:

$$\begin{array}{rcccl}
g(0) & = & x_0 & = & f(X) \\
g(1) & = & x_1 & = & f(x_0^*) > x_0 \\
g(2) & = & x_2 & = & f(x_1^*) > x_1 \\
 & \vdots & & \vdots & \\
g(\alpha^+) & = & x_{\alpha^+} & = & f(x_\alpha^*) > x_\alpha \\
\text{If } \lambda = \text{limit ordinal, } g(\lambda) & = & x_\lambda & = & \text{upper bound of the chain } \{x_\alpha : \alpha \in \lambda\} \\
 & \vdots & & \vdots &
\end{array}$$

In the case where λ is a limit ordinal, the hypothesis guarantees that the upper bound of the chain $\{x_\alpha : \alpha \in \lambda\}$ exists in X. Since the function g is strictly increasing it is one-to-one. Since X has no maximal element, the function g maps $\phi = \{\alpha : \alpha \in \phi\}$ one-to-one into X, contradicting the fact that X's Hartog number ϕ is the least ordinal

which cannot be mapped one-to-one into X. The source of the contradiction is our assumption that X has no maximal element. We must conclude that X has a maximal element, as required.

Theorem 30.6 [ZL] Suppose that those partially ordered sets $(X, <)$ in which every chain has an upper bound must have a maximal element. Then, given any subset $\mathscr{S} \subseteq \mathscr{P}(S) - \varnothing$, there exists a choice function $f : \mathscr{S} \to S$ which maps each set in $\mathscr{S}$ to one of its elements.

Proof:

What we are given: That partially ordered sets $(X, <)$ in which every chain has an upper bound have a maximal element. That $\mathscr{S} \subseteq \mathscr{P}(S) - \varnothing$.
What we are required to show: That there exists a choice function $f : \mathscr{S} \to S$ which maps each set in $\mathscr{S}$ to one of its elements.
Let

$$\mathscr{F} = \{f : \mathscr{D} \to S : f \text{ is a choice function with domain } \mathscr{D} \subset \mathscr{S}\}$$

Note that $\mathscr{F}$ is non-empty since it has been shown that all finite sets of sets have a choice function. We will partially order the functions in $\mathscr{F}$ by inclusion "$\subset$". That is,

$$[f \subset g] \Leftrightarrow [\{(S, s) \in \mathscr{S} \times S : f(S) = s\} \subset \{(S, s) \in \mathscr{S} \times S : g(S) = s\}]$$

Let C be a chain in $(\mathscr{F}, \subset)$. Then $\cup\{f : f \in C\}$ is an upper bound of C which is contained in $\mathscr{F}$. Then every chain in the partially ordered set $(\mathscr{F}, \subset)$ has an upper bound. By hypothesis, $(\mathscr{F}, \subset)$ has a maximal element, say h.

We claim that $h : \mathscr{S} \to S$ is a choice function on $\mathscr{S}$
Suppose not. That is, suppose there exists a non-empty element $S^* \in \mathscr{S}$ which does not belong to the domain of h. Let $h^* = h \cup \{(S^*, s^*)\}$. Since S^* is non-empty, then there exists an element s^* in S^*. Then $h \subset h^* \in \mathscr{F}$. This is a contradiction, since h was declared to be a maximal element of $\mathscr{F}$. We conclude that $h : \mathscr{S} \to S$ is a choice function for $\mathscr{S}$.

Hence, the Axiom of choice follows from Zorn's lemma.

30.7 A consequence of Zorn's lemma

We now provide a proof of a statement normally encountered in a course of linear algebra. In first year courses, this statement is often stated without proof. Postulating this statement is easier than postulating the more abstract Zorn's lemma. We begin by recalling a few facts.

A basis B of a vector space V is a non-empty subset of V with special properties. This set B need not be finite. But it must satisfy two properties:

1) Every *finite* linear combination of elements in B which equals zero must have only zeroes as coefficients. This is called the *linear independence property.*
2) Every vector in V is a linear combination of finitely many elements of B. In such cases we say that "*B spans V*".

There are algorithms we can use to find a basis for a finite dimensional vector space. But some vector spaces do not have a finite basis. For example, the vector space of all countably infinite sequences of real numbers, $(a_1, a_2, a_3, \dots,)$ does not have a finite basis. But if we assume Zorn's lemma, we can prove that it has a basis. But this proof does not show how to construct it or produce an explicit basis for vector spaces which have no finite spanning family. In fact, for this vector space, no basis can be found. We prove below that a basis exists for any vector space. The acronym [ZL] next to the theorem statement informs the reader that the proof invokes Zorn's lemma.

Theorem 30.7 [ZL] Every vector space has a basis.

Proof:

Let V be a vector space and let $(\mathscr{F}, \subseteq_{\mathscr{F}})$ be the set of all linearly independent subsets of the vector space V ordered by inclusion $\subseteq_{\mathscr{F}}$. The set $\mathscr{F}$ is non-empty since non-zero singleton sets are linearly independent.

Let $\mathscr{C}$ be a chain of linearly independent subsets in $\mathscr{F}$.
We claim that the union $\cup_{C\in\mathscr{C}}C$ is also linearly independent:

- It suffices to show that every *finite* linear combination of elements of $\cup_{C\in\mathscr{C}}C$ which equals zero must have zeroes as coefficients. Let $U = \{v_1, v_2, v_2, \dots, v_n\}$ be a set of vectors in $\cup_{C\in\mathscr{C}}C$.
- Then $U \subseteq C$ for some $C \in \mathscr{C}$ (since C is a chain of subsets).
- Since C is linearly independent, then $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \cdots + \alpha_n v_n = 0$ implies $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.
- So $\cup_{C\in\mathscr{C}}C$ is linearly independent as claimed.

Then every chain $\mathscr{C}$ in $(\mathscr{F}, \subseteq_{\mathscr{F}})$ has an upper bound in $\mathscr{F}$.
By Zorn's lemma, $(\mathscr{F}, \subseteq_{\mathscr{F}})$ has a maximal linearly independent set B^*. That is, B^* is a linearly independent set that is not a subset of any other linearly independent set. We now show that B^* spans V. If $v \in V - B^*$ is not a linear combination of vectors in B^*, then $B^* \cup \{v\}$ is a linearly independent subset of V which properly contains B^*, contradicting the maximality of B^*. So B^* spans V. So B^* is a basis of V. Thus, every vector space has a basis.

Concepts review:

1. What does the Well-ordering principle say?

2. What does the Axiom of choice say?
3. Is the Axiom of choice required to justify the existence of a choice function for finite sets?
4. Is an axiom required to justify the existence of a choice function for countably infinite subsets of $\mathscr{P}(S)$?
5. What does the Axiom of countable choice say?
6. Provide an example of a statement whose proof requires the Axiom of choice.
7. Does the Axiom of countable choice follow from the ZF-axioms?
8. State Zorn's lemma.
9. What linear algebra statement can be proved by invoking Zorn's lemma.

EXERCISES

A.

1. For which of the following sets of sets is the Axiom of choice required to guarantee the existence of a function which selects an element from each set:
   a) An infinite set of sets $\mathscr{S}$ where each set in $\mathscr{S}$ contains one element.
   b) Three sets each containing all elements of $\mathbb{R}$.
   c) A countably infinite number of sets each containing all the rational numbers.
   d) Uncountably many sets each containing three identical golf balls.
   e) A countably infinite number of sets each containing uncountably many golf balls and one marble.
   f) An uncountably infinite number of sets each containing a pair of socks.
   g) An uncountably infinite number of sets each containing two blue marbles and one red one.

B.

2. Let U and V be non-empty sets. Suppose R is a relation in $U \times V$ with domain $T \subseteq U$. Prove that there exists a function $f : T \to V$ such that $f \subseteq R$.
3. Let U and V be non-empty sets. Prove that a function $f : U \to V$ maps U onto V if and only if there exists some function $h : V \to U$ such that $f \circ h$ is the identity function on U.
4. Let S be a non-empty set. Suppose U is a non-empty subset of the set $\mathscr{P}(S)$ partially ordered by "$\subseteq$". Prove that $\operatorname{lub} U = \cup_{x \in U} x$.

C.

5. If S is a set containing more than one element, show that there exists a one-to-one function $f : S \to S$ such that f maps no point x in S to itself. That is, $f(x) \neq x$ for all x in S.
6. Let $\mathscr{S}$ be a set of sets. Let $\mathscr{M} = \{U \in \mathscr{P}(\mathscr{S}) : X, Y \in U \text{ implies } X \cap Y \neq \varnothing\}$. Show that $\mathscr{M}$ contains a maximal element T with respect to the ordering "$\subseteq$". That is, $T \in \mathscr{M}$ and, for any $B \in \mathscr{S} - T$, $T \cup \{B\} \notin \mathscr{M}$.

31 / Regularity and the cumulative hierarchy

Summary. *In this section we state the Axiom of regularity and present some of its equivalent forms. We prove that the Axiom of regularity is equivalent to the statement "Every set has an $\in$-minimal element". We also show that in the presence of the Axiom of regularity, no set can be an element of itself. We define "well-founded sets" and show that in the presence of the Axiom of choice, the Axiom of regularity is equivalent to the statement "Every set is well-founded". The transitive closure of a set is defined. Finally we show how to construct in stages, starting with the empty set, a class of sets whose union, V, contains all sets. The class, V, is referred to as the "Von Neumann's universe". The Axiom of regularity is used to show that this class, V, indeed contains all sets. We then define the "cumulative hierarchy" and the "rank of a set".*

31.1 Equivalent forms of the Axiom of regularity

The Axiom of regularity is the only *ZFC* axiom we have not yet invoked to justify any steps in the proofs of theorems presented up to now. We restate this axiom:

> *Axiom of regularity*: Every non-empty set S has an element x such that $x \cap S = \varnothing$.

To help us better understand the meaning of the Axiom of regularity − also known as the Axiom of Foundation − we investigate some of its equivalent forms and consequences. In what follows, we will say that *m is a minimal element of the set S with respect to the order relation "$<$"* if S is ordered by "$<$" and S contains no element x such $x < m$.

Theorem 31.1 The Axiom of regularity holds true if and only if every non-empty set S contains a *minimal* element with respect to the membership relation "$\in$".

Proof:

($\Rightarrow$)
Given: For any non-empty set S, there exists an element m such that $m \cap S = \varnothing$.
Required to show: That every non-empty set S contains a *minimal* element with respect to the membership relation "$\in$".
Let S be a non-empty set and m be an element of S such that $m \cap S = \varnothing$. Suppose m is not a minimal element of S. That is, suppose S contains an element x such that $x \in m$. Then $x \in m \cap S \neq \varnothing$, contradicting our hypothesis. Then this element m is minimal in

S with respect to "$\in$", as claimed.

($\Leftarrow$)

Given: That every non-empty set S contains a minimal element m with respect to "$\in$".
Required to show: That S contains some element m such that $m \cap S = \varnothing$.

Suppose S is a set such that for every $m \in S$, there exists $x \in m \cap S \neq \varnothing$; then no element m in S is minimal with respect to $\in$. This contradicts our hypothesis. Hence, for every set S there exists m such $m \cap S = \varnothing$.

Up to this point in our study of sets, we have not encountered a set x such that $x \in x$. We thought it would be best to stay away from such "creatures", at least until we better understand the difficulties that they may cause. We will now see that the statement "No set can be an element of itself" is a consequence of the Axiom of regularity.

Theorem 31.2 [Axiom of regularity] No set is an element of itself.

Proof:

Let x be an element and $S = \{x\}$. Suppose $x \in x$. Then, $x \in x \cap S$. Then, for every element, x, of S, $x \in x \cap S \neq \varnothing$, This contradicts the Axiom of regularity. Then, for any set x, $x \notin x$.

So the *ZFC*-set theoretic universe contains no set which is an element of itself. Recall that in the formal definition of "ordinal numbers", we required ordinal numbers to be *strictly* $\in$-well-ordered. Now we see that defining ordinals as being simply $\in$-well-ordered would have been sufficient, in the sense that the strictly ordered property would follow from "regularity". Proceeding as we did allows us to see that ordinals exist as sets even in the absence of the Axiom of regularity.

Definition 31.3 We say that a class S is *well-founded* if S does not contain an infinite descending chain of sets. That is, there does not exist an infinite sequence $\{x_n : n \in \omega\}$ such that $\cdots \in x_4 \in x_3 \in x_2 \in x_1 \in x_0$.

The set-theoretic universe we have explored up to now was not assumed to be a *well-founded* universe. Sets which are not well-founded were simply not considered or raised as a subject for discussion. Attempting to prove that non-well-founded sets exist, or do not exist, was more or less viewed as a digression from the concepts we

were studying at that time.

We will now show that in the presence of the Axiom of choice,[1] the two statements "*Every set is well-founded*" and the *Axiom of regularity* are equivalent. This means that in the absence of the Axiom of regularity, we would have to accept that non-well-founded sets may exist and study what impact the existence of such sets has in our set-theoretic universe.[2]

Theorem 31.4 [AC] The *Axiom of regularity* and the statement "*Every set is well-founded*" are equivalent statements.

Proof:

($\Rightarrow$)

Given: The Axiom of regularity holds true.
Required to show: That all sets are well-founded.

Suppose there exists a set which contains an infinite descending chain of sets $S = \{x_n : n \in \omega\}$. This means that an $\in$-ordered chain such as $\cdots \in x_4 \in x_3 \in x_2 \in x_1 \in x_0$ exists. By hypothesis, S must contain some element m such that $m \cap S = \varnothing$. This element m must be equal to x_k for some $k \in \omega$. Since $x_{k+1} \in m \cap S$, we have a contradiction. So non-well-founded sets cannot exist in the presence of regularity.

($\Leftarrow$)

Given: That every set is well-founded.
Required to show: That every non-empty set S contains an element m such that $m \cap S = \varnothing$.

Suppose there exists a non-empty set S such that for every $m \in S$, $m \cap S \neq \varnothing$. Then there exists a relation $R \subset S \times S$ such that $(m, x) \in R$ if and only if $x \in m \cap S$. Since $m \cap S \neq \varnothing$ for all $m \in S$, the domain of R is all of S. Invoking the Axiom of choice, there exists a "choice function" $f : S \to S$, $f \subseteq R$, where, for each $m \in S$, $f(m) \in m \cap S$. Let $x_0 = f(S)$. We recursively define a function $g : \omega_0 \to S$ as follows:

$$\begin{array}{lcll} g(0) & = & x_0 & = & f(S) \\ g(1) & = & x_1 & = & f(x_0) \in x_0 \cap S \\ g(2) & = & x_2 & = & f(x_1) \in x_1 \cap S \\ & \vdots & & \vdots & \\ g(n+1) & = & x_{n+1} & = & f(x_n) \in x_n \cap S \\ & \vdots & & \vdots & \end{array}$$

[1] The theorem actually uses a weak form of the Axiom of choice, called the Axiom of dependent choice

[2] It was shown in 1929 by von Neumann that if "*ZF* without regularity" is consistent, then "*ZF* with regularity" is also consistent.

We have thus constructed a set $\{x_n : n \in \omega_0\}$ where, for each $n \in \omega_0$, $x_{n+1} \in x_n$. The set $\{x_n : n \in \omega_0\}$ is an infinite descending chain. Its existence contradicts the hypothesis stating that S is a well-founded set, a set in which no such chain can exist. The source of our contradiction is our supposition that S does not contain a minimal element m. Hence, in presence of the Axiom of choice, the Axiom of regularity follows from the statement "Every set is well-founded".

31.2 Transitive closure of a set

Before we discuss another characterization of the Axiom of regularity, we introduce what is known as the *transitive closure* of a set. Recall that a set S is a transitive set if whenever $y \in S$ and $x \in y$, then $x \in S$.

Definition 31.5 Let x be a set. The *transitive closure of* x is a set t_x satisfying the following three properties:

1) The set t_x is a transitive set.

2) $x \subseteq t_x$

3) t_x is the $\subseteq$-least transitive set satisfying properties 1 and 2.

Example − Consider the set $S = \{2\}$. We see that the set S is not transitive, since $1 \in 2 = \{0, 1\}$, $2 \in S$, but $1 \notin S$. So S does not contain all the elements required for it to be transitive. Starting with S, we will construct, step by step, its transitive closure, t_S. We have seen that the element 1 is missing, so let's add it to S: Let $S_1 = \{1, 2\}$. We see that S_1 is not transitive since since $0 \in 1 = \{0\}$, $1 \in S_1$, but $0 \notin S$. We then add to S_1, the element 0: Let $S_2 = \{0, 1, 2\}$. We see that S_2 is the natural number 3 known to be transitive. Then, the transitive closure, t_S, of $S = \{2\}$ is the natural number $3 = \{0, 1, 2\}$.

Completing a non-transitive set, S, to its transitive closure, t_S, means to add to S all the elements which belong to elements of the set.

The following theorem guarantees that every non-empty set, x, has a transitive closure, t_x.

Theorem 31.6 Let x be a set. Then there exists a smallest transitive set, t_x, which contains all elements of x. That is, every set, x, has a transitive closure, t_x.

Proof:

Let $\mathscr{S}$ denote the class of all sets. Let $f : \mathscr{S} \to \mathscr{S}$ be a function defined as: $f(u) = \cup\{y \in \mathscr{S} : y \in u\}$. Let $x_0 \in \mathscr{S}$. We recursively define the function $g : \omega_0 \to \mathscr{S}$ as follows:

$$\begin{array}{lcll} g(0) & = & x_0 & \in \ \mathscr{S} \\ g(1) & = & x_1 & = \ f(x_0) = \cup\{y \in \mathscr{S} : y \in x_0\} \\ g(2) & = & x_2 & = \ f(x_1) = \cup\{y \in \mathscr{S} : y \in x_1\} \\ & \vdots & & \vdots \\ g(n) & = & x_n & = \ f(x_{n-1}) = \cup\{y \in \mathscr{S} : y \in x_{n-1}\} \\ g(n+1) & = & x_{n+1} & = \ f(x_n) = \cup\{y \in \mathscr{S} : y \in x_n\} \\ & \vdots & & \vdots \end{array}$$

Let

$$t_{x_0} = \cup\{x_n : n \in \omega_0\} = x_0 \cup x_1 \cup x_2 \cup \ \cdots$$

Since x_0 is a set, each x_n is the union of a set of sets and so t_{x_0} is itself a set.

Claim: That t_{x_0} is a transitive set.

Suppose $u \in t_{x_0}$ and $v \in u$. We are required to show that $v \in t_{x_0}$. Since $t_{x_0} = \cup\{x_n : n \in \omega_0\}$, $u \in x_k$ for some $k \in \omega$. Since $x_{k+1} = f(x_k) = \cup\{y \in \mathscr{S} : y \in x_k\}$, $u \subseteq x_{k+1}$. Hence, $v \in u \subseteq x_{k+1} \subseteq \cup\{x_n : n \in \omega_0\} = t_{x_0}$. So t_{x_0} is a transitive set, as claimed.

Suppose now that s is some transitive set such that $x_0 \subseteq s$.

Claim: $t_{x_0} \subseteq s$.

It suffices to show that $x_n \subseteq s$ for all n. We will show this by induction.
Base case: That $x_0 \subseteq s$ is given.
Inductive hypothesis: Suppose $x_n \subseteq s$. We are required to show that $x_{n+1} \subseteq s$. If $u \in x_{n+1} = f(x_n) = \cup\{y \in \mathscr{S} : y \in x_n\}$, then $u \in y$, for some $y \in x_n \subseteq s$. Since s is transitive, $u \in y \subseteq s$ implies $u \in s$. Then $x_{n+1} \subseteq s$.

Hence, by mathematical induction, $x_n \subseteq s$, for all n. Since $t_{x_0} = \cup\{x_n : n \in \omega\}$, $t_{x_0} \subseteq s$, as claimed.

We have thus constructed the smallest transitive set t_{x_0} which contains x_0. Then for any set x there exists a smallest transitive set t_x which contains x.

31.3 Constructing the class of all sets in stages

In this section we will construct, by stages, an ordinal-indexed class

$$\{V_\alpha : \alpha \in \mathscr{O}\}$$

of sets, starting with the empty set $V_0 = \varnothing$. This structured class, ordered by inclusion "$\subset$", is referred to as the *Cumulative hierarchy of sets*.[1] The union, $V = \cup\{V_\alpha :$

[1]The terms "rank hierarchy of sets" is sometimes used.

$\alpha \in \mathscr{O}\}$, of all the V_α-sets is referred to as the *Von Neumann universe* of sets or the *Von Neumann hierarchy of sets*. The restrictions imposed by the Axiom of regularity on the type of sets which belong to the *ZFC*-universe will guarantee that every set which belongs to the *ZFC*-universe of sets also belongs to some $V_\alpha \subset V$. That is, V accounts a for all sets whose existence is determined from the *ZFC*-axioms.

We have often used in this text the symbol $\mathscr{S}$ to denote the "class of all sets" in the *ZFC*-universe. The class V of sets which we will now investigate will also denote the class of all sets; but the class V is structured, as we will soon see. It is constructed in stages, starting only with the empty set $\varnothing$. The construction of the class, $\{V_\alpha : \alpha \in \mathscr{O}\}$, is described below.

Definition 31.7 Define the class function, $f : \mathscr{S} \to \mathscr{S}$, as $f(S) = \mathscr{P}(S)$. The elements of the class, $\{V_\alpha : \alpha \in \mathscr{O}\}$, belong to the image of the class function $g(\alpha) = V_\alpha$ recursively defined as follows:

$$\begin{array}{rcccccl}
g(0) & = & V_0 & = & \varnothing & = & 0 \\
g(1) & = & V_1 & = & f(V_0) & = & \mathscr{P}(0) = \{\varnothing\} = 1 \\
g(2) & = & V_2 & = & f(V_1) & = & \mathscr{P}(1) = \{\ \{\varnothing\}, \varnothing\ \} = 2^1 = 2 \\
g(3) & = & V_3 & = & f(V_2) & = & \mathscr{P}(2) = \{\ \{\{\varnothing\}, \varnothing\}, \{\{\varnothing\}\}, \{\varnothing\}, \varnothing\ \} \\
g(4) & = & V_4 & = & f(V_3) & = & \mathscr{P}(V_3) \ \text{(}2^4 = 16\text{ elements)} \\
 & \vdots & & \vdots & & & \\
g(\alpha^+) & = & V_{\alpha^+} & = & f(V_\alpha) & = & \mathscr{P}(V_\alpha) \\
 & \vdots & & \vdots & & & \\
\text{If } \lambda = \text{limit ordinal, } g(\lambda) & = & V_\lambda & = & \cup_{\alpha\in\lambda} V_\alpha & & \\
 & \vdots & & \vdots & & &
\end{array}$$

The class $\{V_\alpha : \alpha \in \mathscr{O}\}$ is called the *Cumulative hierarchy of sets*. The union of all the elements of the cumulative hierarchy of sets is denoted as:

$$V = \cup_{\alpha\in\mathscr{O}} V_\alpha$$

We verify that every single element, V_α of the cumulative hierarchy of sets is indeed a set: Since each V_α is either the image of a set, the power set of a set, or the union of a set of sets, then each V_α is a set (by the axioms of replacement, union and power set). That is,

$$\{V_\alpha : \alpha \in \mathscr{O}\}$$

is a class of sets.

Every V_α is a transitive set : Let $P(\alpha)$ denote the statement "V_α is transitive".

Since $V_0 = \varnothing$, $P(0)$ trivially holds true.
Suppose $P(\alpha)$ holds true. That is, suppose V_α is transitive.

$$\begin{aligned} y \in x \in V_{\alpha^+} = \mathscr{P}(V_\alpha) &\Rightarrow y \in x \subseteq V_\alpha \\ &\Rightarrow y \subseteq V_\alpha \quad \text{(Since } V_\alpha \text{ is transitive.)} \\ &\Rightarrow y \in \mathscr{P}(V_\alpha) = V_{\alpha^+} \\ &\Rightarrow V_{\alpha^+} \text{ is transitive} \\ &\Rightarrow P(\alpha^+) \text{ holds true.} \end{aligned}$$

Suppose γ is a limit ordinal and that $\alpha \in \gamma$ implies $P(\alpha)$ holds true.

$$\begin{aligned} y \in x \in V_\gamma = \cup_{\alpha\in\gamma} V_\alpha &\Rightarrow y \in x \in V_\beta \text{ for some transitive } V_\beta \\ &\Rightarrow y \in V_\beta \subseteq V_\gamma \\ &\Rightarrow V_\gamma \text{ is transitive} \\ &\Rightarrow P(\gamma) \text{ holds true.} \end{aligned}$$

By transfinite induction, every V_α is transitive.

Up to now, we have always represented the class of all sets which has evolved from the ten *ZFC*-axioms by the symbol, $\mathscr{S} = \{x : x \text{ is a set}\}$. Since every element of $\{V_\alpha : \alpha \in \mathscr{O}\}$ is a set, we can write, $\{V_\alpha : \alpha \in \mathscr{O}\} \subseteq \mathscr{S}$. Furthermore,

$$\begin{aligned} x \in V = \cup_{\alpha\in\mathscr{O}} V_\alpha &\Rightarrow x \in V_\gamma, \quad \text{for some } \gamma \in \mathscr{O} \\ &\Rightarrow x \subseteq V_\gamma, \quad \text{Since } V_\gamma \text{ is transitive.} \\ &\Rightarrow x \text{ is a set.} \\ &\Rightarrow x \in \mathscr{S} \end{aligned}$$

Hence

$$V = \cup_{\alpha\in\mathscr{O}} V_\alpha \subseteq \mathscr{S}$$

We now wonder whether $\mathscr{S} \subseteq \cup\{V_\alpha : \alpha \in \mathscr{O}\}$. That is, are all sets accounted for in V. We will prove that this is indeed the case. With this objective in mind we first establish the following lemma.

Lemma 31.8 The class $\{V_\alpha : \alpha \in \mathscr{O}\}$ is a strictly $\in$-increasing ($\subset$-increasing) chain of transitive sets. That is, if $\alpha \in \beta$, then $V_\alpha \in V_\beta$. Hence, $V_\alpha \subset V_\beta$.

Proof:

We have proven above that V_α is transitive for every $\alpha \in \mathscr{O}$.

Let $P(\gamma)$ denote the statement "$\alpha \in \gamma$ implies $V_\alpha \in V_\gamma$".

Suppose $P(\beta)$ holds true all $\beta \in \gamma$. That is, if $\alpha \in \beta \in \gamma$, then $V_\alpha \in V_\beta$.
Suppose $\phi \in \gamma$. We are required to show that $V_\phi \in V_\gamma$.
– Case 1: If γ is a limit ordinal, then $V_\phi \in V_{\phi^+} \subset \cup_{\alpha\in\gamma} V_\alpha = V_\gamma$.
– Case 2: Suppose $\gamma = \psi^+$. If $\phi \in \psi$, then $V_\phi \in V_\psi$, by the inductive hypothesis. If $\phi = \psi$, then $V_\phi = V_\psi \in \mathscr{P}(V_\psi) = V_{\psi^+} = V_\gamma$; hence, $P(\gamma)$ holds true.

By transfinite induction, $\alpha \in \gamma$ implies $V_\alpha \in V_\gamma$ for all γ. Since V_γ is transitive, $V_\alpha \subset V_\gamma$ for all γ

Lemma 31.9 For any non-empty set B, $B \subset V$ implies $B \in V$.

Proof:

Given: B is a non-empty set such that $B \subset V = \cup\{V_\alpha : \alpha \in \mathscr{O}\}$.
Required to show: That $B \in V$.

Let $u \in B$. Then $B \subset V \Rightarrow u \in V_\alpha$ for some $\alpha \in \mathscr{O}$.
Then the set $\{\alpha \in \mathscr{O} : u \in V_\alpha\}$ is non-empty. This ensures that the function $f : B \to \mathscr{O}$ defined as

$$f(u) = \text{least}\{\alpha \in \mathscr{O} : u \in V_\alpha\}$$

is well-defined.

- Since B is a set, by the Axiom of replacement, $f[B]$ is a set of ordinals. Since $f[B]$ is a set, $\beta = \cup_{\alpha\in f[B]}\alpha$ (a union of a set of sets) is a set. By theorems 27.11 and 27.12, β is an ordinal.
- Since $\alpha \subseteq \beta$ for all $\alpha \in f[B]$ and β is transitive, $\alpha \in_= \beta$, for all $\alpha \in f[B]$.[1]
- By lemma 31.8, for every $\alpha \in f[B]$, $[\alpha \in_= \beta] \Rightarrow [V_\alpha \subseteq V_\beta]$.

Then, for every $u \in B$, $u \in V_{f(u)} \subseteq V_\beta$. This implies that $B \subseteq V_\beta$ and so $B \in_= \mathscr{P}(V_\beta) = V_{\beta^+} \subset V$, as claimed.

In the following theorem we refer to the sets in $\mathscr{S}$, the class of all sets.

Theorem 31.10 [Axiom of regularity] For every set x, $x \in V = \cup_{\alpha\in\mathscr{O}} V_\alpha$. That is, $\mathscr{S} \subseteq V = \cup_{\alpha\in\mathscr{O}} V_\alpha$.

Proof:

[1] Recall that "$\in_=$" is read as "is an element of, or is equal to".

Given: That x is a set and that $V = \cup_{\alpha \in \mathscr{O}} V_\alpha$.
Required to show: That $x \in V$.

If $x = \varnothing$, then $x \in V$, we are done. We then suppose $x \neq \varnothing$.
Claim : The transitive closure, t_x, of the set x, is a proper subset of V.
Suppose $t_x \not\subset V$. That is, suppose there is an element u such that $u \in t_x$ and $u \notin V$. Then the set $U = \{u \in t_x : u \notin V\}$ is non-empty. Since U is non-empty, by the Axiom of regularity, U contains a minimal element $m \in U$. That is, $m \in U$ and $m \cap U = \varnothing$. Note that $m \notin V$. If $m = \varnothing$ then $m \in V$, a contradiction; so m is a non-empty set. Suppose $y \in m$. Since t_x is transitive and $m \in t_x$, then $y \in t_x$. But since $m \cap U = \varnothing$, y cannot belong to U. Then $y \in V$. We conclude that $m \subset t_x \cap V$. By the previous lemma, $[m \subset V] \Rightarrow [m \in V]$. This contradicts the fact that $m \in U$. The source of this contradiction is our supposition that $U = \{u \in t_x : u \notin V\}$ is non-empty. Then $U = \{u \in t_x : u \notin V\}$ is empty and so $t_x \subset V$, as claimed.
By the previous lemma, $t_x \in V$. Then $t_x \in V_\alpha$ for some α.

$$\begin{aligned} x \subseteq t_x \in V_\alpha &\Rightarrow x \subseteq t_x \subseteq V_\alpha \quad \text{(Since } V_\alpha \text{ is a transitive set)} \\ &\Rightarrow x \subseteq V_\alpha \\ &\Rightarrow x \in \mathscr{P}(V_\alpha) = V_{\alpha^+} \subset V \end{aligned}$$

So $x \in V$, as required.

Since $\mathscr{S} \subseteq V$ and every element of V is a set, then the class $V = \cup_{\alpha \in \mathscr{O}} V_\alpha$ provides us with a description of the set-theoretic universe, $\mathscr{S}$, which evolves from the ten *ZFC*-axioms (nine *ZF*-axioms plus Choice) listed in chapter one. It is interesting to note the class, $V = \cup_{\alpha \in \mathscr{O}} V_\alpha$, contains all sets only in the presence of the Axiom of regularity. That is, without the restrictions that the Axiom of regularity imposes on sets, some sets in $\mathscr{S}$ would not be accounted for in V. We now show that if $\mathscr{S} \subseteq V$, then the Axiom of regularity holds true on $\mathscr{S}$.

Theorem 31.11 Let $V = \cup_{\alpha \in \mathscr{O}} V_\alpha$ be the class of sets constructed as described above and $\mathscr{S}$ denote the class of all sets. If $\mathscr{S} \subseteq V$, then every set in $\mathscr{S}$ has a $\in$-minimal element.

Proof:

What we are given: That the class V contains all sets.
What we are required to show: That every non-empty set x contains a $\in$-minimal element.

Suppose x is a non-empty set. By hypothesis, x belongs to $V = \cup_{\alpha \in \mathscr{O}} V_\alpha$. Then $x \in V_\alpha$ for some α. Since V_α is transitive $x \subset V_\alpha \subset V$. We can then define a function $f : x \to \mathscr{O}$ as $f(u) = \text{least}\{\alpha \in \mathscr{O} : u \subset V_\alpha\}$. Since x is a non-empty set, by the Axiom of replacement, $f[x]$ is a non-empty set of ordinals and so contains a least element, say ϕ. Since ϕ

is in the image of x under f, there exists an element m in the domain x of f such that $f(m) = \phi$. Then ϕ is the least ordinal such that $m \subset V_\phi$; equivalently, it is the least ordinal such that $m \in V_{\phi^+}$. Then $m \notin V_\phi$.

We claim that m is an $\in$-minimal element of x: Suppose not. Suppose $z \in m \cap x$. Then since $z \in m \subset V_\phi$ implies $z \in V_\phi$ we have $f(z) \in f(m) = \phi$. This contradicts the fact that ϕ is minimal in $f[x]$. Then m is an $\in$-minimal of x as claimed.

Then every set has a minimal element with respect to "$\in$".

The above theorem shows that,

For every ordinal α, V_α satisfies the axiom of regularity.

31.4 The rank of a set.

The following concept will allow the reader to develop some familiarity with the elements of $\{V_\alpha : \alpha \in \mathscr{O}\}$ and how they are defined.

Given any set U, we will define the rank of the set U, denoted as rank(U), as follows:

$$\text{rank}(U) = \text{least}\{\alpha \in \mathscr{O} : U \subseteq V_\alpha\}$$

First note that rank : $V \to \mathscr{O}$ mapping a set, U, to a unique ordinal number, rank(U), is a well-defined function.

Example – We list the sets V_0 to V_4 to help us determine the rank of a few simple sets.

$$\begin{array}{lllllll} g(0) & = & V_0 & = & \varnothing & = & 0 \\ g(1) & = & V_1 & = & f(V_0) & = & \mathscr{P}(0) = \{\varnothing\} = 1 \\ g(2) & = & V_2 & = & f(V_1) & = & \mathscr{P}(1) = \{\ \{\varnothing\}, \varnothing\ \} = 2^1 = 2 \\ g(3) & = & V_3 & = & f(V_2) & = & \mathscr{P}(2) = \{\ \{\{\varnothing\}, \varnothing\}, \{\{\varnothing\}\}, \{\varnothing\}, \varnothing\ \} \\ g(4) & = & V_4 & = & f(V_3) & = & \mathscr{P}(V_3)\ \ (2^4 = 16 \text{ elements}) \end{array}$$

- Suppose $A = 3$. Since $3 = \{0, 1, 2\} = \{\ \varnothing, \{\varnothing\}, \{\{\varnothing\}, \varnothing\}\ \} \subset V_3$ and $3 \notin V_2$, then rank$(A) = 3$.
- Suppose $B = \{\{\varnothing\}\}$. We see that $\{\{\varnothing\}\} \subset V_2$ and $\{\{\varnothing\}\} \notin V_1$; hence, rank$(B) = 2$.
- Suppose $C = \{\ \{\{\{\varnothing\}\}\}, \varnothing\}$. We see that both $\{\{\{\varnothing\}\}\}$ and $\varnothing$ belong to $\mathscr{P}(V_3) = V_4$; hence, $C \subset V_4$. Since $\{\{\{\varnothing\}\}\} \notin V_3$, $C \not\subset V_3$; hence, rank$(C) = 4$.

If we know the rank of the elements of a set U it may help determine the rank of the set itself. We prove a few useful facts about the rank of various sets.

Theorem 31.12 Let $V = \cup\{V_\alpha : \alpha \in \mathscr{O}\}$.

a) The rank of the empty set $\varnothing$ is zero.

b) If $U \in V_\beta$, then $\text{rank}(U) < \beta$; hence, $U \notin V_{\text{rank}(U)}$ for all sets U.

c) If U is a set then $[\text{rank}(U) < \beta] \Rightarrow [U \in V_\beta]$.

d) If U and V are sets such that $U \in V$, then $\text{rank}(U) < \text{rank}(V)$.

e) If γ is an ordinal, then $\text{rank}(\gamma) = \gamma$.

Proof:

a) Note that if $U = \varnothing$, then $\text{rank}(U)$ is the least ordinal number α such that $\varnothing \subseteq \alpha$, namely 0. So $\text{rank}(\varnothing) = 0$.

b) Given: $U \in V_\beta$.
Required to show: $\text{rank}(U) < \beta$ and so, $U \notin V_{\text{rank}(U)}$ for all sets U.

To show this we consider two cases:

Case 1: Suppose $\beta = \alpha + 1$ for some α. Then $V_\beta = V_{\alpha+1} = \mathscr{P}(V_\alpha)$. Then $[U \in V_\beta] \Rightarrow [U \subseteq V_\alpha]$. It then follows that the $\text{rank}(U) \leq \alpha < \alpha + 1 = \beta$, as claimed.
Case 2: Suppose β is a limit ordinal. Since $U \in V_\beta = \cup_{\alpha\in\beta} V_\alpha$, then $U \in V_\alpha$ for some $\alpha \in \beta$. Since V_α is transitive $U \subseteq V_\alpha$ so again, $\text{rank}(U) \leq \alpha < \beta$, as required.

If $U \in V_{\text{rank}(U)}$, then, by case 1 and 2, $\text{rank}(U) < \text{rank}(U)$, a contradiction. Hence, for all sets U, $U \notin V_{\text{rank}(U)}$, as required.

c) Given: $\text{rank}(U) < \beta$.
Required to show: $U \in V_\beta$.
See that $[\text{rank}(U) < \beta] \Rightarrow [U \subseteq V_\alpha]$, for some α where $\alpha+1 \leq \beta$. Then $U \in V_{\alpha+1} \subseteq V_\beta$. Then $U \in V_\beta$, as required.

So for any ordinal α, V_α is precisely the set of all sets U whose rank is less than α.

d) Given: U and V are sets such that $U \in V$.
Required to show: $\text{rank}(U) < \text{rank}(V)$.

$$\begin{aligned} U \in V \subseteq V_{\text{rank}(V)} &\Rightarrow U \in V_{\text{rank}(V)} \\ &\Rightarrow \text{rank}(U) < \text{rank}(V) \quad \text{(By part b).)} \end{aligned}$$

e) Given: γ is an ordinal.
Required to show: That $\text{rank}(\gamma) = \gamma$.

Claim: $\text{rank}(\gamma) \leq \gamma$ for all ordinals γ.
Suppose $\gamma < \text{rank}(\gamma)$.

$$\begin{aligned}
\gamma < \text{ rank}(\gamma) &\Rightarrow \gamma < \beta \leq \text{ rank}(\gamma) \quad \text{For some ordinal } \beta. \\
&\Rightarrow \gamma \in \beta \in_{=} \text{ rank}(\gamma) \\
&\Rightarrow \gamma \in \beta \subseteq V_\beta \subseteq V_{\text{rank}(\gamma)} \quad \text{By part b) of lemma 31.8.} \\
&\Rightarrow \gamma \in V_{\text{rank}(\gamma)} \quad \text{Contradicting part b).}
\end{aligned}$$

The source of the contradiction is our supposition that $\gamma < \text{rank}(\gamma)$. Then $\text{rank}(\gamma) \leq \gamma$, as claimed.

We now show that $\text{rank}(\gamma) = \gamma$ for all ordinals γ. We prove this by transfinite mathematical induction. Suppose $\text{rank}(\alpha) = \alpha$ for all ordinals $\alpha \in \gamma$. It suffices to show that $\text{rank}(\gamma) = \gamma$. Since we have shown that $\text{rank}(\gamma) \leq \gamma$, it suffices to show that $\text{rank}(\gamma) \not< \gamma$. Suppose not. That is, suppose $\text{rank}(\gamma) = \beta < \gamma$. Then, by the inductive hypothesis, $\text{rank}(\beta) = \beta = \text{rank}(\gamma)$. Since, $\beta < \gamma$, $\beta \in \gamma$ and so, by part d), $\text{rank}(\beta) < \text{rank}(\gamma)$, contradicting $\text{rank}(\beta) = \text{rank}(\gamma)$. Hence, $\text{rank}(\gamma) < \gamma$ is impossible. Then $\text{rank}(\gamma) \not< \gamma$ and $\text{rank}(\gamma) \leq \gamma$ implies $\text{rank}(\gamma) = \gamma$. By transfinite induction, $\text{rank}(\gamma) = \gamma$ for all ordinals γ.

Remark: If α is any ordinal then, by part e) of theorem 31.12, $\text{rank}(\alpha) = \alpha$, therefore $\alpha \subseteq V_\alpha$ and $\alpha \not\subset V_\beta$, for any ordinal $\beta \in \alpha$. This means that the set of all natural numbers, ω_0, is a subset of V_{ω_0}. Furthermore, we mentioned earlier that $\mathbb{R} \subseteq \mathscr{P}^7(\omega_0) = V_{\omega_0+7}$ (see page 161); then we can safely say that $\mathbb{R} \subseteq V_{\omega_0+\omega_0}$. Hence, even for relatively small ordinals, α, V_α can contain many of the sets we really care about in mathematics. In fact, as we shall soon see, many of the *ZFC*-axioms hold true in some V_α where α is relatively small.

31.5 The *ZFC*-set-axioms ingrained in the V_α-sets.

In theorem 31.11, we have shown that if $\mathscr{S}$ denotes the class of all sets in the *ZFC*-universe, then $\mathscr{S} = V = \cup_{\alpha \in \mathscr{O}} V_\alpha$. Since all of the *ZFC*-axioms hold true on $\mathscr{S}$, then they must hold true in reference to the sets in $V = \cup_{\alpha \in \mathscr{O}} V_\alpha$. In particular, we have seen that if we exclude the Axiom of regularity from the list of the *ZFC*-axioms, then $\cup_{\alpha \in \mathscr{O}} V_\alpha$ is a proper subclass of $\mathscr{S}$.

We now investigate which of the V_α-sets satisfy each of the *ZFC*-axioms. This will provide us an opportunity to analyze and better understand what each of the *ZFC*-axioms mean. If Axiom A holds true in V_α then we say that "V_α is a *model* of Axiom A".

Axiom of pair. Recall that the Axiom of pair states that if a and b are sets, then $\{a, b\}$ is also a set. To show that a set K satisfies this axiom we must prove that

$a, b \in K \Rightarrow \{a, b\} \in K$.

Proposition 31.13 Let β be a limit ordinal and a and b be two elements of V_β. Then $\{a, b\}$ is an element of V_β. That is, V_β satisfies the property described by the *Axiom of pair*.

Proof:

We are given that a and b are two elements of V_β where β is a limit ordinal.
We are required to show that $\{a, b\} \in V_\beta$.
Since $a, b \in V_\beta = \cup_{\alpha\in\beta} V_\alpha$, then there exists some $\gamma \in \beta$ such that $\{a, b\} \subset V_\gamma \subset V_\beta$. Then $\{a, b\} \in \mathscr{P}(V_\gamma) = V_{\gamma+1} \subset V_\beta$. Then $\{a, b\} \in V_\beta$.

So if β is a limit ordinal, V_β is a model of the Axiom of pair.

Axiom of union. The Axiom of union states that if $A = \{U : U \in A\}$ is a set of sets, then the union $D = \cup\{U : U \in A\}$ is a set.

Proposition 31.14 Let β be any ordinal. If U is an element of V_β, then $\cup\{x : x \in U\} \in V_\beta$. That is, V_β satisfies the property described by the *Axiom of union*.

Proof:

Given: That $U \in V_\beta$.
Required to show: That $\cup\{x : x \in U\} \in V_\beta$.
Since $U \in V_\beta$, then, by property b) in theorem 31.12, $\text{rank}(U) < \beta$.
Let $y \in \cup\{x : x \in U\}$. Then y is an element of some $x \in U$. Since $y \in x \in U$, then $\text{rank}(y) < \text{rank}(x) < \text{rank}(U)$ (by part c) of theorem 31.12 above). By part b) above,

$$\Big[\,[\text{rank}(y) < \text{rank}(U)] \text{ and } [\text{rank}(U) < \beta]\,\Big] \Rightarrow [y \in V_{\text{rank}(U)} \subset V_\beta]$$

So $y \in V_\beta$. We deduce that $\cup\{x : x \in U\} \subseteq V_{\text{rank}(U)} \subset V_\beta$. This means that $\cup\{x : x \in U\} \in \mathscr{P}(V_{\text{rank}(U)}) \subseteq V_\beta$. It follows that $\cup\{x : x \in U\} \in V_\beta$, as required.
So every V_α satisfies the property described in the Axiom of union.

So, for any ordinal β, V_β is a model of the Axiom of union.

Axiom of power set. We now prove that for any limit ordinal β, V_β satisfies the property described by the Axiom of power set. To show that a set K satisfies this property we must show that if $U \in K$, then $\mathscr{P}(U) \in K$.

Proposition 31.15 Let β be a limit ordinal. If U is an element of V_β, then V_β also contains a set $Y = \mathscr{P}(U)$ such that $S \subseteq U$ implies $S \in Y$. That is, V_β satisfies the property described by the *Axiom of power set*.

Proof:

Given: That $U \in V_\beta$. Then by part b) of theorem 31.12, $\text{rank}(U) < \beta$.
Required to show: That $\mathscr{P}(U) \in V_\beta$.
Since $U \in V_\beta$, then, by property b) in theorem 31.12, $\text{rank}(U) < \beta$.
For any $A \subseteq U$, $A \subseteq U \subseteq V_{\text{rank}(U)}$. Then, for all subsets A of U,

$$A \in \mathscr{P}(V_{\text{rank}(U)}) = V_{\text{rank}(U)+1}$$

Then $\mathscr{P}(U) \subseteq V_{\text{rank}(U)+1}$, and so $\mathscr{P}(U) \in \mathscr{P}(V_{\text{rank}(U)+1}) = V_{\text{rank}(U)+2}$. Since β is a limit ordinal, and $\text{rank}(U) < \beta$, then $\text{rank}(U) + 1 < \text{rank}(U) + 2 < \beta$. Then we obtain

$$\mathscr{P}(U) \in \mathscr{P}(V_{\text{rank}(U)+1}) = V_{\text{rank}(U)+2} \subset V_\beta$$

We conclude that $\mathscr{P}(U) \in V_\beta$, as required.
So, for every limit ordinal β, V_β satisfies the property described by the Axiom of power set.

We have just shown that if β is limit ordinal, V_β is a model of the Axiom of power set.

Axiom of extent. We begin with a brief discussion of the Axiom of extent. We remind ourselves about what the Axiom of extent states: Given a pair of sets A and B, if for all elements x, $(x \in A \Leftrightarrow x \in B)$ holds true, then $A = B$. If we want to verify whether a set K satisfies the Axiom of extent we must show that: Given any pair of elements x and y in K, if for all $z \in K (z \in x \Leftrightarrow z \in y)$, then x and y are the same set.

We first provide an example of a set in which the Axiom of extent does *not* hold true. Suppose

$$K = \{\{b, c, d\}, \{a, b, d\}, \{b, c\}, b, d\}$$

For the pair $\{b, c, d\}$, $\{a, b, d\}$ in K, and elements, $b, d \in K$, we see that the conditions

$$\begin{aligned} b \in \{a, b, d\} &\Leftrightarrow b \in \{b, c, d\} \\ d \in \{a, b, d\} &\Leftrightarrow d \in \{b, c, d\} \end{aligned}$$

hold true in K, but in spite of this $\{a, b, d\} \neq \{b, c, d\}$. Although K "recognizes" b and d as its elements it does not "recognize" the elements a and c (meaning that

$a, c \notin K$). So the Axiom of extent does not hold true in the set K.

We now consider the set

$$H = \Big\{ \varnothing,\ \{\varnothing\},\ \{\{\varnothing\},\ \varnothing\},\ \{\{\varnothing\}\},\ \{\varnothing, \{\varnothing\}\} \Big\}$$

Now H satisfies the Axiom of extent since,

$$\begin{aligned} \varnothing \in \{\varnothing, \{\varnothing\}\} &\Leftrightarrow \varnothing \in \{\{\varnothing\}, \varnothing\} \\ \{\varnothing\} \in \{\varnothing, \{\varnothing\}\} &\Leftrightarrow \{\varnothing\} \in \{\{\varnothing\}, \varnothing\} \end{aligned}$$

and the set H "recognizes" all the elements of both sets $\{\varnothing, \{\varnothing\}\}$ and $\{\{\varnothing\}, \varnothing\}$, as elements of itself and so is in a position to confirm that $\{\varnothing, \{\varnothing\}\}$ and $\{\{\varnothing\}, \varnothing\}$ are indeed the same set. This is something that the set K in the previous example could not do. The reason that the Axiom of extent is satisfied in H is that H is a transitive set. The transitive property of the set K allows it to recognize all the elements of its sets and confirm which sets are equal and which sets are not. With these thoughts in mind we confirm that every set V_α satisfies the Axiom of extent.

Proposition 31.16 Let α be any ordinal. For any two elements x and y of V_α, if for all $z \in V_\alpha (z \in x \Leftrightarrow z \in y)$, then x and y are the same set.
Proof:

Given: That for any two elements x and y of V_α, $\forall z \in V_\alpha (z \in x \Leftrightarrow z \in y)$.
Required to show: That x and y are the same set.

Suppose x and y are *not* the same set. Then, without loss of generality there is some element $u \in y - x$ (that is, $u \in y$ but $u \notin x$). Since V_α is a transitive set, then $[u \in y \in V_\alpha] \Rightarrow [u \in V_\alpha]$. So we have, "$u \in y$ and $u \notin x$", for some $u \in V_\alpha$. But this contradicts our hypothesis "for all $z \in V_\alpha (z \in x \Leftrightarrow z \in y)$"; that is, for $u \in V_\alpha$, $(u \in y) \Rightarrow (u \in x)$.

So, for every ordinal α, V_α satisfies the property described by the Axiom of extent.

We now see that for any ordinal β, V_β is a model of the Axiom of extent.

Axiom of infinity. The Axiom of infinity states that there exists a non-empty *set A* (called an inductive set) that satisfies the condition: "$x \in A$" $\Rightarrow$ "$x \cup \{x\} \in A$". The set, ω_0, was defined to be the smallest inductive set. To show that a set K satisfies the Axiom of infinity we need to show that $\omega_0 \in K$.

Proposition 31.17 Let α be an ordinal number such that $\alpha > \omega_0$. Then $\omega_0 \in V_\alpha$.

Proof:

Given: That α be an ordinal number such that $\alpha > \omega_0$.
Required to show: That $\omega_0 \in V_\alpha$.

First note that V_n is finite for all $n \in \omega_0$. This is easily verified by a proof by mathematical induction on ω_0. Since $V_{\omega_0} = \cup\{V_n : n \in \omega_0\}$ and ω_0 is infinite, $\omega_0 \in V_{\omega_0}$ is impossible. Since ω_0 is an ordinal, $\text{rank}(\omega_0) = \omega_0$ (by theorem 31.12). Hence, the least ordinal α such that $\omega_0 \subseteq V_\alpha$ is ω_0. Then $\omega_0 \subseteq V_{\omega_0}$. This implies $\omega_0 \in \mathscr{P}(V_{\omega_0}) = V_{\omega_0+1} \subseteq V_\gamma$, for all $\gamma > \omega_0$.

We conclude that $\omega_0 \in V_\alpha$ for all α such that $\alpha > \omega_0$, as required.

We have shown that for any ordinal $\beta > \omega_0$, V_β is a model of the Axiom of infinity.

Axiom of subsets. The Axiom of subsets states that if S is a given set and ϕ is a formula expressed in the language of set theory, then the class of all elements x of S for which $\phi(x)$ holds true is a set. To show that the Axiom of subsets holds true in a set K, we must show that if M is a class whose elements belong to a set $S \in K$, then $M \in K$.

Proposition 31.18 Let α be an ordinal number. Then the Axiom of subsets holds true in V_α.

Proof:

Given: That α is any ordinal, that $S \in V_\alpha$ and $M = \{x \in S : \phi(x)\} \subseteq S$.
Required to show: That $M \in V_\alpha$.

Since $S \in V_\alpha$, then $\text{rank}(S) < \alpha$ (by theorem 31.12). Since $M = \{x \in S : \phi(x)\} \subseteq S$, then $\text{rank}(M) \leq \text{rank}(S) < \alpha$ (again by theorem 31.12). Given that $\text{rank}(M) < \alpha$, then $V_{\text{rank}(M)} \subset V_\alpha$. It follows that

$$M \subseteq V_{\text{rank}(M)} \in \mathscr{P}(V_{\text{rank}(M)}) = V_{\text{rank}(M)\,+\,1} \subseteq V_\alpha$$

Since M is a subset of $V_{\text{rank}(M)}$, then $M \in \mathscr{P}(V_{\text{rank}(M)}) \subseteq V_\alpha$. We conclude that $M \in V_\alpha$, as required.

We now see that for any ordinal β, V_β is a model of the Axiom of subsets.

Axiom of replacement. The Axiom of replacement can be expressed as follows: Let w be a set. If ϕ is a formula so that for every $x \in w$, $[\phi(x, z)$ and $\phi(x, y)] \Rightarrow [y = z]$, then the class $\{u : \phi(x, u)$ for some $u \in w\}$ is a set.

To show that a set K satisfies this axiom we must show that if $w \in K$ and ϕ is a formula so that for every $x \in w$, $[\phi(x, z)$ and $\phi(x, y)] \Rightarrow [y = z]$, then $\{u : \phi(x, u)$ for some $u \in w\} \in K$.

We first provide an example of a transitive set in which the Axiom of replacement does not hold true. Consider the set $V_{\omega_0\omega_0}$.[1] We first note that $\omega_0 \subseteq V_{\text{rank}(\omega_0)} = V_{\omega_0} \subset V_{\omega_0\omega_0}$; hence, $\omega_0 \subset V_{\omega_0\omega_0}$. Define the function $f : \omega_0 \to \omega_0\omega_0$ as follows: $f(n) = \omega_0 n$. Then $f[\omega_0] = \{\omega_0 n : n \in \omega_0\} = \omega_0\omega_0$.

We claim that $f[\omega_0] \notin V_{\omega_0\omega_0}$: Since $\text{rank}(\omega_0\omega_0) = \omega_0\omega_0$, then $\omega_0\omega_0$ is the least ordinal α such that $f[\omega_0] = \omega_0\omega_0 \subseteq V_\alpha$. Then $\omega_0\omega_0 + 1$ is the the least ordinal such that $f[\omega_0] = \omega_0\omega_0 \in \mathscr{P}(\omega_0\omega_0) = V_{\omega_0\omega_0+1}$. Then $f[\omega_0] = \omega_0\omega_0 \notin V_{\omega_0\omega_0}$. So the images of sets in $V_{\omega_0\omega_0}$ need not necessarily be sets in $V_{\omega_0\omega_0}$, as claimed. (Note that we can prove that $V_{\omega_0+\omega_0}$ does not satisfy the Axiom of replacement by mimicking this proof.)

We will, however, show that the Axiom of replacement holds true in V_{ω_0}.

Proposition 31.19 The set V_{ω_0} satisfies the property described by the Axiom of replacement.

Proof:

Given: That $w \in V_{\omega_0}$ and $f : w \to V$ is a function mapping w into V.
Required to show: That $f[w] \in V_{\omega_0}$.
Since $w \in V_{\omega_0} = \cup\{V_n : n \in \omega_0\}$, then $w \in V_k$ for some $k \in \omega_0$ and so w is a finite set. Since f is a function and w is finite, then $f[w]$ is a finite set. Say, $f[w] = \{b_0, b_1, b_2, \ldots, b_m\}$. Let $\beta = \text{lub}\{\text{rank}(b_i) + 1 : i = 0 \text{ to } m\}$. Then β is the least ordinal such that $b_i \in V_{\text{rank}(b_i)+1} \subseteq V_\beta$, for $i = 1$ to m. Then $f[w] = \{b_0, b_1, b_2, \ldots, b_m\} \subseteq V_\beta$. Then $f[w] = \{b_0, b_1, b_2, \ldots, b_m\} \in \mathscr{P}(V_\beta) = V_{\beta+1} \in V_{\omega_0}$, as required.

We have just shown that V_{ω_0} is a model of the Axiom of replacement.

Axiom of choice. The Axiom of choice states that if A is a set of sets, then there exists a function $f : A \to \cup\{x : x \in A\}$ which maps each subset x of A to an element of $y_x \in x$. Since $\mathscr{S} = V = \cup\{V_n : n \in \omega_0\}$, then we know that the Axiom of choice holds true on V. We now show that provided the ordinal γ is a limit ordinal, then the Axiom of choice also holds true on V_γ.

Proposition 31.20 Let γ be a limit ordinal. Then the set V_γ satisfies the property described by the Axiom of choice.

Proof:

[1] Recall that:
$\omega_0 2 = \{1, 2, 3, \ldots, \omega_0, \omega_0 + 1, \ldots, \omega_0 + n, \ldots, \}$,
$\omega_0 3 = \{1, 2, 3, \ldots, \omega_0, \ldots, \omega_0 2, \omega_0 2 + 1, \ldots, \omega_0 2 + n, \ldots, \}$, and
$\omega_0\omega_0 = \{1, 2, 3, \ldots, \omega_0, \ldots, \omega_0 2, \ldots, \omega_0 3 \ldots, \omega_0 n, \ldots, \}$

Given: That γ is a limit ordinal and that the Axiom of choice holds true on V.
Required to show: That the Axiom of choice holds true on V_γ.

Let $U \in V_\gamma = \cup\{V_\alpha : \alpha \in \gamma\}$ where $U \neq \varnothing$ and $\varnothing \notin U$. We are required to show that there exists a function $f \in V_\gamma$ mapping U onto a set $f[U] \in V_\gamma$ such that $f(x) \in x$ for each $x \in U$.

Since γ is a limit ordinal and $U \in V_\gamma = \cup\{V_\alpha : \alpha \in \gamma\}$, then $U \in V_\beta$ for some ordinal $\beta \in \gamma$. Since V_β is transitive,

$$\begin{aligned} U \in V_\beta &\Rightarrow U \subseteq V_\beta \\ &\Rightarrow x \in V_\beta \text{ for each } x \in U \\ &\Rightarrow x \subseteq V_\beta \text{ for each } x \in U \\ &\Rightarrow \cup\{x : x \in U\} \subseteq V_\beta \end{aligned}$$

Now U and $\cup\{x : x \in U\} \subseteq V_\beta$ are also elements of V. Since the Axiom of choice holds true on V, then there exists a function $f \in V$, $f : U \to \cup\{x : x \in U\}$ mapping each element x of U to an element $y_x \in x \in V_\beta$. We are required to show that $f \in V_\gamma$.

See that $f[U] = \{y_x : x \in U\} \subseteq V_\beta$; hence, $f[U] \in \mathscr{P}(V_\beta) = V_{\beta+1} \in V_\gamma$.

Claim: $f \in V_\gamma$.

We have seen that if $x \in U$, then $y_x = f(x) \in V_\beta \in V_\gamma$. Recall that the ordered pair (x, y_x) is defined as $\{\{x\}, \{x, y_x\}\}$. Given $(x, y_x) \in f$,

$$\begin{aligned} x, y_x \in V_\beta &\Rightarrow \{x, y_x\} \subseteq V_\beta \\ &\Rightarrow \{x, y_x\} \in \mathscr{P}(V_\beta) \\ &\Rightarrow \{x, y_z\} \in V_{\beta+1} \end{aligned}$$

Similarly, $x \in V_\beta$ implies $\{x\} \in V_{\beta+1}$. For the same reasons,

$$\{x\}, \{x, y_x\} \in V_{\beta+1} \Rightarrow (x, y_x) = \{\{x\}, \{x, y_x\}\} \in V_{\beta+2}$$

Then $f = \{(x, y_x) : x \in U\} \subseteq V_{\beta+2}$. This implies $f \in V_{\beta+3} \in V_\gamma$, and so $f \in V_\gamma$ as claimed.

We have shown that there exists a function $f \in V_\gamma$ mapping U onto a set $f[U] \in V_\gamma$ such that $f(x) \in x$ for each $x \in U$, as required.

We have shown that if β is a limit ordinal, V_β is a model of the Axiom of choice.

Axiom of construction. We now show that for any ordinal α, the Axiom of construction also holds true on V_α.

Proposition 31.21 Let α be an ordinal. Then the set V_α satisfies the property described by the Axiom of construction.

Proof:

Given: That α is an ordinal.
Required to show: That the Axiom of construction holds true on V_α.

Let $T \in V_\alpha$ and ϕ be a formula in the language of set theory. Then $\text{rank}(T) < \alpha$ (by theorem 31.12 b)). Let $B = \{x \in T : \phi(x)\}$. Since $B \subseteq T$, $\text{rank}(B) \leq \text{rank}(T) < \alpha$. By theorem 31.12 c), $B \in V_\alpha$, as required.

We now gather together the above results on those V_α's which are models for the ZFC-axioms. If we step back a bit, we obtain the following picture:

a) V_{ω_0} is a model for all *ZFC* axioms, except for the Axiom of infinity.
b) For the limit ordinal $\beta = \omega_0\omega_0$, or $\beta = \omega_0 + \omega_0$, V_β is a model for all *ZFC* axioms, except for the Axiom of replacement.

It can then be shown that the Axiom of infinity cannot be logically derived from the other nine axioms. That is, it is independent of the other axioms. This is of course what we want. It is just that, quite often, given a set of axioms, we don't always know for sure whether a particular axiom is independent of others in this set.

Similarly, for the limit ordinal, $\beta = \omega_0\omega_0$, we have shown that nine of the ten *ZFC*-axioms hold true in V_β. It can be shown that the Axiom of replacement is independent of the other nine axioms. That is, it is impossible to prove the Axiom of replacement from the other axioms.

31.6 The GCH and the cumulative hierarchy.

Note that neither the Continuum hypothesis (CH), nor the Generalized continuum hypothesis (GCH), play a role in the construction of V. Similarly, the proof showing that $\mathscr{S} = V$ does not invoke CH nor GCH. That is, $\mathscr{S} = V$ neither assumes nor negates CH and GCH.

We pause to reflect a bit on this matter. Recall that the Generalized continuum hypothesis declares that for any infinite set S, there does not exist a set T such that $S \subset T \subset \mathscr{P}(S)$ where the cardinality of T is strictly larger than the cardinality of S and strictly less than the cardinality of $\mathscr{P}(S)$ (recalling that "cardinal numbers" are initial ordinals). Equivalently,

$$\text{GCH} \Leftrightarrow [\aleph_{\alpha+1} = 2^{\aleph_\alpha},\ \forall\, \alpha \in \mathscr{O}]$$

Suppose we *don't* assume GCH. Then it *may* be the case, for example, that $\aleph_0 < \aleph_1 < \aleph_2 < \aleph_3 = 2^{\aleph_0}$, where $\aleph_1$ is the least ordinal such that $\aleph_1 \not\sim_e \aleph_0$, $\aleph_2$ is the least ordinal such that $\aleph_2 \not\sim_e \aleph_1$, and $\aleph_3$ is the least ordinal such that $\aleph_3 \not\sim_e \aleph_2$. Will these "extra" sets, $\aleph_1$, $\aleph_2$, for example, be present in some V_α? We have $\omega_0 < \omega_1 < \omega_2 < \omega_3 = 2^{\omega_0}$,

where ω_0, ω_1, ω_2 and ω_3 are the corresponding initial ordinals. We have shown above that $\text{rank}(\omega_1) = \omega_1$. That is, the least α such that $\omega_1 \subseteq V_\alpha$ is ω_1. In this case, $\omega_1 \subseteq V_{\omega_1}$ while $\omega_1 \notin V_{\omega_0}$. This means $\omega_1 \in \mathscr{P}(V_{\omega_1}) = V_{\omega_1+1}$ and so ω_1 appears in $V_{\omega_1+1} \subset V = \cup_{\alpha\in\mathscr{O}} V_\alpha$. Similarly, $\omega_2 \in V_{\omega_2+1} \subset V$. So whether we assume GCH or not, all sets in $\mathscr{S}$ are accounted for in V. So the equality, $\mathscr{S} = V$, is not sensitive to assumptions made on CH or GCH.

Concepts review:

1. State the Axiom of regularity.
2. What does it mean to say that a set S has a minimal element with respect to $\in$?
3. Which statement which refers to a minimal element of sets is equivalent to the Axiom of regularity?
4. What does it mean to say that a set is well-founded?
5. If we assume the Axiom of choice, which ZFC axiom is equivalent to the statement "Every set is well-founded"?
6. Given a non-empty set x, what is the transitive closure of x?
7. Given a non-empty set x, does x necessarily have a transitive closure?
8. How is the class of sets $V = \cup_{\alpha\in\mathscr{O}} V_\alpha$ constructed?
9. The class $\{V_\alpha : \alpha \in \mathscr{O}\}$ has a strict linear ordering with respect to which order relation?
10. The statement "The class $\{V_\alpha : \alpha \in \mathscr{O}\}$ contains all sets" is equivalent to which ZFC axiom?
11. What is the rank of a set U?
12. If α is an ordinal, what is its rank?

EXERCISES

A. 1. Describe the transitive closure t_x for each of the following sets x.
   a) $x = \{1, 2\}$
   b) $x = \{0, \{2, 3\}\}$

c) $x = \{0, 1, 2\}$

d) $x = \{\{\varnothing\}\}$

2. Show that a set S is transitive if and only if for every $x \in S$, $x \cap S = x$.
3. Show that x is transitive if and only if x is its own transitive closure t_x.
4. Show that the V_α's in the expression $V = \cup_{\alpha \in \mathscr{O}} V_\alpha$ are sets.

B.

5. Show that if $\gamma \in \mathscr{O}$, $\cup_{\alpha \in \gamma} \alpha = \sup \{\alpha : \alpha \in \gamma\}$.
6. Write out all the elements of each of the sets V_0 to V_5.
7. Describe V_ω and V_{ω^+}.
8. What is the least ordinal α such that $7 \subset V_\alpha$?

C.

9. If x is a set, then we define the rank of x, rank(x), as follows:

$$\text{rank}(x) = \text{ least}\{\alpha \in \mathscr{O} : x \subseteq V_\alpha\}$$

a) What is the rank of the set $\mathbb{N}$?

b) If the sets A and B have the same rank what is the rank of $A \cup B$?

c) Can you find a set S such that the rank of $\cup_{A \in S} A$ equals the rank of S?

d) Can you find a set S such that the rank of $\cup_{A \in S} A$ is an element of the rank of S?

10. Show that the Axiom of infinity cannot be proven from the other ZFC-axioms.

32 / Martin's axiom

Summary. *In this section we define the "countable chain condition" on a partially ordered set $(P, \leq)$. We then define those subsets of $(P, \leq)$ called "filters". We introduce an axiom which is independent of ZFC called Martin's axiom, of particular interest when $\neg$CH is assumed. We then list a few consequences of this axiom.*

32.1 Introduction

At this point we have discussed 9 basic set theory axioms we called the *ZF-axioms*. To these 9 axioms we have adjoined a tenth axiom called the *Axiom of choice*. When viewed together these axioms are referred to as the *ZFC*-axioms or *ZF*+Choice. Most mathematicians view these ten axioms as constituting a firm and reliable foundation of mathematics. A few mathematicians or logicians, as well as certain philosophers of mathematics, continue to investigate these axioms identifying and analyzing what they consider to be some of their shortcomings or weak points, occasionally questioning the validity of some of these. And this is fine, since no one can prove that the *ZFC*-axioms will not, at some point in time, lead to some contradiction. Many mathematicians view some other axioms, independent of these, as being useful tools to prove certain mathematical statements they consider to be important. Examples of these are the Continuum hypotheses axioms: CH, GCH, $\neg$CH and $\neg$GCH. The Continuum hypothesis, CH, declares that the smallest cardinal number which is larger than the countable cardinal $\aleph_0$ is $\aleph_1 = 2^{\aleph_0}$. The negation, $\neg$CH, of CH declares that there is at least one uncountable cardinal $\aleph_1$ such that $\aleph_0 < \aleph_1 < 2^{\aleph_0}$. The Generalized Continuum hypothesis, GCH, states that, for any cardinal number $\aleph_\alpha$, the smallest cardinal number which is greater than $\aleph_\alpha$ is $\aleph_{\alpha+1} = 2^{\aleph_\alpha}$. It's negation declares that there are cardinal numbers $\aleph_\alpha$ such that $\aleph_\alpha < \aleph_{\alpha+1} < 2^{\aleph_\alpha}$.

In this section we will discuss another axiom called *Martin's axiom*[1]. This is a slightly more advanced topic of set theory since the understanding of some proofs assume some basic knowledge of topology on the part of the reader. Before we state and describe this axiom we must introduce two notions associated to partially ordered sets, $(P, \leq)$: the *countable chain condition* in P and special subsets of P called *filters*.

32.2 The countable chain condition *ccc* on a partially ordered set.

Recall that a partially ordered set is a set P on which we have defined a relation $\leq$ which is reflexive, transitive and anti-symmetric. As an example consider the closed interval $X = [0, 1]$. Let $\tau(X)$ be a subset of the power set, $\mathscr{P}(X)$, of X defined as follows:

[1] Introduced in 1970 by Donald A. Martin and Robert M. Solovay

$\tau(X) = \{U \in \mathscr{P}(X) : U \neq \varnothing$ and U is a union of open intervals $[0, a), (a, b), (a, 1]\}$. The elements of $\tau(X)$ are referred to as the non-empty *open subsets* of X. If we equip $\tau(X)$ with the relation $\subseteq$ then $(\tau(X), \subseteq)$ is an example of a partially ordered set. In what follows, $\wedge P$ denotes the minimal element of P, (if P has one).

Suppose $(P, \leq)$ is a partially ordered set which may or may not contain $\wedge P$ (the minimal element of P). Let A be a subset of P such that 1) $\wedge P \notin A$ (if $\wedge P$ exists), 2) for any pair u, $v \in A$, u and v are not comparable under $\leq$ and 3) there does not exist an element $r \in P$ such that $r \leq u$ and $r \leq v$. When A satisfies simultaneously these three properties we say that "A is *strong antichain* in the partially ordered set P". We normally drop the adjective "strong" and simply say "antichain" when there is no risk of confustion. With this in mind, we introduce the following concept.

Definition 32.1 Let $(P, \leq)$ be a partially ordered set (which may or may not contain $\wedge P$). If P contains no uncountable strong antichain then $(P, \leq)$ is said to satisfy the *countable chain condition*[1]. In this case, we say that $(P, \leq)$ satisfies the *ccc* or that $(P, \leq)$ is a *ccc* partial order.

For example, let $X = [0, 1]$ and $(\tau(X), \subseteq)$ be the partially ordered set defined above. We verify that $(\tau(X), \subseteq)$ is a *ccc* partially ordered set. Suppose A is a strong antichain in $\tau(X)$. The *ccc* property means that the open sets in A are pairwise disjoint. We index the elements of A as follows: $A = \{U_\kappa : \kappa < \aleph_\alpha\}$ for some cardinal $\aleph_\alpha$. From each $U_\kappa \in A$ we choose precisely one rational number q_κ (we can do this since U_κ is a union of open intervals). Since $|\mathbb{Q}| = \aleph_0$ then $|\{q_\kappa : \kappa < \aleph_\alpha\}| \leq \aleph_0$. It follows that $|A| \leq \aleph_0$ and so A is countable. We conclude that $(\tau(X), \subseteq)$ is a *ccc* partial ordered set.

32.3 Dense subsets of a partially ordered set $(P, \leq)$.

Given a partially ordered set $(P, \leq)$ there are special subsets of P said to be *dense in* P. We explain what this means in the following definition.

Definition 32.2 Let $(P, \leq)$ be a partially ordered set. Let D be a subset of P such that for every element p in P there exists an element d in D such that $d \leq p$. A subset D satisfying this property is said to be *dense in the partial ordering* $(P, \leq)$.

---

[1] Although "countable antichain condition" would have probably been a better choice of words to describe this property.

Examples of dense subsets of a partially ordered set.

Example 1. Let $X = [0,1]$ and $(\tau(X), \subseteq)$ be the partially ordered set defined above. We will construct a dense subset $\mathscr{D}$ of $(\tau(X), \subseteq)$ as follows. Let

$$\mathscr{E} = \{(x-\varepsilon, x+\varepsilon) \subseteq [0,1] : x \in (0,1), \varepsilon \in (0,1)\}$$

We claim that $\mathscr{E}$ is dense in $(\tau(X), \subseteq)$. Let U be an element of $\tau(X)$. Then U is a non-empty open set (read "a union of open intervals"). Since U is non-empty there is an $x \in U$ such that x is not 0 or 1. Then there exists ε such that $(x-\varepsilon, x+\varepsilon) \subset (a,b) \subseteq U$. (This is a fundamental property of the real numbers.). Then, for every element U of $\tau(X)$, there is an element of $\mathscr{E}$ which is a subset of U. So $\mathscr{E}$ is dense in $(\tau(X), \subseteq)$ as claimed.

Example 2. Let $(P, \leq)$ be a partially ordered set. If $x \in P$ we define $x^{\downarrow}$ as follows:

$$x^{\downarrow} = \{y \in P : y \leq x\}$$

Let u and v be fixed elements of P. We say that u and v are *compatible* in P if there exists $t \in P$ (where t is not the minimum element of P) such that $t \in u^{\downarrow} \cap u^{\downarrow}$. Let $V_{(u,v)} = \{x \in P : x \text{ is not compatible with } u \text{ or } v\}$, $W_{(u,v)} = u^{\downarrow} \cap v^{\downarrow}$ and $D_{(u,v)} = V_{(u,v)} \cup W_{(u,v)}$.

We claim that $D_{(u,v)}$ is dense in P: Let $z \in P$ and $y \leq z$. If $y^{\downarrow} \cap u^{\downarrow} = \varnothing$ or $y^{\downarrow} \cap v^{\downarrow} = \varnothing$ then $y \in V_{(u,v)}$ so $y \in D_{(u,v)}$ and we are done. Suppose, on the other hand, that $q \in y^{\downarrow} \cap u^{\downarrow}$. Since $y \leq z$, $q \in z^{\downarrow} \cap u^{\downarrow}$ where $q \leq y \leq z$. If $q^{\downarrow} \cap v^{\downarrow} = \varnothing$ then $q \in V_{(u,v)}$ and we are done. If $k \in q^{\downarrow} \cap v^{\downarrow}$ then $k \in u^{\downarrow} \cap v^{\downarrow}$. We then have $k \leq z$ and $k \in W_{(u,v)} \subseteq D_{(u,v)}$. So $D_{(u,v)}$ is dense in P, as claimed.

32.4 A filter in a partially ordered set $(P, \leq)$.

A filter is a non-empty subset of a partially ordered set $(P, \leq)$ which satisfies two properties. Those subsets of partially ordered sets which we call *filters* are defined as follows.

Definition 32.3 Let F be a subset of a partially ordered set $(P, \leq)$. If F is non-empty and satisfies the two properties, 1) If x and y belong to F there exists z in F which is less than or equal to both x and y (i.e., F is a *filter base* or *downward directed*), 2) if x belongs to F and x is less than or equal to an element y of P, then y belongs to F (i.e., F is upward closed). A filter in $(P, \leq)$ is a *proper filter* if it is not all of P. If $x \in P$ then the set of all elements above x is called a *principal filter with principal element x*. Such a filter is the smallest filter which contains x.

The notion of a filter is often seen in the context of $(\mathscr{P}(X), \subseteq)$, the power set for some non-empty set X ordered by inclusion $\subseteq$. In this case, a *filter* $\mathscr{F}$ in $(\mathscr{P}(X), \subseteq)$ is seen as being a non-empty subset $\mathscr{F}$ of $\mathscr{P}(X)$ which satisfies:

1. $\mathscr{F}$ is *closed under supersets* (i.e., if $F \in \mathscr{F}$ and $F \subset T$ then $T \in \mathscr{F}$)
2. $\mathscr{F}$ is *closed under finite intersections.* That is, if $A, B \in \mathscr{F}$ then $A \cap B$ belongs to $\mathscr{F}$. This generalizes to "the intersection of finitely many elements of $\mathscr{F}$ belongs to $\mathscr{F}$ and so is never empty".

When the second condition is satisfied we say that $\mathscr{F}$ satisfies the *finite intersection property.*

For example, if X is an infinite set, the subset $\mathscr{F} = \{U \in \mathscr{P}(X) : X - U \text{ is finite } \} \subset \mathscr{P}(X)$ does not contain the empty set, is closed under supersets and satisfies the *finite intersection property* and so is a proper filter in $(\mathscr{P}(X), \subseteq)$. When a filter is a subset of $(\mathscr{P}(X), \subseteq)$ we will be using the script font, $\mathscr{F}$, to represent it. Note that $\mathscr{F}$ is proper if and only if $\varnothing \notin \mathscr{F}$.

32.5 Martin's axiom.

Martin's axiom is a statement concerning those sets of cardinality $\kappa < 2^{\aleph_0}$ which, when hypothesized, allows sets of cardinality κ to behave more like those sets of cardinality $\aleph_0$ then those sets which are of cardinality $2^{\aleph_0}$. We begin by defining the following statement called Martin's κ-statement, denoted by MA(κ).

> *MA(κ)*: Let κ be an infinite cardinal and $(P, \leq)$ be a non-empty partially ordered set satisfying the *countable chain condition.* Let $\mathscr{D} = \{D \in \mathscr{P}(P) : D \text{ is dense in } P\}$ such that $|\mathscr{D}| \leq \kappa$. Then there is a proper filter $F \subseteq P$ such that, $F \cap D \neq \varnothing$ for every set $D \in \mathscr{D}$.

When referring to a partially ordered set such as $(\mathscr{P}(X), \subseteq)$ the Martin's κ-statement becomes:

> *MA(κ) statement statement for a power set*: Let κ be an infinite cardinal. Suppose $\mathscr{P}(X)$ contains no uncountable family of pairwise disjoint subsets of X (i.e., $\mathscr{P}(X)$ satisfies *ccc*). Suppose $\mathscr{D}^* = \{\mathscr{D} \subseteq \mathscr{P}(X) : \mathscr{D} \text{ is dense in } \mathscr{P}(X)\}$. If $|\mathscr{D}^*| \leq \kappa$ then there is a proper filter $\mathscr{F} \subseteq \mathscr{P}(X)$ such that, for every set $\mathscr{D} \in \mathscr{D}^*$, $\mathscr{F} \cap \mathscr{D} \neq \varnothing$.

We will first show that MA($\aleph_0$) holds true in *ZFC*.

Theorem 32.4 The statement MA($\aleph_0$) holds true in *ZFC*.

Proof: Suppose $(P, \leq)$ is a non-empty partially ordered. (Note that the *ccc* property is not required for the Martin's $\aleph_0$-statement to hold true.) Let $\mathscr{D}$ be a family of dense subsets of P such that $|\mathscr{D}| \leq \aleph_0$. Let $a \in P$.

Case 1: We consider the case where $\mathscr{D}$ is empty. We let $F = \{x \in P : x \geq a\}$ be the principal filter with principal element a. Since F intersects every element of $\mathscr{D}$ then MA($\aleph_0$) holds true.

Case 2: We consider the case where $0 < |\mathscr{D}| \leq \aleph_0$. We can then enumerate the sets in $\mathscr{D}$ as $\mathscr{D} = \{D_1, D_2, D_3, \ldots, \}$. Since, for each i, D_i is dense in P we can choose some $d_1 \in D_1$ such that then $d_1 \leq a$, choose $d_2 \in D_2$ such that $d_2 \leq d_1 \leq a$, and if $d_n \leq d_{n-1} \leq \cdots \leq d_1 \leq a$ choose $d_{n+1} \in D_{n+1}$ such that $d_{n+1} \leq \cdots d_1 \leq a$ (Choice). We now let the set F be one that contains all $\{d_i\}_{i=1,2,3,\ldots,}$ and all elements of P which are above d_1. That is, $q \in F$ if and only if $q = d_i$ or $q \geq a \geq d_i$. We claim that F is a filter in P. 1) Clearly F is non-empty. 2) If b is in P such that $a \leq b$ then $b \in F$, 3) if b, c belong to F, then either $b < c$ or $c < b$. Without loss of generality, suppose $c < b$. Then $c \leq c$ and $c \leq b$. Then F is a filter, as claimed. Furthermore, F intersects every D_i at d_i. So F is the filter which satisfies the condition for MA($\aleph_0$).

We can also show that, in *ZFC*, MA($2^{\aleph_0}$) is false.

Before we prove this we state the following well-known facts about the closed and bounded interval $X = [0, 1]$. If $U \in \tau(X)$ (where $\tau(X)$ is defined above as being the set of all non-empty open subsets of X) then the *closure* of U, denoted by cl(U), is a subset of X where $X -$ cl(U) is the the largest open set which contains elements outside of U. A subset is said to be *closed* if and only if its complement is an open subset. The simplest example is: $\text{cl}(a, b) = [a, b]$, a closed interval X.

The following are well-known facts about closed and bounded subsets of $\mathbb{R}$.

Fact #1: If K is a closed and bounded subset of $\mathbb{R}$ and $\mathscr{F} = \{F : F \text{ is closed in } K\}$ is known to be a filter in $\mathscr{P}(K)$ (i.e, satisfies the finite intersection property) then $\cap\{F : F \in \mathscr{F}\} \neq \varnothing$.

Fact #2: If (a, b) is an open interval in $X = [0, 1]$ and $x \in (a, b) \subset [0, 1]$ then there exists $c, d \in X$ such that $x \in (c, d) \subset \text{cl}(c, d) = [c, d] \subset (a, b)$.

Theorem 32.5 The statement MA($2^{\aleph_0}$) fails in *ZFC*.

Proof: Let X = [0, 1]. For each $x \in X$ let $U_x = X - \{x\}$ (the complement of $\{x\}$ in X). Then $\mathscr{U} = \{U_x : x \in X\} \subset \tau(X)$ where $|\mathscr{U}| = |X| = 2^{\aleph_0}$. For each $x \in X$, let $\mathscr{D}_x = \{D \in \tau(X) : \text{ cl}(D) \subseteq U_x\}$. We claim that, for each $x \in X$, $\mathscr{D}_x$ is dense

in $(\tau(X), \subseteq)$. Suppose $M \in \tau(X)$. Then there exists some $D \in \mathscr{D}_x$ such that $D \subset \text{cl}(D) \subseteq M \cap U_x \in \tau(X)$. Then $\mathscr{D}_x$ is dense in the partially ordered $(\tau(X), \subseteq)$ as claimed. Then $\mathscr{D} = \{\mathscr{D}_x : x \in X\}$ is a set of dense subsets of $(\tau(X), \subseteq)$ such that $|\mathscr{D}| = 2^{\aleph_0}$.

Suppose MA($2^{\aleph_0}$) holds true for the partially ordered set $(\tau(X), \subseteq)$. Then there exists a *proper* filter $\mathscr{F} = \{F : F \in \tau(X)\}$ such that, for each $x \in X$, $\mathscr{F} \cap \mathscr{D}_x$ is non-empty. For each $x \in X$, choose $D_x \in \mathscr{F} \cap \mathscr{D}_x$. Since $\mathscr{F}$ is a filter, finite intersections of the chosen elements $\{D_x : x \in X\}$ are non-empty. Then $\{\text{cl}(D_x) : x \in X\}$ must also satisfy the finite intersection property. Since $D_x \in \mathscr{D}_x$ then $\text{cl}(D_x) \subseteq \mathscr{D}_x$. Since X is closed and bounded (compact) then $\cap\{\text{cl}(D_x) : x \in X\} \neq \varnothing$. This contradicts the fact that

$$\{\text{cl}(D_x) : x \in X\} \subseteq \cap\{\mathscr{D}_x : x \in X\} \subseteq \cap\{U_x : x \in X\} = \varnothing$$

The source of our contradiction is our supposition that MA($2^{\aleph_0}$) holds true. We conclude that MA($2^{\aleph_0}$) does not hold true in *ZFC*.

The two theorems above show that the only Martin κ-statements which are of interest are those where κ is such that $\aleph_0 \leq \kappa < 2^{\aleph_0}$.

We then state Martin's axiom as follows.

Definition 32.6 *Martin's axiom*, MA, is defined as being MA(κ) where κ satisfies $\aleph_0 \leq \kappa < 2^{\aleph_0}$.

Trivially CH $\Rightarrow$ MA. On the other hand, it has been shown (by Solovay and Martin) that Martin's axiom is independent of *ZFC* and is consistent with *ZFC* + $\neg$CH .

32.6 Consequences of Martin's axiom in topology.

There are a few equivalent forms of Martin's axiom. For those readers familiar with point-set topology we present a nicely formulated consequence of Martin's axiom which is of interest. The following statement is in fact equivalent to MA. We prove that it is equivalent to MA in the Appendix A.

Theorem 32.7 [MA] Suppose κ is an infinite cardinal such that $\kappa < 2^{\aleph_0}$. If X is a Hausdorff compact space with *ccc* and $\{U_\alpha : \alpha \leq \kappa\}$ is a family of open dense[1] subsets of X then $\cap\{U_\alpha : \alpha\} \neq \varnothing$.

[1]The subset U is *dense* in X in the topological sense if and only if $\text{cl}(U) = X$

Certain readers may find that this topological statement is reminiscent of the *Baire category theorem* where one version is stated as: "If X is a locally compact Hausdorff space and $\mathscr{D} = \{D_\alpha : \alpha < \aleph_0\}$ is a countable set of open and dense subsets in X, then $\cap\{D_\alpha : \alpha < \aleph_0\}$ is dense in X."[4]

Other well-known topological consequences of Martin's axiom are:

- MA($\aleph_1$) implies that "If X is topological space such every closed subset of X is a G_δ-set then every subspace of X has a countable dense subset."
- MA($\aleph_1$) implies that "A product of *ccc* topological spaces is *ccc*".
- When MA is assumed: "If κ is such that $\aleph_0 \leq \kappa < 2^{\aleph_0}$ then $2^\kappa = 2^{\aleph_0}$."
- *A statement which is equivalent to MA*: Let κ be such that $\aleph_0 \leq \kappa < 2^{\aleph_0}$ and let X be a *ccc* Hausdorff topological space such that the subset $\{x \in X : x$ has a compact neighbourhood$\}$ is dense in X. If $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$ is a family of open dense subsets of X then $\cap\{D_\alpha : \alpha \leq \kappa\}$ is dense in X.

Concepts review:

1. When does a partially ordered set satisfy the "countable chain condition"?
2. What is an open subset of the closed interval $X = [0, 1]$?
3. What subset of $\mathscr{P}([0, 1])$ does $\tau([0, 1])$ represent?
4. What is a strong antichain in a partially ordered set $(P, \leq)$?
5. What is a dense subset of a partially ordered set $(P, \leq)$?
6. What is filter in a partially ordered set $(P, \leq)$? What is proper filter? What is a principal filter?
7. What does it mean to say that a family of subsets satisfies the "finite intersection property"?
8. State the Martin's κ-statement MA(κ).
9. State MA(κ) when it refers specifically to a power set $(\mathscr{P}, \subseteq)$.
10. What can be said about MA($\aleph_0$)?

[4] In fact, it's similarity to MA is such that some may want to refer to MA as an "Enhanced" *Baire category theorem*.

11. What can be said about MA($2^{\aleph_0}$)?
12. State Martin's axiom, MA.
13. State the Baire Category theorem.

Part X

Ordinal arithmetic

33 / Ordinal arithmetic: Addition.

Summary. *In this section we define the operation of addition on the ordinal numbers. We then show its most basic properties and provide a few examples.*

33.1 A definition of ordinal addition.

Just as for cardinal numbers we can define addition of ordinal numbers. Ordinal number addition will be very similar to cardinal number addition. Recall how cardinal number addition was defined:

> Let S and T be two sets such that $S \cap T = \varnothing$ where S is of cardinality κ and T is of cardinality λ. We define: $\kappa + \lambda$ as the cardinality of $S \cup T$.

Now the ordinality of a well-ordered set S is the unique ordinal which is order isomorphic to S. If we simply transpose the "cardinal addition definition" onto the ordinal numbers, then we would obtain: "If $(S, \leq_S)$ and $(T, \leq_T)$ are well-ordered sets of ordinality α and β, respectively, then $\alpha + \beta = {}^{\text{ord}}(S \cup T)$." There is, however, some critical information missing here. We can only speak of the ordinality of a set in reference to a stated well-ordering of that set. We should then begin by defining a well-ordering of $S \cup T$.

Definition 33.1 Let $(S, \leq_S)$ and $(T, \leq_T)$ be two *disjoint* well-ordered sets. We define the relation "$\leq_{S \cup T}$" on $S \cup T$ as follows:

a) $u \leq_{S \cup T} v$ if $\{u, v\} \subseteq S$ and $u \leq_S v$.

b) $u \leq_{S \cup T} v$ if $\{u, v\} \subseteq T$ and $u \leq_T v$.

c) $u \leq_{S \cup T} v$ if $u \in S$, $v \in T$.

Our next step should be a verification that this newly defined order relation actually well-orders the union $S \cup T$. We express this in the form of a theorem, and leave the straightforward proof as an exercise.

Theorem 33.2 Let $(S, \leq_S)$ and $(T, \leq_T)$ be two *disjoint* well-ordered sets. Then the relation $\leq_{S \cup T}$ well-orders the set $S \cup T$.

Proof: The proof is left as an exercise.

We can now define addition of two ordinal numbers as the ordinality of the union of disjoint well-ordered sets.

Definition 33.3 Let α and β be two ordinal numbers. Let $(S, \leq_S)$ and $(T, \leq_T)$ be two *disjoint* well-ordered sets of order type α and β respectively.[1] We define $\alpha + \beta$ as follows:

$$\alpha + \beta = {}^{\text{ord}}(S \cup T,\ \leq_{S\cup T})$$

The stated definition of addition of ordinal numbers α and β applies only to order types of disjoint well-ordered sets $(S, \leq_S)$ and $(T, \leq_T)$. If we wish to add the order types α and β of two sets $(S, \leq_S)$ and $(T, \leq_T)$ with non empty intersection, we can construct the sets $(S \times \{0\},\ \leq_0)$ and $(T \times \{1\}, \leq_1)$ of the same order types each equipped with the lexicographical ordering. That is,

$$\begin{aligned} (u, 0) \leq_0 (v, 0) \quad &\text{if} \quad u \leq_S v \\ (u, 1) \leq_1 (v, 1) \quad &\text{if} \quad u \leq_T v \end{aligned}$$

We then add the ordinals α and β as defined.

Theorem 33.4 Let $(S, \leq_S)$, $(T, \leq_T)$ and $(U, \leq_U)$, $(V, \leq_V)$ be two pairs of *disjoint* well-ordered sets such that

$$\begin{aligned} {}^{\text{ord}}S &= \alpha = {}^{\text{ord}}U \\ {}^{\text{ord}}T &= \beta = {}^{\text{ord}}V \end{aligned}$$

Then ${}^{\text{ord}}(S \cup T, \leq_{S\cup T}) = \alpha + \beta = {}^{\text{ord}}(U \cup V, \leq_{U\cup V})$. Hence, addition of ordinal numbers is well-defined.

Proof: The proof is left as an exercise.

33.2 Examples.

Addition of natural numbers, when viewed as ordinals, should agree with results obtained when adding natural numbers the usual way. We have already verified that this is the case for cardinal numbers. The following example illustrates that this is the case for addition of natural numbers if viewed as ordinals. Note that in the examples and theorem below, "$<$" represents the "ordinal inclusion" order relation $\in$.

[1] Addition can also be defined inductively as follows: For all α and β, a) $\beta + 0 = \beta$, b) $\beta + (\alpha + 1) = (\beta + \alpha) + 1$, c) $\beta + \alpha = \text{lub}\{\beta + \gamma : \gamma < \alpha\}$ whenever α is a limit ordinal.

a) *Example.* Determine the sum $3 + 7$ when these natural numbers are viewed as ordinals. Also determine the sum $7 + 3$.

Solution: We see that $3 = {}^{\text{ord}}\{7, 8, 9\}$ and $7 = {}^{\text{ord}}\{0, 1, 2, 3, 4, 5, 6\}$. The choice of the natural numbers used is arbitrary. The chosen well-ordered set representatives are disjoint. See that ${}^{\text{ord}}\{7, 8, 9, 0, 1, 2, \ldots, 6\} = {}^{\text{ord}}\{0, 1, 2, \ldots, 9\}$ since $\{7, 8, 9, 0, 1, 2, \ldots, 6\}$ (with the ordering defined on unions) and $\{0, 1, 2, \ldots, 9\}$ (with the usual natural number ordering) are order isomorphic. By definition,

$$3 + 7 = {}^{\text{ord}}\{7, 8, 9, 0, 1, 2, \ldots, 6\} = {}^{\text{ord}}\{0, 1, 2, \ldots, 9\} = 10$$

$$7 + 3 = {}^{\text{ord}}\{0, 1, 2, \ldots, 9\} = 10$$

b) *Example.* Determine both sums $\omega_0 + 7$ and $7 + \omega_0$.

Solution: So that we obtain disjoint well-ordered set representatives we will use

$$\begin{aligned} 7 &= {}^{\text{ord}}\{0, 1, 2, 3, 4, 5, 6\} \\ \omega_0 &= {}^{\text{ord}}\{7, 8, 9, 10, \ldots\} \end{aligned}$$

Then, by definition,

$$\begin{aligned} 7 + \omega_0 &= {}^{\text{ord}}\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, \ldots, \} = \omega_0 \\ \omega_0 + 7 &= {}^{\text{ord}}\{7, 8, 9, 10, \ldots, 0, 1, 2, 3, 4, 5, 6\} = \omega_0 + 7 \end{aligned}$$

It is worth noting that $7 + \omega_0 = \omega_0 < \omega_0 + 7$ and so, even if addition on finite ordinals is commutative, this is not always the case for addition of transfinite ordinals.

c) *Example.* Show that $\omega_0 + 1$, when viewed as the sum of the ordinal ω_0 and the ordinal $1 = \{0\}$, is the immediate successor, ω_0^+, of ω_0.

Solution:

$$\begin{aligned} \omega_0 + 1 &= {}^{\text{ord}}\left(\omega_0 \times \{0\} \cup \{(0, 1)\}\right)_{\leq_{\omega_0 \times \{0\} \cup \{(0,1)\}}} \\ &= {}^{\text{ord}}\left(\{0, 1, 2, 3, \ldots\} \cup \{\omega_0\}\right) \\ &= \{0, 1, 2, 3, \ldots\} \cup \{\omega_0\} \\ &= \omega_0 \cup \{\omega_0\} \\ &= \omega_0^+ \end{aligned}$$

33.3 Basic properties of ordinal addition.

Many of the addition properties generalize from finite ordinal addition to transfinite ordinal addition. But we should watch out for those properties that do not.

Theorem 33.5 Let α, β and γ be three (non-zero) ordinal numbers. Then:

a) $(\alpha+\beta)+\gamma = \alpha+(\beta+\gamma)$ (Addition is associative.)

b) For any ordinal $\gamma > 0$, $\alpha < \alpha+\gamma$

c) For any ordinal γ, $\gamma \leq \alpha+\gamma$

d) $\alpha < \beta \Rightarrow \alpha+\gamma \leq \beta+\gamma$

e) $\alpha < \beta \Rightarrow \gamma+\alpha < \gamma+\beta$

f) $\alpha+\beta = \alpha+\gamma \Rightarrow \beta = \gamma$ (Left term cancellation is valid.)

g) $\alpha+0 = \alpha$

Proof: Let A, B and C be three pairwise disjoint well-ordered set representatives of the ordinals α, β and γ respectively.

a) We are required to show that $(\alpha+\beta)+\gamma = \alpha+(\beta+\gamma)$. By definition

$$\begin{aligned}(\alpha+\beta)+\gamma &= {}^{\text{ord}}(A\cup B)+{}^{\text{ord}}C\\ &= {}^{\text{ord}}[(A\cup B)\cup C] \text{ (By definition of ordinal addition.)}\\ &= {}^{\text{ord}}[A\cup(B\cup C)] \text{ (This is justified below at *.)}\\ &= {}^{\text{ord}}A+{}^{\text{ord}}(B\cup C)\\ &= \alpha+(\beta+\gamma)\end{aligned}$$

* *Justification of* ${}^{ord}[(A\cup B)\cup C] = {}^{ord}[A\cup(B\cup C)]$: The two well-ordered sets $(A\cup B)\cup C$ and $A\cup(B\cup C)$ contain the same elements and so are equal as sets. To conclude that they have the same ordinality we must show that their elements are ordered in the same way. We will proceed by cases:

1) If both u and v belong to one of A, B, or C, then $u < v$ in $(A\cup B)\cup C$ if and only if $u < v$ in $A\cup(B\cup C)$.
2) If $u \in A$ and v belongs either to B or C, then $u < v$ in $(A\cup B)\cup C$ if and only if $u < v$ in $A\cup(B\cup C)$.
3) If $u \in B$ and $v \in C$, then $u < v$ in $B\cup C$ and so $u < v$ in $(A\cup B)\cup C$ if and only if $u < v$ in $A\cup(B\cup C)$.

In all cases the order of u and v is respected in both sets $(A\cup B)\cup C$ and $A\cup(B\cup C)$. So not only are they equal sets, they are order isomorphic and so have the same ordinality.

b) We are given that $\alpha+\gamma = {}^{\text{ord}}(A\cup C)$ and $\gamma > 0$. We are required to show that $\alpha < \alpha+\gamma$. Since C is well-ordered and non-empty it contains a least element, say x. Then for every $y \in A$, $y < x$. Then $A = S_x = \{u \in A\cup B : u < x\}$, an initial segment of $A\cup B$ equipped with the ordering $\leq_{A\cup B}$. Then α is an initial segment of $\alpha+\gamma$ and so $\alpha \in \alpha+\gamma$. Equivalently, $\alpha < \alpha+\gamma$.

c) We are given that $\alpha + \gamma = {}^{\text{ord}}(A \cup C)$. We are required to show that $\gamma \leq \alpha + \gamma$. Since both C and $A \cup C$ are well-ordered they are either order isomorphic or one is order isomorphic to an initial segment of the other (by theorem 26.7). If C and $A \cup C$ are order isomorphic, then $\gamma = \alpha + \gamma$. Suppose they are not order isomorphic. Then $\alpha \neq 0$ and so C is a proper subset of $A \cup C$. If $A \cup C$ is order isomorphic to a subset of C, then $A \cup C$ is order isomorphic to a proper subset of itself. Since no well-ordered set can be order isomorphic to a proper subset of itself, then $A \cup C \not<_{\text{WO}} C$; hence, $\alpha + \gamma \not< \gamma$, so $\gamma < \alpha + \gamma$.

d) We are given that $\alpha < \beta$. We are required to show that $\alpha + \gamma \leq \beta + \gamma$.
Since $\alpha < \beta$ there exists an order isomorphism $f : A \to B$ which maps A to an initial segment B^* of B. Consider the function $g : A \cup C \to B \cup C$ such that $g|_A = f$ and $g|_C$ is the identity map on C. Note that $a < c$ for all $a \in A$ and $c \in C$ and $b < c$ for all $b \in B$ and $c \in C$ and so $g : A \cup C \to B^* \cup C$ is an order isomorphism mapping $A \cup C$ onto $B^* \cup C$. Since ${}^{\text{ord}}(B^* \cup C) \leq {}^{\text{ord}}(B \cup C) = \beta + \gamma$, $\alpha + \gamma \leq \beta + \gamma$.[1]

e) We are given that $\alpha < \beta$. We are required to prove that $\gamma + \alpha < \gamma + \beta$.
Since $\alpha < \beta$ there exists an order isomorphism $f : A \to B$ which maps A to an initial segment B^* of B. Consider the function $g : C \cup A \to C \cup B^*$ such that $g|_C$ is the identity map and and $g|_A = f$. Note that $c < a$ for all $c \in C$ and $a \in A$ and $c < b$ for all $c \in C$ and $b \in B^*$ and so $g[C \cup A] = C \cup B^*$ is an initial segment of $C \cup B$. Since ${}^{\text{ord}}(C \cup A) = {}^{\text{ord}}(C \cup B^*) < {}^{\text{ord}}(C \cup B) = \gamma + \beta$, then $\gamma + \alpha < \gamma + \beta$.

f) We are given that $\alpha + \beta = \alpha + \gamma$. We are required to show that $\beta = \gamma$.
Suppose $\beta < \gamma$. Then by part e), $\alpha + \beta < \alpha + \gamma$, a contradiction. Similarly, $\gamma < \beta$ implies $\alpha + \gamma < \alpha + \beta$, again a contradiction. We conclude that $\beta = \gamma$.

g) We are given that α is an ordinal number. We are required to show that $\alpha + 0 = \alpha$. Simply see that $\alpha + 0 = {}^{\text{ord}}(A \cup \{\ \}) = {}^{\text{ord}}A = \alpha$.

Remark: In part f) of the above theorem we show that "left cancellation" on addition applies just like for natural numbers. However "right cancellation" does not work. For example,

$$\begin{aligned} 2 + \omega_0 &= {}^{\text{ord}}\{0, 1\} + {}^{\text{ord}}\{6, 7, 8, 9, \ldots, \} \\ &= {}^{\text{ord}}(\{0, 1\} \cup \{6, 7, 8, 9, \ldots, \}) \\ &= {}^{\text{ord}}\{0, 1, 6, 7, 8, \ldots, \} \\ &= \omega_0 \end{aligned}$$

Similarly, $3 + \omega_0 = \omega_0$. But $2 + \omega_0 = 3 + \omega_0 \not\Rightarrow 2 = 3$.

[1] Note that $g[A \cup C] = B^* \cup C$ need not be an initial segment of $B \cup C$. So even if $B^* \cup C \subset B \cup C$ equality ${}^{\text{ord}}(B^* \cup C) = {}^{\text{ord}}(B \cup C)$ is possible. For example, for $U = \{0, 1, 3, 4, 5, \ldots\}$ and $V = \{0, 1, 2, 3, 4, 5, \ldots\}$ we have both $U \subset V$ and ${}^{\text{ord}}U = \omega_0 = {}^{\text{ord}}V$.

33.4 Addition of limit ordinals.

When adding limit ordinals, determining the simplest form for the sum requires some thought. For example,

$$\begin{aligned}(\omega_1 + 7) + \omega_1 &= \omega_1 + (7 + \omega_1) \\ &= \omega_1 + \omega_1 = \omega_1 2\end{aligned}$$

We provide another approach to addition, for cases where the second term is a limit ordinal. Recall that given a non-empty subset A of a well-ordered set W, an upper bound of A is any element u of W such that $a \leq u$ for all $a \in A$. Suppose the element, s, is the least upper bound of A. That is, s is an upper bound of A, and, for any upper bound u, $s \leq u$. In this case we write $s = \text{lub}(A)$ (or $\sup A$).

For example, the ordinal $5 = \{0, 1, 2, 3, 4\}$ has least upper bound, $\text{lub}(5) = 4$, since 4 is greater than or equal to all of the elements of 5 and it is the least of all upper bounds. The limit ordinal $\omega_0 = \{0, 1, 2, 3, \ldots, \}$ has as a least upper bound, $\text{lub}(\omega_0) = \omega_0$, itself. Note that in this case, $\text{lub}(\omega_0)$ is not an element of ω_0. In fact, α is a limit ordinal if and only if $\text{lub}\,\alpha = \alpha \notin \alpha$.

Another property characterizes limit ordinals. The ordinal γ is a limit ordinal if and only if $\cup_{\alpha\in\gamma}\alpha = \gamma$. In the case where γ has an immediate predecessor, say, β, then $\gamma = \{0, 1, 2, 3, \ldots, \beta\}$ and so

$$\cup_{\alpha\in\gamma}\alpha = 0 \cup 1 \cup 2 \cup \cdots \cup \beta = \beta$$

So even if it is always true that $\text{lub}\,\gamma = \cup_{\alpha\in\gamma}\alpha$, we have $\text{lub}\,\gamma = \cup_{\alpha\in\gamma}\alpha = \gamma$ only in the cases where γ is a limit ordinal.

For example, the least upper bound of ω_1 is ω_1 while the least upper bound of $\omega_1 + 3 = \{0, 1, 2, \ldots, \omega_1, \omega_1 + 1, \omega_1 + 2\}$ is $\omega_1 + 2$. We now show a useful property involving addition of limit ordinals.

Theorem 33.6 Let β be a limit ordinal. Then, for any ordinal, α,

$$\alpha + \beta = \text{lub}\,\{\alpha + \gamma : \gamma < \beta\}$$

Proof:

We are given that β is a limit ordinal and α is any ordinal.
We are required to show that $\alpha + \beta$ is the least upper bound of the set $\{\alpha + \gamma : \gamma < \beta\}$.
We claim that $\alpha + \beta$ is an upper bound of the set $\{\alpha + \gamma : \gamma < \beta\}$:

- For $\delta < \beta$, by theorem 33.5 part e), $\alpha + \delta < \alpha + \beta$. So $\alpha + \beta$ is an upper bound of the set $\{\alpha + \gamma : \gamma < \beta\}$ as claimed.

We claim that $\alpha + \beta$ is the least such upper bound:

– Suppose δ is any upper bound of the set $\{\alpha + \gamma : \gamma < \beta\}$. Then $\alpha + \gamma \leq \delta$ for all $\gamma < \beta$. Suppose $\delta < \alpha + \beta$. Then for all γ, $\alpha + \gamma \leq \delta < \alpha + \beta$. Then there exist a least ordinal $\mu \in \beta$ such that $\delta \leq \alpha + \mu < \alpha + \beta$. Since β is a limit ordinal $\mu^+ < \beta$, then $\delta < \alpha + \mu^+ < \beta$. This contradicts the fact that $\alpha + \gamma \leq \delta$ for all $\gamma < \beta$. Then $\delta \geq \alpha + \beta$. So $\alpha + \beta$ is the least such upper bound of $\{\alpha + \gamma : \gamma < \beta\}$ as required.

Example: Compute the sum $(\omega_0 + 7) + 100 + (2 + \omega_0)$ to its simplest form.

Solution:

$$\begin{aligned}
(\omega_0 + 7) + 100 + (2 + \omega_0) &= (\omega_0 + 7) + (100 + 2) + \omega_0 \\
&= (\omega_0 + 7) + (102 + \omega_0) \\
&= (\omega_0 + 7) + \text{lub}\,\{102 + n : n \in \omega_0\} \\
&= (\omega_0 + 7) + \omega_0 \\
&= \omega_0 + (7 + \omega_0) \\
&= \omega_0 + \text{lub}\,\{7 + n : n \in \omega_0\} \\
&= \omega_0 + \omega_0
\end{aligned}$$

Concepts review:

1. Given two disjoint well-ordered sets $(S, \leq_S)$ and $(T, \leq_T)$ define a well-ordering on $S \cup T$.

2. For any two ordinals α and β how is $\alpha + \beta$ defined?

3. For which ordinals does the given property hold true.

   a) $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$

   b) $\alpha < \alpha + \gamma$

   c) $\gamma \leq \alpha + \gamma$

   d) $\alpha + \gamma \leq \beta + \gamma$

   e) $\gamma + \alpha < \gamma + \beta$

   f) $\alpha + \beta = \alpha + \gamma \Rightarrow \beta = \gamma$

   g) $\alpha + 0 = \alpha$

4. If β is a limit ordinal simplify the expression $\sup\{\alpha + \gamma : \gamma < \beta\}$.

EXERCISES

A.
1. Suppose $4 + \beta = \beta$. Prove that $\omega_0 \leq \beta$.
2. Show that if $\alpha < \gamma$, then $\alpha + 1 \leq \gamma$.

B.
3. Suppose that α and δ are ordinals such that $\alpha \leq \delta$. Show that there can only be one ordinal β such that $\alpha + \beta = \delta$.
4. Compute or simplify the sum $(50 + \omega_0) + (\omega_0 + \omega_1)$.
5. Show that if α is a finite ordinal and γ is a limit ordinal, then the least upper bound of $\alpha + \gamma$ is γ.
6. Show that for any ordinal α and limit ordinal γ, $\alpha + \gamma$ is a limit ordinal.
7. Provide a concrete example of ordinals such that $\alpha < \beta$ and $\alpha + \gamma = \beta + \gamma$ simultaneously hold true.

C.
8. Let $(S, \leq_S)$ and $(T, \leq_T)$ be two *disjoint* well-ordered sets. Show that the relation $\leq_{S \cup T}$ well-orders the set $S \cup T$.
9. Let $(S, \leq_S)$, $(T, \leq_T)$ and $(U, \leq_U)$, $(V, \leq_V)$ be two pairs of *disjoint* well-ordered sets such that

$$\begin{aligned} {}^{\text{ord}}S &= \alpha = {}^{\text{ord}}U \\ {}^{\text{ord}}T &= \beta = {}^{\text{ord}}V \end{aligned}$$

Show that ${}^{\text{ord}}(S \cup T, \leq_{S \cup T}) = \alpha + \beta = {}^{\text{ord}}(U \cup V, \leq_{U \cup V})$. Hence, addition of ordinal numbers is well-defined.

34 / Ordinal arithmetic: Multiplication and Exponentiation.

Summary. *In this section we define the "lexicographic ordering" of the Cartesian product of two well-ordered sets. We then define the multiplication of two ordinals α and β as $\alpha \times \beta = {}^{ord}B \times A$ where A and B are their respective set representatives. We then list a few basic properties of ordinal multiplication. This is followed by a definition of exponentiation and a presentation of a few exponentiation properties.*

34.1 Well-ordering the Cartesian product of well-ordered sets.

The definition of multiplication for cardinal numbers will serve as a model for the definition of multiplication on ordinal numbers. Recall how cardinal number multiplication was defined:

> Let S and T be two sets where S is of cardinality κ and T is of cardinality λ. We define: $\kappa \times \lambda$ as the cardinality of $S \times T$.

We cannot simply substitute the words "set" with "well-ordered set" and "cardinality" with the word "ordinality" since the ordinality of a set is always expressed in terms of a well-ordering on that set. We will order the elements of $S \times T$ lexicographically. We have discussed this ordering before, but since it is essential in the definition of ordinal multiplication we define it formally.

Definition 34.1 Let $(S, \leq_S)$ and $(T, \leq_T)$ be two well-ordered sets. We define the *lexicographic ordering* on the Cartesian product $S \times T$ as follows:

$$(s_1, t_1) \leq_{S\times T} (s_2, t_2) \text{ provided } \begin{cases} s_1 <_S s_2 \\ \quad \text{or} \\ s_1 = s_2 \quad \text{and} \quad t_1 \leq_T t_2 \end{cases}$$

Theorem 34.2 Let $(S, \leq_S)$ and $(T, \leq_T)$ be two well-ordered sets. The lexicographic ordering of the Cartesian product $S \times T$ is a well-ordering.

Proof: The proof is left as an exercise.

Theorem 34.3 If the well-ordered sets, S_1 and S_2, are order isomorphic and the well-ordered sets, T_1 and T_2, are order isomorphic, then the lexicographically ordered Cartesian products, $S_1 \times T_1$ and $S_2 \times T_2$, are order isomorphic.

Proof:

We are given onto order isomorphisms $f : S_1 \to S_2$ and $g : T_1 \to T_2$.
We are required to produce an onto order isomorphism $h : S_1 \times T_1 \to S_2 \times T_2$.
We define the function $h : S_1 \times T_1 \to S_2 \times T_2$ as $h(s,t) = (f(s), g(t))$.
We show that h is a well-defined one-to-one function on $S_1 \times T_1$: Since

$$\begin{aligned} h(s,t) = h(a,b) &\Rightarrow (f(s), g(t)) = (f(a), g(b)) \\ &\Rightarrow f(s) = f(a) \text{ and } g(t) = g(b) \\ &\Rightarrow s = a \text{ and } t = b \text{ (Since both } f \text{ and } g \text{ are one-to-one.)} \\ &\Rightarrow (s,t) = (a,b) \end{aligned}$$

then h is one-to-one.
The function h is onto $S_2 \times T_2$: If $(s,t) \in S_2 \times T_2$, then, since f and g are "onto" S_2 and T_2 respectively, $s = f(a)$ and $t = g(b)$ for some $a \in S_1$ and $b \in T_1$; hence, $h(a,b) = (s,t)$. Hence, h is onto $S_2 \times T_2$.
The function h respects the ordering of the sets:

$$\begin{aligned} (s_1,t_1) \leq_{S\times T} (s_2,t_2) &\Leftrightarrow \begin{cases} s_1 <_S s_2 \\ \quad \text{or} \\ s_1 = s_2 \quad \text{and} \quad t_1 \leq_T t_2 \end{cases} \\ &\Leftrightarrow \begin{cases} f(s_1) <_S f(s_2) \\ \quad \text{or} \\ f(s_1) = f(s_2) \quad \text{and} \quad g(t_1) \leq_T g(t_2) \end{cases} \\ &\Leftrightarrow \begin{cases} (f(s_1), g(t_1)) <_{S\times T} (f(s_2), g(t_2)) \\ \quad \text{or} \\ (f(s_1), g(t_1)) \leq (f(s_2), g(t_2)) \quad \text{(Equality} \Leftrightarrow (s_1,t_1) = (s_2,t_2)) \end{cases} \end{aligned}$$

So h is order isomorphic.

34.2 A definition of ordinal multiplication.

We are now set to define ordinal multiplication. At least for finite products, multiplication is closely linked to addition. For example, $3 \times 2 = 3 + 3 = 2 + 2 + 2$. We would expect multiplication of ordinals to be so that $\omega_0 \times 2$ equals $\omega_0 + \omega_0$, for example. We propose the following definition.

Definition 34.4 Let α and β be two ordinals with set representatives A and B respectively. We define the multiplication, $\alpha \times \beta$, as:

$$\alpha \times \beta = {}^{\text{ord}}(B \times A)$$

The product, $\alpha \times \beta$, is equivalently written as, $\alpha\beta$, (respecting the order). Note the order of the terms in the Cartesian product, $B \times A$, is different from the order, $\alpha \times \beta$, of their respective ordinalities.

---

a) *Example.* Compute both $\omega_0 \times 2$ and $2 \times \omega_0$.

$$\begin{aligned} \omega_0 \times 2 &= {}^{\text{ord}}(\{0,1\} \times \mathbb{N}) \\ &= {}^{\text{ord}}\{(0,0),(0,1),(0,2),\ldots,(0,n),\ldots,(1,0),(1,1),(1,2),\ldots,(1,n),\ldots\} \\ &= \omega_0 + \omega_0 \end{aligned}$$

$$\begin{aligned} 2 \times \omega_0 &= {}^{\text{ord}}(\mathbb{N} \times \{0,1\}) \\ &= {}^{\text{ord}}\{(0,0),(0,1),(1,0),(1,1),(2,0),(2,1),\ldots,(n,0),(n,1),\ldots,\} \\ &= \omega_0 \end{aligned}$$

We see that this multiplication is not commutative. At least in one case of $\omega_0 \times 2$ it is linked to the notion of sums.

b) *Example.* Order the following pairs.

i) $((\omega_0, \omega_0 \times 2), 7)$ and $((\omega_0, 2 \times \omega_0), 2 + \omega_0)$,

ii) $(\omega_0, (\omega_0 \times 2, 7))$ and $(\omega_0, (2 \times \omega_0, 2 + \omega_0))$.

Solution of i): We compare $((\omega_0, \omega_0 \times 2), 7)$ and $((\omega_0, 2 \times \omega_0), 2 + \omega_0)$:
We first compare the two first coordinates $(\omega_0, \omega_0 \times 2)$ and $(\omega_0, 2 \times \omega_0)$.

$$\begin{array}{rcl} (\omega_0, \omega_0 \times 2) = (\omega_0, \omega_0 + \omega_0) & \text{and} & (\omega_0, 2 \times \omega_0) = (\omega_0, \omega_0) \\ & \Rightarrow & \\ (\omega_0, 2 \times \omega_0) & < & (\omega_0, \omega_0 \times 2) \\ & \Rightarrow & \\ ((\omega_0, 2 \times \omega_0), 2 + \omega_0) & < & ((\omega_0, \omega_0 \times 2), 7) \end{array}$$

Solution of ii): We compare $(\omega_0, (\omega_0 \times 2, 7))$ and $(\omega_0, (2 \times \omega_0, 2 + \omega_0))$:

$$\begin{array}{rcl} (\omega_0 \times 2, 7) = (\omega_0 + \omega_0, 7) & \text{and} & (2 \times \omega_0, 2 + \omega_0) = (\omega_0, \omega_0) \\ & \Rightarrow & \\ (2 \times \omega_0, 2 + \omega_0) & < & (\omega_0 \times 2, 7) \\ & \Rightarrow & \\ (\omega_0, (2 \times \omega_0, 2 + \omega_0)) & < & (\omega_0, (\omega_0 \times 2, 7)) \end{array}$$

34.3 Basic properties of ordinal multiplication.

We show that some of the multiplication properties generalize from finite ordinal multiplication to transfinite ordinal multiplication.

Theorem 34.5 Let α, β and γ be three ordinal numbers. Then:

a) $(\gamma\beta)\alpha = \gamma(\beta\alpha)$ (Multiplication is associative.)

b) For any $\gamma > 0$, $\alpha < \beta \Rightarrow \gamma\alpha < \gamma\beta$

c) $\gamma(\alpha + \beta) = \gamma\alpha + \gamma\beta$ (Left-hand distribution is acceptable.)

d) For any $\gamma > 0$, $\gamma\alpha = \gamma\beta \Rightarrow \alpha = \beta$ (Left-hand cancellation is acceptable.)

e) $\gamma 0 = 0$

f) For any limit ordinal $\beta \neq 0$, $\alpha\beta = \text{lub}\,\{\alpha\gamma : \gamma < \beta\}$

Proof: Let A, B and C be three pairwise disjoint well-ordered set representatives of the ordinals α, β and γ respectively.

a) We are required to show that $(A \times B) \times C$ is order isomorphic to $A \times (B \times C)$.

$$\begin{aligned}
(\alpha \times \beta) \times \gamma &= {}^{\text{ord}}(B \times A) \times {}^{\text{ord}}C \\
&= {}^{\text{ord}}(C \times (B \times A)) \\
&= {}^{\text{ord}}((C \times B) \times A) \text{ (This is justified below at } {}^{*}\text{.)} \\
&= {}^{\text{ord}}A \times {}^{\text{ord}}(C \times B) \\
&= \alpha \times (\beta \times \gamma)
\end{aligned}$$

* That the sets $C \times (B \times A)$ and $(C \times B) \times A$ are equipotent follows from theorem 4.9. We now show that the one-to-one function $f(a,(b,c)) = ((a,b),c)$ which maps $A \times (B \times C)$ onto $(A \times B) \times C$ respects the ordering:

$$\text{If } \begin{cases} a < c \end{cases} \quad \begin{cases} \text{in } A \times (B \times C): & \longrightarrow & (a,(b,c)) < (c,(d,e)) \\ \text{in } (A \times B) \times C: & (a,b) < (c,d) \Rightarrow & ((a,b),c) < ((c,d),e) \end{cases}$$

$$\text{If } \begin{cases} a = c \\ \text{and} \\ b < d \end{cases} \quad \begin{cases} \text{in } A \times (B \times C): & (b,c) < (d,e) \Rightarrow & (a,(b,c)) < (c,(d,e)) \\ \text{in } (A \times B) \times C: & (a,b) < (c,d) \Rightarrow & ((a,b),c) < ((c,d),e) \end{cases}$$

$$\text{If } \begin{cases} a = c \\ b = d \\ c < e \end{cases} \quad \begin{cases} \text{in } A \times (B \times C): & (b,c) < (d,e) \Rightarrow & (a,(b,c)) < (c,(d,e)) \\ \text{in } (A \times B) \times C: & \longrightarrow & ((a,b),c) < ((c,d),e) \end{cases}$$

We have considered all cases and see that the ordering is respected.

b) We are given that $\gamma > 0$ and $\alpha < \beta$. We are required to show that $\gamma\alpha < \gamma\beta$.
Since $\alpha < \beta$, there exists an order isomorphism $f : A \to B$ mapping A to an initial segment in B. It suffices to show that $C \times A$ is isomorphic to an initial segment of $C \times B$. Define the function $g : C \times A \to C \times B$ as follows: $g((c,a)) = (c, f(a))$.
Since f is one-to-one into B, g is easily seen to be one-to-one.
We show that g respects the order:

- Suppose $(u, v) \leq_{C \times A} (s, t)$.
- If $u = s$, then $v \leq_A t$ and so $(u, f(v)) \leq_{C \times B} (u, f(t)) = (s, f(t))$
- If $u <_C s$, then $(u, f(v)) <_{C \times B} (s, f(t))$.
- So g respects the order.

c) The left-distributive property proof is left as an exercise.

d) We are given that $\gamma > 0$ and $\gamma\alpha = \gamma\beta$. We are required to show that $\alpha = \beta$.
Suppose not. Suppose, without loss of generality that $\alpha < \beta$. Then by part b), $\gamma\alpha < \gamma\beta$, a contradiction. So $\alpha = \beta$, as required.

e) We are required to show that $\gamma \times 0 = 0$.
Simply note that $\gamma \times 0 = {}^{\text{ord}}(\{\ \} \times C) = {}^{\text{ord}}\{\ \} = 0$.

f) We are given that β is a limit ordinal not equal to 0. We are required to show that $\alpha\beta = \text{lub}\,\{\alpha\gamma : \gamma < \beta\}$.
We claim that $\alpha\beta$ is an upper bound of the set $\{\alpha\gamma : \gamma < \beta\}$:
- Note that $\gamma < \beta \Rightarrow \alpha\gamma < \alpha\beta$ (by part b). So $\alpha\beta$ is an upper bound of the set $\{\alpha\gamma : \gamma < \beta\}$ as claimed.

We claim that if δ is an upper bound of the set $\{\alpha\gamma : \gamma < \beta\}$, then $\alpha\beta \leq \delta$:
- Let δ be an upper bound of the set $\{\alpha\gamma : \gamma < \beta\}$.
- We claim that $\delta \geq \alpha\,\text{lub}\,\{\gamma : \gamma < \beta\}$:
  - Suppose $\delta < \alpha\,\text{lub}\,\{\gamma : \gamma < \beta\}$.

$$\begin{aligned} \delta < \alpha\,\text{lub}\,\{\gamma : \gamma < \beta\} &\Rightarrow \delta \in \alpha \bigcup_{\gamma<\beta} \gamma \\ &\Rightarrow \delta \in \alpha\gamma \quad {\scriptstyle \text{For some } \gamma < \beta} \\ &\Rightarrow \delta < \alpha\gamma \end{aligned}$$

  - But $\delta < \alpha\gamma$ contradicts the fact that δ is an upper bound of the set $\{\alpha\gamma : \gamma < \beta\}$.
  - So $\delta \geq \alpha\,\text{lub}\,\{\gamma : \gamma < \beta\}$ as claimed.
- Since β is a limit ordinal, then $\beta = \text{lub}\,\{\gamma : \gamma < \beta\}$.
- So $\delta \geq \alpha\beta$ as claimed.

Combining the two claim statements, we obtain $\alpha\beta = \text{lub}\,\{\alpha\gamma : \gamma < \beta\}$.

34.4 Examples.

When performing ordinal arithmetic it is not always obvious how to simplify expressions. In the following examples we show how some of the properties shown above can be used to simplify expressions.

a) By mathematical induction, show that for any ordinal α and finite ordinal $n < 0$

$$\alpha n = \alpha + \alpha + \cdots + \alpha \text{ } (n \text{ times})$$

Solution: Induction on the natural numbers. Let A be a set such that ${}^{\text{ord}}A = \alpha$.
Let $P(n)$ denote the statement "$\alpha n = \alpha + \alpha + \cdots + \alpha$ $(n \text{ times})$"
Base case: Since $\alpha 1 = {}^{\text{ord}}(\{0\} \times A) = {}^{\text{ord}}A = \alpha$, $P(1)$ holds true.
Inductive hypothesis: Suppose $P(n)$ holds true. Then, by left-hand distributivity of ordinals, $\alpha(n+1) = \alpha n + \alpha 1 = \alpha n + \alpha = \alpha + \alpha + \cdots + \alpha$ $(n+1 \text{ times})$. By mathematical induction $\alpha n = \alpha + \alpha + \cdots + \alpha$ $(n \text{ times})$ for all non-zero finite ordinals n.

b) Define $\alpha + \alpha + \alpha + \cdots +$ (Countably infinite times) $= \text{lub}\{\alpha,\ \alpha+\alpha,\ \alpha+\alpha+\alpha, \ldots\}$. Show that for any α,

$$\alpha\omega_0 = \alpha + \alpha + \alpha + \cdots + \text{ (Countably infinite times)}$$

Solution:

$$\begin{aligned} \alpha\omega_0 &= \text{lub}\{\alpha n : n < \omega_0\} \\ &= \text{lub}\{\alpha,\ \alpha 2,\ \alpha 3, \alpha 4, \ldots\} \\ &= \text{lub}\{\alpha,\ \alpha+\alpha,\ \alpha+\alpha+\alpha, \ldots\} \\ &= \alpha + \alpha + \alpha + \cdots \text{ (Countably infinite times)} \end{aligned}$$

c) Show that $m\omega_0 = \omega_0$ for any non-zero finite ordinal m.

Solution:

$$\begin{aligned} m\omega_0 &= \text{lub}\{mn : n < \omega_0\} \\ &= \omega_0 \end{aligned}$$

d) We define $\omega_0^2 = \omega_0\omega_0$. Express $\omega_0^2 = \omega_0\omega_0$ as an infinite sum illustrating what this means.

Solution:

$$\begin{aligned} \omega_0\omega_0 &= \text{lub}\{\omega_0 n : n < \omega_0\} \\ &= \text{lub}\{\omega_0,\ \omega_0+\omega_0,\ \omega_0+\omega_0+\omega_0, \ldots\} \\ &= \omega_0 + \omega_0 + \omega_0 + \cdots \text{ (Countably infinite times)} \end{aligned}$$

34.5 Some ordinal comparisons.

The following table may be of some help in seeing how countably infinite ordinals are ordered. The ordinals are increasing with respect to $\in$ as we go down the table. All the ordinals in the third column are limit ordinals. There is no end to this process, so there is an unlimited source of ordinal numbers.

$$\begin{array}{ccccc}
0 & \leq & 0,1,2,3,\ldots & < & \omega_0 \\
\omega_0 & \leq & \omega_0,\ \omega_0+1,\ \omega_0+2,\ \omega_0+3,\ldots, & < & \omega_0+\omega_0 \\
\omega_0 2 & \leq & \omega_0 2,\ \omega_0 2+1,\ \omega_0 2+2,\ \omega_0 2+3,\ \ldots, & < & \omega_0 2+\omega_0 \\
\omega_0 3 & \leq & \omega_0 3,\ \omega_0 3+1,\ \ldots,\ \omega_0 4,\ \ldots,\ \omega_0 5,\ \ldots, & < & \omega_0\omega_0=\omega_0^2 \\
\omega_0\omega_0=\omega_0^2 & \leq & \omega_0^2,\ \omega_0^2+1,\ \ \ldots,\ \omega_0^2+\omega_0,\ \ldots,\ \ldots,\ \omega_0^2+\omega_0\omega_0 & = & \omega_0^2+\omega_0^2=\omega_0^2 2 \\
\omega_0^2 2 & \leq & \omega_0^2 2,\ \omega_0^2 2+1\ \ldots,\ \omega_0^2 3,\ \ldots,\ \omega_0^2 4,\ \ldots,\ \omega_0^2\omega_0 & = & \omega_0^3 \\
\omega_0^3 & \leq & \omega_0^3,\ \omega_0^3+1,\ \ldots,\ \omega_0^4,\ \ldots,\omega_0^5,\ \ldots,\omega_0^6,\ \ldots & < & \omega_0^{\omega_0} \\
\omega_0^{\omega_0} & \leq & \omega_0^{\omega_0},\ \omega_0^{\omega_0}+1,\ \ldots,\ (\omega_0^{\omega_0})^{\omega_0},\ \ldots,\ ((\omega_0^{\omega_0})^{\omega_0})^{\omega_0},\ \ldots, & < & \omega_1,\ldots
\end{array}$$

Many of the ordinals listed above may seem incredibly large. In spite of this, note that every one of these ordinals is countable!

Observe that we go from one limit ordinal to the next by adding a countably infinite set, remembering (from theorem 19.6), that adding a countably infinite set to some infinite set S does not change the cardinality of S. We see this in more detail in the following table.

$$\begin{array}{l}
\omega_0 \to \omega_0 2 \to \omega_0 3 \to \cdots \to \omega_0^2 \\
\qquad \to \omega_0^2+1 \to \omega_0^2+2 \to \cdots \to \omega_0^2+\omega_0 \\
\qquad\qquad \to \omega_0^2+\omega_0+1 \to \omega_0^2+\omega_0+2 \to \cdots \to \omega_0^2+\omega_0 2 \\
\qquad\qquad\qquad \to \omega_0^2+\omega_0^2=\omega_0^2 2 \to \cdots \to \omega_0^2\omega_0=\omega_0^3 \\
\qquad\qquad\qquad\qquad \to \omega_0^3+1=\omega_0^3+2 \to \cdots \to \omega_0^4
\end{array}$$

Recall that all of these are elements of the first uncountable ordinal is ω_1 (defined as the least ordinal which cannot be embedded in ω_0). The ordinal ω_1 is uncountable. It is not constructed (other than being the union of all countable ordinals), but its existence is guaranteed by Hartogs' lemma (see 28.9, 28.10 and 28.11).

34.6 Ordinal Exponentiation.

We now define, by transfinite recursion, exponentiation of ordinals .

Definition 34.6 Let γ be any non-zero ordinal. We define the γ-based exponentiation function $g_\gamma : \mathscr{O} \to \mathscr{O}$ as follows:

1) $g_\gamma(0) = 1$

2) $g_\gamma(\alpha^+) = g_\gamma(\alpha)\gamma$

3) $g_\gamma(\alpha) = \text{ lub}\{g_\gamma(\beta) : \beta < \alpha\}$ whenever α is a limit ordinal.

Whenever $\gamma \neq 0$ we represent $g_\gamma(\alpha)$ as γ^α. Then $\gamma^{\alpha+1} = \gamma^\alpha\gamma$. If $\gamma = 0$ we define $\gamma^\alpha = 0^\alpha = 0$.

Many of the principles used in the computation of expressions involving exponentiation of natural numbers extend to large ordinal numbers. We begin with expressions involving inequalities.

Theorem 34.7 Let α, β and γ be three ordinal numbers. Then, assuming $\gamma > 1$,

$$\alpha < \beta \Leftrightarrow \gamma^\alpha < \gamma^\beta$$

Proof: In what follows, $1 < \gamma$.

($\Rightarrow$) Proof by transfinite induction. Let $P(\beta)$ denote the statement "$\alpha < \beta \Rightarrow \gamma^\alpha < \gamma^\beta$". Then $P(0)$ holds (vacuously) true. Suppose $P(\delta)$ holds true for all $\delta < \beta$. That is, suppose $\alpha < \delta \Rightarrow \gamma^\alpha < \gamma^\delta$ for any $\delta < \beta$.
Case 1 : β is a successor ordinal. That is, $\beta = \mu + 1$ for some ordinal μ. Suppose $\delta < \beta = \mu + 1$. It suffices to show that $\gamma^\delta < \gamma^\beta$.

$$\begin{array}{rcl} \delta = \mu & \Rightarrow & \gamma^\delta = \gamma^\mu \\ & \Rightarrow & \gamma^\delta(1) < \gamma^\mu\gamma \quad \text{(By theorem 34.5)} \\ & \Rightarrow & \gamma^\delta < \gamma^{\mu+1} = \gamma^\beta \quad \text{(By definition)} \\ \delta < \mu & \Rightarrow & \gamma^\delta < \gamma^\mu \text{ (By induction hypothesis))} \\ & \Rightarrow & \gamma^\delta < \gamma^\mu(1) < \gamma^\mu\gamma = \gamma^{\mu+1} = \gamma^\beta \end{array}$$

In both cases, $\delta < \beta$ implies $\gamma^\delta < \gamma^\beta$.
Case 2 : β is a limit ordinal. That is, $\beta = \text{ lub}\{\delta : \delta < \beta\}$. We are given that $\delta < \mu < \beta \Rightarrow \gamma^\delta < \gamma^\mu$. Suppose $\delta < \beta$. It suffices to show that $\gamma^\delta < \gamma^\beta$.

$$\begin{array}{rcl} \delta < \beta & \Rightarrow & \delta < \delta + 1 < \beta \\ & \Rightarrow & \gamma^\delta < \gamma^{\delta+1} \leq \text{ lub}\{\gamma^\delta : \delta < \beta\} = \gamma^\beta \quad \text{(By 3) in Definition 34.6)} \\ & \Rightarrow & \gamma^\delta < \gamma^\beta \end{array}$$

In both cases 1 and 2, $\delta < \beta$ implies $\gamma^\delta < \gamma^\beta$.

($\Rightarrow$) Suppose $\gamma^\delta < \gamma^\beta$. If $\beta < \delta$ then by the first part above we have both $\gamma^\beta < \gamma^\delta$ and $\gamma^\delta < \gamma^\beta$, a contradiction. Then $\delta \leq \beta$. Since $\delta = \beta$ implies $\gamma^\delta = \gamma^\beta$, we must have $\delta < \beta$, as required.

Theorem 34.8 Let α, β and γ be three ordinal numbers where $\alpha \neq 0$.

a) $\gamma^\beta \gamma^\alpha = \gamma^{\beta+\alpha}$

b) $(\gamma^\beta)^\alpha = \gamma^{\beta\alpha}$

Proof:

a) Proof by transfinite induction. Let $P(\delta)$ denote the statement "$\gamma^\beta\gamma^\delta = \gamma^{\beta+\delta}$". Then $P(0)$ holds true. Suppose $P(\delta)$ holds true for all $\delta < \alpha$. That is, suppose $\gamma^\beta\gamma^\delta = \gamma^{\beta+\delta}$ whenever $\delta < \alpha$.
Case 1 : α is a successor ordinal. That is, $\alpha = \mu + 1$ for some ordinal μ. Then

$$\begin{aligned}
\gamma^\beta\gamma^\alpha &= \gamma^\beta\gamma^{\mu+1} \\
&= \gamma^\beta\gamma^\mu\gamma \\
&= \gamma^{\beta+\mu}\gamma \quad \text{(By the inductive hypothesis)} \\
&= \gamma^{(\beta+\mu)+1} \\
&= \gamma^{\beta+(\mu+1)} \\
&= \gamma^{\beta+\alpha}
\end{aligned}$$

Case 2 : α is a limit ordinal. That is, $\alpha = \text{ lub}\{\delta : \delta < \alpha\}$. We are given that $\gamma^\beta\gamma^\delta = \gamma^{\beta+\delta}$ whenever $\delta < \alpha$. We are required to show that $\gamma^\beta\gamma^\alpha = \gamma^{\beta+\alpha}$. We know that $\gamma^\alpha = \text{ lub}\{\gamma^\delta : \delta < \alpha\}$.

We claim that γ^α must be a limit ordinal: Suppose not. Then $\gamma^\alpha = \mu + 1$ for some ordinal μ. Then $\mu < \gamma^\delta$ for some $\delta < \alpha$ (for, if not, $\mu < \gamma^\alpha = \text{ lub}\{\gamma^\delta : \delta < \alpha\} \leq \mu$, a contradiction). Then (by 34.7) $\mu + 1 \leq \gamma^\delta < \gamma^\alpha = \mu + 1$, a contradiction. So γ^α is a limit ordinal as claimed.

We claim that $\gamma^{\beta+\alpha} \leq \gamma^\beta\gamma^\alpha$:
By definition 34.6 c), $\gamma^{\beta+\alpha} = \text{ lub}\{\gamma^\delta : \delta < \beta+\alpha\}$ ("α is limit ordinal" $\Rightarrow$ "$\beta + \alpha$ is a limit ordinal").

$$\begin{aligned}
\delta < \beta + \alpha &\Rightarrow \delta < \beta + \mu < \beta + \alpha \text{ for some } \mu < \alpha \\
&\Rightarrow \gamma^\delta < \gamma^{\beta+\mu} = \gamma^\beta\gamma^\mu < \gamma^\beta\gamma^\alpha \text{(By induction hypothesis and theorem 34.7)}
\end{aligned}$$

Hence $\gamma^{\beta+\alpha} = \text{ lub}\{\gamma^\delta : \delta < \beta + \alpha\} \leq \gamma^\beta\gamma^\alpha$, as claimed.

We claim that $\gamma^\beta\gamma^\alpha \leq \gamma^{\beta+\alpha}$:
Since γ^α is a limit ordinal $\gamma^\beta\gamma^\alpha = \text{ lub}\{\gamma^\beta\delta : \delta < \gamma^\alpha\}$. (Theorem 34.5 f))

$$\begin{aligned}
\delta < \gamma^\alpha &\Rightarrow \delta < \gamma^\mu < \gamma^\alpha \text{ for some } \mu < \alpha \\
&\Rightarrow \gamma^\beta\delta < \gamma^\beta\gamma^\mu = \gamma^{\beta+\mu} < \gamma^{\beta+\alpha} \text{(By induction hypothesis.)}
\end{aligned}$$

Then $\gamma^\beta\gamma^\alpha = \text{ lub}\{\gamma^\beta\delta : \delta < \gamma^\alpha\} \leq \gamma^{\beta+\alpha}$, as claimed.
We conclude that $\gamma^\beta\gamma^\alpha = \gamma^{\beta+\alpha}$, as required.

b) This can be proved by transfinite induction as in part a). Let $P(\delta)$ denote the statement "$(\gamma^\beta)^\delta = \gamma^{\beta\delta}$". Then $P(0)$ holds true. Suppose $P(\delta)$ holds true for all $\delta < \alpha$. That is, suppose $(\gamma^\beta)^\delta = \gamma^{\beta\delta}$ whenever $\delta < \alpha$. As in part a) consider the two cases, 1) α has an immediate predecessor, and 2) α is a limit ordinal, separately. The details are left as an exercise.

Concepts review:

1. When defining multiplication of ordinals, what kind of ordering is used on the Cartesian product of the ordinals being multiplied?
2. Define the multiplication of two ordinals α and β.
3. How does the definition of ordinal multiplication compare with the definition of cardinal multiplication?
4. Construct a set representation for the ordinal product $\omega_0 2$.
5. How does the ordinal $\omega_0 3$ compare with the ordinal $3\omega_0$?
6. Is left-hand distribution acceptable in ordinal multiplication?
7. If we start with the countably infinite ordinal ω_0 and gradually increase its ordinality, one at a time, by an endless process of ordinal addition and multiplication, can we reach ω_1 with this process?

EXERCISES

A.
1. Simplify or describe the following expressions.
   a) $\omega_0 + \omega_0^2$
   b) $(\omega_0 + 2) + \omega_0$
   c) $\omega_0 2 + 1$
2. Which ordinal is larger, $3\omega_0$ or $\omega_0 3$? Explain.
3. Does the ordinal $\omega_0^2 + 3\omega_0$ have an immediate predecessor? If it does describe it.
4. Find two ordinals whose ordinal product is $\omega_0 + \omega_0^2$.

B.
5. Prove that if $\delta < \alpha\beta$, then $\delta = \mu\gamma$ for some $\mu \leq \alpha$ and some $\gamma \leq \beta$.
6. Show that $(\omega_0 + \omega_0)\omega_0 = \omega_0\omega_0$

C.

7. Prove that $\alpha 1 = 1\alpha = \alpha$.
8. Prove that if α and β are both finite ordinals, then $\alpha\beta = \beta\alpha$.
9. Prove that if $\alpha\beta = 0$, then either α or β is 0.
10. Prove that $\gamma(\alpha + \beta) = \gamma\alpha + \gamma\beta$.

Appendix

A / Boolean algebras and Martin's axiom.

Summary. *We discuss those partially ordered sets called lattices and Boolean algebras. Any Boolean algebra B is then shown to have a topological representation $\mathscr{B}(\mathscr{S}(B))$. Finally characterizations of Martin's axiom are given in terms of these concepts.*

A.1 Lattices.

Given a partially ordered set $(P, \leq)$ there may be pairs of elements a, b in P such that a and b do not have a common upper bound or a common lower bound in P. Those partially ordered sets in which every pair of elements in P have lower and upper bounds play an important role in mathematics. We refer to these sets as *lattices*.

Definition 0.1 A partially ordered set $(P, \leq)$ is called a *lattice* if $a \vee b = \max\{a, b\}$ and $a \wedge b = \min\{a, b\}$ both exist in P for all pairs a, b in P.

Definition 0.2 If B is a subset of a partially ordered set, $(P, \leq)$, $\vee B$ denotes the least upper bound of B and $\wedge B$ denotes the greatest lower bound of B (both with respect to $\leq$). Note that $\vee B$ and $\wedge B$ may or may not be an element of B. A lattice $(P, \leq)$ is said to be a *complete lattice* if for any non-empty subset B of P, both $\vee B$ and $\wedge B$ exist and belong to P.

Examples of lattices.

1) Let X be a set $(\mathscr{P}(X), \subseteq)$ be a partially ordered set ordered by inclusion. Let A and B be any pair of subsets of X (read, $A, B \in \mathscr{P}(X)$). We define $A \vee B = A \cup B$ and $A \wedge B = A \cap B$. For $\mathscr{B} \subseteq \mathscr{P}(X)$, $\vee\mathscr{B} = \cup\{A \in \mathscr{P}(X) : A \in \mathscr{B}\}$ (read, "$\vee\mathscr{B}$ is the union of all subsets, A, of X such that $A \in \mathscr{B}$"). Also define $\wedge\mathscr{B} = \cap\{A \in \mathscr{P}(X) : A \in \mathscr{B}\}$ (read, "$\wedge\mathscr{B}$ is the intersection of all subsets, A, of X such that $A \in \mathscr{B}$"). In this case, with $\vee$ and $\wedge$ defined as $\cup$ and $\cap$, respectively, $(\mathscr{P}(X), \subseteq, \cup, \cap)$ is a complete lattice.
2) Let $\tau(X)$ be a topology on the set X. Let A and B be any pair of open subsets of X (read, $A, B \in \tau(X)$). We define $A \vee B = A \cup B$ and $A \wedge B = A \cap B$. For $\mathscr{B} \subseteq \mathscr{P}(X)$, $\vee\mathscr{B} = \cup\{A \in \tau(X) : A \in \mathscr{B}\}$. Also define $\wedge\mathscr{B} = \mathrm{int}_X(\cap\{A \in \mathscr{P}(X) : A \in \mathscr{B}\})$. In this case, with $\vee$ and $\wedge$ defined as $\cup$ and $\cap$, respectively, $(\tau(x), \subseteq, \cup, \cap)$ is a complete lattice in $(\mathscr{P}(X), \subseteq)$.

3) Let X be topological space. Let $\mathscr{B}(X) = \{A \subseteq X : A \text{ is clopen in } X\}$[5] partially ordered by inclusion $\subseteq$. If A and B are clopen subsets of X we define $A \vee B = A \cup B$ and $A \wedge B = A \cap B$. Then $(\mathscr{B}(X), \subseteq, \cup, \cap)$ is a lattice in $(\mathscr{P}(X), \subseteq)$. But it is not necessarily complete. Verify that $(\mathscr{B}(\mathbb{Q}), \subseteq, \cup, \cap)$ is *not* a complete lattice.

Definition 0.3 Let X be a topological space. A subset B is said to be *regular open in* X if $B = \text{int}_X(\text{cl}_X(B))$. The set of all regular open subsets of X will be denoted as $\mathscr{R}o(X)$.

For example, $(-5, 0) \cup (0, 5)$ is open in $\mathbb{R}$ but *not* regular open in $\mathbb{R}$.

A particular example of a complete lattice. Let X be a topological space with a Hausdorff topology $\tau(X)$. By definition, $\mathscr{R}o(X) \subseteq \tau(X)$. For the partially ordered set $(\mathscr{R}o(X), \subseteq)$ we define $\wedge$ and $\vee$ as follows: For $A, B \in \mathscr{R}o(X)$,

$$\begin{aligned} A \wedge B &= A \cap B \\ A \vee B &= \text{int}_X(\text{cl}_X(A \cup B)) \end{aligned}$$

If we want to form a lattice $(\mathscr{R}o(X), \subseteq, \vee, \cap)$ in the partially ordered set $(\tau(X), \subseteq)$, we cannot define $A \vee B$ as $A \cup B$ since it is not true, in general, that $\text{int}_X(\text{cl}_X(A)) \cup \text{int}_X(\text{cl}_X(B)) \in \mathscr{R}o(X)$. To see this consider, for example, the open intervals $A = (a, b)$ and $B = (b, c)$ both elements of $\mathscr{R}o(\mathbb{R})$. See that

$$\text{int}_{\mathbb{R}}(\text{cl}_{\mathbb{R}}(A \cup B)) = (a, c) \neq (a, b) \cup (b, c) = A \cup B$$

Hence $\mathscr{R}o(X)$ is not closed with respect to the union, $\cup$, of finitely many sets. The following statement confirms that $(\mathscr{R}o(X), \subseteq, \vee, \cap)$ is a complete lattice in the partially ordered set $(\tau(X), \subseteq)$.

Theorem 0.4 Let X be a topological space. Then $(\mathscr{R}o(X), \subseteq, \vee, \cap)$ is a complete lattice in $(\tau(X), \subseteq)$.

Proof: It immediately follows from our definition of $\vee$ that, for $A, B \in \mathscr{R}o(X)$, $A \vee B \in \mathscr{R}o(X)$. We leave it as an exercise to show that

$$A \cap B = \text{int}_X(\text{cl}_X(A)) \cap \text{int}_X(\text{cl}_X(B)) = \text{int}_X(\text{cl}_X(A \cap B)) \in \mathscr{R}o(X)$$

Suppose $\mathscr{D} \subseteq \mathscr{R}o(X)$. Showing that

$$\begin{aligned} \cap\{B : B \in \mathscr{D}\} &= \text{int}_X(\text{cl}_X(\cap\{B : B \in \mathscr{D}\})) \in \mathscr{R}o(X) \\ \vee\{B : B \in \mathscr{D}\} &= \text{int}_X(\text{cl}_X(\cup\{B : B \in \mathscr{D}\})) \in \mathscr{R}o(X) \end{aligned}$$

[5] A set is clopen if it is simultaneously open and closed in X

is also left as an exercise. Then $(\mathscr{Ro}(X), \subseteq, \vee, \cap)$ is a complete lattice in $(\tau(X), \subseteq)$ as required.

The partially ordered set $\mathscr{Ro}(X)$ forms a base for a topology on X. Since $\varnothing, X \in \mathscr{Ro}(X)$ and $\mathscr{Ro}(X)$ is closed under intersections, $\cap$, then $\mathscr{Ro}(X)$ forms a base for some topology. That is, $\mathscr{Ro}(X)$ generates some topology $\tau^*(X) \subseteq \tau(X)$. If $(X, \tau(X))$ is assumed to be Hausdorff, $\tau(X)$ separates points of X; it then easily follows that $\mathscr{Ro}(X)$ also separates points of X. We have shown that $(X, \tau^*(X))$ a Hausdorff on X.[6]

A.2 Lattice filters.

We will consider a lattice $(L \subseteq, \vee, \wedge)$. A subset $\mathscr{F}$ of $(L \subseteq, \vee, \wedge)$ is called an *L-filter* if

1) $\mathscr{F}$ is non-empty,
2) $\mathscr{F}$ is such that, for non-empty $A, B \in \mathscr{F}$ there exists $D \neq \varnothing$ such that $D \leq A \wedge B \in \mathscr{F}$)
3) $\mathscr{F}$ is such that, if $A \in \mathscr{F}$ and $A \leq C \in L$ then $C \in \mathscr{F}$).

When only conditions 1 and 2 are satisfied we say that $\mathscr{F}$ is an *L-filter base*. We say that the L-filter base, $\mathscr{F}$, generates the L-filter, $\mathscr{F}^{\uparrow}$. If $F \in L$, then $\{F\}^{\uparrow}$ is the L-filter generated by the singleton $\{F\}$. A filter $\mathscr{F}$ is said to be a *proper L-filter* if $\mathscr{F} \neq L$. A filter $\mathscr{F}$ is proper if and only if $\varnothing \notin \mathscr{F}$.

Definition 0.5 Let $(L, \leq, \vee, \wedge)$ be a lattice. An *L-ultrafilter* is a proper filter $\mathscr{F}$ in L which is not properly contained in any other proper filter in L. If the filter $\mathscr{F}$ is such that $\cap\{F : F \in \mathscr{F}\} \neq \varnothing$ then we say that the filter $\mathscr{F}$ is a *fixed ultrafilter*. Ultrafilters which are not fixed are said to be *free ultrafilters*.

Theorem 0.6 Let X be a topological space.

a) Suppose $\mathscr{F}$ is a proper L-filter where $(L, \subseteq, \vee, \wedge)$ is a lattice in $(\mathscr{P}(X), \subseteq)$. Then $\mathscr{F}$ can be extended to an L-ultrafilter.

b) Suppose $\mathscr{F}$ is an L-filter in $(L, \subseteq, \vee, \wedge)$ a lattice in $(\mathscr{P}(X), \subseteq)$. Then $\mathscr{F}$ is an L-ultrafilter if and only if for every $A \subseteq X$, either $A \in \mathscr{F}$ or $X - A \in \mathscr{F}$.

[6] Given a topological space $(X, \tau(X))$ the set $(X, \tau^*(X))$ is referred to as the *semiregularization* of $(X, \tau(X))$.

Proof:

a) Let $\mathscr{F}$ be a proper L-filter. Let $\mathscr{H} = \{\mathscr{M} : \mathscr{M}$ is a proper L-filter such that $\mathscr{F} \subseteq \mathscr{M}\}$. We partially order $\mathscr{H}$ with $\subseteq$. Let $\mathscr{C}$ be a chain in $(\mathscr{H}, \subseteq)$. Then $\cup\{C : C \in \mathscr{C}\}$ is an upper bound of $\mathscr{C}$ with respect to $\subseteq$. So every chain in $\mathscr{H}$ has an upper bound. By Zorn's lemma, $(\mathscr{H}, \subseteq)$ has a maximal element. That is, $\mathscr{H}$ contains a filter, $\mathscr{F}^*$, which is not properly contained in any other filter. Since $\mathscr{F}^* \in \mathscr{H}$, $\mathscr{F} \subseteq \mathscr{F}^*$. Then $\mathscr{F}$ can be extended to an L-ultrafilter, as required.

b) ($\Rightarrow$) We are given that $\mathscr{F}$ is an L-ultrafilter in $L \subseteq \mathscr{P}(X)$ and that $A \subseteq X$. Suppose neither A nor $X - A$ belongs to $\mathscr{F}$. Let $\mathscr{H} = \{B \in L : A \cap F \subseteq B$ for some $F \in \mathscr{F}\}$. Clearly, $\mathscr{F} \subseteq \mathscr{H}$ and $A \in \mathscr{H} - \mathscr{F}$. Then $\mathscr{F}$ is a proper subset of $\mathscr{H}$. Note that $\varnothing \notin \mathscr{H}$ for, if it was, $A \cap F = \varnothing$ for some F and so $F \subseteq X - A$ which would imply $X - A \in \mathscr{F}$, a contradiction. It is a straightforward exercise to show that $\mathscr{H}$ is closed under finite applications of $\wedge$. Hence $\mathscr{H}$ is an L-filter base which will generate a filter $\mathscr{H}^*$. Then $\mathscr{F} \subset \mathscr{H}^*$ which contradicts the fact that $\mathscr{F}$ is an L-ultrafilter.
($\Leftarrow$) Conversely, suppose that for every $A \subseteq X$, either $A \in \mathscr{F}$ or $X - A \in \mathscr{F}$. We are required to show that $\mathscr{F}$ is an ultrafilter. Suppose not. That is, suppose that there exists a proper filter, $\mathscr{H}$, such that $\mathscr{F} \subset \mathscr{H}$. Then there exists some non-empty A such $A \in \mathscr{H} - \mathscr{F}$. Then $X - A$ cannot belong to $\mathscr{F}$, for, if it did, then $A \cap (X - A) = \varnothing$ must belong to $\mathscr{H}$, contradicting the fact that $\mathscr{H}$ is a proper filter. Hence $\mathscr{F}$ must be an L-ultrafilter.

Theorem 0.7 Let X be a topological space. Then $(\mathscr{R}o(X), \subseteq, \vee, \cap)$ is a complete lattice in $(\tau(X), \subseteq)$. An $\mathscr{R}o(X)$-filter, $\mathscr{F}$, is an $\mathscr{R}o(X)$-ultrafilter if and only if, for any $A \in \mathscr{R}o(X)$, either A or $X - \mathrm{cl}_X(A)$ belongs to $\mathscr{F}$.

Proof:

($\Rightarrow$) Suppose $\mathscr{F}$ is an $\mathscr{R}o(X)$-ultrafilter and $A \in \mathscr{R}o(X)$. Let $F \in \mathscr{F}$. Case 1: If $F \cap A = \varnothing$ then (since F is open) $F \subseteq X - \mathrm{cl}_X A$. Since

$$\begin{aligned} X - \mathrm{cl}_X A &= \mathrm{int}_X(X - A) \\ &= \mathrm{int}_X(X - \mathrm{int}_X \mathrm{cl}_X A) \\ &= \mathrm{int}_X \mathrm{cl}_X(X - \mathrm{cl}_X A) \end{aligned}$$

$F \subseteq X - \mathrm{cl}_X A \in \mathscr{R}o(X)$ which implies $X - \mathrm{cl}_X A \in \mathscr{F}$. Case 2: Suppose $F \cap A \neq \varnothing$ for all $F \in \mathscr{F}$. Suppose $A \notin \mathscr{F}$. Let $\mathscr{H} = \{B \in \mathscr{R}o(X) : A \cap F \subseteq B\}$. Then $\mathscr{F} \subseteq \mathscr{H}$ and $A \in \mathscr{H} - \mathscr{F}$. As shown in the theorem above, $\mathscr{H}$ is a filter base which generates a filter $\mathscr{H}^*$ in $(\mathscr{R}o(X), \subseteq)$. Since $\mathscr{F} \subset \mathscr{H}^*$ this contradicts the fact that $\mathscr{F}$ is an $\mathscr{R}o(X)$-ultrafilter. So A must belong to $\mathscr{F}$.
($\Leftarrow$) Suppose that for any $A \in \mathscr{R}o(X)$, either A or $X - \mathrm{cl}_X(A)$ belongs to $\mathscr{F}$. See that the filter $\mathscr{F}$ extends to an $\mathscr{R}o(X)$-ultrafilter $\mathscr{F}^*$. Suppose $A \in \mathscr{F}^*$. If $A \notin \mathscr{F}$ then

$X - \mathrm{cl}_X A \in \mathscr{F} \subseteq \mathscr{F}^*$ implying that $A \cap (X - \mathrm{cl}_X A) = \varnothing \in \mathscr{F}^*$, a contradiction. So $A \in \mathscr{F}$. Then $\mathscr{F}^* \subseteq \mathscr{F}$ which implies that $\mathscr{F}$ is an $\mathscr{R}o(X)$-ultrafilter.

It can similarly be shown that a $\tau(X)$-filter, $\mathscr{F}$, is a $\tau(X)$-ultrafilter if and only if, for any $A \in \tau(X)$, either A or $X - \mathrm{cl}_X(A)$ belongs to $\mathscr{F}$.

A.3 Boolean algebras.

We now define certain lattices with special properties.

Definition 0.8 A lattice $(L, \vee, \wedge)$ is said to be a *distributive lattice* if, for any x, y, and z in L, $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$ and $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$.

- The lattice $(L, \vee, \wedge)$ is said to be a *complemented lattice* it has a maximum element, denoted by 1, and a minimum element, denoted by 0, and for every $x \in L$ there exists a unique x' such that $x \vee x' = 1$ and $x \wedge x' = 0$.
- A complemented distributive lattice is referred to as being a *Boolean algebra*. A Boolean algebra is denoted as $(B, \leq, \vee, \wedge, 0, 1, ')$ when we explicitly want to express what the maximum and minimum elements are.

Examples. The lattices $(\mathscr{P}(X), \subseteq, \cup, \cap, X, \varnothing, ')$ and $(\mathscr{B}(X), \subseteq, \cup, \cap, X, \varnothing, ')$ are the simplest examples of Boolean algebras. In both case $A' = X - A$.

The lattice $(\mathscr{R}o(X), \subseteq, \vee, \cap, X, \varnothing, ')$ is also a Boolean algebra, although this is not at all obvious. To prove this one must show that, for any $A, B, C \in \mathscr{R}o(X)$,

- $X - \mathrm{cl}_X A = \mathrm{int}_X \mathrm{cl}_X (X - \mathrm{cl}_X A) \in \mathscr{R}o(X)$,
- $A \cap (X - \mathrm{cl}_X A) = \varnothing$
- $A \vee (X - \mathrm{cl}_X A) = \mathrm{int}_X(\mathrm{cl}_X(A \cup (X - \mathrm{cl}_X A))) = X$
- $A \vee (B \cap C) = (A \vee B) \cap (A \vee C)$ and $A \cap (B \vee C) = (A \cap B) \vee (A \cap C)$

Proving these is left as an exercise.

Note that in the case where $A \in \mathscr{R}o(X)$, $A' = X - \mathrm{cl}_X A$. It is erroneous to interpret A' as meaning $X - A$.

Like lattices $(L, \subseteq, \vee, \wedge)$ in $\mathscr{P}(X)$ a Boolean algebra $(B, \leq, \vee, \wedge, 0, 1, ')$ contains B-filters and B-ultrafilters. The following concepts and properties are slight generalizations of ones we have already seen, so they will seem familiar to readers.

Boolean filters and ultrafilters. If $(B, \leq, \vee, \wedge, 0, 1, ')$ is a Boolean algebra then a *B-filter*, $\mathscr{F}$, is a non-empty subset of B which is closed under finite applications of $\wedge$ and, for any $x \in \mathscr{F}$, $(x \leq y) \Rightarrow (y \in \mathscr{F})$. We say that the B-filter, $\mathscr{F}$, is a *B-ultrafilter* if $\mathscr{F}$ is a proper filter and is not properly contained in any other proper B-ultrafilter.

We will emphasize three important B-ultrafilter properties. For any B-ultrafilter, $\mathscr{F}$,

- whenever $x \vee y \in \mathscr{F}$, either $x \in \mathscr{F}$ or $y \in \mathscr{F}$.
- if $x \in B$ and $x \wedge y \neq \varnothing$ for all $y \in \mathscr{F}$ then $y \in \mathscr{F}$.
- for any $x \in B$, either $x \in \mathscr{F}$ or $x' \in \mathscr{F}$. To see this suppose $x \notin \mathscr{F}$.

Proving these is left as an exercise.

Definition 0.9 Suppose we are given two lattices $(B_1, \leq_1, \vee_1, \wedge_1, 0, 1, ')$ and $(B_2, \leq_2, \vee_2, \wedge_2, 0, 1, ')$ and a function f which maps elements of B_1 to elements of B_2. We say that $f : B_1 \to B_2$ is a *Boolean homomorphism* if, for any $x, y \in B$,

1) $f(x \vee_1 y) = f(x) \vee_2 f(y)$,
2) $f(x \wedge_1 y) = f(x) \wedge_2 f(y)$
3) $f(x') = f(x)'$.

The function $f : B_1 \to B_2$ is a *Boolean isomorphism* if f is a bijection and both f and $f^{\leftarrow}$ are Boolean homomorphisms.

We say that $f : B_1 \to B_2$ is an *order homomorphism* if $(x \leq_1 y) \Rightarrow (f(x) \leq_2 f(y))$. It can be shown that:

1) If $f : B_1 \to B_2$ is a Boolean homomorphism then f must be an order homomorphism.
2) If $f : B_1 \to f[B_1]$ is a bijective Boolean homomorphism then f must be a Boolean isomorphism.

Proving these is left as an exercise.

Topological representations. We will say that a Boolean algebra $(B, \leq, \vee, \wedge, 0, 1, ')$ has a *topological representation* if the there exists a Boolean isomorphism $f : B \to \mathscr{B}(X)$ mapping B onto $\mathscr{B}(X)$ where $(\mathscr{B}(X), \subseteq, \cup, \cap))$ is the Boolean algebra of all clopen sets on some topological space X.

Definition 0.10 Let $(B, \leq, \vee, \wedge, 0, 1, ')$ be a Boolean algebra. Let $\mathscr{S}(B) = \{\mathscr{U} : \mathscr{U}$ is a B-ultrafilter$\}$. We define the function $f_B : B \to \mathscr{P}(\mathscr{S}(B))$ as follows: $f_B(x) = \{\mathscr{F} \in \mathscr{S}(B) : x \in \mathscr{U}\}$.

Theorem 0.11 Let $(B, \leq, \vee, \wedge, 0, 1, ')$ be a Boolean algebra. Then the set $\{f_B(x) : x \in B\}$ is a base for the open sets of some topology, $\tau(\mathscr{S}(B))$, on the set $\mathscr{S}(B)$ of all B-ultrafilters.

Proof:

Let $\mathscr{B}_B = \{f_B(x) : x \in B\}$. We first show that the sets in $\mathscr{B}_B$ cover all of $\mathscr{S}(B)$. Since B-ultrafilters are proper filters no ultrafilter can contain 0; then $f_B(0) = \varnothing \in \mathscr{B}_B$. Also, 1 belongs to all B-ultrafilters hence $f_B(1) = \mathscr{S}(B) \in \mathscr{B}_B$. Suppose $f_B(x)$ and $f_B(y)$ belong to $\mathscr{B}_B$. Any B-ultrafilter $\mathscr{F}$ contains $x \wedge y$ if and only if it contains both x and y. Then $\mathscr{F} \in f_B(x \wedge y)$ if and only if $\mathscr{F} \in f_B(x) \cap f_B(y)$. We can then write $f_B(x \wedge y) = f_B(x) \cap f_B(y)$. Then the set $\mathscr{B}_B = \{f_B(x) : x \in B\} \subseteq \mathscr{P}(\mathscr{S}(B))$ is a base for the open sets of some topology on $\mathscr{S}(B)$.

For a given Boolean algebra $(B, \leq, \vee, \wedge, 0, 1, ')$ we can now speak of a topological space $(\mathscr{S}(B), \tau(\mathscr{S}(B)))$ which is associated to B. Its elements are B-ultrafilters. When equipped with this topology the set $\mathscr{S}(B)$ is referred to as the *Stone space*. We now describe a few properties of the function f_B which associates B to subsets of the topological space $\mathscr{S}(B)$. We refer to this important theorem as the *Stone representation theorem.*

Theorem 0.12 Let $(B, \leq, \vee, \wedge, 0, 1, ')$ be a Boolean algebra.

1) The function $f_B : B \to \mathscr{P}(\mathscr{S}(B))$ is a Boolean homomorphism mapping B into $\mathscr{P}(\mathscr{S}(B))$.

2) For every $x \in B$, $f_B(x)$ is clopen in $\mathscr{S}(B)$. Hence $f_B[B] \subseteq \mathscr{B}(\mathscr{S}(B))$ (the set of all clopen sets in $\mathscr{S}(B)$).

3) The function $f_B : B \to \mathscr{S}(B)$ is a Boolean isomorphism mapping B into $\mathscr{B}(\mathscr{S}(B))$ (the set of all clopen sets in $\mathscr{S}(B)$).

4) The topological space $(\mathscr{S}(B), \tau(\mathscr{S}(B)))$ where $\tau(\mathscr{S}(B))$ is the topology generated by the open base $\{f_B(x) : x \in B\}$ is a compact zero-dimensional Hausdorff topological space.

5) The Boolean isomorphism $f_B : B \to \mathscr{S}(B)$ maps B onto $\mathscr{B}(\mathscr{S}(B))$ (the set of all clopen sets in $\mathscr{S}(B)$).

Proof:

We are given that $(B, \leq, \vee, \wedge, 0, 1, ')$ is a Boolean algebra.

1) We have shown in the previous theorem that $f_B(x \wedge y) = f_B(x) \cap f_B(y)$.

   We now show that $f_B(x \vee y) = f_B(x) \cup f_B(y)$: If $\mathscr{F} \in f_B(x) \cup f_B(y)$ then either x or y belongs to $\mathscr{F}$. By definition of a filter, $x \vee y \in \mathscr{F}$, hence $\mathscr{F} \in f_B(x \vee y)$. Then $f_B(x) \cup f_B(y) \subseteq f_B(x \vee y)$. On the other hand, if $\mathscr{F} \in f_B(x \vee y)$, then $x \vee y \in \mathscr{F}$. By a property of B-ultrafilters described above, either x or y belongs to $\mathscr{F}$. Hence $\mathscr{F}$ either belongs to $f_B(x)$ or to $f_B(y)$. Then $f_B(x \vee y) \subseteq f_B(x) \cup f_B(y)$. So $f_B(x \vee y) = f_B(x) \cup f_B(y)$.

   We show that $f_B(x') = f_B(x)' = \mathscr{S}(B) - f_B(x)$: We know that $f_B(x) \cap f_B(x') = f_B(x \wedge x') = f_B(0) = \varnothing$. We also know that $f_B(x) \cup f_B(x') = f_B(x \vee x') = f_B(1) = \mathscr{S}(B)$. We conclude that $f_B(x') = \mathscr{S}(B) - f_B(x)$.

   From this we conclude that $f_B : B \to \mathscr{P}(\mathscr{S}(B))$ is a Boolean homomorphism.

2) We are required to show that, for every $x \in B$, $f_B(x)$ is both an open and closed subset of $\mathscr{S}(B)$ (with respect to the topology generated by the base $\{f_B(x) : x \in B\}$). For each x in B, $f_B(x)$ is open in $\mathscr{S}(B)$ (since it is an open base element). Since $f_B(x') = f_B(x)'$ is open and $f_B(x) = \mathscr{S}(B) - f_B(x)'$ (shown above), then $f_B(x)$ is also closed. So every element of $\{f_B(x) : x \in B\}$ is clopen. It follows that $f_B[B] = \{f_B(x) : x \in B\} \subseteq \mathscr{B}(\mathscr{S}(B)) \subseteq \mathscr{P}(\mathscr{S}(B))$.

3) Having already shown that $f_B : B \to f_B[B] \subseteq \mathscr{B}(\mathscr{S}(B))$ is a homomorphism. To show it is an isomorphism it suffices to show that f_B is one-to-one.

   The function f_B *is one-to-one*: Suppose x and y are distinct points in B. Since $x \neq y$ then one of the two statements $x \leq y$, $y \leq x$ must be false. We will assume, without loss of generality, that $x \not\leq y$. We define $x - y = x \wedge y'$. We claim that since $x \not\leq y$ then $x - y \neq 0$. For suppose that $x \not\leq y$ and $x \wedge y' = 0$. See that

$$\begin{aligned} x' = x' \vee 0 &\Rightarrow x' \vee (x \wedge y') \\ &\Rightarrow x' \vee (x \wedge y') \\ &\Rightarrow (x' \vee x) \wedge (x' \vee y') \quad (B \text{ is distributive}) \\ &\Rightarrow 1 \wedge (x' \vee y') \\ &\Rightarrow (x' \vee y') \\ &\Rightarrow y' \leq x' \\ &\Rightarrow x \leq y \end{aligned}$$

   We have a contradiction. The source of the contradiction is our supposition that $x \wedge y' = 0$. So $x - y \neq 0$ as claimed. Let $\mathscr{F}$ be the B-filter $\{x - y\}^{\uparrow}$ generated by

$x - y \in B$. Then $\mathscr{F}$ extends to a B-ultrafilter $\mathscr{F}^*$. Then $\mathscr{F}^* \in f_B(x - y)$. But x and y' are both above $x - y = x \wedge y'$ and so must both belong to $\mathscr{F}^*$. Then $\mathscr{F}^* \in f_B(x) \cap f_B(y')$. Then $\mathscr{F}^*$ cannot belong to $f_B(y)$ (for if it did, y and y' would both belong to $\mathscr{F}^*$). Then $f_B(x) \neq f_B(y)$. We conclude that f_B is one-to-one on B.

4) Let $(\mathscr{S}(B), \tau(\mathscr{S}(B)))$ be the topological space where $\tau(\mathscr{S}(B))$ is the topology generated by the open base $\{f_B(x) : x \in B\}$. A topological space is zero-dimensional if it has an open base of clopen sets. We have shown above that $\{f_B(x) : x \in B\}$ is a set of clopen subsets of $\mathscr{S}(B)$. So $\mathscr{S}(B)$ is zero-dimensional.

The topological space $(\mathscr{S}(B), \tau(\mathscr{S}(B)))$ *is Hausdorff*: Let $\mathscr{F}_1$ and $\mathscr{F}_2$ be distinct B-ultrafilters in $\mathscr{S}(B)$. Then there exists some element $x \in \mathscr{F}_2$ which does not belong to $\mathscr{F}_1$. Then x' must belong to $\mathscr{F}_1$. It follows that $\mathscr{F}_1 \in f_B(x)$ and $\mathscr{F}_2 \in f_B(x')$. Since $x \wedge x' = 0$ implies $f_B(x) \cap f_B(x') = \varnothing$ it follows that $\mathscr{S}(B)$ is a Hausdorff topological space.

The topological space $(\mathscr{S}(B), \tau(\mathscr{S}(B)))$ *is compact*: To show that $\mathscr{S}(B)$ is compact it suffices to show that a set $\{F_i\}_{i \in I}$ of closed subsets which satisfies the finite intersection property has non-empty intersection. Let $\mathscr{H} = \{F_i\}_{i \in I}$ be a collection of closed subsets of $\mathscr{S}(B)$ which satisfies the finite intersection property. Then $\mathscr{H}$ is a filter base of closed sets. Let $\mathscr{H}^*$ be the filter of closed subsets which is generated by the filter base $\mathscr{H}$. Let

$$T = \{x \in B : F \subseteq f_B(x) \text{ for some } F \in \mathscr{H}^*\}$$

Since $\{f_B(x) : x \in B\}$ is a base of clopen sets in $\mathscr{S}(B)$ it is a base for closed sets in $\mathscr{S}(B)$ and so every closed set in $\mathscr{S}(B)$ is the intersection of elements in $\{f_B(x) : x \in B\}$. Then $\cap\{F : F \in \mathscr{H}^*\} = \cap\{f_B(x) : x \in T\}$. Suppose $x, y \in T$. We claim that $x \wedge y \in T$: There exists F_i and F_j in $\mathscr{H}^*$ such that $F_i \subseteq f_B(x)$ and $F_j \subseteq f_B(y)$. Then $F_i \cap F_j \subseteq f_B(x) \cap f_B(y) = f_B(x \wedge y)$. Since $F_i \cap F_j$ must be non-empty and belong to $\mathscr{H}^*$ then $x \wedge y \in T$, as claimed. From this we deduce that T is a filter in B. Now T extends to the B-ultrafilter $\mathscr{T}$. Since $x \in \mathscr{T}$ for all $x \in T$ then $\mathscr{T} \in \cap\{f_B(x) : x \in T\} = \cap\{F : F \in \mathscr{H}^*\}$. Then $\cap\{F : F \in \mathscr{H}\} \neq \varnothing$. We conclude that $\mathscr{S}(B)$ is compact, as required.

5) Let $A \in \mathscr{B}(\mathscr{S}(B))$, the Boolean algebra of all clopen sets in $\mathscr{S}(B)$. Since A is open then $A = \cup\{f_B(x) : x \in K\}$ for some $K \subseteq B$. Now A is also closed and so is compact in $\mathscr{S}(B)$. The collection $\{f_B(x) : x \in K\}$ is an open cover of A and so A has a finite cover. That is, there is a finite subset $M \subseteq K$ such that $A = \cup\{f_B(x) : x \in M\} = f_B(\vee(x \in M))$. We conclude that $A \in f_B[B]$. That is, f_B maps B onto $\mathscr{B}(\mathscr{S}(B))$, as required.

A.4 Characterizations of Martin's axiom.

We can now present and prove a few characterizations of Martin's axioms.

Theorem 0.13 Let κ be an infinite cardinal number such that $\kappa < 2^{\aleph_0}$. Then the following are equivalent:

1) (Martin's axiom MA) If $(P, \leq)$ is a partially ordered set satisfying *ccc* and $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$ is a family of dense subsets of P, then there exists a filter $\mathscr{F}$ on P such that $\mathscr{F} \cap D_\alpha \neq \varnothing$ for each $\alpha \leq \kappa$.

2) If X is a compact Hausdorff topological space satisfying *ccc* and $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$ is a family of dense open subsets of X, then $\cap\{D_\alpha : \alpha \leq \kappa\} \neq \varnothing$.

3) If $(B, \leq, \vee, \wedge, ')$ is a Boolean algebra with the *ccc* property and $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$ is a family of dense subsets of B, then there exists a filter $\mathscr{F}$ on B such that $\mathscr{F} \cap D_\alpha \neq \varnothing$ for each $\alpha \leq \kappa$.

4) If $(P, \leq)$ is a partially ordered set satisfying *ccc* and $|P| \leq \kappa$ and $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$ is a family of dense subsets of P, then there exist a filter $\mathscr{F}$ on P such that $\mathscr{F} \cap D_\alpha \neq \varnothing$ for each $\alpha \leq \kappa$.

Proof:

We are given that κ be an infinite cardinal number such that $\kappa < 2^{\aleph_0}$.

$(1 \Rightarrow 2)$: We begin with the trivial case. Suppose X is finite. If $X = \{x_1, x_2, \ldots, x_n\}$ then every element of X is clopen and so the only dense subset of X is X. Hence the intersection of all dense subsets of X is $X \neq \varnothing$. We are done.

Suppose now that X is an infinite set and that $\tau(X)$ denotes the set of all non-empty open subsets of X. Then $(\tau(X), \subseteq)$ is a partially ordered set of subsets of X. Suppose X does not contain an uncountable family of pairwise disjoint open subsets of X. That is, suppose $(\tau(X), \subseteq)$ satisfies the *ccc*. Suppose $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\} \subseteq \tau(X)$ where D_α is dense in X (i.e., $\mathrm{cl}(D_\alpha) = X$). For each $\alpha < \kappa$, we define $\mathscr{U}_\alpha = \{U \in \tau(X) : \mathrm{cl}(U_\alpha) \subseteq D_\alpha\}$. (Since X is compact Hausdorff and none of D_α's are empty, none of the $\mathscr{U}_\alpha$'s are empty.)

We claim that, for each α, $\mathscr{U}_\alpha$ is a dense subset of the partially ordered set $(\tau(X), \subseteq)$: Suppose $M \in \tau(X)$. It suffices to show that there is element of $\mathscr{U}_\alpha$ which is a subset of M. See that, for any $\alpha < \kappa$, $M \cap D_\alpha \in \tau(X)$ and so there exists an element x and open set U such that $x \in U \subseteq \mathrm{cl}(U) \subseteq M \cap D_\alpha \subseteq D_\alpha$. Then $U \in \mathscr{U}_\alpha$. We have shown that an element of $\mathscr{U}_\alpha$ is a subset of M. So $\mathscr{U}_\alpha$ is a dense subset of $(\tau(X), \subseteq)$ as claimed.

The set $\mathscr{E} = \{\mathscr{U}_\alpha : \alpha \leq \kappa\}$ is then a family of dense subsets of $(\tau(X), \subseteq)$ satisfying *ccc*. By Martin's axiom, $(\tau(X), \subseteq)$ contains a filter $\mathscr{F}$ such that $\mathscr{F} \cap \mathscr{U}_\alpha \neq \varnothing$ for all $\alpha \leq \kappa$. For each

α, choose $F_\alpha \in \mathscr{F} \cap \mathscr{U}_\alpha$. Since $F_\alpha \in \mathscr{U}_\alpha$, then $\mathrm{cl}(F_\alpha) \subseteq D_\alpha$. Since $\mathscr{F}$ is a filter of non-empty open subsets of X which satisfies the *finite intersection property*, then $\{\mathrm{cl}(F_\alpha) : \alpha \leq \kappa\}$ satisfies the finite intersection property inside compact X. Then there must be some $a \in X$ such that $a \in \cap\{\mathrm{cl}(F_\alpha) : \alpha \leq \kappa\} \subseteq \cap\{D_\alpha : \alpha \leq \kappa\}$. Thus $\cap\{D_\alpha : \alpha \leq \kappa\} \neq \varnothing$. This is what we were required to prove.

$(2 \Rightarrow 3)$: We are given that $(B, \leq, \vee, \wedge, ')$ is a Boolean algebra with the *ccc* property and $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$ is a family of dense subsets of B.[7] We are required to find a B-filter, $\mathscr{F}$, such that $\mathscr{F} \cap D_\alpha \neq \varnothing$ for all $\alpha \leq \kappa$. By the *Stone representation theorem* there exists a one-to-one isomorphism f_B which maps B onto $\mathscr{B}(\mathscr{S}(B))$, the set of all clopen subsets of the Stone space $\mathscr{S}(B)$ shown to be compact and zero-dimensional.

The Stone space $\mathscr{S}(B)$ satisfies the ccc: Suppose $\mathscr{A} = \{A_\alpha : \alpha \leq \mu, A_\alpha \in \tau(\mathscr{S}(B))\}$ be a strong antichain of non-empty open subsets of $\mathscr{S}(B)$ of cardinality μ. We claim that $\mu \leq \aleph_0$. For each α we can choose $f_B(a_\alpha) \subseteq A_\alpha$. Since the elements of $\mathscr{A}$ are pairwise disjoint and f_B is one-to-one, then the set $\{f_B(a_\alpha) : \alpha \leq \mu\}$ is of cardinality μ. Since f_B is an isomorphism then $\{a_\alpha : \alpha \leq \mu\}$ is an antichain in B of cardinality μ. Since B is *ccc* then $\mu \leq \aleph_0$. This establishes the claim.

For each $\alpha \leq \kappa$ choose $d_\alpha \in D_\alpha \subseteq B$. If $M_\alpha = \cup\{f_B(d_\alpha) : d_\alpha \in D_\alpha\}$ we see that M_α is an open subset of $\mathscr{S}(B)$ (since $\{f_B(x) : x \in B\}$ is an open base for the topology $\tau(\mathscr{S}(B))$ on $\mathscr{S}(B)$). Suppose u is an ultrafilter in $\mathscr{S}(B) - M_\alpha$ and f_B maps $w \in B$ to a basic open neighbourhood $f_B(w)$ of u. Then there exists $d \in D_\alpha$ such that $0 < d \leq w$; this implies $f_B(d) \subseteq f_B(w) \cap M_\alpha$. So every basic neighbourhood of u intersects M_α. Then M_α is dense in $\mathscr{S}(B)$. So $\{M_\alpha : \alpha \leq \kappa\}$ is collection of open dense subsets of the compact set $\mathscr{S}(B)$.

Our hypothesis guarantees that there exists $z \in \cap\{M_\alpha : \alpha \leq \kappa\}$. Let $\mathscr{B}_z$ denote the set of all basic neighbourhoods of $\mathscr{S}(B)$ which contain z. Now $\mathscr{B}_z$ satisfies the finite intersection property. Then the set $F_z = f_B^{\leftarrow}(B_z)$ is a B-filter.

We claim that the subset F_z in B is the B-filter we seek: Choose an arbitrary $\alpha_1 \leq \kappa$. Since $z \in M_{\alpha_1} = \cup\{f_B(d_\alpha) : d_\alpha \in D_{\alpha_1}\}$ then $z \in f_B(d_\alpha)$ for some α. This implies $d_\alpha \in F_z \cap D_{\alpha_1}$. Then the B-filter, $F_z = f_B^{\leftarrow}(B_z)$, intersects every dense set D_α in B. This concludes the proof.

$(3 \Rightarrow 4)$: We are given that $(P, \leq)$ is a partially ordered set such that $|P| \leq \kappa$ which satisfies *ccc*. Also suppose that the elements of $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$ are known to be dense subsets of P. We are required to show that, given our hypothesis, there exists a P-filter, $\mathscr{H}$, such that $\mathscr{H} \cap D_\alpha \neq \varnothing$ for all $\alpha \leq \kappa$. We will suppose, without loss of generality, that $\wedge P$ does not exist. (We can do this since, if we prove the existence of a P-filter, $\mathscr{H}$, in $P - \{\wedge P\}$ then $\mathscr{H}$ is a P-filter in P.). If $x \in P$ we define "$x^\downarrow$" as

$$x^\downarrow = \{y \in P : y \leq x\}$$

Consider the subset $\mathscr{B}_P = \{x^\downarrow : x \in P\}$. Note that if $a \in x^\downarrow \cap y^\downarrow$ then $a^\downarrow \subseteq x^\downarrow \cap y^\downarrow$, hence $\mathscr{B}_P$ forms an open base for some topology $\tau(P)$ on P. We have previously shown that

[7] Recall that "D_α is dense in B" means "if $0 < x \in B$ there exists $d \in D_\alpha$ such $0 < d \leq x$"

the set of all regular open subsets of P, $(\mathscr{R}o(P), \subseteq, \vee, \cap, P, \varnothing')$ forms a Boolean algebra in $(\tau(P), \subseteq, \cup, \cap, P, \varnothing')$

We claim that since P satisfies ccc then so does $\mathscr{R}o(P)$:[8] Let $\mathscr{A}$ be an antichain in $\mathscr{R}o(P)$ and $A \in \mathscr{A}$. Then $A = \text{int}_P\text{cl}_P(A)$. Since A is open in P then it is the union of base elements of the form $x^{\downarrow}$. We can then choose $a_A{}^{\downarrow} \subseteq A$ such that $\text{int}_P\text{cl}_P(a_A{}^{\downarrow}) \subseteq \text{int}_P\text{cl}_P(A) = A$. Then there is a well-defined one-to-one function $h : \mathscr{A} \to P$ such that $h(A) = a_A$ (Choice!). If $A, B \in \mathscr{A}$ then $A \cap B = \varnothing$ hence $h(A) \cap g(B) = \varnothing$. We conclude that $h[\mathscr{A}] = \{a_A \in P : A \in \mathscr{A}\}$ is an antichain in P. Since P satisfies *ccc* then $h[\mathscr{A}]$ must be countable. Since h is one-to-one, $\mathscr{A}$ must be countable. So $\mathscr{R}o(P)$ satisfies *ccc*, as claimed.

We increase the size of the set $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$ of dense subsets of P, as follows. Let

$$\mathscr{D}^* = \{D_\alpha : \alpha \leq \kappa\} \cup \{D_{(x,y)} : x, y \in P\}$$

where each $D_{(x,y)}$ is a dense subset of P previously defined after definition 32.2.[9] So every element of $\mathscr{D}^*$ is a dense subset of P. Since $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\} \Rightarrow |\mathscr{D}| \leq \kappa$ and $|P| \leq \kappa \Rightarrow |\cup\{D_{(x,y)} : x, y \in P\}| \leq \kappa$ then $|\mathscr{D}^*| \leq \kappa$ (by theorem 23.4 and 29.8). We define the function $g : P \to \mathscr{R}o(P)$ as $g(x) = \text{int}_P\text{cl}_P(x^{\downarrow})$. For $\alpha \leq \kappa$ and $x, y \in P$ we define

$$\begin{aligned} A_\alpha &= g[D_\alpha] = \{\text{int}_P\text{cl}_P(x^{\downarrow}) : x \in D_\alpha\} \\ A_{(x,y)} &= g[D_{(x,y)}] = \{\text{int}_P\text{cl}_P(x^{\downarrow}) : x \in D_{(x,y)}\} \end{aligned}$$

Let $\mathscr{A} = \{A_\alpha : \alpha \leq \kappa\} \cup \{A_{(x,y)} : x, y \in P\} \subseteq \mathscr{R}o(P)$.[10] Let A be an arbitrary element in $\mathscr{A}$.

We claim that A is dense in $\mathscr{R}o(P)$. We first consider the case where A is of the form A_α. To see this, let $H \in \tau(P) - \{\varnothing\}$ such that $H = \text{int}_P\text{cl}_P H \in \mathscr{R}o(P)$. Then there exists $t \in P$ and a basic element $t^{\downarrow} \in \tau(P)$ such that $t \in t^{\downarrow} \subseteq H$. Since D_α is dense in P there exists $d \in D_\alpha$ such that $d \leq t$. So $\text{int}_P\text{cl}_P(d^{\downarrow}) \subseteq \text{int}_P\text{cl}_P(t^{\downarrow}) \subseteq \text{int}_P\text{cl}_P H = H$. We have found an element $\text{int}_P\text{cl}_P(d^{\downarrow})$ in A such that $\text{int}_P\text{cl}_P(d^{\downarrow}) \subseteq H$. So A is dense in $\mathscr{R}o(P)$. If A is of the form $A_{(x,y)}$ the proof that A is dense proceeds identically. So $\mathscr{F}$ intersects every set in $\mathscr{A}$.

By hypothesis, $\mathscr{R}o(P)$ contains a filter, $\mathscr{F}$, such that $\mathscr{F} \cap A \neq \varnothing$ for all $A \in \mathscr{A}$. Let $g : P \to \mathscr{R}o(P)$ be defined as above. That is, $g(x) = \text{int}_P\text{cl}_P(x^{\downarrow})$ and let

$$\mathscr{H} = g^{\leftarrow}[\mathscr{F}] = \{x \in P : \ \text{int}_P\text{cl}_P(x^{\downarrow}) \in \mathscr{F}\}$$

The set $\mathscr{H}$ intersects all elements of $\mathscr{D}^*$: Consider $D_\beta \in \mathscr{D}^*$. Then $g[D_\beta] = A_\beta \in \mathscr{A}$. There must exist $u \in \mathscr{F} \cap A_\beta$. Then $g^{\leftarrow}(u) \in D_\beta \cap \mathscr{H}$. We now consider $D_{(x,y)} \in \mathscr{D}^*$ for some x, y in P. Then $g[D_{(x,y)}] = A_{(x,y)} \in \mathscr{A}$. There must exist $u \in \mathscr{F} \cap A_{(x,y)}$. Then $g^{\leftarrow}(u) \in D_{(x,y)} \cap \mathscr{H}$. Then $\mathscr{H}$ meets all elements of $\mathscr{D}^*$, as claimed.

---

[8] The reader is left to verify the following facts: 1) If $A \subseteq B$ then $\text{int}_X\text{cl}_X(A) \subseteq \text{int}_X\text{cl}_X(B)$ 2) $\text{int}_X\text{cl}_X(A) \cap \text{int}_X\text{cl}_X(B) = \text{int}_X\text{cl}_X(A \cap B)$

[9] The subset $D_{(x,y)}$ is previously defined in example 2 following the definition 32.2 as: $D_{(x,y)} = V_{(x,y)} \cup W_{(x,y)}$ where $V_{(x,y)} = \{z \in P : z \text{ is not compatible with } x \text{ or with } y\}$ and $W_{(x,y)} = x^{\downarrow} \cap y^{\downarrow}$.

[10] The reader is left to verify that, if A is a subset of X then $\text{int}_X\text{cl}_X A$ is regular open in X.

We claim that $\mathscr{H}$ is a proper P-filter: The set $\mathscr{H}$ is non-empty since $\mathscr{F} \cap A_\alpha \neq \varnothing$. The set $\mathscr{H}$ does not contain a minimum element since P was assumed not to have one, hence $\mathscr{H} \neq P$. Suppose $x \in \mathscr{H}$ and $x \leq y$. Then $g(x) = \text{int}_P\text{cl}_P(x^\downarrow) \in \mathscr{F}$ and, since $\mathscr{F}$ is a $\mathscr{R}o(P)$-filter and $g(x) \subseteq g(y) = \text{int}_P\text{cl}_P(y^\downarrow)$, then $g(y) \subseteq \mathscr{F}$, hence $y \in \mathscr{H}$.
Let $x,\ y \in \mathscr{H}$. We are now required to find $q \in \mathscr{H}$ such that $q \leq x$ and $q \leq y$. Since $x,\ y \in \mathscr{H}$ then $g(x),\ g(y) \in \mathscr{F}$. Since $D_{(x,y)} \cap \mathscr{H} \neq \varnothing$ there exists $s \in D_{(x,y)} \cap \mathscr{H}$. Then $g(s) \in \mathscr{F}$. It is not possible that $s^\downarrow \cap x^\downarrow = \varnothing$, for if it was, we would have $\varnothing = g(s^\downarrow \cap x^\downarrow) = g(s^\downarrow) \cap g(x^\downarrow) \in \mathscr{F}$, a contradiction. Also it is impossible that $s^\downarrow \cap y^\downarrow = \varnothing$ for this would not allow $g(s)$ and $g(y)$ to both belong to $\mathscr{F}$. Then $s \in x^\downarrow \cap y^\downarrow$. Then $s \leq x$ and $s \leq y$. We conclude that $\mathscr{H}$ is a P-filter.

This concludes the proof of 3) $\Rightarrow$ 4).

(4 $\Rightarrow$ 1): Suppose $(P \leq)$ be a partially ordered set satisfying *ccc*. Let $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$ be a set of dense subsets of P. We will assume that P does not have a minimum element. Since D_α is dense in P, for each $x \in P$ there exists $q \in D_\alpha$ such that $q \leq x$; hence $x^\downarrow \cap D_\alpha \neq \varnothing$ for all $x \in P$. For each α we define a function $g_\alpha : P \to D_\alpha$ as

$$g_\alpha(x) = x_\alpha \text{ where } x_\alpha \in x^\downarrow \cap D_\alpha \quad \text{(Choice!)}$$

We also define the function $g : P \times P \to P$ as

$$g(x,y) = z \text{ where } z \in x^\downarrow \cap y^\downarrow, \quad g(x,y) \text{ arbitrary if } x^\downarrow \cap y^\downarrow \neq \varnothing \quad \text{(Choice!)}$$

Let $T_0 = \{k\}$ for some $k \in P$. We inductively define the elements of $\{T_n : n < \aleph_0\}$ as follows:

$$T_{n+1} = T_n \cup g[T_n \times T_n] \ \cup \ [\cup\{g_\alpha[T_n] : \alpha \leq \kappa\}]$$

and let $T = \cup\{T_n : n < \aleph_0\}$. Since $T \subseteq P$, T inherits "$\leq$" from P to form a partially ordered set $(T, \leq)$. Also note that since $|\{T_n : n < \aleph_0\}| < \aleph_0$ and $|T_n| \leq \kappa$ then $|T| = |\cup\{T_n : n \leq \aleph_0\}| \leq \aleph_0 \times \kappa = \kappa$. Also see that $g_\alpha[T_n] \subseteq T_{n+1} \subseteq T$ for all n and so $g_\alpha[T] \subseteq T$. Similarly, $g[T \times T] \subseteq T$.

Let $A = \{a_\alpha : \alpha \leq \lambda\}$ be an antichain in $T \subseteq P$. Then ${a_\alpha}^\downarrow \cap {a_\beta}^\downarrow = \varnothing$ in T if $\alpha \neq \beta$. If $p \in P - T$ such that $p \in {a_\alpha}^\downarrow \cap {a_\beta}^\downarrow$ then $g(a_\alpha, a_\beta) = k \in {a_\alpha}^\downarrow \cap {a_\beta}^\downarrow$. Since $k \in T$ then ${a_\alpha}^\downarrow \cap {a_\beta}^\downarrow \neq \varnothing$ in T, a contradiction. Then ${a_\alpha}^\downarrow \cap {a_\beta}^\downarrow = \varnothing$ in P. We conclude that A is an antichain in P. Since P satisfies *ccc* then $\mathscr{A}$ cannot be uncountable. Therefore T satisfies *ccc*.

We claim that, for any given α, $T \cap D_\alpha$ is dense in T: Let $t \in T$ and $\alpha \leq \kappa$. Since D_α is dense in P there exists $z \in D_\alpha$ such that $z \leq t \in T$. Then $g_\alpha(z) = z_\alpha \in z^\downarrow \cap D_\alpha$. Then $z_\alpha \in T$ and $z_\alpha \leq z \leq t$. Since $z_\alpha \in T \cap D_\alpha$ we can conclude that $T \cap D_\alpha$ is dense in T, as claimed.

By hypothesis, there exists a T-filter, F, such that $F \cap (T \cap D_\alpha) \neq \varnothing$ for all $\alpha \leq \kappa$. If $q \in P$ let $\{q\}^\uparrow = \{x \in P : x \geq q\}$ be the principal P-filter generated by q. We let

$F^* = \cup\{\{q\}^\uparrow : q \in F\}$. It easily verified that F^* is a P-filter which contains F. Since $F \subseteq F^*$ and F intersects every D_α then so does F^*. Then F^* is P-filter we were required to find.

We finally have the following topological statement which is equivalent to Martin's axiom.

Theorem 0.14 Let κ be a cardinal such that $\aleph_0 \leq \kappa < 2^{\aleph_0}$. Let X be a Hausdorff topological space satisfying *ccc* such that $\{x \in X : x \text{ has a compact neighbourhood}\}$ is dense in X. Suppose that $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$ is a family of dense open subsets of X. Then $\cap\{D_\alpha : \alpha \leq \kappa\}$ is dense in X if and only if Martin's axiom holds true.

Proof:

$(\Rightarrow)$ Let X be a compact Hausdorff space which satisfies *ccc* where X contains a family of dense open subsets of X, $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$. Since X is compact every point $x \in X$ contains a compact neighbourhood. By hypothesis, $\cap\{D_\alpha : \alpha \leq \kappa\}$ is dense in X, so $\cap\{D_\alpha : \alpha \leq \kappa\}$ is not empty. Then by 2) $\Leftrightarrow$ 1) in the previous theorem, Martin's axiom holds true.

$(\Leftarrow)$ Suppose Martin's axiom holds true. Let X be a Hausdorff topological space satisfying *ccc* such that $T = \{x \in X : x \text{ has a compact neighbourhood}\}$ is dense in X. Suppose that $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$ is a family of dense open subsets of X (where $\aleph_0 \leq \kappa < 2^{\aleph_0}$). We are required to show that $\cap\{D_\alpha : \alpha \leq \kappa\}$ is dense in X. For any non-empty open subset U there exists a point $x \in U \cap T$ and some open neighbourhood S of x with compact closure, $\mathrm{cl}_X S$, such that $x \in \mathrm{cl}_X(S \cap U) \subseteq \mathrm{cl}_X S$. Since X satisfies *ccc* its compact subset $\mathrm{cl}_X(S \cap U)$ must also satisfy *ccc*. For any $D_\alpha \in \mathscr{D}$, $D_\alpha \cap U \cap S$ is open and dense in $\mathrm{cl}_X(S \cap U)$. By the topological equivalent form of MA, there exists $q \in \cap\{D_\alpha \cap U \cap S : \alpha \leq \kappa\}$. Since $\cap\{D_\alpha \cap U \cap S : \alpha \leq \kappa\} \subseteq U \cap (\cap\{D_\alpha : \alpha \leq \kappa\})$ then $q \in \cap\{D_\alpha : \alpha \leq \kappa\} \cap U$. Not only is $\cap\{D_\alpha : \alpha \leq \kappa\}$ non-empty but it also intersects every open subset U of X. So $\cap\{D_\alpha : \alpha \leq \kappa\}$ is dense in X.

Appendix B : List of definitions and statements.

I Axioms and classes : 1 / Classes, sets and axioms

Axiom A1 (Axiom of extent) : For the classes x, A and B, $[A = B] \Leftrightarrow [x \in A \Leftrightarrow x \in B]$

Axiom A2 (Axiom of class construction): Let $P(x)$ designate a statement about x which can be expressed entirely in terms of the symbols $\in, \vee, \wedge, \neg, \Rightarrow, \forall$, brackets and variables x, y, z, ..., A, B, ... Then there exists a class C which consists of all the elements x which satisfy $P(x)$.

Axiom A3 (Axiom of pair) : If A and B are sets , then the doubleton $\{A, B\}$ is a set.

Axiom A4 (Axiom of subsets) : If S is a set and ϕ is a formula describing a particular property, then the class of all sets in S which satisfy this property ϕ is a set. More succinctly, every subclass of a set of sets is a set. (Also called the Axiom of comprehension, Axiom of separation or Axiom of specification).

Axiom A5 (Axiom of power set): If A is a set then the power set $P(A)$ is a set.

Axiom A6 (Axiom of union): If $\mathscr{A}$ is a set of sets then $\bigcup_{C \in \mathscr{A}} C$ is a set.

Axiom A7 (Axiom of replacement): Let A be a set. Let $\phi(x, y)$ be a formula which associates to each element x of A a set y in such a way that, whenever both $\phi(x, y)$ and $\phi(x, z)$ hold true, $y = z$. Then there exists a set B which contains all sets y such that $\phi(x, y)$ holds true for some $x \in A$.

Axiom A8 (Axiom of infinity): There exists a non-empty set A that satisfies the condition: "$X \in A$" $\Rightarrow$ "$X \cup \{X\} \in A$". (A set satisfying this condition is called a *successor set* or an *inductive set*.)

Axiom A9 (Axiom of regularity) Every non-empty set A contains an element x whose intersection with A is empty.

Axiom of choice : For every set $\mathscr{A}$ of non-empty sets there is a rule f which associates to every set A in $\mathscr{A}$ an element $a \in A$.

I Axioms and classes : 2 / Constructing classes and sets

Theorem 2.1 For any class C, $C = C$. If it is not true that $A = B$ we will write $A \neq B$.

Definition 2.2 If A and B are classes (sets) we define $A \subseteq B$ to mean that every element of A is an element of B. That is, $A \subseteq B$ iff $x \in A \Rightarrow x \in B$ If $A \subseteq B$ we will say that A is a *subclass* (*subset*) of B. If $A \subseteq B$ and $A \neq B$ we will say that A is a *proper subclass* (*proper subset*) of B and write $A \subset B$ when we explicitly want to say $A \neq B$.

Theorem 2.3 If C, D, and E are classes (sets) then:

a) $C = C$.
b) $C = D \Rightarrow D = C$.
c) $C = D$ and $D = E \Rightarrow C = D$.
d) $C \subseteq D$ and $D \subseteq C \Rightarrow C = D$.
e) $C \subseteq D$ and $D \subseteq E \Rightarrow C \subseteq E$.

Theorem 2.4 There exists a class which is not an *element*.

Definition 2.5 The Axiom 2 states that $C = \{x : x \neq x\}$ is a class. It contains no elements. We will call the class with no elements the *empty class* and denote it by $\varnothing$.

Theorem 2.6 For any class C, $\varnothing \subseteq C$.

Theorem 2.7 Let S be a set. Then:

a) $\varnothing \subseteq S$ and so $\varnothing$ is a set.
b) The set S is an element. Hence all sets are elements.

Definition 2.8 If A is a set then we define the *power set of* A as being the class $\mathscr{P}(A)$ of all subsets of A. It can be described as follows: $\mathscr{P}(A) = \{X : X \subseteq A\}$.

II Class operations 3 / Operations on classes and sets

Definition 3.1 Let A and B be classes (sets). We define the *union* $A \cup B$ of the class A and the class B as

$$A \cup B = \{x : (x \in A) \vee (x \in B)\}$$

That is, the element $x \in A \cup B$ iff $x \in A$ or $x \in B$. If $\mathscr{A}$ is a non-empty class of classes then we define the *union of all classes in* $\mathscr{A}$ as

$$\bigcup_{C \in \mathscr{A}} C = \{x : x \in C \text{ for some } C \in \mathscr{A}\}$$

That is, the element $x \in \bigcup_{C \in \mathscr{A}} C$ iff there exists $C \in \mathscr{A}$ such that $x \in C$.

Definition 3.2 Let A and B be classes (sets). We define the *intersection* $A \cap B$ of the class A and the class B as

$$A \cap B = \{x : (x \in A) \wedge (x \in B)\}$$

That is, the element $x \in A \cap B$ iff $x \in A$ and $x \in B$. If $\mathscr{A}$ is a non-empty class of classes then we define the *intersection of all classes in* $\mathscr{A}$ as

$$\bigcap_{C \in \mathscr{A}} C = \{x : x \in C \text{ for all } C \in \mathscr{A}\}$$

That is, the element $x \in \bigcap_{C \in \mathscr{A}} C$ iff $x \in C$ for every class C in $\mathscr{A}$.

Definition 3.3 We will say that two classes (sets) C and D are *disjoint* if the two classes have no elements in common. That is, the classes C and D are disjoint if and only if $C \cap D = \varnothing$.

Definition 3.4 The *complement*, C', of a class (set) C is the class of all elements which are not in C. That is, if C is a class, then

$$C' = \{x : x \notin C\}$$

Hence $x \in C'$ iff $x \notin C$. Given two classes (sets) C and D, the *difference* $C - D$, of C and D, is the class

$$C - D = C \cap D'$$

The *symmetric difference*, $C \triangle D$, is the class

$$C \triangle D = (C - D) \cup (D - C)$$

Theorem 3.5 Let C and D be classes (sets). Then,

a) $C \subseteq C \cup D$
b) $C \cap D \subseteq C$

Theorem 3.6 Let C and D be classes (sets). Then,

a) $C \cup (C \cap D) = C$
b) $C \cap (C \cup D) = C$

Theorem 3.7 Let C be a class (a set). Then $(C')' = C$.

Theorem 3.8 *DeMorgan's laws.* Let C and D be classes (sets). Then,

a) $(C \cup D)' = C' \cap D'$
b) $(C \cap D)' = C' \cup D'$

Theorem 3.9 Let C, D and E be classes (sets). Then,

a) $C \cup D = D \cup C$ and $C \cap D = D \cap C$ (*Commutative laws*)
b) $C \cup C = C$ and $C \cap C = C$ (*Idempotent laws*)
c) $C \cup (D \cup E) = (C \cup D) \cup E$ and $C \cap (D \cap E) = (C \cap D) \cap E$ (*Associative laws*)
d) $C \cup (D \cap E) = (C \cup D) \cap (C \cup E)$ and $C \cap (D \cup E) = (C \cap D) \cup (C \cap E)$ (*Distribution*)

Theorem 3.10 Let A be a class (a set) and $\mathscr{U}$ denote the class of all elements.

a) $\mathscr{U} \cup A = \mathscr{U}$
b) $A \cap \mathscr{U} = A$
c) $\mathscr{U}' = \varnothing$
d) $\varnothing' = \mathscr{U}$
e) $A \cup A' = \mathscr{U}$

Theorem 3.11 Let $\mathscr{A}$ be a non-empty class (set).

a) $\left(\bigcup_{C\in\mathscr{A}} C\right)' = \bigcap_{C\in\mathscr{A}} C'$

b) $\left(\bigcap_{C\in\mathscr{A}} C\right)' = \bigcup_{C\in\mathscr{A}} C'$

Theorem 3.12 Let D be a class and $\mathscr{A}$ be a non-empty class (set) of classes.

a) $D \cap \left(\bigcup_{C\in\mathscr{A}} C\right) = \bigcup_{C\in\mathscr{A}} (D \cap C)$

b) $D \cup \left(\bigcap_{C\in\mathscr{A}} C\right) = \bigcap_{C\in\mathscr{A}} (D \cup C)$

Theorem 3.13 Let $\{B_{(i,j)} : i = 1, 2, 3, \ldots,\ j = 1, 2, 3, \ldots\}$ be a set of sets Then $\cup_{i=1}^{\infty}(\cap_{j=1}^{\infty} B_{(i,j)}) = \cap_{j=1}^{\infty}(\cup_{i=1}^{\infty} B_{(i,j)})$.

II Class operations 4 / Cartesian products

Definition 4.1 Let c and d be elements. We define the *ordered pair* (c, d) as $(c, d) = \{\{c\}, \{c, d\}\}$.

Theorem 4.2 Let a, b, c and d be classes (which are elements). Then $(a, b) = (c, d)$ iff $a = c$ and $b = d$.

Alternate definition 4.3 If c and d are classes define (c, d) as follows: $(c, d) = \{\ \{c, \varnothing\}, \{d, \{\varnothing\}\}\ \}$.

Definition 4.4 Let C and D be two classes (sets). We define the *Cartesian product*, $C \times D$, as follows: $C \times D = \{(c, d) : c \in C \text{ and } d \in D\}$.

Lemma 4.5 Let C and D be two classes (sets). Then the *Cartesian product*, $C \times D$, of C and D satisfies the property: $C \times D \subseteq \mathscr{P}(\mathscr{P}(C \cup D))$.

Corollary 4.6 If C and D are classes, then the *Cartesian product*, $C \times D$, is a class. If C and D are sets, then $C \times D$ is a set.

Theorem 4.7 Let C, D, E and F be a classes. Then

a) $C \times (D \cap E) = (C \times D) \cap (C \times E)$

b) $C \times (D \cup E) = (C \times D) \cup (C \times E)$

c) $(C \cap E) \times D = (C \times D) \cap (E \times D)$

d) $(C \cup E) \times D = (C \times D) \cup (E \times D)$

e) $(C \cup D) \times (E \cup F) = (C \times E) \cup (D \times E) \cup (C \times F) \cup (D \times F)$

f) $(C \cap D) \times (E \cap F) = (C \times E) \cap (D \times E) \cap (C \times F) \cap (D \times F)$

Theorem 4.8 If $C \subseteq D$ and $E \subseteq F$, then $C \times E \subseteq D \times F$.

Theorem 4.9 Given three classes (sets) S, U and V there is a one-to-one correspondence between the two classes (sets) $S \times (U \times V)$ and $(S \times U) \times V$.

Theorem 4.10 For classes c, d, e and f, if $(c,d) = \{\{c, \varnothing\}, \{d, \{\varnothing\}\}\}$ and $(e,f) = \{\{e, \varnothing\}, \{f, \{\varnothing\}\}\}$ then $(c,d) = (e,f)$ iff $c = e$ and $d = f$.

III Relations 5 / Relations on a class or set

Definition 5.1 a) We will call any subset R of ordered pairs in $\mathscr{U} \times \mathscr{U}$ a *binary relation.*

b) We will say that R is *binary relation on a class* C if R is a subclass (subset) of $C \times C$. In such cases we will simply say that R is a *relation in* C or *on* C.

c) If A and B are classes (sets) and R is a subclass (subset) of $A \times B$ then R can be viewed as a relation on $A \cup B$.

Definitions 5.2 Let C be a class (a set). a) The relation $\in_C = \{(x,y) : x \in C, y \in C, x \in y\}$ is called the *membership relation on* C. b) The relation

$$Id_C = \{(x,y) : x \in C, y \in C, x = y\}$$

is called the *identity relation on* C.

Definitions 5.3 Let R be a relation on a class (set) C. The *domain of* R is the class, $\operatorname{dom} R = \{x : x \in C \text{ and } (x,y) \in R \text{ for some } y \in C\}$. The *image of* R is the class, $\operatorname{im} R = \{y : y \in C \text{ and } (x,y) \in R \text{ for some } x \in C\}$. The word *range of* R is often used instead of "the image of R". If $R \subseteq A \times B$ is viewed as a relation on $A \cup B$, then $\operatorname{dom} R \subseteq A$ and $\operatorname{im} R \subseteq B$.

Definition 5.4 Let C be a class (a set) and let R be a relation defined in C. The inverse, R^{-1}, of the relation R is defined as follows:

$$R^{-1} = \{(x,y) : (y,x) \in R\}$$

Definition 5.5 Let C be a class (a set) and let R and T be two relations in C. We define the relation $T \circ R$ as follows: $T \circ R = \{(x,y) :$ there exists some $z \in \operatorname{im} R$ such that $(x,z) \in R$ and $(z,y) \in T\}$

III Relations 6 / Equivalence relations and order relations.

Definition 6.1 Let S be a class and R be a relation on S.

a) We say that R is a *reflexive* relation on S if, for every $x \in S$, $(x,x) \in R$.

b) We say that R is a *symmetric* relation on S if, whenever $(x,y) \in R$ then $(y,x) \in R$.

c) We say that R is an *anti-symmetric* relation on S if, whenever $(x,y) \in R$ and $(y,x) \in R$ then $x = y$.

d) We say that R is an *asymmetric* relation on S if, whenever $(x,y) \in R$ then $(y,x) \notin R$.

e) We say that R is a *transitive* relation on S if, whenever $(x,y) \in R$ and $(y,z) \in R$ then $(x,z) \in R$.

f) We say that R is an *irreflexive* relation on S if, for every $x \in S$, $(x, x) \notin R$.

g) We say that R satisfies the property of *comparability* on S if, for every $x, y \in S$ where $x \neq y$, either $(x, y) \in R$ or $(y, x) \in R$.

Definition 6.2 Let S be a class and R be a relation on S. We say that R is an *equivalence relation* on S if R is simultaneously 1) reflexive, 2) symmetric and 3) transitive on S.

Definition 6.3 Let S be a class.

a) *Non-strict order relation.* The relation R is a non-strict order relation on S if it is simultaneously reflexive (aRa holds true for any a in S), antisymmetric (if aRb and bRa then $a = b$) and transitive (aRb and bRc implies aRc) on S. A non-strict order relation, R, on S is said to be a *non-strict linear order relation* if, for every pair of elements a and b in S, either $(a, b) \in R$, $(b, a) \in R$ or $a = b$. That is, every pair of elements are comparable under R.[11] A non-strict ordering, R, on S which is not linear is said to be a *non-strict partial ordering relation* on S.[12]

b) *Strict order relation.* The relation R is a strict order relation if it is simultaneously irreflexive ($(a, a) \notin R$), asymmetric ($(a, b) \in R \Rightarrow (b, a) \notin R$) and transitive on S. If every pair of distinct elements, a and b, in S are comparable under a strict order relation, R, then R is a *strict linear ordering* on S. Those strict orderings which are not linear are called *strict non-linear orderings* or, more commonly, *strict partial ordering relation.*
A non-strict partial order R on S always induces a strict partial order R^* by defining $aR^*b \Rightarrow [aRb \text{ and } a \neq b]$. Similarly, a strict partial order R on S always induces a non-strict partial order $R^\dagger$ by defining $aR^\dagger b \Rightarrow [aRb \text{ or } a = b]$.

Definition 6.4 Let S be a class and R be is a partial ordering or a strict ordering relation on S. If R is a partial ordering relation $(a, b) \in R$ is represented as $a \leq b$ and if R is a strict ordering relation $(a, b) \in R$ is represented as $a < b$. If $a \leq b$ and $a \neq b$, we will simply write $a < b$.

a) A subset of S which is linearly ordered by R is called a *chain* in S. If R linearly orders S then S is a linearly ordered subset of itself and so is a chain.

b) An element a of S is called a *maximal* element of S if there does not exist an element b in S such that $a < b$. An element a of S is called a *minimal* element of S if there does not exist an element b in S such that $b < a$.

c) An element m in S is called the *minimum element of* S if $m \leq a$ ($m < a$) for all $a \in S$. An element M in S is called the *maximum element of* S if $a \leq M$

[11] A class on which is defined a linear ordering R is also said to be *fully ordered* or *totally ordered* by R. In certain branches of mathematics "linearly ordered set" is abbreviated as *l.o.set* or simply called *loset.*

[12] In certain branches of mathematics "partially ordered set" is abbreviated as *p.o.set* or simply called a *poset*

$(a < M)$ for all $a \in S$.

III Relations 7 / The partition of a set induced by an equivalence relation.

Notation 7.1 Let R be an equivalence relation on a set S and let $x \in S$. Then the set S_x is defined as follows: $S_x = \{y : (x, y) \in R\}$. That is, S_x is the set of all elements y in S such that y is related to x under R.

Theorem 7.2 Let R be an equivalence relation on a set S. Let x and y be two elements in S which are not related under R. Then any element z in S which is related to x cannot be related to y.

Theorem 7.3 Let R be an equivalence relation on a set S. Let x and y be two elements in S which are not related under R. Then $S_x \cap S_y = \varnothing$.

Theorem 7.4 Let R be an equivalence relation on a set S. Let x and y be two elements in S which are related under R. Then $S_x = S_y$.

Theorem 7.5 Let R be an equivalence relation on a set S. For every $x \in S$ there exists some $y \in S$ such that $x \in S_y$.

Theorem 7.6 Let R be an equivalence relation on a set S. Then $\bigcup_{x \in S} S_x = S$.

III Relations 8 / On partitions and quotient sets of a set.

Definition 8.1 Let S be a set. We say that a set $\mathscr{C} \subseteq \mathscr{P}(S)$ forms a *partition of* S if $\mathscr{C}$ satisfies the 3 properties:

1) $\bigcup_{A \in \mathscr{C}} A = S$
2) If A and $B \in \mathscr{C}$ and $A \neq B$ then $A \cap B = \varnothing$.
3) $A \neq \varnothing$ for all $A \in \mathscr{C}$.

Definition 8.2 Let S be a set on which an equivalence relation R is defined.

a) Each element S_x of $\mathscr{S}_R = \{S_x : x \in S\}$ is called an *equivalence class of x under R* or an *equivalence class induced by the relation R.*
b) The set $\mathscr{S}_R = \{S_x : x \in S\}$ of all equivalence classes induced by the relation R is called the *quotient set of S induced by R.* The set $\mathscr{S}_R$ is more commonly represented by the symbol S/R. So $S/R = \{S_x : x \in S\}$. From here on we will use the more common notation, S/R.

Theorem 8.3 Let S be a set and $\mathscr{C}$ be a partition of S. Let $R_{\mathscr{C}}$ be the relation such that $(x, y) \in R_{\mathscr{C}}$ iff $\{x, y\} \subseteq S$ for some S in $\mathscr{C}$. Then $R_{\mathscr{C}}$ is an equivalence relation on S.

IV Functions 9 / Functions: A set-theoretic definition.

Definition 9.1 A function from A to B is a triple $\langle f, A, B\rangle$ satisfying the following properties:

1) A and B are classes and $f \subseteq A \times B$
2) For every $a \in A$ there exists $b \in B$ such that $(a, b) \in f$.
3) If $(a, b) \in f$ and $(a, c) \in f$ then $a = c$.

Definition 9.2 If $f : A \to B$ is a function and $D \subseteq A$ then we say that the function $f : D \to C$ is *a restriction of f to D*. In this case we will use the symbol $f|_D$ to represent the restriction of f to D. Note that, if $D \subseteq A$, then we can write $f|_D \subseteq f$ since $f|_D = \{(x, y) : x \in D \text{ and } (x, y) \in f\} \subseteq f$.

Theorem 9.3 Let $f : A \to B$ be a function and suppose $A = C \cup D$. Then $f = f|_c \cup f|_D$.

Theorem 9.4 Two functions $f : A \to B$ and $g : A \to B$ are equal if and only if $f(x) = g(x)$ for all $x \in A$.

Definitions 9.5 Let $f : A \to B$ be a function.

a) We say that "f maps A *onto* B" if $\operatorname{im} f = B$. We often use the expression "$f : A \to B$ is *surjective*" instead of the words *onto B*.
b) We say that "f maps A *one-to-one into* B" if whenever $f(x) = f(y)$ then $x = y$. We often use the expression "$f : A \to B$ is *injective*" instead of the words *one-to-one into B*.
c) If the function $f : A \to B$ is both one-to-one and onto B then we can simply say that f is "one-to-one and onto". Another way of conveying this is to say that f is *bijective*, or f is a *bijection*. So "injective + surjective $\Rightarrow$ bijective".
d) Two classes (or sets) A and B for which there exists some bijective function $f : A \to B$ are said to be in *one-to-one correspondence*.

IV Functions 10 / Compositions of function.

Definition 10.1 Suppose $f : A \to B$ and $g : B \to C$ are two functions such that the image of the function f is contained in the domain of the function g. Let $h = \{(x, z) : y = f(x) \text{ and } z = g(y) = g(f(x))\ \}$. Thus $(x, z) \in h$ if and only if $(x, z) = (x, g(f(x))$. We will call h the *composition of g and f*, and denote it as $g \circ f$.

Theorem 10.2 Let $f : A \to B$ and $g : B \to C$ be two functions such that the image of the function f is contained in the domain of the function g. Then the composition of g and f, $(g \circ f) : A \to C$, is a function.

Theorem 10.3 Let $f : A \to B$, $g : B \to C$ and $h : C \to D$ be three functions. Then $h \circ (g \circ f) = (h \circ g) \circ f$.

Theorem 10.4 Let $f : A \to B$. Then $I_B \circ f = f$ and $f \circ I_A = f$.

Definition 10.5 Let $f : A \to B$. If $g : B \to A$ is a function satisfying $g \circ f = I_A$ then we will call g an "*inverse of* f" and denote it as f^{-1}.

Theorem 10.6 Let $f : A \to B$ be a one-to-one onto function.

a) An inverse function $f^{-1} : B \to A$ of f exists.
b) The function f^{-1} is one-to-one and onto.
c) The function $f^{-1} : B \to A$ satisfies the property $f \circ f^{-1} = I_B$.
d) The inverse function, f^{-1}, of f is unique.

Definition 10.7 A function $f : A \to B$ which is one-to-one and onto is called an *invertible function.*

Theorem 10.8 Let $f : A \to B$ and $g : B \to C$ be two onto-to-one and onto functions.

a) The function $g \circ f$ is also one-to-one and onto.
b) The inverse, $g \circ f$, is $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$

IV Functions 11 / Images and inverse images of sets.

Definition 11.1 Let A and B be sets and suppose $f : A \to B$ is a function acting on A. If S is a subset of $A = \operatorname{dom} f$ we define the expression $f[S]$ as follows: $f[S] = \{y : y = f(x) \text{ for some } x \in S\}$. We will say that $f[S]$ is the *image of the set* S *under* f.

Definitions 11.2 Let $f : A \to B$ be a function where A and B are sets. We define $f^{\leftarrow} : \mathscr{P}(B) \to \mathscr{P}(A)$ as $f^{\leftarrow}(X) = Y$ iff $Y = \{y : y \in A, f(y) \in X\}$. In particular, $f^{\leftarrow}(\{x\}) = Y$ iff $Y = \{y : y \in A, f(y) = x\}$. We will refer to it as the *set-valued inverse function* $f^{\leftarrow}$.

Theorem 11.3 Let $f : A \to B$ be a function. Then $f^{\leftarrow} : \mathscr{P}(\operatorname{im} f) \to \mathscr{P}(A)$ is a one-to-one function on its domain $\mathscr{P}(\operatorname{im} f)$.

Theorem 11.4 Let $f : A \to B$ be a function mapping the set A to the set B. Let $\mathscr{A}$ be a set of subsets of A and $\mathscr{B}$ be a set of subsets of B. Let $D \subseteq A$ and $E \subseteq B$. Then:

a) $f\left[\bigcup_{S\in\mathscr{A}} S\right] = \bigcup_{S\in\mathscr{A}} f[S]$
b) $f\left[\bigcap_{S\in\mathscr{A}} S\right] \subseteq \bigcap_{S\in\mathscr{A}} f[S]$ with equality only if f is one-to-one.
c) $f[A - D] \subseteq B - f[D]$ with equality only if f is one-to-one and onto B.
d) $f^{\leftarrow}\left(\bigcup_{S\in\mathscr{B}} S\right) = \bigcup_{S\in\mathscr{B}} f^{\leftarrow}(S)$
e) $f^{\leftarrow}\left(\bigcap_{S\in\mathscr{B}} S\right) = \bigcap_{S\in\mathscr{B}} f^{\leftarrow}(S)$
f) $f^{\leftarrow}(B - E) = A - f^{\leftarrow}(E)$

IV Functions 12 / Equivalence relations defined by functions.

Definition 12.1 Let $f : A \to B$ be a function which maps a set A into a set B. We define an equivalence relation R_f on A as follows: Two elements a and b are related under R_f if and only if $\{a, b\} \subseteq f^{\leftarrow}(x)$ for some x in im f. The quotient set of A induced by R_f is then $A/R_f = \mathscr{A}_{R_f} = \{f^{\leftarrow}(\{x\}) : x \in f[A]\}$ We will refer to R_f as the *equivalence relation determined by f* and A/R_f (or $\mathscr{A}_{R_f}$) as the *quotient set of A determined by f.*

Theorem 12.2 Let $f : S \to T$ be a function where S and T are sets. There exists an onto function $g_f : S \to S/R_f$ and a one-to-one function $h_f : S/R_f \to T$ such that $h_f \circ g_f = f$. The function, $h_f \circ g_f = f$, is called the *canonical decomposition of f.*

V From sets to numbers 13 / The natural numbers.

Definition 13.1 For any set x, we define the *successor* x^+, of x as $x^+ = x \cup \{x\}$.

Definition 13.2 If x is a set then $x^+ = x \cup \{x\}$. A set A is called an *inductive set* if it satisfies the following two properties:

a) $\varnothing \in A$.

b) $x \in A \Rightarrow x^+ \in A$.

Definition 13.3 We define the *set of all natural numbers*, $\mathbb{N}$, as the intersection of all inductive sets. That is $\mathbb{N} = \{x : x \in I$ for any inductive set $I\}$.

Theorem 13.4 Let A be a subset of $\mathbb{N}$. If A satisfies the two properties:

a) $0 \in A$

b) $m \in A \Rightarrow m^+ \in A$

then $A = \mathbb{N}$.

Corollary 13.5 (The Principle of mathematical induction.) Let P denote a particular set property. Suppose $P(n)$ means "the property P is satisfied depending on the value of the natural number n". Let

$$A = \{n \in \mathbb{N} : P(n) \text{ holds true } \}$$

If A satisfies the two properties:

a) $0 \in A$. That is $P(0)$ holds true,

b) $(n \in A) \Rightarrow (n^+ \in A)$. That is, $P(n)$ holds true $\Rightarrow P(n^+)$ holds true.

then $A = \mathbb{N}$. That is, $P(n)$ holds true for all natural numbers n.

Definition 13.6 A set S which satisfies the property "$x \in S \Rightarrow x \subset S$" is called a *transitive set.*

Theorem 13.7 The non-empty set S is a transitive set if and only if the property "$x \in y$ and $y \in S$" $\Rightarrow$ "$x \in S$".

Theorem 13.8 The set $\mathbb{N}$ of natural numbers is a transitive set.

Theorem 13.9 a) For natural numbers n, m, $m \in n \Rightarrow m \subseteq n$. Hence every natural number is a transitive set.

b) For any natural number n, $n \neq n^+$.

c) For any natural number n, $n \notin n$.

d) For any distinct natural numbers n, m, $m \subset n \Rightarrow m \in n$.

Theorem 13.10 Let m and n be distinct natural numbers.

a) If $m \subset n$ then $m^+ \subseteq n$.

b) *All natural numbers are comparable.* Either $m \subset n$ or $n \subset m$. Equivalently, $m \in n$ or $n \in m$. Hence both "$\subset$" and "$\in$" linearly order $\mathbb{N}$.

c) There is no natural number m such that $n \subset m \subset n^+$.

Theorem 13.11 Every natural number has an immediate predecessor. If k and n are natural numbers such that $k^+ = n$ then k is called an *immediate predecessor* of n. For any non-zero natural number n, $k = \bigcup_{m \subset n} m$ is an immediate predecessor of n.

Theorem 3.12 *Unique immediate predecessors.* Any non-zero natural number has a *unique* immediate predecessor.

Theorem 3.13 (The Principle of mathematical induction: second version.) Let P denote a particular property. Suppose $P(n)$ means "the property P is satisfied depending on the value of the natural number n". Let

$$A = \{n \in \mathbb{N} : P(n) \text{ holds true } \}$$

Suppose that, for any natural number n,

$$P(k) \text{ is true for all } k < n \Rightarrow P(n) \text{ is true}$$

Then $P(n)$ holds true for all natural numbers n.

V From sets to numbers 14 / The natural numbers as a well-ordered set.

Notation 14.1 We define the relation "$\in_=$" on $\mathbb{N}$ as follows:

$$m \in_= n \text{ if and only if } m = n \text{ or } m \in n$$

If $m \in_= n$ and we want to state explicitly that $m \neq n$ we write $m \in n$.

Theorem 14.2 Let $(S, \leq)$ be a linearly ordered set. Suppose $T \subseteq S$. The element q is *a least element of T with respect to "$\leq$"* if $q \in T$ and $q \leq m$ for all $m \in T$. If S is equipped with a strict linear ordering "$<$" a *least element of T with respect to $<$* is an element $q \in T$ such that $q < m$ for all $m \in T$ where $m \neq q$. The set $(S, \leq)$ is said to be *well-ordered* with respect to "$\leq$" if every non-empty subset T of S contains its least element with respect to $\leq$. Similarly, the set $(S, <)$ is said to be *well-ordered* with respect to "$<$" if every non-empty subset T of S contains its least element with respect to $<$.

Theorem 14.3 The natural numbers $\mathbb{N}$ is a strict $\in$-well-ordered set.

Corollary 14.4 Every natural numbers n is a $\in$-well-ordered set.

Theorem 14.5 Any bounded non-empty subset of $(\mathbb{N}, \in)$ has a maximal element.

Definition 14.6 Consider the set $\{1, 2\}^{\mathbb{N}}$ of all functions mapping natural numbers to 1 or 2. We define the *lexicographic order* "$<$" on $\{1, 2\}^{\mathbb{N}}$ as follows: For any two elements $f = \{a_0, a_1, a_2, a_3, \ldots, \}$ and $g = \{b_0, b_1, b_2, b_3, \ldots, \}$ in $\{1, 2\}^{\mathbb{N}}$, $f = g$ if and only if $a_i = b_i$ for all $i \in \mathbb{N}$ and $f < g$ if and only if for the first two unequal corresponding terms a_i and b_i, $a_i \in b_i$. A lexicographic ordering can similarly be defined on $S^{\mathbb{N}}$ where S is any subset of $\mathbb{N}$.

V From sets to numbers 15 / Arithmetic of the natural numbers.

Definition 15.1 Let m be a fixed natural number. Addition of a natural number n with m is defined as the function $r_m : \mathbb{N} \to \mathbb{N}$ satisfying the two conditions

$$\begin{aligned} r_m(0) &= m \\ r_m(n^+) &= [r_m(n)]^+ \end{aligned}$$

The expression $m + n$ as simply another way of writing $r_m(n)$. Thus

$$r_m(0) = m \quad \Leftrightarrow \quad m + 0 = m \tag{5}$$

$$r_m(n^+) = [r_m(n)]^+ \quad \Leftrightarrow \quad m + n^+ = (m + n)^+ \tag{6}$$

Theorem 15.2 Let m be a fixed natural number and let $r_m : \mathbb{N} \to \mathbb{N}$ be a function satisfying the two properties

$$\begin{cases} r_m(0) &= m \\ r_m(n^+) &= [r_m(n)]^+ \end{cases}$$

Then r_m is a well-defined function on $\mathbb{N}$.

Definition 15.3 For any natural number m, multiplication with the natural number m is defined as the function $s_m : \mathbb{N} \to \mathbb{N}$ satisfying the two conditions

$$\begin{aligned} s_m(0) &= 0 \\ s_m(n^+) &= s_m(n) + m \end{aligned}$$

We define the expression mn and $m \times n$ as alternate ways of writing $s_m(n)$. Thus

$$s_m(0) = 0 \quad \Leftrightarrow \quad m0 = m \times 0 = 0 \tag{7}$$

$$s_m(n^+) = s_m(n) + m \quad \Leftrightarrow \quad mn^+ = mn + m = m \times n + m \tag{8}$$

Theorem 15.4 Let m be a fixed natural number and let $s_m : \mathbb{N} \to \mathbb{N}$ be a function satisfying the two properties

$$\begin{cases} s_m(0) &= 0 \\ s_m(n^+) &= s_m(n) + m \end{cases}$$

Then s_m is a well-defined function on $\mathbb{N}$.

Theorem 15.5 For any two natural numbers m and n, $m \in_= n$ if and only if there exists a *unique* natural number k such that $n = m + k$.

Definition 15.6 For any two natural numbers m and n such that $m \leq n$, the *unique* natural number k satisfying $n = m + k$ is called the *difference between n and m* and is denoted by $n - m$. The operation "$-$" is called *subtraction*.

V From sets to numbers 16 / The integers $\mathbb{Z}$ and the rationals $\mathbb{Q}$.

Theorem 16.1 Let $Z = \mathbb{N} \times \mathbb{N}$. Let R_z be a relation on Z that is defined as follows: $(a, b)R_z(c, d)$ if and only if $a + d = b + c$. Then R_z is an equivalence relation on Z.

Corollary 16.2 Let $Z = \mathbb{N} \times \mathbb{N}$ be equipped with the equivalence relation R_z defined as:

$$(a, b)R_z(c, d) \Leftrightarrow a + d = b + c$$

For each $n \in \mathbb{N}$ let $[(0, n)]$ and $[(n, 0)]$ denote the R_z-equivalence classes containing the elements $(0, n)$ and $(n, 0)$ respectively. Then the quotient set induced by R_z can be expressed as $Z/R_z = \{[(0, n)] : n \in \mathbb{N}\} \cup \{[(n, 0)] : n \in \mathbb{N}\}$

Definitions 16.3 The set of *integers*, $\mathbb{Z}$, is defined as:

$$\mathbb{Z} = Z/R_z = \{[(a, b)] : a, b \in \mathbb{N}\} = \{[(0, n)] : n \in \mathbb{N}\} \cup \{[(n, 0)] : n \in \mathbb{N}\}$$

a) *Negative integers*: The set of *negative integers* is defined as being the set

$$\mathbb{Z}^- = \{[(0, n)] : n \in \mathbb{N}\}$$

Positive integers: The set of *positive integers* is defined as being the set

$$\mathbb{Z}^+ = \{[(n, 0)] : n \in \mathbb{N}\}$$

The elements of the form $[(0, n)]$ can be represented by $-n = [(0, n)]$ while the elements of the form $[(n, 0)]$ can be represented as $n = [(n, 0)]$.

b) *Order relation on* $\mathbb{Z}$: We define a relation $\leq_z$ on $\mathbb{Z}$ as follows: $[(a, b)] \leq_z [(c, d)]$ if and only if $a + d \leq b + c$. It is a routine exercise to show that $\leq_z$ is a linear ordering of $\mathbb{Z}$.

c) *Addition on* $\mathbb{Z}$: We must sometimes distinguish between addition of natural numbers and addition of integers. Where there is a risk of confusion we will use the following notation: "$+_n$" means addition of natural numbers while "$+_z$" means addition of integers.
Addition $+_z$ on $\mathbb{Z}$ is defined as:

$$[(a, b)] +_z [(c, d)] = [(a +_n c,\ b +_n d)]$$

d) *Opposites of integers*: The *opposite* $-[(a, b)]$ of $[(a, b)]$ is defined as

$$-[(a, b)] = [(b, a)]^1$$

e) *Subtraction on integers*: Subtraction "$-_z$" on $\mathbb{Z}$ is defined as:

$$[(a, b)] -_z [(c, d)] = [(a, b)] + (-[(c, d)])^2$$

f) *Multiplication of integers*: Multiplication $\times_z$ on $\mathbb{Z}$ is defined as

$$[(a, b)] \times_z [(c, d)] = [(ac + bd, ad + bc)].^3$$

In particular, $[(0, n)] \times_z [(m, 0)] = [(0 + 0, 0 + nm)] = [(0, nm)] = -[(nm, 0)]$ and $[(n, 0)] \times_z [(m, 0)] = [(nm, 0)]$.

g) *Absolute value of an integer*: The absolute value, $|n|$, of an integer n is defined as

$$|n| = \begin{cases} n & \text{if} \quad 0 \leq_z n \\ -n & \text{if} \quad n <_z 0 \end{cases}$$

[1]Note that $-[(n, 0)] = [(0, n)] = -n$.

[2]When there is no risk of confusion with subtraction of other types of numbers we will simply use "$-$".

[3]Note that the "center dot" can be used instead of the "$\times''_z$ symbol. When there is no risk of confusion with multiplication of other types of numbers we will simply use "$\times$".

h) *Equality of two integers*: If (a, b) and (c, d) are ordered pairs which are equivalent under the relation R_z, then the R_z-equivalence classes $[(a, b)]$ and $[(c, d)]$ are equal sets. To emphasize that they are equal sets under the relation R_z we can write

$$[(a, b)] =_z [(c, d)]$$

i) *Distribution properties*: If $[(a, b)]$, $[(c, d)]$ and $[(e, f)]$ are integers then

$$[(a, b)] \times_z ([(c, d)] +_z [(e, f)]) =_z [(a, b)] \times_z [(c, d)] +_z [(a, b)] \times_z [(e, f)]$$

and

$$([(c, d)] +_z [(e, f)]) \times_z [(a, b)] =_z [(c, d)] \times_z [(a, b)] +_z [(e, f)] \times_z [(a, b)]$$

Theorem 16.4 Let $Q = \mathbb{Z} \times \mathbb{Z}^*$ where $\mathbb{Z}^* = \mathbb{Z} - \{0\}$. Let R_q be a relation on Q defined as follows: $(a, b)R_q(c, d)$ if and only if $a \times_z d = b \times_z c$. Then R_q is an equivalence relation on Q.

Definitions 16.5 The set of *rational numbers*, $\mathbb{Q}$, is defined as:

$$\mathbb{Q} = Q/R_q = \{[(a, b)] : a \in \mathbb{Z}, b \in \mathbb{Z}^*\}^1$$

The expression $[(a, b)]$ is normally written in the form $\frac{a}{b}$.

a) We define a relation $\leq_q$ on $\mathbb{Q}$ as follows: If b and d are both positive, $[(a, b)] \leq_q [(c, d)]$ if and only if $a \times_z d \leq_z b \times_z c$.

b) Addition $+_q$ on $\mathbb{Q}$ is defined as:

$$[(a, b)] +_q [(c, d)] = [(ad +_z bc,\ b \times_z d)]$$

c) Subtraction $-_q$ on $\mathbb{Q}$ is defined as:

$$[(a, b)] -_q [(c, d)] = [(a, b)] +_q [(-c, d)]$$

d) Multiplication $\times_q$ on $\mathbb{Q}$ is defined as

$$[(a, b)] \times_q [(c, d)] = [(a \times_z c, b \times_z d)]$$

e) *Equality of two rational numbers*: If (a, b) and (c, d) are ordered pairs of integers $(b, d \neq 0)$ which are equivalent under the relation R_q, then the R_q-equivalence classes $[(a, b)]$ and $[(c, d)]$ are equal sets. To emphasize that they are equal sets under the relation R_q we can write

$$[(a, b)] =_q [(c, d)]$$

[1] Recall that a and b is shorthand for expressions of the form $[(0, a)]$ or $-[(0, b)]$

f) *Opposites of rational numbers.* If (a, b) is an ordered pair of integers $(b \neq 0)$ and $[(a, b)]$ is its R_q-equivalence class then the opposite of the rational number $[(a, b)]$ is defined as $[(-a, b)]$ and is denoted as $-[(a, b)] =_q [(-a, b)]$.

Theorem 16.6 Suppose a and b are positive integers where $b \neq 0$ and $[(a, b)]$ is an R_q equivalence class. Then

a) $[(-a, -b)] = \frac{-a}{-b} = \frac{a}{b} = [(a, b)]$.

b) $-[(a, b)] =_q [(-a, b)] = \frac{-a}{b} = \frac{a}{-b} = [(a, -b)]$

V From sets to numbers 17 / Dedekind cuts: "Real numbers are us!"

Definition 17.1 For any real number r, let ${}_{\mathbb{R}}S_r$ denote the interval $(-\infty, r)$ in $\mathbb{R}$. It is the subset of all real numbers strictly smaller than r. The subset ${}_{\mathbb{R}}S_r$ is called an *initial segment in* $\mathbb{R}$. For any real number r, the subset ${}_{\mathbb{Q}}S_r = (-\infty, r) \cap \mathbb{Q} = {}_{\mathbb{R}}S_r \cap \mathbb{Q}$ is called an *initial segment in* $\mathbb{Q}$. For each of ${}_{\mathbb{R}}S_r$ and ${}_{\mathbb{Q}}S_r$ the real number r is the least upper bound of $(-\infty, r)$ and $(-\infty, r) \cap \mathbb{Q}$ respectively. Note that r may be a non-rational number even for ${}_{\mathbb{Q}}S_r$ a proper subset of $\mathbb{Q}$.

Definition 17.2 The elements ${}_{\mathbb{Q}}S_r = (-\infty, r) \cap \mathbb{Q}$ of the set $\mathscr{D} = \{{}_{\mathbb{Q}}S_r : r \in \mathbb{R}\}$ are called *Dedekind cuts*

Definition 17.3 The set of all Dedekind cuts $\mathscr{D}$, linearly ordered by inclusion with addition $+$ and multiplication $\times$ as described above is called the *real numbers*. Those Dedekind cuts which do not have a least upper bound in $\mathbb{Q}$ are called *irrational numbers*.

Lemma 17.4 The union of a set of Dedekind cuts is either $\mathbb{Q}$ or a Dedekind cut.

Theorem 17.5 Every non-empty subset of $\mathbb{R}$ which has an upper bound has a least upper bound.

VI Infinite sets 18 / Infinite sets versus finite sets.

Definition 18.1 A set S is said to be an *infinite set* if S has a *proper* subset X such that a function $f : S \to X$ maps S one-to-one onto X. If a set S is not infinite then we will say that it is a *finite set*.

Theorem 18.2 Basic properties of infinite and finite sets.

a) The empty set is a finite set.

b) Any singleton set is a finite set.

c) Any set which has a subset which is infinite must itself be infinite.

d) Any subset of a finite set must be finite.

Theorem 18.3 Let $f : X \to Y$ be a one-to-one function mapping X onto Y. The set Y is infinite if and only if the set X is infinite.

Corollary 18.4 The one-to-one image of a finite set is finite.

Lemma 18.5 Let S be an infinite set and $x \in S$. Then $S - \{x\}$ is an infinite set.

Theorem 18.6 Every natural number n is a finite set.

Corollary 18.7 [AC] A set S is finite if and only if S is empty or in one-to-one correspondence with some natural number n.

Theorem 18.8 *The recursively constructed function theorem.* Let S be a set, $k : \mathscr{P}(S) \to S$ be a well-defined function on $\mathscr{P}(S)$ and $f \subseteq \mathbb{N} \times S$ be a relation. We write $f(n) = a$ if and only if $(n, a) \in f$. Let $m \in S$. Suppose the relation f satisfies the two properties

$$\begin{cases} f(0) = m & \Rightarrow \quad (0, m) = (0, f(0)) \in f \\ (n, f(n)) \in f & \Rightarrow \quad (n+1, k(S - \{f(0), f(1), \ldots, f(n)\}) = (n+1, f(n+1)) \in f \end{cases}$$

Then f is a well-defined function on $\mathbb{N}$.

Theorem 18.9 [AC] A set S is an infinite set if and only if it contains a one-to-one image of the $\mathbb{N}$.

Theorem 18.10 If the set S is a finite set and $f : S \to X$ is a function, then $f[S]$ is finite.

Theorem 18.11 If a set S contains n elements then $\mathscr{P}(S)$ contains 2^n elements. Hence, if a set S is a finite set, then the set $\mathscr{P}(S)$ is finite.

VI Infinite sets 19 / Countable and uncountable sets.

Definition 19.1 Two sets A and B are said to be *equipotent* sets if there exists a one-to-one function $f : A \to B$ mapping one onto the other. If A and B are *equipotent* we will say that "A is equipotent to B" or "A is equipotent with B".

Definition 19.2 *Countable sets* are those sets that are either finite or equipotent to $\mathbb{N}$. All infinite sets which are not countable are called *uncountable sets.*

Theorem 19.3 A subset of a countable set is countable.

Theorem 19.4 Any finite product, $\mathbb{N} \times \mathbb{N} \times \cdots \times \mathbb{N}$, of $\mathbb{N}$ is a countable set.

Lemma 19.5 Suppose f maps an infinite countable set A onto a set $B = f[A]$. Then B is countable.

Theorem 19.6 Let $\{A_i : i \in S \subseteq \mathbb{N}\}$ be a countable set of non-empty countable sets A_i. Then $\bigcup_{i \in S} A_i$ is countable.

Theorem 19.7 The set of all real numbers $\mathbb{R}$ is uncountable.

VI Infinite sets 20 / Properties of the equipotence relation.

Theorem 20.1 Let $\mathscr{S}$ be a class of sets. The equipotence relation R_e on $\mathscr{S}$ is an equivalence relation on $\mathscr{S}$.

Theorem 20.2 Suppose A, B, C and D are sets such that $A \sim_e B$ and $C \sim_e D$ where $A \cap C = \varnothing = B \cap D$. Then $(A \cup C) \sim_e (B \cup D)$.

Theorem 20.3 Suppose A, B, C and D are sets such that $A \sim_e B$ and $C \sim_e D$. Then $A \times C \sim_e B \times D$.

Corollary 20.4 Suppose A and B are infinite sets.

1) If $\{A, B\} \subset [\mathbb{N}]_e$ then $A \times B \in [\mathbb{N}]_e$.
2) If $\{A, B\} \subset [\mathbb{R}]_e$ then $A \times B \in [\mathbb{R}]_e$.

Theorem 20.5 If $\{A_i : i = 0, 1, 2, ..., n\}$ is a set of n non-empty countable sets then $\prod_{i=0}^{n} A_i$ is countable for all n.

Theorem 20.6 Suppose S is and infinite set and T is a countable set such that $S \cap T = \varnothing$. Then $S \sim_e S \cup T$.

Theorem 20.7 If the sets A and B are equipotent, then so are their associated power sets $\mathscr{P}(A)$ and $\mathscr{P}(B)$.

Theorem 20.8 Any non-empty set S is embedded in its power set $\mathscr{P}(S)$. But *no* subset of S is equipotent with $\mathscr{P}(S)$.

Definition 20.9 We will say that the non-empty set A is *properly embedded* in the set B if A is equipotent to a proper subset of B but B *is not equipotent to* A *or any of its subsets*. To represent the relationship "A is properly embedded in B" we will write

$$A \hookrightarrow_e B$$

If A and B are sets such that A is equipotent to a subset of B (where B may, or may not, be embedded in A), we will say that A is *embedded* in B. To represent the relationship "A is embedded in B" we will write

$$A \hookrightarrow_{e\sim} B$$

Definition 20.10 Let $\mathscr{S} = \{S : S \text{ is a set}\}$ and $\mathscr{E} = \{[S]_e : S \in \mathscr{S}\}$. Let $[A]_e$ and $[B]_e$ be elements of $\mathscr{E}$. We write

$$[A]_e <_e [B]_e$$

if and only if $A \hookrightarrow_e B$. We write

$$[A]_e \leq_e [B]_e$$

if and only if $A \hookrightarrow_{e\sim} B$.

Proposition 20.11 Let S be any set. Suppose $\mathscr{P}^0(S) = S$, $\mathscr{P}^1(S) = \mathscr{P}(S)$ and $\mathscr{P}^n(S) = \mathscr{P}(\mathscr{P}^{n-1}(S))$ for all $n \geq 1$. The set

$$\{[S]_e, [\mathscr{P}(S)]_e, [\mathscr{P}^2(S)]_e, [\mathscr{P}^3(S)]_e, \ldots, [\mathscr{P}^n(S)]_e \ldots, \}$$

forms an infinite $<_e$-ordered chain of distinct classes in $\mathscr{E}$.

Theorem 20.12 For every any non-empty set S, $2^S = \{\chi_T : T \in \mathscr{P}(S)\} \sim_e \mathscr{P}(S)$

Theorem 20.13 The set $\mathbb{R}$ is embedded in $\mathscr{P}(\mathbb{N})$.

Theorem 20.14 The set $\mathscr{P}(\mathbb{N})$ is embedded in $\mathbb{R}$.

VI Infinite sets 21 / The Shröder-Bernstein theorem.

Theorem 21.1 (The Schröder-Bernstein theorem) If S and T are infinite subsets where S is embedded in T and T is embedded in S then S and T are equipotent.

Lemma 21.2 Let T be a proper subset of the set S and $f : S \to T$ be a one-to-one function mapping S into T. Then there exists a one-to-one function $h : S \to T$ mapping S **onto** T.

Theorem 21.3 The set $\mathbb{R}$ of all real numbers is equipotent to $\mathscr{P}(\mathbb{N})$.

Theorem 21.4 The sets $\mathbb{N}^{\mathbb{N}}$ and $\mathbb{R}$ are equipotent.

Definition 21.5 If A and B are two sets, then the symbol B^A refers to the set of all functions mapping A into B.

VII Cardinal numbers 22 / An introduction to cardinal numbers.

Postulate 22.1 There exists a class of sets $\mathscr{C}$ which satisfies the following properties:

1. Every natural number n is an element of $\mathscr{C}$.
2. Any set $S \in \mathscr{S}$ is equipotent with precisely one element in $\mathscr{C}$

The sets in $\mathscr{C}$ are called *cardinal numbers*. When we say that a set S *has cardinality* κ we mean that $\kappa \in \mathscr{C}$ and that $S \sim_e \kappa$. If the set S has cardinality κ, we will write $|S| = \kappa$.

Definition 23.2 If S and T are sets and $\kappa = |S|$ and $\lambda = |T|$ then we define *addition*"+", *multiplication* "×" and *exponentiation* of two cardinal numbers as follows:

a) If $S \cap T = \varnothing$,
$$\kappa + \lambda = |S \cup T|$$

b)
$$\kappa \times \lambda = |S \times T|$$

c)
$$\kappa^\lambda = |S^T|$$

where S^T represents the set of all functions mapping T into S (as previously defined). That is, $|S|^{|T|} = |S^T|$. For convenience we define $0^\lambda = 0$ and $\kappa^0 = 1$.

Theorem 22.3 The class $\mathscr{C}$ of all cardinal numbers is a proper class.

VII Cardinal numbers 23 / Arithmetic of cardinal numbers.

Theorem 23.1 Addition on $\mathscr{C}$ is well-defined. That is, if S_1, S_2, T_1 and T_2 are sets such that $\kappa = |S_1| = |S_2|$ and $\lambda = |T_1| = |T_2|$, then $|S_1 \cup T_1| = \kappa + \lambda = |S_2 \cup T_2|$.

Theorem 23.2 Let κ, λ, ϕ and ψ be any four cardinal numbers. Then

a) $\kappa + \lambda = \lambda + \kappa$ (Commutativity of addition)

b) $(\kappa + \lambda) + \phi = \kappa + (\lambda + \phi)$ (Associativity of addition)

c) $\kappa \leq \kappa + \lambda$

d) $\kappa \leq \lambda$ and $\phi \leq \psi \Rightarrow \kappa + \phi \leq \lambda + \psi$.

Theorem 23.3 Multiplication on $\mathscr{C}$ is well-defined. That is, if S_1, S_2, T_1 and T_2 are sets such that $\kappa = |S_1| = |S_2|$ and $\lambda = |T_1| = |T_2|$, then $|S_1 \times T_1| = \kappa \times \lambda = |S_2 \times T_2|$.

Theorem 23.4 Let κ, λ, ϕ and ψ be any three cardinal numbers. Then

a) $\kappa \times \lambda = \lambda \times \kappa$ (Commutativity of multiplication)

b) $(\kappa \times \lambda) \times \phi = \kappa \times (\lambda \times \phi)$ (Associativity of multiplication)

c) $\kappa \times (\lambda + \phi) = (\kappa \times \lambda) + (\kappa \times \phi)$ (Left-hand distributivity)

d) $\lambda > 0 \Rightarrow \kappa \leq (\kappa \times \lambda)$

e) $\kappa \leq \lambda$ and $\phi \leq \psi \Rightarrow \kappa \times \phi \leq \lambda \times \psi$.

f) $\kappa + \kappa = 2 \times \kappa$.

g) $\kappa + \kappa \leq \kappa \times \kappa$ when $\kappa \geq 2$.

VII Cardinal numbers 24 / Exponentiation of cardinal numbers.

Theorem 24.1 Exponentiation on $\mathscr{C}$ is well-defined. That is, if S, S^*, T and T^* are sets such that $|S| = |S^*|$ and $|T| = |T^*|$, then $|S^T| = |S^{*T^*}|$.

Theorem 24.2 Let κ, λ and ϕ be any three cardinal numbers. Then

a) $\kappa^{\lambda+\phi} = \kappa^\lambda \times \kappa^\phi$

b) $(\kappa^\lambda)^\phi = \kappa^{\lambda\times\phi}$

c) $(\kappa \times \lambda)^\phi = \kappa^\phi \times \lambda^\phi$.

Theorem 24.3 Let κ, λ, and α be infinite cardinal numbers. Then

a) $\kappa \leq \kappa^\lambda$

b) $\alpha \leq \kappa \Rightarrow \alpha^\lambda \leq \kappa^\lambda$

c) $\alpha \leq \lambda \Rightarrow \kappa^\alpha \leq \kappa^\lambda$

VII Cardinal numbers 25 / Sets of cardinality c

Theorem 25.1 Let $\mathbb{C}$ denote the set of all complex numbers and $\mathbb{J}$ denote the set of all irrational numbers. Let n denote the cardinality of a non-empty finite set.

a) The cardinality of $\mathbb{R}^n$ is c.

b) The cardinality of $\mathbb{C}$ is c.

c) The cardinality of $\mathbb{J}$ is c.

Theorem 25.2

a) Let $\mathscr{S}_\mathbb{R}$ denote the set of all countably infinite sequences of real numbers. The cardinality of $\mathscr{S}_\mathbb{R}$ is c.

b) Let $\mathscr{S}_\mathbb{N}$ denote the set of all countably infinite sequences of natural numbers. The cardinality of $\mathscr{S}_\mathbb{N}$ is c.

c) Let $\mathbb{N}^\mathbb{N}_{(1-1)}$ denote the set of all one-to-one functions mapping $\mathbb{N}$ to $\mathbb{N}$. The cardinality of $\mathbb{N}^\mathbb{N}_{(1-1)}$ is c.

d) Let $\mathbb{R}^\mathbb{N}_{(1-1)}$ denote the set of all one-to-one functions mapping $\mathbb{N}$ to $\mathbb{R}$. Then the cardinality of $\mathbb{R}^\mathbb{N}_{(1-1)}$ is c.

Proposition 25.3 The Cantor set has cardinality c.

VIII Ordinal numbers 26 / Well-ordered sets.

Theorem 26.1 Let $f : T \to S$ be a one-to-one function mapping T onto S. If T is a well-ordered then T induces a well-ordering on S. In particular, every countable set can be well-ordered.

Definition 26.2 Given a *well-ordered* set $(S, \leq)$, a proper subset U satisfying the property

$$[u \in U \text{ and } x \leq u] \Rightarrow [x \in U]$$

is called an *initial segment* of S. In this definition the strict order relation $<$ can be used instead of $\leq$ without altering the meaning of "initial segment".

Theorem 26.3 If $(S, \leq)$ is a well-ordered set then every initial segment in S is of the form $S_a = \{x \in S : x < a\}$ for some $a \in S$.

Definition 26.4 Let $f : (S, \leq_S) \to (T, \leq_T)$ be a function mapping a well-ordered class, $(S, \leq_S)$, onto a well-ordered class, $(T, \leq_T)$. Note that the symbols $\leq_S$ and $\leq_T$ will allow us to distinguish between the order relations applied to each set S and T.

a) We will say that the function f is *increasing* on $(S, \leq_S)$ if

$$(x \leq_S y) \Rightarrow (f(x) \leq_T f(y))$$

b) We will say that the function f is *strictly increasing* on $(S, \leq_S)$ if

$$(x <_S y) \Rightarrow (f(x) <_T f(y))$$

A strictly increasing function must be one-to-one.

c) If $f : (S, \leq_S) \to (T, \leq_T)$ is strictly increasing then f is said to be an *order isomorphism* mapping S into T.

If there exists an *onto* order isomorphism between the two well-ordered classes, $(S, \leq_S)$ and $(T, \leq_T)$, we will say that the classes are *order isomorphic*, or that a function maps S *order isomorphically* onto T.

If there exists an *onto* order isomorphism between the two well-ordered classes $(S, \leq_S)$ and $(T, \leq_T)$ we will say that the classes are *order isomorphic* or that a function maps S *order isomorphically* onto T.

Theorem 26.5 Let $(S, \leq_S)$ and $(T, \leq_T)$ be a well-ordered sets.

a) The inverse of an order isomorphism is an order isomorphism.

b) If $f : (S, \leq_S) \to (S, \leq_S)$ is a *strictly increasing* function mapping S into itself then $f(x) \geq x$ for all $x \in S$.

c) The set S cannot be order isomorphic to an initial segment of itself.

d) If $f : (S, \leq_S) \to (S, \leq_S)$ is an order isomorphism then f is the identity function.

e) If $f : (S, \leq_S) \to (T, \leq_T)$ and $g : (S, \leq_S) \to (T, \leq_T)$ are order isomorphisms mapping S onto T then $f = g$.

f) Suppose $f : (S, \leq_S) \to (T, \leq_T)$ is an order isomorphism mapping S onto an initial segment of T. Then S and T cannot be order isomorphic.

Notation 26.6 Let S and T be two well-ordered sets. Then the expression

$$S \sim_{\text{WO}} T$$

means "S and T are order isomorphic". The expression

$$S <_{\text{WO}} T$$

means "$S \sim_{\text{WO}} T_a$" where T_a is some *initial segment* of T. The expression

$$S \leq_{\text{WO}} T$$

means "$S \sim_{\text{WO}} T$ or $S <_{\text{WO}} T$".

Theorem 26.7 Let $(S, \leq_S)$ and $(T, \leq_T)$ be two well-ordered sets. Then either $S \leq_{\text{WO}} T$ or $T \leq_{\text{WO}} S$.

Proposition 26.8 For every natural number n, the lexicographically ordered set $S = \{1, 2, \ldots, n\} \times \mathbb{N}$ is well-ordered.

VIII Ordinal numbers 27 / Ordinal numbers: Definition and properties.

Definition 27.1 Let S be a set. If S satisfies the two properties,

1) S a transitive set,
2) S is strictly $\in$-well-ordered

then S is called an *ordinal number*.

Notation 27.2 When viewed as an ordinal number, $\mathbb{N}$ will be represented by the lower-case Greek letter ω.

Theorem 27.3 If α is an ordinal number then so is its successor $\alpha^+ = \alpha \cup \{\alpha\}$.

Definition 27.4 Suppose the set S is $<$-ordered. We say that an element y in S is an *immediate successor* of the element x if $x < y$ and there does not exist any element z in S such that $x < z < y$. We say that x is an *immediate predecessor* of y if y is an immediate successor of x.

Theorem 27.5 Let α be an ordinal number greater than zero.

a) Every element x of the ordinal α is an initial segment of α.

b) The ordinal α is an initial segment of some ordinal.

c) Every initial segment x in α is an ordinal number.

d) Every element of the ordinal α is an ordinal number.

Proposition 27.6 Any infinite ordinal not equal to ω contains ω.

Proposition 27.7 Let α and β be distinct ordinal numbers. If $\alpha \subset \beta$, then $\alpha \in \beta$.

Lemma 27.8 If the ordinals α and β are order isomorphic, then $\alpha = \beta$.

Theorem 27.9 The relation "$\in$" linearly orders the class of all ordinals.

Definition 27.10 An ordinal α which does not contain a maximal element is called a *limit ordinal*.

Proposition 27.11 If U is a non-empty set of ordinals which contains a maximal element β with respect to $\in$, then the union, $\cup\{\alpha : \alpha \in U\}$, is equal to the maximal ordinal, β, of U.

Theorem 27.12 If U is a set of ordinals which does *not* contain a maximal element with respect to "$\in$", then $\gamma = \cup\{\alpha : \alpha \in U\}$ is a limit ordinal which is not contained in U. Furthermore, γ is the $\in$-least ordinal which contains all elements of U.

Corollary 27.13 Let U be a non-empty set of ordinals which contains no maximal element. If U satisfies the "initial segment property", then U is the limit ordinal $\cup\{\alpha : \alpha \in U\}$.

Definition 27.14 Let T be a non-empty subset of an ordered set $(S, <)$. If u is an upper bound of the set T and, for any other upper bound v of T, $u \leq v$, then we say that u is the *least upper bound* of T. We also abbreviate the expression by writing $u = \text{lub}\, T$ or $u = \text{lub}(T)$.

Theorem 27.15 Let γ be a non-zero ordinal number. The following are equivalent:

1) The ordinal γ is a limit ordinal.

2) The ordinal γ is such that $\gamma = \cup\{\alpha : \alpha \in \gamma\}$.

3) The ordinal γ is such that $\text{lub}(\gamma) = \gamma$.

VIII Ordinal numbers 28 / Properties of the class of ordinal numbers.

Theorem 28.1 The class, $\mathscr{O}$, of ordinal numbers is a strict $\in$-linearly ordered class.

Theorem 28.2 The class $\mathscr{O}$ of all ordinal numbers is $\in$-well-ordered.

Theorem 28.3 A set S is an initial segment of $\mathscr{O}$ if and only if S is an ordinal number.

Theorem 28.4 The class $\mathscr{O}$ of all ordinal numbers is not a set.

Theorem 28.5 Principle of transfinite induction. Let $\{x_\alpha : \alpha \in \mathscr{O}\}$ be a class whose elements are indexed by the ordinals. Let P denote a particular element property. Suppose $P(\alpha)$ means "the element x_α satisfies the property P". Suppose that, for any $\beta \in \mathscr{O}$,

$$\text{"}P(\alpha) \text{ is true } \forall\ \alpha \in \beta\text{" implies "}P(\beta) \text{ is true"}$$

Then $P(\alpha)$ holds true for all ordinals $\alpha \in \mathscr{O}$.

Corollary 28.6 Transfinite induction. Version 2. Let $\{x_\alpha : \alpha \in \mathscr{O}\}$ be a class whose elements are indexed by the ordinals. Let P denote a particular element property. Suppose $P(\alpha)$ means "the element x_α satisfies the property P". Suppose that:

1) $P(0)$ holds true,
2) $P(\alpha)$ holds true implies $P(\alpha + 1)$ holds true,
3) If β is a limit ordinal, "$P(\alpha)$ is true for all $\alpha \in \beta$ implies $P(\beta)$ is true".

Then $P(\alpha)$ holds true for all ordinals α.

Theorem 28.7 Let S be a $<$-well-ordered set. Then S is order isomorphic to some ordinal number $\alpha \in \mathscr{O}$. Furthermore this order isomorphism is unique.

Definition 28.8 Let S be a $<$-well-ordered set. If α is the unique ordinal which is order isomorphic to S then we will say that S is of *order type* α, or of *ordinality* α. of S is α. If S is of ordinality α, we will write ${}^{\text{ord}}S = \alpha$.

Lemma 28.9 Hartogs' lemma. Let S be any set. Then there exists an ordinal α which is not equipotent with S or any of its subsets.

Theorem 28.10 There exists an uncountable ordinal.

Corollary 28.11 The class $\omega_1 = \{\alpha \in \mathscr{O} : \alpha \text{ is a countable ordinal } \}$ is the least uncountable ordinal.

Definition 28.12 Let S be any set. Let

$$U = \{\alpha \in \mathscr{O} :\ \alpha \text{ not equipotent to any subset of } S\}$$

By Hartogs' lemma the class U is non-empty. Since $\mathscr{O}$ is $\in$-well-ordered, U contains a unique least ordinal $h(S)$. We will call the ordinal $h(S)$ the *Hartogs number of* S. Then h can be viewed as a class function which associates each set S in the class of all sets to a unique ordinal number α in the class of all ordinals $\mathscr{O}$.

Theorem 28.13 There exists a strictly increasing class $\{\omega_\alpha : \alpha \in \mathscr{O}\}$ of pairwise non-equipotent infinite ordinals all of which are uncountable except for $\omega_0 = \omega$.

Theorem 28.14 Let $\{\omega_\alpha : \alpha \in \mathscr{O}\}$ be the class of ordinals as defined in the previous theorem.

a) Every element of $\{\omega_\alpha : \alpha \in \mathscr{O}\}$ is a limit ordinal.

b) For every ordinal α, either $\alpha \in \omega_\alpha$ or $\alpha = \omega_\alpha$.

Theorem 28.15 The Transfinite recursion theorem. Let W be a well-ordered class and $f : W \to W$ be a class function mapping W into W. Let $u \in W$. Then there exists a unique class function $g : \mathscr{O} \to W$ which satisfies the following properties:

a) $g(0) = u$

b) $g(\alpha^+) = f(g(\alpha)),\ \forall \alpha \in \mathscr{O}$

c) $g(\beta) = \text{lub}\{g(\alpha) : \alpha \in \beta\},\ \forall$ limit ordinals β

VIII Ordinal numbers 29 / Initial ordinals: "Cardinal numbers are us!"

Definition 29.1 We say that β is an *initial ordinal* if it is the least of all ordinals equipotent with itself. That is, β is an *initial ordinal* if $\alpha \in \beta \Rightarrow \alpha \not\sim_e \beta$.

Lemma 29.2 The class of all initial ordinals is a subclass of $\mathscr{I}$.

Theorem 29.3 The class $\mathscr{I}$ is precisely the class of all initial ordinals.

Theorem 29.4 [AC] The Well-ordering theorem. Every set can be well-ordered.

Theorem 29.5 The class of all initial ordinals $\mathscr{I} = \omega_0 \cup \{\omega_\alpha : \alpha \in \mathscr{O}\}$ satisfies the following properties :

1. Every set S is equipotent to exactly one element in $\mathscr{I}$.
2. Two sets S and T are equipotent if and only if they are equipotent to the same element of $\mathscr{I}$.
3. The class $\mathscr{I}$ is $\in$-linearly ordered.

Definition 29.6 Cardinal numbers. An ordinal is called a *cardinal number* if and only if this ordinal is an initial ordinal. The class, $\mathscr{I}$, is also referred to as the class, $\mathscr{C}$, of all cardinal numbers.

Lemma 29.7 For any infinite cardinal number κ, define a relation $<_*$ on $\kappa \times \kappa$ as follows. For pairs (α, β) and (γ, ψ) of *ordinal* pairs in $\kappa \times \kappa$,

$$(\alpha, \beta) <_* (\gamma, \psi) \begin{cases} \alpha \cup \beta \in \gamma \cup \psi \\ \text{or} \\ \beta \in \psi \text{ when } \alpha \cup \beta = \gamma \cup \psi \\ \text{or} \\ \alpha \in \gamma \text{ when } \alpha \cup \beta = \gamma \cup \psi \text{ and } \beta = \psi \end{cases}$$

Then $<_*$ well-orders $\kappa \times \kappa$.

Theorem 29.8 [AC] For any ordinal α, $\aleph_\alpha \times \aleph_\alpha = \aleph_\alpha$.

Corollary 29.9 Let κ be an infinite cardinal and $\{A_\alpha : \alpha \in \beta\}$ be a set of non-empty sets indexed by the elements of the ordinal $\beta \in_= \kappa$ where $|A_\alpha| \in_= \kappa$ for all $\alpha \in \beta$. Then $|\cup \{A_\alpha : \alpha \in \beta\}| \in_= \kappa$.

Corollary 29.10 For any infinite cardinal number κ, $\kappa^\kappa = 2^\kappa$.

Definition 29.11

a) We say that an infinite cardinal number, $\aleph_\gamma$, is a *successor cardinal* if the index, γ, has an immediate predecessor (i.e., $\gamma = \beta + 1$, for some β). The expression, $\aleph_{\alpha^+} = \aleph_{\alpha+1}$, denotes a successor cardinal. We say that an infinite cardinal number, $\aleph_\gamma$, is a *limit cardinal* if γ is a limit ordinal (i.e., $\gamma = \text{lub}\{\alpha : \alpha \in \gamma\}$).

b) We say that a limit cardinal $\aleph_\gamma$ is a *strong limit cardinal* if $\aleph_\gamma$ is uncountable and $\{2^{\aleph_\alpha} : \alpha \in \gamma\} \subseteq \aleph_\gamma$.

Theorem 29.12 [GCH] Every uncountable limit cardinal is a strong limit cardinal.

Theorem 29.13 [AC] There exists a strong limit cardinal number.

Definition 29.14

a) We say that an infinite cardinal $\aleph_\gamma$ is a *singular cardinal number* if $\aleph_\gamma$ is the least upper bound of a strictly increasing sequence of ordinals, $\{\alpha_\kappa : \kappa \in \beta\}$, indexed by the elements of some ordinal β in $\aleph_\gamma$.

b) An infinite cardinal $\aleph_\gamma$ is said to be a *regular cardinal number* if it is not a singular cardinal number. That is, there does not exist an ordinal, β, in $\aleph_\gamma$ such that $\aleph_\gamma = \text{lub}\{\alpha_\kappa : \kappa \in \beta\}$.

Theorem 29.15 Every infinite successor cardinal, $\aleph_{\alpha^+} = \aleph_{\alpha+1}$, is a regular cardinal.

Theorem 29.16 Let γ be an infinite cardinal number. If γ is a singular cardinal then the cardinal number $\aleph_\gamma$ is a singular cardinal.

Definition 29.17 A regular cardinal number which a limit cardinal is called an *inaccessible cardinal*. A regular cardinal number which is a strong limit cardinal is called a *strongly inaccessible cardinal*.

Definition 29.18 Let $\aleph_\gamma$ be an infinite cardinal. We say the *cofinality of* $\aleph_\gamma$ is β and write $\text{cf}(\aleph_\gamma) = \beta$ if β is the least ordinal in $\aleph_\gamma$ which indexes an increasing set of ordinals $\{\theta_\alpha : \alpha < \beta\}$ such that $\aleph_\gamma = \text{lub}\{\theta_\alpha : \alpha < \beta\}$. If no such β exists in $\aleph_\gamma$ then we say the cofinality $\text{cf}(\aleph_\gamma)$ of is $\aleph_\gamma$ and write $\text{cf}(\aleph_\gamma) = \aleph_\gamma$.

Theorem 29.19 The cofinality $\mathrm{cf}(\aleph_\gamma)$ of an infinite cardinal number $\aleph_\gamma$ is a cardinal number. Hence $\mathrm{cf}(\aleph_\gamma)$ is the smallest cardinality of all sets which are cofinal in $\aleph_\gamma$.

Theorem 29.20 The cofinality $\mathrm{cf}(\varphi)$ of an infinite cardinal number φ is a regular cardinal.

Theorem 29.21 If κ is an infinite cardinal and $\mathrm{cf}(\kappa) \leq \lambda$, then $\kappa < \kappa^\lambda$.

IX More on axioms: Choice, regularity and Martin's axiom 30 / Axiom of choice

Theorem 30.1 Suppose $\mathscr{S}$ is a finite set of non-empty sets whose union is the set M. Then there exists a function $f : \mathscr{S} \to M$ which maps each set to one of its elements.

Theorem 30.2 Let AC* denote the statement:

"For any set $\mathscr{S} = \{S_\alpha : \alpha \in \gamma\}$ of non-empty sets, $\Pi_{\alpha\in\gamma} S_\alpha$ is non-empty."

The Axiom of choice holds true if and only if AC* holds true. The Axiom of choice holds true if and only if AC* holds true.

Theorem 30.3 The statement "Every set is well-orderable" holds true if and only if the Axiom of choice holds true.

Theorem 30.4 [AC] Any infinite set can be expressed as the union of a pairwise disjoint set of infinite countable sets.

Theorem 30.5 [AC] Let $(X, <)$ be a partially ordered set. If every chain of X has an upper bound then X has a maximal element.

Theorem 30.6 Suppose that those partially ordered sets $(X, <)$ in which every chain has an upper bound must have a maximal element. Then given any subset $\mathscr{S} \subseteq \mathscr{P}(S) - \varnothing$ there exists a choice function $f : \mathscr{S} \to S$ which maps each set in $\mathscr{S}$ to one of its elements.

Theorem 30.7 [ZL] Every vector space has a basis.

IX More on axioms: Choice, regularity and Martin's axiom 31 / Axiom of regularity and cumulative hierarchy.

Theorem 31.1 The Axiom of regularity holds true if and only if every non-empty set S contains a *minimal* element with respect to the membership relation "$\in$".

Theorem 31.2 [Axiom of regularity] No set is an element of itself.

Definition 31.3 We say that a class is *well-founded* if it does not contain an infinite descending chain of sets. That is, there does not exist an infinite sequence $\{x_n : n \in \omega\}$ such that $\cdots \in x_4 \in x_3 \in x_2 \in x_1 \in x_0$.

Theorem 31.4 [AC] The *Axiom of regularity* and the statement "*Every set is well-founded*" are equivalent statements.

Definition 31.5 Let x be a set. The transitive closure of x is a set t_x satisfying the following three properties:

1) The set t_x is a transitive set.
2) $x \subseteq t_x$
3) t_x is the $\subseteq$-least transitive set satisfying properties 1 and 2.

Theorem 31.6 Let x be a set. Then there exists a smallest transitive set t_x which contains all elements of x. That is, if s is a transitive set such that $x \subseteq s$, then $x \subseteq t_x \subseteq s$.

Definition 31.7 Define the class function $f : \mathscr{S} \to \mathscr{S}$ as $f(S) = \mathscr{P}(S)$. The elements of the class $\{V_\alpha : \alpha \in \mathscr{O}\}$ belong to the image of the class function $g(\alpha) = V_\alpha$ recursively defined as follows:

$$\begin{array}{rcccccl}
g(0) & = & V_0 & = & \varnothing & = & 0 \\
g(1) & = & V_1 & = & f(V_0) & = & \mathscr{P}(0) = \{\varnothing\} = 1 \\
g(2) & = & V_2 & = & f(V_1) & = & \mathscr{P}(1) = \{\{\varnothing\}, \varnothing\} = 2^1 = 2 \\
g(3) & = & V_3 & = & f(V_2) & = & \mathscr{P}(2) = \{\{\{\{\varnothing\}, \varnothing\}\}, \{\{\varnothing\}\}, \{\varnothing\}, \varnothing\} \\
g(4) & = & V_4 & = & f(V_3) & = & \mathscr{P}(4) \text{ (16 elements)} \\
 & & \vdots & & \vdots & & \\
g(\alpha^+) & = & V_{\alpha^+} & = & f(V_\alpha) & = & \mathscr{P}(V_\alpha) \\
 & & \vdots & & \vdots & & \\
\text{If } \lambda = \text{limit ordinal, } g(\lambda) & = & V_\lambda & = & \cup_{\alpha\in\lambda} V_\alpha & & \\
 & & \vdots & & \vdots & &
\end{array}$$

The class $\{V_\alpha : \alpha \in \mathscr{O}\}$ is called the *Cumulative hierarchy of sets*. We define $V = \cup_{\alpha\in\mathscr{O}} V_\alpha$.

Lemma 31.8 a) Each set V_α is a transitive set.

b) If $\alpha \in \beta$ then $V_\alpha \in V_\beta$. Hence $V_\alpha \subset V_\beta$.

Lemma 31.9 For any non-empty set B, $B \subset V$ implies $B \in V$.

Theorem 31.10 For every set x, $x \in V = \cup_{\alpha\in\mathscr{O}} V_\alpha$.

Theorem 31.11 Let $V = \cup_{\alpha\in\mathscr{O}} V_\alpha$ be the class of sets constructed as described above. If V contains all sets then every set has a $\in$-minimal element.

Theorem 31.12 Let $V = \cup_{\alpha\in\mathscr{O}} V_\alpha$ be the class of sets constructed as described above.

a) The rank of the empty set $\varnothing$ is zero.

b) If $U \in V_\beta$ then $\text{rank}(U) < \beta$; hence $U \notin V_{\text{rank}(U)}$ for all sets U. Conversely, $\text{rank}(U) < \beta \Rightarrow U \in V_\beta$.

c) If U and V are sets such that $U \in V$ then $\text{rank}(U) < \text{rank}(V)$.

d) If γ is an ordinal then $\text{rank}(\gamma) = \gamma$.

Proposition 31.13 Let β be a limit ordinal. If a and b are two elements of V_β. Then $\{a, b\}$ is an element of V_β. That is, V_β satisfies the property described by the Axiom of pair.

Proposition 31.14 Let β be any ordinal. If U is an element of V_β then $\cup\{x : x \in U\} \in V_\beta$. That is, V_β satisfies the property described by the Axiom of union.

Proposition 31.15 Let β be a limit ordinal. If U is an element of V_β then V_β also contains a set $Y = \mathscr{P}(U)$ such that $S \subseteq U$ implies $S \in Y$. That is, V_β satisfies the property described by the Axiom of power set.

Proposition 31.16 Let α be any ordinal. For any two elements x and y of V_α, if for all $z \in V_\alpha (z \in x \Leftrightarrow z \in y)$, then x and y are the same set.

Proposition 31.17 Let α be an ordinal number such that $\alpha > \omega_0$. Then $\omega_0 \in V_\alpha$.

Proposition 31.18 Let α be an ordinal number. Then the Axiom of subsets holds true in V_α.

Proposition 31.19 The set V_{ω_0} satisfies the property described by the Axiom of replacement.

Proposition 31.20 Let γ be a limit ordinal. Then the set V_γ satisfies the property described by the Axiom of choice.

Proposition 31.21 Let α be an ordinal. Then the set V_α satisfies the property described by the Axiom of construction.

IX More on axioms: Choice, regularity and Martin's axiom 32 / Martin's axiom.

Definition 32.1 Let $(P, \leq)$ be a partially ordered set. If P contains no uncountable strong antichain then $(P, \leq)$ is said to satisfy the *countable chain condition*. In this case, we say that $(P, \leq)$ satisfies the *ccc* or that $(P, \leq)$ is a *ccc* partial order.

Definition 32.2 Let $(P, \leq)$ be a partially ordered set. Let D be a subset of P such that for every element p in P there exists an element d in D such that $d \leq p$. A subset D satisfying this property is said to be *dense in the partial ordering* $(P, \leq)$.

Definition 32.3 Let F be a subset of a partially ordered set $(P, \leq)$. If F is non-empty and satisfies the two properties, 1) If x and y belong to F there exists z in F which is less than or equal to both x and y (i.e., F is a *filter base* or *downward directed*), 2) if x belongs to F and x is less than or equal to an element y of P, then y belongs to F (i.e., F is upward closed). A filter in $(P, \leq)$ is a *proper filter* if it is not all of P. If $x \in P$ then the set of all elements above x is called a *principal filter with principal element* x. Such a filter is the smallest filter which contains x.

Theorem 32.4 MA(κ): Let κ be an infinite cardinal and $(P, \leq)$ be a non-empty partially ordered set satisfying the *countable chain condition*. Let $\mathscr{D} = \{D \in \mathscr{P}(P) : D \text{ is dense in } P\}$ such that $|\mathscr{D}| \leq \kappa$. Then there is a proper filter $F \subseteq P$ such that, $F \cap D \neq \varnothing$ for every set $D \in \mathscr{D}$. The statement MA($\aleph_0$) holds true in *ZFC*.

Theorem 32.5 The statement MA($2^{\aleph_0}$) fails in *ZFC*.

Definition 32.6 *Martin's axiom*, MA, is defined as being MA(κ) where κ satisfies $\aleph_0 \leq \kappa < 2^{\aleph_0}$

Theorem 32.7 [MA] Suppose κ is an infinite cardinal such that $\kappa < 2^{\aleph_0}$. If X is a Hausdorff compact space with *ccc* and $\{U_\alpha : \alpha \leq \kappa\}$ is a family of open dense subsets of X then $\cap\{U_\alpha : \alpha\} \neq \varnothing$.

X Ordinal numbers arithmetic 33 / Addition.

Definition 33.1 Let $(S, \leq_S)$ and $(T, \leq_T)$ be two *disjoint* well-ordered sets. We define the relation "$\leq_{S\cup T}$" on $S \cup T$ as follows:

a) $u \leq_{S\cup T} v$ if $\{u, v\} \subseteq S$ and $u \leq_S v$.

b) $u \leq_{S\cup T} v$ if $\{u, v\} \subseteq T$ and $u \leq_T v$.

c) $u \leq_{S\cup T} v$ if $u \in S$, $v \in T$.

Theorem 33.2 Let $(S, \leq_S)$ and $(T, \leq_T)$ be two *disjoint* well-ordered sets. Then the relation $\leq_{S\cup T}$ well-orders the set $S \cup T$.

Definition 33.3 Let α and β be two ordinal numbers. Let $(S, \leq_S)$ and $(T, \leq_T)$ be two *disjoint* well-ordered sets of order type α and β respectively.[1] We define $\alpha + \beta$ as follows:

$$\alpha + \beta = {}^{\text{ord}}(S \cup T,\ \leq_{S\cup T})$$

Theorem 33.4 Let $(S, \leq_S)$, $(T, \leq_T)$ and $(U, \leq_U)$, $(V, \leq_V)$ be two pairs of *disjoint* well-ordered sets such that

[1] Addition can also be defined inductively as follows: For all α and β, a) $\beta + 0 = \beta$, b) $\beta + (\alpha + 1) = (\beta + \alpha) + 1$, c) $\beta + \alpha = \text{lub}\{\beta + \gamma : \gamma < \alpha\}$ whenever α is a limit ordinal.

$$\begin{aligned} {}^{\text{ord}}S &= \alpha = {}^{\text{ord}}U \\ {}^{\text{ord}}T &= \beta = {}^{\text{ord}}V \end{aligned}$$

Then ${}^{\text{ord}}(S \cup T, \leq_{S \cup T}) = \alpha + \beta = {}^{\text{ord}}(U \cup V, \leq_{U \cup V})$. Hence addition of ordinal numbers is well-defined.

Theorem 33.5 Let α, β and γ be three ordinal numbers. Then:

a) $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ (Addition is associative.)
b) For any ordinal $\gamma > 0$, $\alpha < \alpha + \gamma$
c) For any ordinal γ, $\gamma \leq \alpha + \gamma$
d) $\alpha < \beta \Rightarrow \alpha + \gamma \leq \beta + \gamma$
e) $\alpha < \beta \Rightarrow \gamma + \alpha < \gamma + \beta$
f) $\alpha + \beta = \alpha + \gamma \Rightarrow \beta = \gamma$ (Left term cancellation is acceptable.)
g) $\alpha + 0 = \alpha$

Theorem 33.6 Let β be a limit ordinal. Then, for any ordinal, α,

$$\alpha + \beta = \sup \{\alpha + \gamma : \gamma < \beta\}$$

X Ordinal numbers arithmetic 34 / Multiplication

Definition 34.1 Let $(S, \leq_S)$ and $(T, \leq_T)$ be two well-ordered sets. We define the *lexicographic ordering* on the Cartesian product $S \times T$ as follows:

$$(s_1, t_1) \leq_{S \times T} (s_2, t_2) \text{ provided } \begin{cases} s_1 <_S s_2 \\ \quad \text{or} \\ s_1 = s_2 \quad \text{and} \quad t_1 \leq_T t_2 \end{cases}$$

Theorem 34.2 Let $(S, \leq_S)$ and $(T, \leq_T)$ be two well-ordered sets. The lexicographic ordering of the Cartesian product $S \times T$ is a well-ordering.

Theorem 34.3 If the well-ordered sets S_1 and S_2 are order isomorphic and the well-ordered sets T_1 and T_2 are order isomorphic then the lexicographically ordered Cartesian products $S_1 \times T_1$ and $S_2 \times T_2$ are order isomorphic.

Definition 34.4 Let α and β be two ordinals with set representatives A and B respectively. We define the multiplication $\alpha \times \beta$ as:

$$\alpha \times \beta = {}^{\text{ord}}(B \times A)$$

The product $\alpha \times \beta$ is equivalently written as $\alpha\beta$, (respecting the order). Note the order of the terms in the Cartesian product $B \times A$ is different from the order $\alpha \times \beta$ of their respective ordinalities.

Theorem 34.5 Let α, β and γ be three ordinal numbers. Then:

a) $(\gamma\beta)\alpha = \gamma(\beta\alpha)$ (Multiplication is associative.)
b) For any $\gamma > 0$, $\alpha < \beta \Rightarrow \gamma\alpha < \gamma\beta$
c) $\gamma(\alpha + \beta) = \gamma\alpha + \gamma\beta$ (Left-hand distribution is acceptable.)
d) For any $\gamma > 0$, $\gamma\alpha = \gamma\beta \Rightarrow \alpha = \beta$ (Left-hand cancellation is acceptable.)
e) $\gamma 0 = 0$
f) For any limit ordinal $\beta \neq 0$, $\alpha\beta = \sup\{\alpha\gamma : \gamma < \beta\}$

Definition 34.6 Let γ be any non-zero ordinal. We define the γ-based exponentiation function $g_\gamma : \mathscr{O} \to \mathscr{O}$ as follows:

1) $g_\gamma(0) = 1$
2) $g_\gamma(\alpha^+) = g_\gamma(\alpha)\gamma$
3) $g_\gamma(\alpha) = \text{ lub}\{g_\gamma(\beta) : \beta < \alpha\}$ whenever α is a limit ordinal.

Whenever $\gamma \neq 0$ we represent $g_\gamma(\alpha)$ as γ^α. Then $\gamma^{\alpha+1} = \gamma^\alpha\gamma$. If $\gamma = 0$ we define $\gamma^\alpha = 0^\alpha = 0$.

Theorem 34.7 Let α, β and γ be three ordinal numbers. Then, assuming $\gamma > 1$,

$$\alpha < \beta \Leftrightarrow \gamma^\alpha < \gamma^\beta$$

Theorem 34.8 Let α, β and γ be three ordinal numbers where $\alpha \neq 0$.

a) $\gamma^\beta\gamma^\alpha = \gamma^{\beta+\alpha}$
b) $(\gamma^\beta)^\alpha = \gamma^{\beta\alpha}$

Appendix A / Boolean algebras and Martin's axiom.

Definition 0.1 A partially ordered set $(P, \leq)$ is called a *lattice* if $a \vee b = \max\{a, b\}$ and $a \wedge b = \min\{a, b\}$ both exist in P for all pairs a, b in P.

Definition 0.2 If B is a subset of a partially ordered set, $(P, \leq)$, $\vee B$ denotes the least upper bound of B and $\wedge B$ denotes the greatest lower bound of B (both with respect to $\leq$). Note that $\vee B$ and $\wedge B$ may or may not be an element of B. A lattice $(P, \leq)$ is said to be a *complete lattice* if for any non-empty subset B of P, both $\vee B$ and $\wedge B$ exist and belong to P.

Definition 0.3 Let X be a topological space. A subset B is said to be *regular open in* X if $B = \text{int}_X(\text{cl}_X(B))$. The set of all regular open subsets of X will be denoted as $\mathscr{R}o(X)$.

Theorem 0.4 Let X be a topological space. Then $(\mathscr{Ro}(X), \subseteq, \vee, \cap)$ is a complete lattice in $(\tau(X), \subseteq)$.

Definition 0.5 Let $(L, \leq, \vee, \wedge)$ be a lattice. An *L-ultrafilter* is a proper filter $\mathscr{F}$ in L which is not properly contained in any other proper filter in L. If the filter $\mathscr{F}$ is such that $\cap\{F : F \in \mathscr{F}\} \neq \varnothing$ then we say that the filter $\mathscr{F}$ is a *fixed ultrafilter*. Ultrafilters which are not fixed are said to be *free ultrafilters*.

Theorem 0.6 Let X be a topological space.

a) Suppose $\mathscr{F}$ is a proper L-filter where $(L, \subseteq, \vee, \wedge)$ is a lattice in $(\mathscr{P}(X), \subseteq)$. Then $\mathscr{F}$ can be extended to an L-ultrafilter.

b) Suppose $\mathscr{F}$ is an L-filter in $(L, \subseteq, \vee, \wedge)$ a lattice in $(\mathscr{P}(X), \subseteq)$. Then $\mathscr{F}$ is an L-ultrafilter if and only if for every $A \subseteq X$, either $A \in \mathscr{F}$ or $X - A \in \mathscr{F}$.

Theorem 0.7 Let X be a topological space. Then $(\mathscr{Ro}(X), \subseteq, \vee, \cap)$ is a complete lattice in $(\tau(X), \subseteq)$. An $\mathscr{Ro}(X)$-filter, $\mathscr{F}$, is an $\mathscr{Ro}(X)$-ultrafilter if and only if, for any $A \in \mathscr{Ro}(X)$, either A or $X - \mathrm{cl}_X(A)$ belongs to $\mathscr{F}$.

Definition 0.8 A lattice $(L, \vee, \wedge)$ is said to be a *distributive lattice* if, for any x, y, and z in L, $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$ and $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$.

· The lattice $(L, \vee, \wedge)$ is said to be a *complemented lattice* it has a maximum element, denoted by 1, and a minimum element, denoted by 0, and for every $x \in L$ there exists a unique x' such that $x \vee x' = 1$ and $x \wedge x' = 0$.

· A complemented distributive lattice is referred to as being a *Boolean algebra*. A Boolean algebra is denoted as $(B, \leq, \vee, \wedge, 0, 1, ')$ when we explicitly want to express what the maximum and minimum elements are.

Definition 0.9 Suppose we are given two lattices $(B_1, \leq_1, \vee_1, \wedge_1, 0, 1, ')$ and $(B_2, \leq_2, \vee_2, \wedge_2, 0, 1, ')$ and a function f which maps elements of B_1 to elements of B_2. We say that $f : B_1 \to B_2$ is a *Boolean homomorphism* if, for any $x, y \in B$,

1) $f(x \vee_1 y) = f(x) \vee_2 f(y)$,

2) $f(x \wedge_1 y) = f(x) \wedge_2 f(y)$

3) $f(x') = f(x)'$.

The function $f : B_1 \to B_2$ is a *Boolean isomorphism* if f is a bijection and both f and $f^{\leftarrow}$ are Boolean homomorphisms.

Definition 0.10 Let $(B, \leq, \vee, \wedge, 0, 1, ')$ be a Boolean algebra. Let $\mathscr{S}(B) = \{\mathscr{U} : \mathscr{U}$ is a B-ultrafilter$\}$. We define the function $f_B : B \to \mathscr{P}(\mathscr{S}(B))$ as follows: $f_B(x) = \{\mathscr{F} \in \mathscr{S}(B) : x \in \mathscr{U}\}$.

Theorem 0.11 Let $(B, \leq, \vee, \wedge, 0, 1, ')$ be a Boolean algebra. Then the set $\{f_B(x) : x \in B\}$ is a base for the open sets of some topology, $\tau(\mathscr{S}(B))$, on the set $\mathscr{S}(B)$ of all B-ultrafilters.

Theorem 0.12 Let $(B, \leq, \vee, \wedge, 0, 1, ')$ be a Boolean algebra.

1) The function $f_B : B \to \mathscr{P}(\mathscr{S}(B))$ is a Boolean homomorphism mapping B into $\mathscr{P}(\mathscr{S}(B))$.
2) For every $x \in B$, $f_B(x)$ is clopen in $\mathscr{S}(B)$. Hence $f_B[B] \subseteq \mathscr{B}(\mathscr{S}(B))$ (the set of all clopen sets in $\mathscr{S}(B)$).
3) The function $f_B : B \to \mathscr{S}(B)$ is a Boolean isomorphism mapping B into $\mathscr{B}(\mathscr{S}(B))$ (the set of all clopen sets in $\mathscr{S}(B)$).
4) The topological space $(\mathscr{S}(B), \tau(\mathscr{S}(B)))$ where $\tau(\mathscr{S}(B))$ is the topology generated by the open base $\{f_B(x) : x \in B\}$ is a compact zero-dimensional Hausdorff topological space.
5) The Boolean isomorphism $f_B : B \to \mathscr{S}(B)$ maps B onto $\mathscr{B}(\mathscr{S}(B))$ (the set of all clopen sets in $\mathscr{S}(B)$).

Theorem 0.13 Let κ be an infinite cardinal number such that $\kappa < 2^{\aleph_0}$. Then the following are equivalent:

1) (Martin's axiom MA) If $(P, \leq)$ is a partially ordered set satisfying *ccc* and $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$ is a family of dense subsets of P, then there exists a filter $\mathscr{F}$ on P such that $\mathscr{F} \cap D_\alpha \neq \varnothing$ for each $\alpha \leq \kappa$.
2) If X is a compact Hausdorff topological space satisfying *ccc* and $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$ is a family of dense open subsets of X, then $\cap\{D_\alpha : \alpha \leq \kappa\} \neq \varnothing$.
3) If $(B, \leq, \vee, \wedge, ')$ is a Boolean algebra with the *ccc* property and $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$ is a family of dense subsets of B, then there exists a filter $\mathscr{F}$ on B such that $\mathscr{F} \cap D_\alpha \neq \varnothing$ for each $\alpha \leq \kappa$.
4) If $(P, \leq)$ is a partially ordered set satisfying *ccc* and $|P| \leq \kappa$ and $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$ is a family of dense subsets of P, then there exist a filter $\mathscr{F}$ on P such that $\mathscr{F} \cap D_\alpha \neq \varnothing$ for each $\alpha \leq \kappa$.

Theorem 0.14 Let κ be a cardinal such that $\aleph_0 \leq \kappa < 2^{\aleph_0}$. Let X be a Hausdorff topological space satisfying *ccc* such that $\{x \in X : x \text{ has a compact neighbourhood}\}$ is dense in X. Suppose that $\mathscr{D} = \{D_\alpha : \alpha \leq \kappa\}$ is a family of dense open subsets of X. Then $\cap\{D_\alpha : \alpha \leq \kappa\}$ is dense in X if and only if Martin's axiom holds true.

Bibliography

1. Enderton, H.B. *Elements of set theory*, Academic Press, Inc., 1977.
2. Fremlin, D.H. *Consequences of Martin's axiom*, Cambridge University press, London, 1984.
3. Hrbacek, K., Jech, T. *Introduction to set theory*, Marcel Dekker, Inc., New York and Basel, 1984.
4. Holz, M., Steffens, K., Weitz, E. *Introduction to Cardinal Arithmetic*, Birkhäuser, Basel, 1999
5. Monk, J.D., *Introduction to set theory*, McGraw-Hill, New York, 1969
6. Roitman, J. *Introduction to modern set theory*, John Wiley & sons, New York, 1990.
7. Pinter, C. *Set theory*, Addison-Wesley Publishing Company, 1971,
8. Potter, M., *Set Theory and Its Philosophy: A Critical Introduction*, Oxford University Press, 2004
9. Porter, J. R., Woods, R. G. *Extensions and absolutes of Hausdorff spaces*, Springer-Verlag, 1988.
10. Willard, S. *General Topology*, Addison-Wesley Publishing Company, 1968.
11. Weese, M., Winfried, J. *Discovering Modern set theory*, American mathematical society, 1998.

ISBN 978-0-9938485-0-6

Index