

**Bertrand Russell and Numbers:
*Introduction to Mathematical
Philosophy: Chapters 1-7***

Bradley M. Cardona

May 9, 2023

Submitted to the Department of Mathematics in partial fulfillment of the requirements for the degree of Bachelor of Science.

First Reader: Dr. Anthony Lo Bello
Second Reader: Dr. Brent Carswell

I hereby recognize and pledge to fulfill my responsibilities, as defined in the Honor Code, and to maintain the integrity of both myself and the College community as a whole.

Pledge:

Bradley M. Cardona

Acknowledgements

My undergraduate years would not have been possible without the guidance, mentorship, and support from many individuals.

I am indebted, first and foremost, to my advisor, Professor Anthony Lo Bello, for his unwavering support, encouragement, and patience—not to mention playful humor—throughout this senior project process. His insightful feedback and guidance were invaluable in directing and shaping my research, which only helped to elevate my already-lofty respect for Bertrand Russell.

I owe a special word of thanks to Professor Bradley Hersh, with whom I spent the Summer of 2021 studying the genetics of fruit flies, and to Professor Caryn Werner, with whom I spent the Summer of 2022 exploring systems of algebraic curves in the projective plane.

I wish to extend many thanks to the several other mathematics professors under whom I studied during my time at Allegheny; namely, Professor Brent Carswell, Professor Harald Ellers, Professor Tamara Lakins, and Professor Rachel Weir, each of whose time, expertise, and valuable insights I will always be greatly appreciative of.

I also owe a great many thanks to my friends for their support, encouragement, and valuable feedback throughout my studies. For her constant support and ineffable friendship, I am forever indebted to Isabella James.

Finally, I would like to express my heartfelt gratitude to my family—my mother, Adrienne, and my dear siblings, Donna and Xavier—and loved ones, without whom none of this would be possible.

Abstract

We investigate *Introduction to Mathematical Philosophy* by Bertrand Russell, first published in 1919. This book is an accessible introduction to what Russell and Alfred North Whitehead wrote in *Principia Mathematica*, the famous three-volume work on the foundations of mathematics.

Contents

1	Bertrand Russell: Mathematician, Philosopher, Humanitarian, and Writer	5
2	Part I: Russell's <i>Introduction to Mathematical Philosophy</i>	7
2.1	Chapter 1: The Series of Natural Numbers	7
2.2	Chapter 2: Definition of Number	10
2.3	Chapter 3: Finitude and Mathematical Induction	14
2.4	Chapter 4: The Definition of Order	20
2.5	Chapter 5: Kinds of Relations	27
2.6	Chapter 6: Similarity of Relations	34
2.7	Chapter 7: Rational, Real, and Complex Numbers	40
3	Part II: The Modern Approach	50
3.1	Outline of the Set Theory needed for the Study of Numbers . .	50
4	Conclusion	89
	References	92

1 Bertrand Russell: Mathematician, Philosopher, Humanitarian, and Writer

Born into a family of the British aristocracy in Monmouthshire, United Kingdom, Bertrand Russell (1872–1970) is rightly hailed as a polymath of the twentieth century. A mathematician, philosopher, humanitarian, and writer, his influence spanned (and today still spans) a breadth of subjects.

As a mathematician, Russell is well-known for having discovered a paradox in set theory, aptly named Russell’s Paradox (See Chapter 3 below). This paradox was immediately recognized by mathematicians as a significant obstacle to naive set theory (the early version of set theory in which Russell found his paradox), and was one of several paradoxes that impelled them to create axiomatic set theory.

Russell is most celebrated in mathematics for having co-written, alongside fellow British mathematician Alfred North Whitehead, a three-volume mathematical text entitled *Principia Mathematica*. Written as a defense of logicism (the idea that mathematics is reducible to logic), this weighty tome was an impetus for research in the foundations of mathematics, leading eventually to the development of modern mathematical logic. Due to its being densely-worded and symbology-heavy, this work is sometimes taunted for having probably been read in its entirety by no one. It takes 762 pages before Russell and Whitehead prove definitively, for instance, that $1 + 1 = 2$.

Being someone who instinctively championed his own beliefs, Bertrand Russell was also a humanitarian, frequently criticizing political causes that he thought wrong. A crystal-clear instance in which he displayed his contrarian disposition occurred during World War I. During a lecture in 1918, Russell publicly denounced the idea of the United States entering the war on the United Kingdom’s side, thereby earning himself a six months’ stay at Brixton Prison. (It was at Brixton, in fact, where Russell wrote *Introduction to Mathematical Philosophy*, the primary book on which this senior project is based.) Though not a complete pacifist, Russell, during the Second World War, fervently advocated for the total abolition of nuclear weapons.

In the world of philosophy, Russell is no doubt most well-known for hav-

ing written *A History of Western Philosophy*, in which he spans the timeline from the philosophy of the early Greeks to that of John Dewey (1859–1952). Russell won the Nobel Prize in Literature for this book in 1950. Though this book has since been heavily critiqued by scholars of philosophy, its residual effect on me has remained deeply profound and influential. One chapter I was deeply struck by two years ago concerned the philosophy of Socrates, in which Russell explains the Socratic definition of philosophers as those who are “lovers of the vision of truth.” Believing then that mathematics is the means by which one can get closest to truth, I decided to switch my major to mathematics.

In the following pages I present, in Part I, a detailed summary, chapter by chapter, of Russell’s ideas about number in his *Introduction to Mathematical Philosophy*. Then, in Part II, I give an outline of the modern treatment of the topic following Robert Stoll and Paul Halmos. For the sake of brevity I omit the proofs, which are standard and can be easily found in the sources indicated. The purpose here is to show the influence of Russell in the way the modern treatment of the subject is organized. In both parts I add whatever comments and critical observations I find necessary or useful. Finally, in the conclusion, I summarize what I have discovered and indicate the most obvious instances of Russell’s influence.

2 Part I: Russell's *Introduction to Mathematical Philosophy*

2.1 Chapter 1: The Series of Natural Numbers

In Chapter 1 of *Introduction to Mathematical Philosophy*, Russell explains the aim of this book, in addition to why Giuseppe Peano's treatment of the natural numbers is not as complete as it first seems.

The familiar way of studying mathematics is in the “constructive” manner: from natural numbers and integers to rationals and real numbers; from addition and multiplication to differentiation and integration, and so forth to higher mathematics. In this book, Russell begins from the opposite direction. Rather than first building on top of the natural numbers, Russell attempts to reduce mathematics to its logical components. Although it is tempting to define or deduce from an initial assumption, Russell urges us to “ask instead what more general principle can be found, in terms of which what was our starting point can be defined or deduced” [3, 1]. Two approaches are therefore required to expand the scope our logical abilities, one “to take us backward to the logical foundations of the things that we are inclined to take for granted in mathematics,” the other “to take us forward to the higher mathematics.” [3, 2].

Russell begins the “backward” approach by discussing the notable work of the Italian mathematician Giuseppe Peano, who published his famous Peano axioms in the treatise *Arithmetices principia, nova methodo exposita*, Turin: Bocca Brothers, 1889. In this publication, Peano attempted to show that the entire theory of natural numbers could be derived from three “primitive ideas” (undefined terms) and five “primitive propositions” (postulates).

Peano's three primitive ideas are: 0, number, and successor.* By “successor,” he means the number that follows another number. By “number,” he means the “class”—or, more commonly, the “set”—of natural numbers.

Peano's five primitive propositions can be stated as follows:

*Peano actually started with 1 but we are starting with 0 because Russell does.

1. 0 is a number.
2. The successor of any number is a number.
3. No two numbers have the same successor.
4. 0 is not the successor of any number.
5. Any property which belongs to 0, and also to the successor of every number which has the property, belongs to all numbers.

Russell explains why Peano's treatment of the natural numbers is not as complete as it first seems. Why Peano's treatment is not the last word is explained by Russell in the quotation I give on page 9.

To begin, he observes that Peano's three primitive ideas are capable of an infinite number of different concrete interpretations. Consider, for example, the three following cases.

First, since "0" is not strictly defined, "0" can be taken as *any* other natural number, allowing us to arbitrarily pick the first natural number. For instance, let "0" (the number we commonly think of as the first natural number) be taken to mean 42, and let "number" be taken to mean all numbers from 42 onward in the series of natural numbers. Here, we find that all five of Peano's propositions are satisfied.

Second, since "number" is not strictly defined, let "number" mean what we normally call "even numbers", and let "successor" be what results from adding two to it. Here, all five of Peano's propositions are satisfied still.

Third, since "successor" is not strictly defined, let "0" mean the number 1, let "number" mean the set $\{1, \frac{1}{3}, \frac{1}{9}, \frac{1}{27}, \dots\}$, and let "successor" mean a "third." Again, all five of Peano's propositions are satisfied.

It is clear from these examples that "0" and "number" and "successor" are ideas that each have many different concrete interpretations, casting much doubt that Peano's five propositions should be taken as definite arithmetic truth. Russell refines this point by providing a generalization, proving that, given *any* series that is "endless, contains no repetitions, has a beginning,

and has no terms that cannot be reached from the beginning in a finite number of steps’—a series he calls a **progression**—we will have a set of terms satisfying Peano’s axioms [3, 8]. Peano’s five propositions, therefore, cannot be definitive arithmetically, since “each different progression will give rise to a different interpretation of all the propositions of traditional pure mathematics; [and] all these possible interpretations will be equally true” [3, 9].

Although Peano’s system assumes that we know what is meant by “0” and “number” and “successor”, Russell has shown that this is not so. It is true that this discovery might not impair pure mathematics, but it most certainly will impair arithmetic in daily life. Believing then that mathematics should lead us to pragmatic conclusions in addition to theoretical ones, Russell rightly notes, “We want ‘0’ and ‘number’ and ‘successor’ to have meanings which will give us the right allowance of fingers and eyes and noses” [3, 9]. We thus do not yet have an adequate basis for arithmetic: we do not know if there are any definite sets of terms verifying Peano’s axioms; moreover, we do not have numbers that can be used for counting common objects, which requires that they have a definite meaning.

2.2 Chapter 2: Definition of Number

The second chapter of Russell's *Introduction to Mathematical Philosophy* is dedicated to the definition of number.

In 1884, the German logician Friedrich Ludwig Gottlob Frege published *The Foundations of Arithmetic* (German: *Die Grundlagen der Arithmetik*), which investigates the philosophical foundations of arithmetic. Although this publication was largely ignored by his contemporaries, Russell believed that the correct definition of number was contained therein.

It is not uncommon, when attempting to define “number,” to mistakenly define “plurality,” which is altogether something different. A key detail about plurality is that it is not an instance of *number*, but of *some particular number*. For example, a pair of women is an instance of the number 2, and the number 2 is an instance of number; the pair, however, is not an instance of number. That is, the number 2 is *not* the pair comprised of Bella and Laine; rather, it is something that all pairs have in common, and which distinguishes them from other sets.

A set may be defined in two ways: (1) by **enumeration** or (2) by **a defining property**. A set would be defined by enumeration—that is, “by extension”—if we were to say “This set consists of Bella and Laine.” And a set would be defined by a defining property—i.e., by *intension*—if we were to say “students of Allegheny College” or “blonde-haired women.” Of these two types of definitions, the one by intension, as emphasized by Russell, is logically more fundamental. This is for two reasons; namely, that (I) “the extensional definition can always be reduced to an intensional one”; and that (II) “the intensional one often cannot even theoretically be reduced to the extensional one” [3, 24]. It is clear that (I) must be true, since the enumeration of the set consisting of Bella and Laine can be reduced to the defining property “ x is Bella or x is Laine,” where x is contained in the set. (In other words, this defining property is true for two x 's, namely, Bella and Laine.) Moreover, (II) must also be true, since a set may be *impossible* to enumerate.

Russell gives three reasons why it is important that a definition by intension is logically more fundamental than one by extension. First, numbers themselves form an infinite set, and hence cannot be defined by enumeration.

Second, the sets having a given number of elements themselves presumably form an infinite set. Third, we want to define ‘number’ in such a way that we can speak of the number of elements in an infinite set; and it necessarily follows that such a number must be defined by intension [3, 13].

A set is often interchangeable with a defining property of it. One difference between the two, however, is that there is only *one* set having a given set of elements, whereas there are always *many* different defining properties by which a given set may be defined. Knowing that defining properties are never unique is useful, since any defining property can be used in place of the set whenever uniqueness is not important.

A **family of sets** is a set whose elements are themselves sets. When deciding whether two sets should belong to the same family of sets, our first reaction might be to put them in the same family of sets if they have the same number of elements. But this way of thinking is incorrect. Although we are all used to the operation of counting, counting in itself is, logically speaking, a complex operation. Furthermore, counting the number of elements in a set is only possible when the set itself is finite. When we define number, then, we cannot assume that all numbers are finite—and even if we did, we still could not use counting to define numbers, since numbers themselves are used in counting.

Hence, we must invoke the concept of one-to-one relations. “A relation is said to be **one-to-one**,” says Russell, “when, if x has the relation in question to y , no other element x' has the same relation to y , and x does not have the same relation to any term other than y [3, 15]. Using this relation allows us to discover whether two sets have the same number of elements, even when we do not know what that number is. Consider, for instance, a world in which there is no polygamy or polyandry; in such a world there must necessarily be a one-to-one relation of husband and wife. This implies that the number of husbands must be equal to the number of wives, even though the exact number of husbands and wives is unknown.

The set of elements that have a given relation to something (i.e., the set of input values for which a function is defined) is called the **domain** of the relation; hence, husbands are the domain of the “husband to wife” relation. Conversely, the “wife to husband” relation is called the **converse** of the

“husband to wife” relation. The **converse domain (range)** of a relation is the domain of its converse; the set of wives, therefore, is the range of the “husband to wife” relation. Using these definitions, we may say that one set is **similar** to another when there is a one-to-one relation in which the one set is the domain, and the other is the range.

The similarity relation is **reflexive** (“every set is similar to itself”), **symmetrical** (“if a set α is similar to a set β , then β is similar to α ”), and **transitive** (“if α is similar to β , and β is similar to γ , then α is similar to γ ”) [3, 16].

A key detail that Russell points out is that the act of counting is only applicable to finite sets, and “depends upon and assumes the fact that two [sets] that are similar have the same number of [elements]” [3, 17]. (If we were to count 20 elements, for example, we would simply be showing that the set of these elements is similar to the set of numbers 1 to 20.) Hence, the notion of similarity is “logically presupposed” in the operation of counting, and the notion of similarity is for several reasons “logically simpler” than the operation of counting. For one, the notion of similarity does not require an order. (Above it was established that the number of husbands must be equal to the number of wives, even though the exact number of husbands and wives is unknown.) Moreover, the notion of similarity does not require that the sets which are similar should be finite. By way of illustration, if we had the natural numbers (excluding 0) on the one hand, and their respective reciprocals on the other hand, it is clear that we could map 2 to $\frac{1}{2}$, 3 to $\frac{1}{3}$, 4 to $\frac{1}{4}$, and so on, thus showing that these two sets are similar.

We can consequently use the notion of similarity to decide when two sets should belong to the same family of sets. Regardless of the number of elements a set may have, the sets that are similar to it will have the same number of elements. We may thus use *similarity* as a definition of “having the same number of elements.” Naturally we might think that the set of couples (say) is something different from the number 2. But, as Russell reassuringly adds, “there is no doubt about the set of couples: it is indubitable and not difficult to define, whereas the number 2, in any other sense, is a metaphysical entity about which we can never feel sure that it exists or that we have tracked it down” [3, 18]. It is therefore more sensible to use the set of couples, which we are sure of, than to use a slippery definition of the number 2. Accordingly,

we may state the following definition: “The **number of a set** is the set of all those sets that are similar to it” [3, 18]. It follows from this definition that the set of all couples will *itself* be the number 2. And we can thus say that “a **number** is anything which is the number of some set” [3, 19].

Hence, the main result of this chapter is the finding that the investigation into the meaning of number. The question “What is a number?” leads to the development of the Theory of Sets, relations, and equivalence relations.

2.3 Chapter 3: Finitude and Mathematical Induction

It was established in Chapter I that the theory of natural numbers can be defined if we know what is meant by “0,” “number,” and “successor.” But the natural numbers can actually be defined even if we only know what is meant by “0” and “successor.” To explain how this can be done, Russell differentiates “finite” from “infinite,” demonstrating why the method by which it is done cannot be applied in the case of the infinite.

Russell begins by stating that if we start with 0 and proceed stepwise from each number to its successor, then it is clear we can reach any specific number. For example, to reach the number 3 from 0, we could say “1 is the successor of 0, 2 is the successor of 1, and 3 is the successor of 2.” However, this is not enough to prove the general proposition that *all* such numbers can be reached in this way [3, 20]. Therefore, we must see whether there is another way by which this proposition can be proved. In answering this question, Russell first considers the numbers that can be reached using “0” and “successor.” If we say something such as “1 is the successor of 0, 2 is the successor of 1, and so on,” it is tempting to say that “and so on” means that the process of proceeding to the successor may be repeated a finite number of times. But this definition assumes that we know what is meant by a “finite number”—which has not yet been defined.

The answer to this problem, says Russell, lies in *mathematical induction*. Although mathematical induction was previously presented as a *principle*, Russell now shows that it is in fact a *definition*. To do so, he provides several new definitions.

A **hereditary property** in the natural number series is a property such that, whenever it belongs to a number n , it also belongs to $n + 1$, the successor of n .

A **hereditary set** is a set such that, whenever n is an element of that set, so is $n + 1$.

An **inductive property** is a hereditary property which belongs to 0.

An **inductive set** is a hereditary set of which 0 is an element.

The **posterity of a given natural number with respect to the relation ‘immediate predecessor’ (which is the converse of ‘successor’)** is the set of all those elements that belong to every hereditary set to which the given number belongs [3, 22]. The ‘posterity of 0,’ for instance, is the set which consists of those elements which belong to every inductive set. (Notice that ‘0’ is an element of the inductive set, and is thus an element of the ‘posterity of 0.’)

The ‘posterity of 0’ then is ‘the set of those elements (including 0) that can be reached from 0 by successive steps from next to next.’ However, Russell outlines the distinction between these two sets thus: “the notion of ‘the set of elements that can be reached from 0 by successive steps from next to next’ is *vague*, though it seems as if it conveyed a definite meaning; on the other hand, ‘the posterity of 0’ is *precise* and *explicit* just where the other idea is hazy” [3, 22]. Russell thus defines the “natural numbers” as **the posterity of 0 with respect to the relation ‘immediate predecessor.’**

Russell has thus defined one of Peano’s three primitive ideas in terms of the other two. Specifically, he has arrived at “number” using only “0” and “successor.” (Since “posterity of 0” is what we *meant* to mean when we spoke of “the elements that can be reached from 0 by successive steps from next to next,” it might be more specific to say that we have arrived at “number” using only “0” and “*immediate predecessor*.”) Additionally, two of Peano’s primitive propositions—namely, the one asserting that 0 is a number (proposition 1) and the one asserting mathematical induction (proposition 5)—have become unnecessary, since they result from the definition of “number.”

Using postulates 2 to 4 of Peano’s postulates, with the relation “immediate predecessor” in place of the relation “successor,” we may summarize the preceding discussion as follows. Russell has emended Peano’s Postulates to read: There exists a set X and a relation P (“immediate predecessor”) defined in X and an element $0 \in X$ such that

- If $y \in X$ and $y \neq 0$, there exists some $x \in X$ for which xPy .
- xPz and $yPz \implies x = y$, where $x, y, z \in X$.
- There is no $x \in X$ such that $sP0$.

The natural numbers, symbolized by \mathbb{N} , is the posterity of 0 with respect to P .

Russell has thus established that Peano's primitive idea of "number" can be defined in terms of the other two primitive ideas, namely, "0" and "successor," both of which can be defined by the general definition of number.

By the general definition of number, we can say "0" is the number of elements in a set that has no elements (this is often called the "empty set"), which is the set of all sets that are similar to the empty set; that is, the set whose only element is the empty set. (Hence, "0" is the set whose only element is the empty set.)

Using this same definition, we can also define "successor." Given any number n , let A be a set that has n elements, and let x be an element that is not an element of A . Then the set consisting of A with x added on will have $n + 1$ elements. We thus can state the following definition: "The **successor** of the *number of elements* in the set A is the number of elements in the set consisting of A together with x , where x is any element not belonging to the set" [3, 23]. In modern terminology, this defines successor as an operation in the set of cardinal numbers.

Above we established that two of Peano's primitive propositions—namely, propositions 1 and 5—become unnecessary, since they result from the definition of "number." This leaves us to prove the three remaining primitive propositions; specifically, that (2) the successor of any number is a number; that (3) no two numbers have the same successor; and that (4) 0 is not the successor of any number. According to Russell, (2) and (4) are easily proven; however, proving (3) is difficult if we assume that the total number of things in the universe is finite.

Let us first consider if the total number of things in the universe was *not finite*. We could say that for two numbers, say a and b , neither of which is the total number of things in the universe, that it is easy to prove that we cannot have $a + 1 = b + 1$ unless we have $a = b$. Hence, proving (3) poses no problem.

Let us now consider if the total number of things in the universe was *finite*. If the number of things in the universe was (say) 10, then there would be no set of 11 things, and the number 11 would be the empty set. (This is to

be contrasted with the number 0; the number 0 would be the set containing the empty set, whereas the number 11 would be the empty set.) Likewise, there would be no set of 12 things, and the number 12 would also be the empty set. Hence, the successor of 10 would be the same as the successor of 11, but 10 is clearly not the same as 11. Hence, proving (3) poses a problem.

We now know then that if we assume the number of things in the universe to be *not finite*, then we can define Peano's three primitive ideas, in addition to proving his five primitive propositions, "by means of primitive ideas and propositions belonging to logic" [3, 25]. "It follows," Russell says satisfyingly, "that all pure mathematics, in so far as it is deducible from the theory of the natural numbers, is only a prolongation of logic" [3, 25].

We have shown that the process of mathematical induction can be used to define the natural numbers; it will be useful to recognize, however, that this type of induction is generalizable. Recall that the natural numbers were defined as the posterity of 0 with respect to the relation of a number to its "immediate predecessor" (the converse of "successor"). If N is the "immediate predecessor" relation, then clearly any number a will have the relation N to $a + 1$. A property is "hereditary with respect to N ," or **N -hereditary**, if, whenever the property belongs to a number a , it also belongs to $a + 1$. Moreover, a number b will be said to belong to the "posterity of a with respect to the relation N " if b has every N -hereditary property belonging to a .

These definitions can be generalized to any other relation. Thus if R is any relation whatsoever, we can state the following definitions.

An **R -hereditary property** is a property such that, if it belongs to a term c , and c has the relation R to d , then it belongs to d . This is made more precise by the next definition, since "property" has not been defined.

An **R -hereditary set** is a set whose defining property is R -hereditary. (Recall that a "defining property" of a set is a property shared between all elements of that set.) That is, A is an R -hereditary set if $x \in A$ and $xRy \implies y \in A$.

An element c is an **R -ancestor** of the term d if d has every R -hereditary property that c has, provided c is a term which has the relation R to some-

thing *or* to which something has the relation R . (Such a definition helps us to avoid the situation where c has the relation to nothing and where nothing has the relation to c , in which case we do not want to say that c is an R -ancestor of d .)

The **R -posterity** of c is the set of all terms to which c is an R -ancestor.

From the foregoing definitions, it is clear that if an element is the ancestor of anything, then it is its own ancestor and belongs to its own posterity.

Let us now take R to be the relation “parent.” It is of note that, prior to Frege developing his generalized theory of induction, no one could define “ancestor” precisely in terms of “parent.” It would have involved the assumption that the number of things in consideration is finite. For instance, suppose we were given the following series:

$$A, B, C, \dots, X, Y, Z.$$

A beginner’s definition of “ancestor” in terms of “parent” would “naturally say that A is an ancestor of Z if, between A and Z , there are a certain number of people, $B, C \dots$, of whom B is a child of A , each is a parent of the next, until the last, who is a parent of Z ” [3, 26]. But this definition is not satisfactory unless we add that the number of intermediate terms is finite. This series begins with a series of letters with no end, and then ends with a series of letters with no beginning. Is C an ancestor of X ? It will be so, according to the beginner’s definition of ancestor suggested above. (The beginner’s definition allows us to have a series with an “infinite” number of intermediate terms.) However, it will not be so according to any definition which will give the idea of “finite” that we would like to define. For this reason, it is essential that the number of intermediaries (between any two terms in a series) is “finite.” But, as we saw, “finite” can be defined by means of mathematical induction.

Using Frege’s generalized theory of induction, therefore, we now can concretely define “ancestor” in terms of some ancestral relation. It is clear that it is simpler to define the ancestral relation generally, instead of defining it strictly initially for the case of the relation of n to $n + 1$, and then extending it to other cases (as in the case of mathematical induction).

Hence, we now understand that mathematical induction is a definition—not a principle. There are some numbers to which mathematical induction can be applied (for example, the natural numbers), and there are other numbers to which it cannot be applied (for example, the cardinal numbers). If “natural numbers” are defined as numbers that possess all inductive properties, it will follow that all numbers that possess all inductive properties are natural numbers.

Mathematical induction enables us to differentiate the “finite” from the “infinite,” and might be stated simply in the following way: “What can be inferred from next to next can be inferred from first to last” [3, 27]. But this statement is true only when the number of intermediate steps between first and last is finite. To elucidate the argument from “next to next,” and its connection with the idea of “finite,” Russell uses a perceptive analogy of the jerks of a goods train: “When a train is very long, it is a very long time before its last truck moves. If the train were infinitely long, there would be an infinite succession of jerks, and the time would never come when the whole train would be in motion. Nevertheless, if there were a series of trucks no longer than the series of [natural] numbers . . . every truck would begin to move sooner or later if the engine persevered, though there would always be other trucks further back which had not yet begun to move” [3, 28].

2.4 Chapter 4: The Definition of Order

In Chapter 4 of *Introduction to Mathematical Philosophy*, Russell seeks a definition of order.

When we think of the natural numbers, it is not uncommon to think of them in terms of their order of magnitude (0, 1, 2, 3, ...), but they are actually capable of an infinite number of other arrangements. Although one order—for instance, the order of magnitude—might be more familiar, others are equally valid. But whichever order we may choose, the resulting order will be one which the elements of the set certainly have, whether we choose to notice it or not.

It is important to recognize that order lies in a relation among the elements of the set, in respect of which some appear as “earlier” and some as “later.” If a set has many orders, then there are many relations among the elements of that set. But are there certain properties a relation must have in order to give rise to an order?

With respect to an ordering relation, we must be able to say, of two elements in a set, that one “precedes” and the other “follows.” In order to use “precedes” and “follows” in the way in which we should normally understand them, we require that such a relation is *asymmetrical*, *transitive*, and *connected* [3, 31].

An **asymmetrical** relation is one such that, if x precedes y , then y must not also precede x . For example, the relation “taller” is asymmetrical: if x is taller than y , then y is not taller than x .

A **transitive** relation is one such that, if x precedes y and y precedes z , then x precedes z . The relation “taller” is also transitive: if x is taller than y and y is taller than z , then x is taller than z . It should be noted that some relations are asymmetrical but not transitive, while other relations are transitive but not asymmetrical. An example of the former case is the relation “father”, and an example of the latter case is the relation “sameness of height.”

A **connected** relation is one such that, given *any* two elements of the

set which is to be ordered, there must be one which precedes and the other which follows. For instance, of any two integers, one is smaller and the other greater; but of any two complex numbers this is not true [3, 32].

Russell claims that whenever an order exists, some relation having these three properties can be found generating it [3, 32]. To demonstrate why this must be true, Russell introduces a few definitions:

A relation is an **aliorelative**, or “is contained in (or implies) diversity,” if no term has this relation to itself. The relation “greater” is an aliorelative; the relation “equal” is not.

If a given relation holds between x and y and between y and z , then the **square** of that relation is the one which holds between x and z . For instance, if the relation “father” holds between x and y and between y and z , then the square of that relation is “grandfather,” since x is the grandfather of z .

“The **domain** of a relation consists of all those terms that have the relation to something or other, and the **converse domain** consists of all those terms to which something or other has the relation” [3, 32].

“The **field** of a relation consists of its domain and converse domain together” [3, 32].

“One relation is said to **contain** or **be implied by** another if it holds whenever the other holds” [3, 32].

An asymmetrical relation is the same thing as a relation whose square is an aliorelative. (Take the asymmetrical relation “father.” The square of this relation is “grandfather,” and no term is the grandfather of itself.) An asymmetrical relation is always an aliorelative. (The relation “father” is aliorelative, since no term is the father of itself.) But an aliorelative is often not asymmetrical. (The aliorelative relation “is a sibling of” is not asymmetrical. If Donna is a sibling of Xavier, for instance, then it does not follow that Xavier is not a sibling of Donna.)

A transitive relation is one which is implied by its square. Therefore, the relation “ancestor” is transitive; but the relation “father” is not.

“A relation is ‘connected’ when, given any two different terms in its field, the relation holds between the first and the second or between the second and the first (not excluding the possibility that both may happen, though both cannot happen if the relation is asymmetrical.)” [3, 33].

The three properties of being (1) aliorelative, (2) transitive, and (3) connected, are mutually independent, since a relation may have two without having the third. For example:

- The relation “ancestor” satisfies (1) and (2), but not (3). (The field of the “ancestor” relation is all people, but it is not uncommon for person 1 not to be the ancestor of person 2, and, moreover, for person 2 not to be the ancestor of person 1.)
- The relation “less than or equal to,” among numbers, satisfies (2) and (3), but not (1).
- The relation “greater or less,” among numbers, satisfies (1) and (3), but not (2).

A **serial relation** is aliorelative, transitive, and connected; or, equivalently, asymmetrical, transitive, and connected. (Recall an asymmetrical relation is always an aliorelative.)

A **series** is the same thing as a serial relation.

Russell notes briefly that a series is the serial relation itself and not the field of a serial relation. It would be a mistake to consider the field of the relation as *the* series, as a field can have *multiple* series with different ordering relations. The serial relation determines both the field and the order, making it the series, but the field cannot be considered the series.

If P is a serial relation, then the phrase “ x precedes y ” refers to the relation between x and y , written as xPy . The relation P , then, emphasizing once more what was said above, must abide by three properties:

1. x cannot precede itself (P is aliorelative).
2. If x precedes y and y precedes z , then x must precede z (P is transitive).

3. If x and y are in the field of P , then either x precedes y or y precedes x (P is connected).

These three properties ensure that the characteristics of a series will also be present in the ordering relation, and vice versa.

The definition is purely logical and applies to any serial relation. Although a serial relation always exists where there is a series, it may not always be the most natural relation to consider as the generator of the series. For example, in the case of the natural number series, the relation of “immediate succession” between consecutive numbers is asymmetrical but *not* transitive or connected. (Hence, the relation of “immediate succession” is not serial.) However, *from* immediate succession we can *derive* the “ancestral” relation (considered in Chapter 3) by mathematical induction; and this relation is the same as the relation “less than or equal to” among the natural numbers. And the relation “less than,” excluding “equal to,” is what is needed to generate the series of natural numbers. This relation is defined as “**m is less than n**” when **n possesses every hereditary property possessed by the successor of m**. This relation is asymmetrical, transitive, and connected and orders the natural numbers. This order is known as the “natural order” or “order of magnitude.”

The generation of series by means of relations resembling that of n to $n + 1$ is very common. The generation of a series can be understood as the passing from one term to the next, as long as there is a next, or back to the one before, as long as there is one before [3, 35].

The **proper posterity of x with respect to R** is the set of all terms that possess every R -hereditary property possessed by *every* term to which x has the relation R . (This definition is slightly different from that of **R-posterity of x^\dagger** , so as to account for cases where there may be *many* terms to which x has the relation R . For example, there may be many children to whom one father has the relation “father of.”)

A term x is a **proper ancestor of a term y with respect to R** (or a **proper R-ancestor of y**) if y belongs to the proper posterity of x with

[†]The R -posterity of x is the set of all terms to which x is an R -ancestor (from page 19 above), i.e., the set of all terms that have every R -hereditary property that x has.

respect to R .

For the generation of series by the relation R between consecutive terms to be possible, the relation “proper R -ancestor” must be an aliorelative, transitive, and connected. This relation will always be transitive, but it may not always be aliorelative or connected, which would prevent the generation of a series.

For instance, let R be the relation of sitting on someone’s left at a round table at which there are twelve people. Then the proper R -posterity of each sitting person consists of everyone who can be reached by going around the table from left to right. Specifically, the proper R -posterity of each person includes everyone at the table, including the person himself. In such a case, though the relation “proper R -ancestor” is connected (given any two people sitting at the table, there must be one person that is the proper R -ancestor of the other), and the relation R itself is aliorelative (no person is seated to the left of himself), a series is not generated because the relation “proper R -ancestor” is not an aliorelative (each person belongs to the proper R -posterity of himself) [3, 36].

The question of when series can be generated by ancestral relations derived from relations of consecutiveness is important. If the relation R is a one-to-one (or many-to-one) relation, then the “proper R -ancestor” must be connected, and all that remains is to ensure that it is aliorelative.

There are several ways to generate series, but all of them require the identification of a serial relation. For example, let us consider the three-term relation “between,” which allows for the ordering of points in a straight line.

To define the relation “between,” we first need to consider three points on a straight line in ordinary space. There must be one of these points that lies between the other two. This is not true for points on a closed curve, like a circle, as we can travel from one point to another without passing through the third. The relation “between” is thus unique to open series (as opposed to cyclic series) and allows us to arrange points in a line in an ordered fashion.

Let’s suppose that we have two points a, b , such that the line (ab) consists of three parts (besides a and b themselves):

1. Points between a and b .
2. Points x such that a is between x and b .
3. Points y such that b is between a and y .

To ensure that the relation “between” can arrange the points on the line in a meaningful way, we need to make certain assumptions. These assumptions are:

1. If anything is between a and b , then a and b cannot be the same point.
2. Anything between a and b must also be between b and a .
3. Anything between a and b cannot be identical to either a or b .
4. If x is between a and b , then anything between a and x must also be between a and b .
5. If x is between a and b and b is between x and y , then b must be between a and y .
6. If x and y are between a and b , then they must be the same or x must be between a and y or between y and b .
7. If b is between a and x and also between a and y , then x and y must be the same or x must be between b and y or y must be between b and x .

Therefore, the concept of order can be generated by means of a three-term relation, such as the “between” relation. To effectively use the “between” relation to arrange points on a straight line, these seven properties must be made to ensure that the relationship is meaningful and can be used to order the points in a specific manner.

Russell observes that any three-term relation which verifies these properties give rise to series. By way of illustration, Russell considers the relation “to the left of.” If a is to the left of b , then the points on the line (ab) are defined as follows:

1. Those between which and b lies a —which we will call those to the left of a .

2. The point a itself.
3. Those between a and b .
4. The point b itself.
5. Those between which and a lies b —which we will call those to the right of a .

The definition of “to the left of” is given as follows: For two points x, y on a line (ab) , x is said to be **to the left of** y if one of the following cases holds:

1. Both x and y are to the left of a , and y is between x and a .
2. x is to the left of a , and y is a or b or between a and b or to the right of b .
3. x is a and y is between a and b or is b or is to the right of b .
4. Both x and y are between a and b , and y is between x and b .
5. x is between a and b , and y is b or to the right of b .
6. x is b and y is to the right of b .
7. Both x and y are to the right of b , and x is between b and y .

From the seven properties that were assigned to the relation “between,” says Russell, “it can be deduced that the relation ‘to the left of,’ as above defined, is a serial relation as we defined the term” [3, 43].

Cyclic order, such as that of the points on a circle, cannot be generated by three-term relations of “between.” In fact, a relation of four terms, called “separation of couples,” is needed to generate cyclic order. Given any four points on a circle— e.g., $a, b, x,$ and y —it is possible to separate them into two couples, say (a, b) and (x, y) , such that in order to get from “ a to b ” one must pass through either x or y , and in order to get from x to y one must pass through either a or b [3, 43]. This relation can generate a cyclic order, but the process is more complicated than generating an open order from “between.”

2.5 Chapter 5: Kinds of Relations

In Chapter 5 of *Introduction to Mathematical Philosophy*, Russell discusses the significance of the different types of relations.

Russell begins by emphasizing the importance of having a clear understanding of the various kinds of relations and their properties, as some properties may only be relevant for specific types of relations. One relation he considers is the serial relation, whose three properties, as discussed in Chapter 4, are asymmetry, transitivity, and connexity.

Asymmetry refers to the property of a relation that is incompatible with its converse, that is, if $x \rightarrow y$, then $y \not\rightarrow x$. Russell explains that it is possible to separate a symmetrical relation into two asymmetrical relations. Consider the symmetrical relation “spouse.” If we assume the spouse of a male is always female and the spouse of a female is always male, then the relation “spouse” can be separated into two asymmetrical relations, as follows:

1. By limiting the domain of “spouse” to males or by limiting the converse of “spouse” to females, we obtain the relation “husband.”
2. By limiting the domain of “spouse” to females or by limiting the converse of “spouse” to males, we obtain the relation “wife.”

The symmetrical relation “spouse” can be separated into two asymmetrical relations because there are two mutually exclusive sets, namely, “males” and “females,” such that, whenever the relation “spouse” holds between two people, one person is a member of “males” and one person is a member of “females.” Hence, the relation “spouse” with its domain confined to “males” will be asymmetrical, and so will the relation when its domain is confined to “females.” But such cases are rare. If we have a series of more than two terms, for instance, then all terms, except “the first and last (if these exist),” belong both to the domain and to the converse domain of the generating relation, so that a relation like husband, where the domain and converse domain do not overlap, is excluded.” [3, 43].

Russell then discusses the important question of how to construct relations that have certain useful properties by using operations on relations that only have rudimentary versions of these properties. It is relatively easy, for

instance, to construct transitivity and connexity in many cases when the original relation does not have these properties. For example, if R is any relation whatsoever, the “ancestral relation derived from R by generalized induction is transitive, and if R is a many-one relation, the ancestral relation will be connected if it is confined to the posterity of a given term” [3, 43].

However, it is much more difficult to construct asymmetry. The method used to derive the relation “husband” from the relation “spouse,” as mentioned above, cannot be used in the cases where the domain and converse domain overlap—in cases such as “greater,” “before,” or “to the right of.” [3, 43]. In these cases, a symmetrical relation can be obtained by adding the original relation and its converse, but it is not possible to go back to the original asymmetrical relation without the help of some asymmetrical relation. For example, the “greater” relation can be combined with its converse (the “less” relation) to form the “greater or less”—i.e., “unequal”—relation, which is symmetrical, but there is nothing in this relation to indicate that it is the sum of two asymmetrical relations.

From a classification perspective, asymmetry is a more important characteristic than being aliorelative. Asymmetrical relations are aliorelative, but the reverse is not true. (The aliorelative relation “unequal,” for example, is symmetrical.)

Russell then notes that it is possible to replace relational propositions with predicates so long as the relations are symmetrical. Symmetrical relations that are not aliorelative, if they are not transitive, may be regarded as asserting a *common predicate*; whereas symmetrical relations that are aliorelative may be regarded as asserting *incompatible predicates*’ [3, 44]. For example, the relation “similarity between sets,” used to define “numbers” in Chapter 2, is symmetrical and transitive yet not aliorelative. It is possible, although less simple, to regard the “number” of a collection as a predicate of the collection. In this case, two similar sets will have the same numerical predicate, while two sets that are not similar will have different numerical predicates. This method of replacing relations with predicates is not possible when the relations are asymmetrical, because “both sameness and difference of predicates are symmetrical” [3, 44]. Hence, asymmetrical relations are, according to Russell, “the most characteristically relational of relations, and the most important to the philosopher who wishes to study

the ultimate logical nature of relations” [3, 45].

Russell next provides a comprehensive overview of one-many relations. A one-many relation is a relation where at most one term is related to a given term. (Hence, *one-one* relations are a subset of *one-many* relations.) Examples of one-many relations include “father,” “mother,” and “square of.” Relations like “parent” and “square root of” are not one-many. (“Parent” is many-one or many-many; and “square root of” is many-one.)

In theory, all relations can be converted into one-many relations. For instance, consider the “less” relation among the natural numbers. For any number greater than 1, there will not be just one number that has the “less” relation to it, but a whole set of numbers that are less than it. This set, known as the **proper ancestry** of the number, is a one-many relation since each number determines a unique set of numbers that constitutes its proper ancestry. The proper ancestry of the number 2, for example, would be the set of numbers $\{0, 1\}$. But “proper ancestry” is a one-many relation (recall that a one-one relation *is* a one-many relation.), since each number determines a single set of numbers as constituting its proper ancestry. Therefore, says Russell, “the relation *less than* can be replaced by *being a member of the proper ancestry of*” [3, 45]. Sticking with the previous example, then, we may write

0 is a member of the proper ancestry of 2.

1 is a member of the proper ancestry of 2.

According to Russell, though, this reduction of a relation to a one-many relation does not provide a technical simplification and is not considered a philosophical analysis due to the notion that sets are “logical fictions.” Therefore, one-many relations will continue to be regarded as a special type of relation.

The concept of one-many relations is present in all phrases of the form “the so-and-so of such-and-such.” For instance, “the mother of John Stuart Mill” describes a person by means of a one-many relation to a specific term. As a person cannot have more than one mother, the phrase “the mother of John Stuart Mill” refers to a specific person, even if her identity is unknown.

It is worth noting that all mathematical functions arise from one-many relations; terms such as the “sine of x ,” are described through a one-many relation (in this case, “sine”) to a given term x , similar to “the mother of x ”:

$$x \xrightarrow{\text{sine}} \text{“sine of } x\text{”};$$

$$x \xrightarrow{\text{mother}} \text{“mother of } x\text{”}.$$

These functions are known as **descriptive functions**, which can be represented as “the term having the relation R to x ” or simply “the R of x ,” where R represents any one-many relation [3, 46]:

$$x \xrightarrow{R} \text{“}R \text{ of } x\text{”}.$$

The use of “the R of x ” as a descriptive term requires that x is a term to which something has the relation R , and that only one term has the relation R to x , because the use of “the” implies uniqueness. For example, we can talk about “the father of x ” if x refers to a human being except Adam and Eve, but not if x refers to a table or chair or any other object without a father.

Therefore, the existence of “the R of x ” is determined by there being only *one* term with the relation R to x . This occurs when x is part of the converse domain of R , but not otherwise. In mathematical terms, x is the “argument” of the function, and the term with the relation R to x —i.e., “the R of x ”—is the “value” of the function for the argument x . For a one-many relation R , the range of possible arguments for the function is the converse domain of R , and the range of possible values is the domain.

Important concepts in relation logic, such as converse, domain, converse domain, and field, are examples of descriptive functions. Russell introduces more examples as the discussion continues.

Above it was noted that one-one relations are a subset of one-many relation; in addition to knowing their formal definition, one-one relations are crucial to understand. The formal definition of one-one relations can be derived from that of one-many relations. On the one hand, one-one relations are defined as relations that are both one-many and many-one, i.e., “one-many relations which are also the converses of one-many relations” [3, 46]. One-many relations, on the other hand, can be defined as those such that,

if x has some relation to y , then there is no other term that has that same relation to y . Or, they can be defined as relations such that, given two terms x and x' , the terms to which x has the given relation and those to which x' has the given relation have no member in common.

The **relative product** of two relations, R and S , is a relation that holds between x and z when there is an intermediate term y , such that x has the relation R to y and y has the relation S to z :

$$x \xrightarrow{R} y \xrightarrow{S} z.$$

In the case of one-many relations, the relative product of the relation and its converse implies identity. For instance, if we take R to be the one-many relation “father” and we take S to be its converse (say, the relation “son”) it follows that x must be identical to z . For one-one relations, the relative product of the relation and its converse, *as well as* the converse and the relation, implies identity.

When a relation R exists, it is helpful to think of y as being reached from x through an “R-step” or “R-vector.” And in the same manner, x can be reached from y through a “backward R-step.” For one-many relations, then, an R -step followed by a backward R -step should bring you back to your starting point. However, this is not always the case for other relations, such as the relation of child to parent or grandchild to grandparent.

It should be noted that the relative product of two relations is not always *commutative*, meaning the relative product of R and S is not the same as the relative product of S and R . For example, the relative product of parent and sister is aunt, but the relative product of sister and parent is parent.

One-one relations establish a correspondence between two sets, term by term. This means that every term in one set has a corresponding term in the other set. The concept is easiest to understand when the two sets have no overlapping members, such as the set of husbands and the set of wives. In this case, it is clear which term represents the **referent** (“the term from which the relation goes”) and which represents the **relatum** (“the term to which the relation goes”) [3, 48]. For example, if x and y represent husband and wife, respectively, then with respect to the relation “husband,” x is the

referent and y is the relatum, but with respect to the relation “wife,” y is the referent and x is the relatum.

Relations can have a **sense**, which refers to the direction in which the relation goes. The sense of a relation that goes from x to y is opposite to the sense of the corresponding relation from y to x . This concept of a relation having a sense is fundamental and helps explain why order can be created through relation. The set of all possible referents in a relation is referred to as its **domain**, and the set of all possible relata is its **converse domain**.

However, it is not uncommon for the domain and converse domain of a one-one relation to overlap. For example, the relation between the first 10 positive integers (excluding 0) and the result of adding 1 to each, results in the same 10 positive integers but with 1 removed from the beginning and 11 added to the end. This relation of n to $n + 1$ is a one-one relation. Another example is the relation between the first 10 positive integers and their double, which results in 5 of the original 10 integers. The relation between a number and its double is also a one-one relation.

An especially interesting case occurs when the converse domain is only a part of the domain. For example, consider the relation “ $n + 1$ ” where the domain is all the natural numbers n , instead of just the first 10 positive integers. If we arrange two rows of numbers, such that the numbers in the domain are in the top row and the number in the converse domain are in the bottom row, we have:

$$\begin{array}{l} 1, 2, 3, 4, 5, \dots, n, \dots \\ 2, 3, 4, 5, 6, \dots, n + 1, \dots \end{array}$$

In each of these cases, all natural numbers are in the top row, but only some are in the bottom row. These types of relations, where the converse domain is a **proper part** (i.e., a part but not the whole) of the domain, are explored later by Russell when dealing with the concept of infinity.

Another type of relation is called a “permutation,” where the domain and converse domain are identical. For example, the six possible arrangements of (a, b, c) illustrate permutations. Each arrangement can be transformed into another by a correlation. For example, (a, b, c) can be transformed to (c, b, a) ,

if a is correlated with c , b is correlated with itself, and c is correlated with a . The combination of two permutations results in another permutation, and the permutations of a given set form a **group**.

These different types of correlations are important in different contexts. The uses of one-one correlations are especially important, and will be explored in the next chapter.

2.6 Chapter 6: Similarity of Relations

In Chapter 2, Russell defined two sets to be *similar* if they have the same number of terms, meaning there is a one-to-one correlation between them. In this chapter, Russell seeks to define a comparable relation between relations called “likeness.”

To define likeness, Russell employs the notion of correlation, assuming that “the domain of one relation can be correlated with the domain of the other, and the converse domain with the converse domain” [3, 52]. However, this is not enough for the desired resemblance between the two relations. What is desired is that whenever one relation holds between two terms, the other relation should hold between the **correlates** of those terms. (By “correlate,” what Russell means is that if x has some relation R to y , then, with respect to R , the correlate of x is y , and the correlate of y is x .)

Russell uses the example of a map to illustrate the concept of “likeness” between relations. If we say one place is *north* of another, then the place on the map corresponding to the one is *above* the place on the map corresponding to the other. Hence, writes Russell, the “space-relations in the map have ‘likeness’ to the space-relations in the country mapped” [3, 53]. And it is this connection between relations that he wishes to define.

In defining “likeness,” Russell imposes a constraint on the types of relations he will consider. Specifically, he considers only those relations that have **fields**, i.e., those that allow for the creation of a single set by combining the domain and converse domain. Russell uses the example of the relation “domain” to illustrate a case where this constraint does *not* hold, as it has all sets as its domain (every set is the domain of some relation) and all relations as its converse domain (every relation has a domain). However, sets and relations cannot be combined to create a single set since they are of different **logical types**. Russell does not delve into the complex topic of types, but emphasizes that it is important to recognize when we are avoiding it.

This raises the question: When does a relation have a field? Without being pedantic, Russell asserts that a relation only has a “field” if its domain and converse domain belong to the same logical type; in other words, if it is **homogeneous**. To provide a general sense of what he means by “logical

types,” Russell says that “individuals, [sets] of individuals, relations between individuals, relations between [sets], and relations of [sets] to individuals, and so on, are different types.” [3, 53]. The concept of likeness—which has yet to be defined—is not particularly useful, according to Russell, when applied to relations that are not homogeneous. Therefore, when defining likeness, he will simplify our task by referring to the “field” of one of the relations involved. (In other words, he will restrict himself to relations that *are* homogenous.) While this restriction limits the generality of our definition, Russell notes that it is not of any practical significance, and once mentioned, need not be remembered [3, 53].

The concept of likeness between two relations P and Q is defined as the existence of a one-to-one relation S that has the *field* of P as its domain and the *field* of Q as its converse domain. For every instance in which the relation P holds, there must be a corresponding instance in which the relation Q holds, and *vice versa*. Figure 1 makes this clearer.

This definition can be simplified by introducing the concept of a **correlator**. We say that S is a correlator of P and Q if S is a one-to-one relation that “has the field of Q as its converse domain, and is such that P is the relative product of S and Q and the converse of S ” [3, 54]. “Two relations P and Q ,” therefore, “are said to be similar, or to have **likeness**, if there exists at least one correlator of P and Q ” [3, 54].

Russell explains that relations that have likeness share *all* properties that are independent of the terms in their fields. For instance, if one relation is transitive, then the other is also transitive; and the same holds for other general properties of relations. Statements involving the *actual terms* in the field of a relation may not hold when applied to a similar relation; however, these statements can always be translated into *analogous statements* that do hold.

We are thus led to a problem in mathematical philosophy related to the interpretation of statements, where we may know the grammar and syntax of the statement, but not the vocabulary. The problem is stated as follows: What are the possible meanings of a statement whose vocabulary is unknown but whose grammar and syntax are known, and what are the meanings of the unknown words that would make it true? This question is significant be-

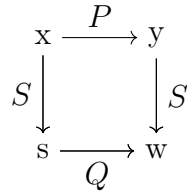


Figure 1: Likeness between two relations P and Q .

cause it reflects the state of our knowledge of nature, where we have a better understanding of the *form* of nature than the *matter*, and we only know that there is likely some interpretation of the terms used in scientific propositions that will make them approximately true. This question will be answered in a later chapter; for now, Russell says that we must further investigate the subject of likeness.

It is noted by Russell that the properties of similar relations are identical except for those dependent on the specific terms in their fields. To group these similar relations, the term **relation-number of a given relation** is introduced, which refers to the set of all relations similar to a given relation. (This definition reflects that of a “number” of a set in Chapter 2, which is the set of sets similar to that set.) More broadly, we may say that the **relation-numbers** are the set of all those sets of relations that are relation-numbers of various relations; in other words, a relation number is a set of relations consisting of all those relations that are similar to one element of the set. By establishing this terminology, says Russell, we can create a system for grouping and studying similar relations in a more structured way.

To avoid confusion with relation-numbers, Russell now uses **cardinal number** in place of the “number” of a set. Hence, the cardinal number of a set is the set of all sets that are similar to that set [3, 56].

The relation-numbers are applicable to series. Two series are considered equally long if they have the same relation-number. If we have two *finite* series whose fields have the same cardinal number of terms, then they will have the same relation-number. Hence, in the case of finite series, “there is a parallelism between cardinal and relation-numbers” [3, 56]. Russell uses the term “serial numbers” to refer to relation-numbers that are applicable to

series. Therefore, a finite serial number can be determined when the cardinal number of terms in the field of a series having that serial number is known. “If n is a finite cardinal number,” says Russell, “the relation-number of a series which has n terms is called the **ordinal number n** ” [3, 57]. But when the cardinal number of terms in the field of a series is *infinite*, the relation-number of the series cannot be determined by the cardinal number alone. In fact, an *infinite number* of relation-numbers can exist for *one* infinite cardinal number. This is because the “length” or relation-number of an infinite series can vary without a change in the cardinal number of terms. For example, the cardinal numbers of \mathbb{N} and \mathbb{Z} . (The cardinal number of both sets is \aleph_0 , but they[‡] have different relation-numbers because a correlator from \mathbb{N} to \mathbb{Z} cannot map 0 to an integer that has no predecessor.) In contrast, for finite series, the relation-number is uniquely determined by the cardinal number of terms in the field.

Arithmetic operations can be defined for relation-numbers just as they are for cardinal numbers. Russell considers the sum of two non-overlapping series. To define the sum of their relation-numbers as the sum of the relation-numbers of the two series, we must first order the series by placing one before the other. Let P and Q be the generating relations of the two series. In the sum of P and Q , with P preceding Q , every element of the field of P comes before every element of the field of Q . Therefore, the serial relation that we need to define as the sum of P and Q is not solely “ P or Q ,” but “ P or Q or the relation of any [element] of the field of P to any [element] of the field of Q ” [3, 57]. Assuming that P and Q do not overlap, this relation is serial, but “ P or Q ” is not serial because it is not connected.[§]

Series are not the only application of the idea of likeness. Russell has already mentioned maps, but extends our thoughts to geometry generally. When the system of relations applied by geometry to one set of terms can be brought fully into relations of likeness with a system applied to another set of terms, the resulting geometries are indistinguishable from a mathematical standpoint, i.e., “all the propositions are the same, except for the fact that

[‡]The successor relation in \mathbb{N} and the successor relation in \mathbb{Z} .

[§]A “connected” relation, as discussed in Chapter 4, is a relation such that, given any two different terms in its field, the relation holds between the first and the second or between the second and the first, not excluding the possibility that both may happen, though both cannot happen if the relation is asymmetrical.

they are applied in one case to one set of terms and in the other to another” [3, 58].

That said, Russell says that a mathematician need not be preoccupied with the specific nature or essence of points, lines, and planes, even when engaging in applied mathematics. While there is empirical evidence supporting some aspects of geometry that are not definitional, there is no empirical evidence regarding the true nature of a “point.” A point should satisfy our axioms as closely as possible, but need not necessarily be “very small” or “without parts” [3, 59]. As long as a logical structure, no matter how complex, can be constructed from empirical material that satisfies our geometrical axioms, it may legitimately be called a “point.” This illustrates the general principle that what is important in mathematics, and to a large extent in physical science, is not the *intrinsic nature* of our terms, but rather the *logical nature* of their interrelations.

We can describe two similar relations as having the same “structure.” For mathematical purposes, what matters about a relation is not its intrinsic nature, but the instances in which it holds. Just as a set can be defined by different but co-extensive concepts—Russell gives the example of “man” and “featherless biped”—two relations that are conceptually distinct can hold in the same set of instances. An “instance” in which a relation holds is a pair of terms with an order, such that one term comes first and the other second, and the first term has the relation in question to the second. If we consider the relation “mother,” for instance, we can define its “extension” as the set of all ordered pairs (x, y) in which x is the mother of y . Mathematically speaking, the only thing that matters about the relation mother is that it defines this set of ordered pairs. In general, we can say that the “extension” of a relation is the set of ordered pairs (x, y) in which x has the relation in question to y .

Russell takes a further step in the process of abstraction and examines what is meant by “structure.” If we are given a sufficiently simple relation, we can create a map of it. For example, we can consider a relation whose extension includes the following ordered couples: $ab, ac, ad, bc, ce, dc, de$, where a, b, c, d, e are five terms, regardless of what they are. We can make a “map” of this relation (see Figure 2) by placing five points on a plane and connecting them with arrows.

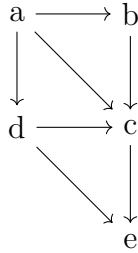


Figure 2: A map of the relation whose ordered couples are ab , ac , ad , bc , ce , dc , and de .

The map reveals the “structure” of the relation. It is evident that the “structure” of the relation does not depend on the specific terms that compose the field of the relation. The field can be altered without altering the structure, and the structure can be altered without altering the field. (Replacing a with f being an example of the former case. Adding the ordered couple ae being an example of the latter case.) Generally speaking, then, the “structure” of a relation is independent of the specific terms in its field. If two relations can be mapped onto each other or each can be its own map, they have the same “likeness” or relation-number. And the relation-number is equivalent to what is meant by the vague term “structure.”

Russell concludes this chapter by emphasizing the importance of understanding structure in philosophy to avoid speculation. The dismissal of objective counterparts to subjective concepts such as space, time, and phenomena suggests limited knowledge of the objective world. However, if objective counterparts exist, they must have the same structure as the phenomenal world; in other words, all true propositions about phenomena must also be true for the objective world, differing only in irrelevant individuality. While some philosophers avoid asserting objective counterparts, others are reserved on the subject to prevent excessive convergence between the real and phenomenal worlds. For these reasons, the notion of structure or relation-number is significant in many ways.

2.7 Chapter 7: Rational, Real, and Complex Numbers

Russell has so far provided definitions for cardinal numbers and relation-numbers, which has allowed us to define “ordinal” numbers. But these definitions are not sufficient to cover other types of numbers, such as negative, fractional, irrational, and complex numbers. Hence, in this chapter, Russell provides logical definitions for each of these types of numbers.

To begin, Russell notes that one of the reasons for the delay in discovering accurate definitions of number extensions is the mistaken belief that each type of extension includes the “previous sorts as special cases” [3, 63]. For instance, it was previously believed that positive integers could be identified with signless integers, and that a fraction with a denominator of 1 could be identified with its numerator. Similarly, rational numbers were thought to belong to the set of real numbers, and the complex numbers were thought to include real numbers with an imaginary part of zero. Russell, however, believes all these suppositions are incorrect and must be discarded for accurate definitions to be established.

Russell starts with **positive and negative integers**. He notes that $+1$ and -1 must both be relations, and must be each other’s converses. Specifically, $+1$ is the relation of n to $n + 1$, while -1 is the relation of $n + 1$ to n . Likewise, $+m$ is the relation of n to $n + m$, while $-m$ is the relation of $n + m$ to n , where m is any natural number. Russell emphasizes the distinction between $+m$ and m : $+m$ is a one-to-one relation so long as n is a cardinal number (finite or infinite), and m is a natural cardinal number. That is, $+m$ is distinct from m in that it is a *relation*—not a set of sets.

Russell then discusses **fractions**. Although Russell’s close friend Dr. Alfred North Whitehead developed a theory of fractions for their application to measurement, an easier method can be adopted for defining fractions that have the required mathematical properties. Russell defines the fraction $\frac{m}{n}$ as the relation that holds between two natural numbers x and y when $xn = ym$. (Russell defines fractions this way so that $\frac{x}{y} = \frac{m}{n}$.) For clarity, we can define $\frac{m}{n}$ as the relation

$$x \xrightarrow{\frac{m}{n}} y, \text{ when } xn = ym.$$

This definition allows us to prove that $\frac{m}{n}$ is a one-to-one relation if neither

m nor n is zero. It thus follows that the relation $\frac{m}{1}$ is a relation between two integers x and y , provided that $x = my$. That is,

$$x \xrightarrow{\frac{m}{1}} y, \text{ when } x = my.$$

However, like $+m$, the fraction $\frac{m}{1}$ is distinct from the natural cardinal number m , since “a relation and a [set] of [sets] are objects of utterly different types” [3, 64]. Note that Russell has thus defined what we call positive fractions.

Russell also mentions that the relation $\frac{0}{n}$ is always the same, regardless of the natural number n , and may be called the “zero of rational numbers.” That is,

$$x \xrightarrow{\frac{0}{n}} y, \text{ when } xn = 0, \text{ since } y \cdot 0 = 0.$$

However, $\frac{0}{n}$ is not identical to the cardinal number 0, since they are of different types.

Conversely, the relation $\frac{m}{0}$ is always the same, regardless of the natural number m , and may be called “the infinity of rationals.” That is,

$$x \xrightarrow{\frac{m}{0}} y, \text{ when } 0 = ym, \text{ since } x \cdot 0 = 0.$$

However, this type of infinity is different from the “Cantorian infinite,” which Russell discusses in the next chapter. While the infinity of rationals is not too important and could be dispensed with if necessary, the Cantorian infinite, according to Russell, “opens the way to whole new realms of mathematics and philosophy” [3, 65].

Russell notes that among fractions, zero and infinity are unique in that they are not one-one. Zero is one-many (since y can be any natural number), whereas infinity is many-one (since x can be any natural number).

Russell next defines **greater** and **less** among fractions. Given two fractions, $\frac{a}{b}$ and $\frac{c}{d}$, we say that $\frac{a}{b}$ is less than $\frac{c}{d}$ when ad is less than cb . That is,

$$\frac{a}{b} \xrightarrow{\text{less than}} \frac{c}{d}, \text{ when } ad \text{ is less than } cb.$$

The relation “less than” is serial, and therefore the fractions form a series in order of magnitude. The smallest term in this series is zero and the largest is infinity. However, if we omit zero and infinity from the series, there is no longer a smallest or largest fraction.

Any fraction other than zero and infinity can be shown to have a smaller and a larger fraction, and this implies there are always other fractions between any two fractions. For example, if $\frac{a}{b}$ and $\frac{c}{d}$ are two fractions, and $\frac{a}{b}$ is less than $\frac{c}{d}$, then $\frac{a+c}{b+d}$ will be greater than $\frac{a}{b}$ and less than $\frac{c}{d}$. This means that the series of fractions is “compact,” as there are always other terms between any two. A **compact series** is one in which there are always other terms between any two, and no two terms are consecutive. The fractions in order of magnitude, therefore, form a compact series, which is “generated purely logically without any appeal to space, time, or any other empirical datum” [3, 66].

Positive and negative fractions can be defined in a way similar to positive and negative integers. The sum of two fractions, $\frac{a}{b}$ and $\frac{c}{d}$, is defined as $\frac{ad+cb}{bd}$. We define $+\frac{c}{d}$ as the relation of $\frac{a}{b}$ to $\frac{a}{b} + \frac{c}{d}$, where $\frac{a}{b}$ is any fraction; and $-\frac{c}{d}$ is the converse of $+\frac{c}{d}$. That is,

$$\frac{a}{b} \xrightarrow[\frac{-c}{d}]{\frac{+c}{d}} \frac{a}{b} + \frac{c}{d}.$$

There are other possible ways to define positive and negative fractions, but this method is a clear adaptation of the way positive and negative integers are defined.

Russell next introduces the concept of **real numbers** as an extension of the idea of number, which includes irrational numbers. He references the discovery of **incommensurables** by Pythagoras—for example, the lengths of the diagonal and the side of a square are incommensurable, because they cannot be expressed as a ratio of two integers—and how, through geometry, this led to the idea of irrational numbers. He then discusses the proof in Euclid’s tenth book (Book X, Proposition 10) that there is no fraction whose square is 2. The inability to express the length of the diagonal of a square with a one-inch side in terms of a fraction, he argues, seems like a challenge to arithmetic by nature itself. The scope of this problem is not limited to

geometry; it also encompasses algebra when solving equations.

Russell then explains that it is possible to find fractions whose squares approach *closer and closer* to 2. We can construct an **ascending series** of fractions whose squares are all less than 2, but differing from 2 in their “later fractions” by less than any assigned amount. For instance, if we choose one-trillionth as the assigned amount, then after a certain term, say the fifteenth, all terms in the series will have squares that differ from 2 by less than this amount. Using the standard arithmetic rule to extract the $\sqrt{2}$, an infinite decimal can be obtained that satisfies these conditions. Similarly, it is possible to construct a **descending series** of fractions whose squares are all greater than 2, but differing from 2 in their “later fractions” by less than any assigned amount

However, this method does not lead to $\sqrt{2}$. By way of illustration, Russell divides all fractions into two sets: those whose squares are *less* than 2, and those whose squares are *greater* than 2. The former set has no maximum, and the latter set has no minimum. Hence, all fractions can be divided into two sets such that “all the terms in one set are less than all in the other, there is no maximum to the one set, and there is no minimum to the other” [3, 68]. This implies that there is nothing between these two sets, where the $\sqrt{2}$ should be located. Thus, although a tight “cordon” (as Russell calls it) has been drawn, it has not captured $\sqrt{2}$.

The foregoing method of dividing all the terms of a series into two sets, where one entirely precedes the other, is known as a **Dedekind cut**. With respect to what happens at the point of section, there are four cases:

1. The lower section has a maximum, the upper section a minimum.
2. The lower section has a maximum, the upper section no minimum.
3. The lower section has no maximum, the upper section a minimum.
4. The lower section has no maximum, the upper section no minimum.

Case 1 only occurs in series with consecutive terms (like the integers) and can be neglected. In case 2, the maximum of the lower section is the **lower limit** of the upper section. In case 3, the minimum of the upper section is the

upper limit of the lower section. In case 4, there is a **Dedekind gap**; that is, neither the lower section nor the upper section has a limit or a last term. In this case, Russell says there is an **irrational section**, “since sections of the series of fractions have ‘gaps’ that correspond to irrationals” [3, 70].

The delay in developing the true theory of irrationals, says Russell, was due to a mistaken belief that series of fractions must have **limits**. Therefore, the term “limit” must be defined thoroughly.

According to Russell, a set S has an **upper limit** L with respect to a relation R if three conditions are satisfied:

1. S has no maximum in R . (See definition of “maximum” below.)
2. Every element of S that belongs to the field of R precedes L . By “precedes,” Russell means “has the relation R to.”
3. Every element of the field of R that is before L precedes some element of S .

A term m is considered a **maximum** of a set S with respect to a relation R if m is an element of S and of the field of R , and does not have the relation R to any other element of S [3, 70].

Russell emphasizes that these definitions do not require the terms to which they are applied to be quantitative. In a series of moments of time arranged by earlier and later, for example, the “maximum” (if it exists) would be the last moment. However, if arranged by later and earlier, the “maximum” (if it exists) would be the first moment.

The **minimum** of a set with respect to R is defined as the maximum of S with respect to the *converse* of R , while the **lower limit** with respect to R is the upper limit with respect to the *converse* of R [3, 70].

Although the ideas of limit and maximum do not require the relation with respect to which they are defined to be serial, they are most often applied in serial or quasi-serial cases.

Another important idea is the “upper boundary,” which is the “maximum or upper limit” of a set of terms chosen from a series. That is, the **upper boundary** of a set is the maximum (“the last element”) if it has one, and if not, it is the upper limit (“the first term after all of them”), if it exists [3, 70]. If there is neither a maximum nor an upper limit, then there is no upper boundary. Similarly, the **lower boundary** is the minimum or lower limit.

Regarding the four kinds of Dedekind cuts, Russell notes that in the first three cases, each section has a boundary (either upper or lower), while in case 4, neither section has a boundary. Moreover, he notes that if the lower section has an upper boundary, then the upper section has a lower boundary. In cases 2 and 3, the two boundaries are identical, while in case 1, they are consecutive terms of the series.

Russell defines a series as **Dedekindian** if *every* section of the series has a boundary, whether upper or lower. The series of fractions in order of magnitude has “gaps” (for example, $\sqrt{2}$); hence, this series is not Dedekindian. Russell surmises that people have been influenced by spatial imagination and have believed that series must have limits “in cases where it seems odd if they do not” [3, 71]. For example, some people allowed themselves to “postulate” an irrational limit to fill the Dedekind gap when they realized there was no rational limit to the fractions whose square is less than 2. Dedekind himself, in fact, wrote the axiom that “the gap must always be filled,” meaning that every section has a boundary. Therefore, series that satisfy this axiom are called “Dedekindian.” There are an infinite number of series, however, for which this axiom is not verified.

Comparing the advantages of the method of “postulating” to those of theft, Russell wisely suggests using “honest toil” to find a precise definition of an irrational Dedekind cut. To do this, he says, we must rid ourselves of the idea “that an irrational must be the limit of a set of fractions” [3, 71]. Instead, Russell proposes defining a new kind of number called “real numbers,” which will include both rational and irrational numbers. Rational numbers correspond to fractions, “in the same kind of way in which $\frac{n}{1}$ corresponds to the integer n ; but they are not the same as fractions” [3, 72]. To represent an irrational number, we may use an irrational cut, which is represented by its lower section. In order to define rational numbers, then, Russell suggests confining ourselves to cuts in which the lower section has no

maximum—such cuts he calls **segments**. Since no segment has a maximum, we care only whether it has an upper limit; if it has an upper limit, then, by definition, it has an upper boundary.

On the one hand, a segment that corresponds to a *fraction* has an upper boundary (the fraction itself), since its upper limit is the fraction itself. Such a segment consists of all fractions less than their upper boundary. For example, consider the fraction $\frac{2}{3}$. The corresponding segment would be the set of all fractions that are less than $\frac{2}{3}$. This segment would include fractions like $\frac{1}{2}$, $\frac{3}{4}$, and $\frac{6}{10}$ —but not fractions like $\frac{2}{3}$, $\frac{7}{6}$, or 1.

On the other hand, a segment that corresponds to an *irrational number* has no upper boundary, since it has no upper limit. (Recall from above that we must rid ourselves of the idea “that an irrational must be the limit of a set of fractions.”)

Segments, whether they have a boundary or not, are such that, of any two associated with one series, one must be part of the other; hence, all segments can be arranged in a series by the relation of whole and part. A series with Dedekind gaps, where there are “segments without boundaries” (e.g., the irrational numbers), “will give rise to more segments than it has terms, since each term will define one segment having that term for boundary, and then the ‘segments without boundaries’ will be extra” [3, 72]. For example, consider the series

$$3, 3.1, 3.14, 3.141, 3.1415, 3.14159, \dots$$

If x is a term in this series, then x will define one segment having x as its boundary. However, there will be a gap at the area of a circle of radius 1, thus giving rise to segments without boundaries that will be extra. Russell has thus reached a point where he can define “real number,” “irrational number,” and “rational real number.”

A **real number** is “a segment of the series of fractions in order of magnitude” [3, 72]. This definition is equivalent to “a segment of the series of fractions which has (or does not have) an upper boundary.”

An **irrational number** is “a segment of the series of fractions that has

no [upper] boundary” [3, 72].

A **rational real number** is “a segment of the series of fractions which has [an upper] boundary” and thus “consists of all fractions less than a certain fraction, which corresponds to the rational real number. For example, the real number 1 is the set of all proper fractions [3, 72].

We might think intuitively that an irrational number is the limit of a set of fractions, but Russell declares that “it is actually the limit of the corresponding set of rational real numbers in the series of segments ordered by whole and part” [3, 73]. For instance, $\sqrt{2}$ —although it is not a segment of the series of fractions that has an upper limit—is the upper limit of all those segments of the series of fractions that correspond to fractions whose square is less than 2. In other words, says Russell, “ $\sqrt{2}$ is the segment *consisting of* all those [fractions] whose square is less than 2” [3, 73].

Russell notes that it is easy to prove that the *series of segments* of any series is Dedekindian. Given any set of segments, their boundary will be their “logical sum,” i.e., the set of all those terms that belong to at least one segment of the set.

Russell’s definition of real numbers is an example of **construction** as opposed to **postulation**, which was used to define the cardinal numbers. This method has the advantage of requiring no new assumptions and allows us to “proceed deductively from the original apparatus of logic” [3, 73].

The above definition allows us to easily define addition and multiplication. Given two real numbers u and v , each of which is a set of fractions, we take any element of u and any element of v and add them together according to the rule for the addition of fractions. The set of all such sums that can be obtained by varying the selected elements of u and v forms a new set of fractions, which is easy to prove is a segment of the series of fractions. This new set is defined as the sum of u and v . Hence, the **arithmetical sum of two real numbers** is “the set of the arithmetical sums of an element of the one and an element of the other chosen in all possible ways” [3, 73].

According to Russell, the arithmetical product of two real numbers can be defined in a similar way to the arithmetical sum, except that we instead

multiply an element of one by an element of the other in all possible ways. The set of fractions generated by this process is defined as the **product of the two real numbers**. Russell notes, however, that in these definitions, the series of fractions is defined so as to exclude 0 and infinity.

Complex numbers involve the square root of a negative number; hence, the letter i is used to represent the square root of -1 , and any number involving the square root of a negative number can be expressed in the form $x + yi$, where x and y are real numbers. The “real part” of such a number is x , while the “imaginary part” is i .

Complex numbers are less important in geometry than in algebra and analysis. In the latter two cases, they are required for the extraction of roots and the solution of equations. If we were operating in the complex numbers, for instance, the equation $x^2 + 1 = 0$ would have two roots; but if we were confined to real numbers, it would have zero roots. Although the creation of such “non-physical” numbers might at first seem disingenuous, Russell notes that “every generalization of number has presented itself as needed for some simple problem” [3, 74]. For example, negative numbers were needed so that subtraction could be possible, and fractions were needed for division. But “extensions of number are not *created* by the mere need for them: they are created by the definition” [3, 75]. Hence, Russell claims we must now turn our attention to the complex numbers.

Russell claims that a complex number can be defined “as an ordered couple of real numbers” [3, 75]. By defining complex numbers as ordered couples of real numbers, we can ensure that two real numbers are needed to determine a complex number and that two complex numbers are only equal if their corresponding couples of real numbers are equal. Further properties can be achieved by defining addition and multiplication rules as follows:

$$\begin{aligned} \textit{Addition} : (x + yi) + (x' + y'i) &= (x + x') + (y + y')i \\ \textit{Multiplication} : (x + yi)(x' + y'i) &= (xx' - yy') + (xy' + x'y)i \end{aligned}$$

Thus, given two ordered couples of real numbers (x, y) and (x', y') , their sum is the couple $(x + x', y + y')$, and their product is the couple $(xx' - yy', xy' + x'y)$. These definitions ensure that the ordered couples have the

desired properties. For example, the product of $(0, y)$ and $(0, y')$ is $(-yy', 0)$. Therefore, the square of $(0, 1)$ is $(-1, 0)$. The couples with an imaginary part of zero can be identified with real numbers, even though this “is an error in theory”, but “a convenience in practice” [3, 76]. Thus, $(0, 1)$ is represented by i , and $(-1, 0)$ by -1 . These multiplication rules ensure that the square of i is -1 , as desired. Hence, these definitions serves all necessary purposes.

Russell concludes this chapter by providing some other use cases of complex numbers. Complex numbers can be geometrically interpreted in the plane; and complex numbers of higher orders have uses in geometry. In the latter case, complex numbers can be defined as one-many relations whose domain consists of certain real numbers, and whose converse domain consists of integers from 1 to n . This is usually denoted by the notation

$$(x_1, x_2, x_3, \dots, x_n),$$

where “the suffixes denote correlation with the integers used as suffixes, and the correlation is one-many, not necessarily one-one, because x_r and x_s may be equal when r and s are not equal” [3, 76]. This definition, if accompanied by a suitable multiplication rule, serves all purposes for which complex numbers of higher orders are needed.

Russell has thus completed his review on number extensions not involving infinity.

3 Part II: The Modern Approach

3.1 Outline of the Set Theory needed for the Study of Numbers

Chapter 1: Sets and Relations

1.1: Cantor's Concept of a Set

Set and *element of a set* are primitive or undefined terms, and attempts to define them are futile, like Euclid's attempt to define *point*, *line*, and *plane*. Such attempts, such as those of Cantor, characterize what is called *Naive* or *Intuitive Set Theory*.

We say a **set** S is any collection of definite, distinguishable objects of our intuition or of our intellect to be conceived as a whole [4, 2].

Moreover, we say the objects in a set S are called the **elements** or **members** of S [4, 2].

Assumptions about sets are called axioms, which are statements that are taken to be true without proof.

1.2: The Basis of Intuitive Set Theory

Axiom. [4, 4]. **The intuitive principle of extension** states that two sets are equal if and only if they have the same elements.

This intuitive principle of extension is called by Halmos the **Axiom of Extension** [2, 2].

Definition. [4, 5]. The set $\{x\}$, a so-called **unit set**, is the set whose sole member is x .

Definition. [4, 5]. A **collection of sets** is a set whose elements are themselves sets.

Axiom. [4, 6]. **The intuitive principle of abstraction** states that a formula $P(x)$ defines a set A by the convention that the elements of A are

exactly those objects a such that $P(a)$ is a true statement.

This intuitive principle of abstraction is called by Halmos the **Axiom of Specification**. This axiom states that to every set A and to every condition $S(x)$ there corresponds a set B whose elements are exactly those elements x of A for which $S(x)$ holds [2, 6].

Russell's Paradox is a well-known paradox in set theory that arises from the Intuitive Principle of Abstraction (the Axiom of Specification). Using Russell's terminology, we define a set A to be **normal** if $A \notin A$, a set A to be **abnormal** if $A \in A$, and a set \mathcal{A} to be the set of all normal sets. Russell's paradox arises from the question: Is \mathcal{A} normal or abnormal?

If \mathcal{A} is normal, then it is abnormal—a contradiction. Conversely, if \mathcal{A} is abnormal, then it is normal—again a contradiction.

Russell's paradox is a consequence of a fundamental issue in naive set theory, which is that the unrestricted ability to form sets based on arbitrary properties leads to paradoxes. These paradoxes demonstrate the need for axiomatic set theory, which provides a rigorous and formal framework for avoiding such paradoxes. Such an apparatus was formulated by Zermelo as follows.

Zermelo's Axioms for Set Theory: There exist sets (denoted A, B, C, \dots) and there exist elements of sets (denoted a, b, c, \dots) satisfying the following postulates:

Axiom of Extension: If for all x , $x \in A$ implies $x \in B$, and $x \in B$ implies $x \in A$, then $A = B$.

Existence of the Null Set: There exists a set \emptyset with no elements.

Axiom of Pairing: If a and b are elements of sets A and B respectively, there exists the set $\{a, b\}$.

Axiom of Infinity: If A and B are sets and $x \in B$ but $x \notin A$, then there exists the set $A \cup \{x\}$.

Axiom of Complements: If A and B are sets, then there exists the set $A - B$ consisting of all elements of A that are not also elements of B .

Axiom of Union: If A and B are sets, there exists the set $A \cup B$ of all elements that are either elements of A or of B or of both A and B . The strong form of this axiom is: If \mathcal{B} is any collection of sets, there exists the set $\bigcup \mathcal{B}$ consisting of all elements that belong to some member of \mathcal{B} .

Axiom of Intersection: If A and B are sets, there exists the set $A \cap B$ of all elements that are elements both of A and of B . The strong form of this axiom is: If \mathcal{B} is any collection of sets, there exists the set $\bigcap \mathcal{B}$ consisting of all elements that belong to every member of \mathcal{B} .

Axiom of the Power Set: If X is a set, there exists the set $\mathcal{P}(X)$, the set of all subsets of X .

Axiom of Separation (or Abstraction or Specification): Rather, If X is a set, and if ϕ is a property that an element of X either has or does not have, then there exists the subset of X consisting of all elements of X that have property ϕ .

Axiom of the Cartesian Product: If A and B are sets, there exists the set $A \times B$ consisting of all elements $\{a, \{a, b\}\}$ where a is an element of A and b is an element of B . The strong form of this axiom follows from the Axiom of Choice. (See below.)

Axiom of Foundation: Every non-empty set A has an element x such that $A \cap \{x\} = \emptyset$. (This axiom rules out the existence of “the set of all sets.”)

Once one has defined what a function f from a set A to a set B is, one may state the Axioms of Replacement and of Choice.

Axiom of Replacement: If A and B are sets and f is a function from A to B , then there exists the set $f(A)$.

Axiom of Choice: If \mathcal{B} is any collection of sets, then there exists a set B consisting of one element each chosen in any way from every set in \mathcal{B} .

1.3: Inclusion

Definition. [4, 10]. If A and B are sets, then A is **included in** B , symbolized by

$$A \subseteq B,$$

if and only if each element of A is an element of B . This is synonymous with saying A is a **subset** of B . Moreover, this is the same as saying B **includes** A , symbolized by

$$B \supseteq A.$$

Hence, $A \subseteq B$ and $B \supseteq A$ each means that, for all x , if $x \in A$, then $x \in B$.

Definition. [4, 10]. The set A is **properly included** in B , symbolized by

$$A \subset B$$

(or, alternatively, A is a **proper subset** of B , and B **properly includes** A), if and only if $A \subseteq B$ and $A \neq B$.

Definition. [4, 11]. The intuitive principle of extension implies that there can be only one set with no elements. We call this set the **empty set** and symbolize it by

$$\emptyset.$$

The existence of the empty set is really an axiom.

Axiom: There exists the “empty set” \emptyset containing no elements.

Axiom. [4, 11]. The set of all subsets of a set A is the **power set** of A , symbolized by

$$\mathcal{P}(A).$$

Thus, $\mathcal{P}(A)$ is an abbreviation for

$$\{B \mid B \subseteq A\}.$$

The existence of the power set is called by Halmos the **Axiom of Powers**. This axiom states that there exists a collection of sets that contains among its elements all the subsets of a given set [2, 10].

Axiom. [2, 9]. **Axiom of Pairing.** For any two sets there exists a set that they both belong to.

1.4: Operations on Sets

Axiom. [4, 12]. The **union** (**sum**, **join**) of the sets A and B , symbolized by $A \cup B$ and read “ A union B ” or “ A cup B ,” is the set of all objects which are elements of either A or B ; that is,

$$A \cup B = \{x | x \in A \text{ or } x \in B\}.$$

The existence of the union is called by Halmos the **Axiom of Unions**. This axiom states that for every collection of sets there exists a set that contains all the elements that belong to at least one set of the given collection [2, 12].

Zermelo’s Axiom. If A is a set and $x \notin A$, there there exists a set $A \cup \{x\}$.

This postulate allowed Zermelo to present an alternative definition of the natural numbers to that given by Peano. Zermelo’s definition of the **natural numbers** is:

- 0 is the empty set \emptyset .
- 1 is the set $\{\emptyset\}$.
- 2 is the set $\{\emptyset, \{\emptyset\}\}$.
- 3 is the set $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}$.
- \vdots
- etc.

The problem Russell sees with this definition is the symbol $\dot{\cdot}$, which means “we repeat this process indefinitely.”

Definition. [4, 13]. The **intersection** of the sets A and B , symbolized by $A \cap B$ and read “ A intersection B ,” is the set of all objects which are elements of both A and B ; that is,

$$A \cap B = \{x | x \in A \text{ and } x \in B\}.$$

The existence of the intersection is called by Halmos the **Axiom of Intersections**. This axiom states that for every collection of sets there exists a set that contains all the elements that belong to all sets of the given collection.

Definition. [4, 13]. Two sets A and B are **disjoint** if and only if $A \cap B = \emptyset$.

Definition. [4, 13]. Two sets A and B **intersect** if and only if $A \cap B \neq \emptyset$.

Definition. [4, 13]. A collection of sets is a **disjoint collection** if and only if each distinct pair of its element sets is disjoint.

Definition. [4, 13]. A **partition** of a set X is a disjoint collection α of nonempty and distinct subsets of X such that each element of X is an element of some (and, hence, exactly one) element of α .

Definition. [4, 13]. The **relative complement** of A with respect to a set X is $X \cap \bar{A}$; this is usually shortened to $X - A$, read “ X minus A .” Thus,

$$X - A = \{x \in X | x \notin A\},$$

that is, the set of the elements of X which are not elements of A .

The existence of the complement of A in X is really an axiom:

Axiom: If A is a set and $A \subseteq X$, then there exists the complement of A in X containing all elements of X not in A .

Definition. [4, 13]. The **symmetric difference** of sets A and B , symbolized by $A + B$, is defined as follows:

$$A + B = (A - B) \cup (B - A).$$

Definition. [4, 13]. If all sets under consideration in a certain discussion are subsets of a set U , then U is called the **universal set** (for that discussion).

1.5: The Algebra of Sets

Definition. [4, 17]. The basic ingredients of the algebra of sets are various **identities**—equations which are true whatever the universal set U and no matter what particular subsets the letters (other than U and \emptyset) represent.

Theorem (5.1). [4, 17-18] For any subsets A, B, C of a set U the following equations are identities. Here, \bar{A} is an abbreviation for $U - A$.

$$A \cup (B \cup C) = (A \cup B) \cup C. \quad (1)$$

$$A \cap (B \cap C) = (A \cap B) \cap C. \quad (1')$$

$$A \cup B = B \cup A. \quad (2)$$

$$A \cap B = B \cap A. \quad (2')$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C). \quad (3)$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C). \quad (3')$$

$$A \cup \emptyset = A. \quad (4)$$

$$A \cap U = A. \quad (4')$$

$$A \cup \bar{A} = U. \quad (5)$$

$$A \cap \bar{A} = \emptyset. \quad (5')$$

Identities 1 and 1' are referred to as the **associative laws** for union and intersection, respectively, and identities 2 and 2' as the **commutative laws** for these operations. Identities 3 and 3' are the **distributive laws** for unions and intersection, respectively. Notice that each member of a pair is obtainable from the other member by interchanging \cup and \cap and, simultaneously, \emptyset and U .

Definition. [4, 18]. An equation, or an expression, or a statement within the framework of the algebra of sets obtained from another by interchanging \cup and \cap along with \emptyset and U throughout is the **dual** of the original.

Definition. [4, 19]. Accepting the fact that every theorem of the algebra of sets is deducible from 1 – 5 and 1' – 5', we then obtain the **principle of duality** for the algebra of sets: If T is any theorem expressed in terms of \cup , \cap , and $\bar{}$, then the dual of T is also a theorem.

Theorem (5.2). [4, 19-20]. For all subsets A and B of a set U , the following

statements are valid. Here, \bar{A} is an abbreviation for $U - A$.

$$\text{If, for all } A, A \cup B = A, \text{ then } B = \emptyset. \quad (6)$$

$$\text{If, for all } A, A \cap B = A, \text{ then } B = U. \quad (6')$$

$$\text{If } A \cup B = U \text{ and } A \cap B = \emptyset, \text{ then } B = \bar{A}. \quad (7, 7')$$

$$\bar{\bar{A}} = A. \quad (8, 8')$$

$$\bar{\emptyset} = U. \quad (9)$$

$$\bar{U} = \emptyset. \quad (9')$$

$$A \cup A = A. \quad (10)$$

$$A \cap A = A. \quad (10')$$

$$A \cup U = U. \quad (11)$$

$$A \cap \emptyset = \emptyset. \quad (11')$$

$$A \cup (A \cap B) = A. \quad (12)$$

$$A \cap (A \cup B) = A. \quad (12')$$

$$\overline{A \cup B} = \bar{A} \cap \bar{B}. \quad (13)$$

$$\overline{A \cap B} = \bar{A} \cup \bar{B}. \quad (13')$$

10 and 10' are the **idempotent laws**, 12 and 12' are the **absorption laws**, and 13 and 13' the **DeMorgan laws**. The identities 7, 7' and 8, 8' are each numbered twice to emphasize that each is unchanged by the operation which converts it into its dual; such formulas are call **self-dual**.

Theorem (5.3). [4, 20]. The following statements about sets A and B are equivalent to one another.

1. $A \subseteq B$.
2. $A \cap B = A$.
3. $A \cup B = B$.

1.6: Relations

Definition. [4, 24]. The **ordered pair** of x and y , symbolized by

$$(x, y),$$

is the set

$$\{\{x\}, \{x, y\}\},$$

that is, the two-element set one of whose elements, $\{x, y\}$, is the unordered pair involved, and the other, $\{x\}$, determines which element of this unordered pair is to be considered as being “first.”

Definition. [4, 24]. We call x the **first coordinate** and y the **second coordinate** of the ordered pair (x, y) .

Definition. [4, 25]. The **ordered triple** of x, y, z , symbolized by (x, y, z) , is defined to be the ordered pair $((x, y), z)$.

Definition. [4, 25]. Assuming that ordered $(n-1)$ -tuples have been defined, we take the **ordered n -tuple** of x_1, x_2, \dots, x_n , symbolized by (x_1, x_2, \dots, x_n) to be $((x_1, x_2, \dots, x_{n-1}), x_n)$.

Definition. [4, 25]. A **binary relation** is a set of ordered pairs, that is, a set each of whose elements is an ordered pair.

Definition. [4, 25]. If R is a relation, we write $(x, y) \in R$ and xRy interchangeably, and we say that x is **R -related** to y .

Definition. [4, 26]. If R is a relation, then the **domain** of R is

$$\{x \mid \text{for some } y, (x, y) \in R\}.$$

Definition. [4, 26]. If R is a relation, then the **range** of R , is

$$\{y \mid \text{for some } x, (x, y) \in R\}.$$

Definition. [4, 26]. The **Cartesian product**, denoted $X \times Y$, is the set of all pairs (x, y) , such that x is an element of some fixed set X and y is an element of some fixed set Y . Thus,

$$X \times Y = \{(x, y) \mid x \in X \text{ and } y \in Y\}.$$

Definition. [4, 26]. If R is a relation and $R \subseteq X \times Y$, then R is referred to as a **relation from X to Y** .

Definition. [4, 26]. A relation from Z to Z will be called a **relation in Z** .

Definition. [4, 26]. If X is a set, then $X \times X$ is a relation in X which we shall call the **universal relation** in X ; this is a suggestive name, since, for each pair x, y of elements in X , we have $x(X \times X)y$. At the other extreme is the **void relation** in X , consisting of the empty set. Intermediate is the **identity relation** in X , symbolized by ι or ι_x , which is $\{(x, x) | x \in X\}$.

Definition. [4, 27]. If R is a relation and A is a set, then $R(A)$ is defined by

$$R(A) = \{y \mid \text{for some } x \text{ in } A, xRy\}.$$

This set is suggestively called the set of **R -relatives** of elements of A .

1.7: Equivalence Relations

Definition. [4, 29]. A relation R in a set X is **reflexive** if xRx for each x in X .

Definition. [4, 29]. A relation R in a set X is **symmetric** if xRy implies yRx .

Definition. [4, 29]. A relation R in a set X is **transitive** if xRy and yRz implies xRz .

Definition. [4, 29]. A relation in a set is an **equivalence relation** if it is reflexive, symmetric, and transitive.

Definition. [4, 30]. If R is an equivalence relation on the set X , then a subset A of X is an **equivalence class** (R -equivalence class) if and only if there is a member x of A such that A is equal to the set of all y for which xRy .

Theorem (7.1). [4, 30]. Let R be an equivalence relation on X . Then the collection of distinct R -equivalence classes is a partition of X . Conversely, if \mathcal{P} is a partition of X , and a relation R is defined by aRb if and only if there exists A in \mathcal{P} such that $a, b \in A$, then R is an equivalence relation on X . Moreover, if an equivalence relation R determines the partition \mathcal{P} of X , then the equivalence relation defined by \mathcal{P} is equal to R . Conversely, if a partition \mathcal{P} of X determines the equivalence relation R , then the partition of X defined by R is equal to \mathcal{P} .

Definition. [4, 29]. The relation of **congruence modulo n in \mathbb{Z}** is defined for a non-zero integer n as follows: x is congruent to y , symbolized $x \equiv y \pmod{n}$, if and only if n divides $x - y$. This relation is an equivalence relation on \mathbb{Z} .

Definition. [4, 31]. A class of congruent numbers is often called a **residue class modulo n** .

Definition. [4, 31]. If R is an equivalence relation on X , we shall denote the partition of X induced by R by X/R (read “ X modulo R ”) and call it the **quotient set** of X by R .

Theorem (7.2). [4, 32]. A relation R is an equivalence relation if and only if there exists a disjoint collection \mathcal{P} of nonempty sets such that

$$R = \{(x, y) \mid \text{for some } C \text{ in } \mathcal{P}, (x, y) \in C \times C\}.$$

1.8: Functions

Definition. [4, 34]. A **function** is a relation such that no two distinct elements have the same first coordinate. Hence, f is a function if and only if it satisfies the following conditions:

1. The members of f are ordered pairs.
2. If (x, y) and (x, z) are elements of f , then $y = z$.

Synonyms for the word “function” are numerous and include **transformation**, **map** or **mapping**, **correspondence**, and **operator**.

Definition. [4, 34]. If f is a function and $(x, y) \in f$, so that xfy , then x is an **argument** of f .

Definition. [4, 34]. If f is a function and $(x, y) \in f$, so that xfy , there is a great variety of terminology for y ; for example, the **value** of f at x , the **image** of x under f , the element into which f carries x . There are also various symbols for y : xf , $f(x)$ (or, more simply, fx), x^f .

Definition. [4, 35]. A function f with range R_f is **into** Y if and only if $R_f \subset Y$, and f is **onto** Y if and only if $R_f = Y$.

Definition. [4, 35]. For corresponding notation for the domain of a function we shall say that f is **on** X when the domain of f is X . The symbols

$$f : X \rightarrow Y \text{ and } X \xrightarrow{f} Y$$

are commonly used to signify that f is a function on the set X into the set Y .

Definition. [4, 36]. If $f : X \rightarrow Y$, and if $A \subseteq X$, then $f \cap (A \times Y)$ is a function on A into Y (called the **restriction** of f to A and abbreviated $f|A$). Explicitly, $f|A$ is the function on A such that $(f|A)(a) = f(a)$ for a in A .

Definition. [4, 36]. Complementary to the definition of a restriction, the function f is an **extension** of a function g if and only if $g \subseteq f$.

Definition. [4, 36]. We denote the **identity map** on X as i_X .

Definition. [4, 36]. If $A \subseteq X$, then $i_X|A = i_A$. If $i_X|A$ is considered as a function on A into X , then it is the **injection mapping** on A into X .

Definition. [4, 36]. A function is called **one-to-one** if it maps distinct elements onto distinct elements. Symbolically, a function f is one-to-one if and only if

$$x_1 \neq x_2 \text{ implies } f(x_1) \neq f(x_2).$$

That is,

$$f(x_1) = f(x_2) \text{ implies } x_1 = x_2.$$

Because of the symmetrical situation that a one-to-one map on X onto Y portrays, it is often called a **one-to-one correspondence between X and Y** .

Definition. [4, 36]. Introducing the notation X^n for the set of all n -tuples (x_1, x_2, \dots, x_n) , where each x is a member of the set X , a function, whose domain is X^n and whose range is included in X , is an **n -ary operation in X** . In place of “1-ary” we shall say “unary”; in place of “2-ary,” we shall say “binary.”

1.9: Composition and Inversion for Functions

Definition. [4, 38-39]. The **composite** of functions f and g , symbolized

$$g \circ f$$

is the set

$$\{(x, z) \mid \text{there is a } y \text{ such that } xfy \text{ and } ygz\}.$$

This relation is a function, and this operation for functions is called (functional) **composition**.

Definition. [4, 41]. If f is one-to-one, the function resulting from f by interchanging the coordinates of members of f is called the **inverse function** of f , symbolized

$$f^{-1}.$$

This operation, which is defined only for one-to-one functions, is called (functional) **inversion**.

Definition. [4, 42]. A set of the form $f^{-1}[A]$ we call the **inverse** or **counter image** of A under f .

1.10: Operations for Collections of Sets

Axiom. [4, 43]. Let \mathcal{A} be a collection of sets. The **union** of \mathcal{A} is the set of all objects x such that x belongs to at least one set of the collection \mathcal{A} . That is, it is

$$\{x \mid x \in X \text{ for some } X \text{ in } \mathcal{A}\}.$$

This set is symbolized by

$$\bigcup \mathcal{A} \text{ or } \bigcup \{X \mid X \in \mathcal{A}\} \text{ or } \bigcup_{X \in \mathcal{A}} X.$$

As in the case of the union of two sets, the existence of the union of an arbitrary number of sets (possibly infinite) is really an axiom.

Axiom. [4, 44]. The **intersection** of a nonempty collection \mathcal{A} of sets is the set of all objects x such that x belongs to every set of the collection \mathcal{A} . That is, it is

$$\{x \mid x \in X \text{ for all } X \text{ in } \mathcal{A}\}.$$

This set is symbolized by

$$\bigcap \mathcal{A} \text{ or } \bigcap \{X | X \in \mathcal{A}\} \text{ or } \bigcap_{X \in \mathcal{A}} X.$$

Similarly to the situation with regard to arbitrary unions, the existence of the intersection of an arbitrary, possibly infinite, collection of sets is really an axiom.

Definition. [4, 45]. Let y be a function on a set I into a set Y . Let us call an element i of the domain I an **index**; I itself an **index set**; the range of y an **indexed set**; and the function y itself a **family**. We shall denote the value of y at i by y_i and call y_i the **i th coordinate** of the family. Thereby, we may write

$$y = \{(i, y_i) \in I \times Y | i \in I\}.$$

Once we shall have defined the natural numbers we will be able to make the following definition:

Definition. [4, 45]. A **sequence** is a family on the set of positive (or, nonnegative) integers into a set Y . That is, a sequence is a function for which $\{1, 2, \dots, n, \dots\}$ or $\{0, 1, \dots, n, \dots\}$ serves as an index set.

Definition. [4, 45]. By the phrase “a family $\{A_i\}$ of subsets of U ” we shall understand a function A on some set I of indices into $\mathcal{P}(U)$. The union of the range of such a family is called the union of the family $\{A_i\}$ or the union of the sets A_i . The standard notation for it is

$$\bigcup \{A_i | i \in I\} \text{ or } \bigcup_{i \in I} A_i \text{ or } \bigcup_i A_i.$$

Theorem (10.1). [4, 46]. Let $\{A_i\}$ with $i \in I$ be a family of subsets of U and let $B \subseteq U$. Then

$$B \cap \bigcup_i A_i = \bigcup_i (B \cap A_i) \text{ and } B \cup \bigcap_i A_i = \bigcap_i (B \cup A_i). \quad (1)$$

$$U - \bigcup_i A_i = \bigcup_i (U - A_i) \text{ and } U - \bigcap_i A_i = \bigcup_i (U - A_i). \quad (2)$$

If J is a subset of I , then

$$\bigcup_{j \in J} A_j \subseteq \bigcup_{i \in I} A_i \text{ and } \bigcup_{j \in J} A_j \supseteq \bigcup_{i \in I} A_i. \quad (3)$$

Axiom. [4, 47]. If $\{A_i\}$ with $i \in I$ is a family of sets, then the **Cartesian product** of the family, in symbols

$$\times \{A_i | i \in I\} \text{ or } \times_{i \in I} A_i \text{ or } \times_i A_i$$

is the set of all families $\{a_i\}$ with domain I and such that $a_i \in A_i$ for each $i \in I$.

The existence of the Cartesian product of an arbitrary (possibly infinite) collection of sets is called by Halmos the **Axiom of Choice**. This axiom states that the Cartesian product of a non-empty family of non-empty sets is non-empty [2, 59].

Definition. [4, 47]. Let $\{A_i\}$ with $i \in I$ be a family of sets and let A be its Cartesian product. If J is a subset of I , then there is a natural correspondence of the elements of A with those of $\times_{i \in J} A_i$. To formulate this explicitly, we use the fact that an element a of A is a family $\{a_i\}$ with I as domain. Then the element b , let us say, of $\times_{i \in J} A_i$ which is the natural correspondent of a is the restriction of a to J . We shall write b_i for a_i when $i \in J$. The function on A whose value at a is b is called the **projection** on A onto $\times_{i \in J} A_i$. If $J = \{j\}$ and p_j is the projection on A onto A_j , then $p_j(a) = a_j$, which is called the **j-coordinate** of a .

Definition. [4, 48]. A relation R in X is **antisymmetric** if and only if for each x and y in X the validity of xRy and yRx imply that $x = y$.

Definition. [4, 48]. A **partial ordering** in a set X is a reflexive, antisymmetric, and transitive relation in X . Note that it is custom to designate partial orderings by the symbol \leq .

Definition (4, 49). A relation R **partially orders** a set Y if and only if $R \cap (Y \times Y)$ is a partial ordering in Y [4, 49].

Definition. [4, 49]. If the relation \leq partially orders X , and x and y are elements of X , it may or may not be the case that $x \leq y$. If it is not, we write $x \not\leq y$. Additionally, we abbreviate “ $x \leq y$ and $x \neq y$ ” to “ $x < y$ ” and say x is **less than** y , or x **precedes** y , or y is **greater than** x . When it is convenient, we use $y \geq x$ and $y > x$ as alternatives for $x \leq y$ and $x < y$, respectively.

Definition. [4, 49]. A relation R in X is **irreflexive** (Russell’s “aliorelative”) if and only if for no x in X is xRx .

Definition. [4, 50]. A relation R is a **total** (or **simple** or **linear**) **ordering** if and only if it is a partial ordering such that xRy or yRx whenever x and y are distinct members of the domain (which is equal to the range) of R

Definition. [4, 50]. A relation R **totally orders** a set Y if and only if $R \cap (Y \times Y)$ is a total ordering in Y .

Definition. [4, 50]. A **partially ordered set** is an ordered pair (X, \leq) such that \leq partially orders X .

Definition. [4, 50]. A **totally ordered set** or **chain** is an ordered pair (X, \leq) such that \leq totally orders X .

Definition. [4, 52]. A function $f : X \rightarrow X'$ is **order-preserving (isotone)** relative to an ordering \leq for X and an ordering \leq' for X' if and only if $x \leq y$ implies $f(x) \leq' f(y)$.

Definition. [4, 52]. An **isomorphism** between the partially ordered sets (X, \leq) and (X', \leq') is a one-to-one correspondence between X and X' such that both it and its inverse are order-preserving. (Note that if we do not have an inverse function that can reverse an order-preserving function between X and X' , then we call such a function a **homomorphism**.) If such a correspondence exists, then one partially ordered set is an **isomorphic image** of the other, or, more simply, the two partially ordered sets are **isomorphic**.

Definition. [4, 53]. The **least element** of a set X relative to a partial ordering \leq is a y in X such that $y \leq x$ for all x in X . If such an element exists, then it is unique, so one should speak of *the* least element of X .

Definition. [4, 53]. A **minimal element** of a set X relative to \leq is a y in X , such that for no x in X is $x < y$. Such an element is not necessarily unique.

Definition. [4, 53]. The **greatest element** of a set X relative to \leq is a y in X such that $x \leq y$ for all x in X . If such an element exists, then it is unique, so one should speak of *the* greatest element of X .

Definition. [4, 53]. A **maximal element** of a set X relative to \leq is a y in X , such that for no x in X is $x > y$. Such an element is not necessarily unique.

Definition. [4, 53]. A partially-ordered set (X, \leq) is **well-ordered** if and only if each nonempty subset has a least element.

Definition. [4, 53]. If (X, \leq) is a partially ordered set and $A \subseteq X$, then an element x in X is an **upper bound** of A if and only if $a \leq x$ for all a in A . Similarly, an element x in X is a **lower bound** of A if and only if $x \leq a$ for all a in A . A set may have many upper and lower bounds.

Lemma. [2, 62]. **Zorn's Lemma.** If X is a partially ordered set such that every chain in X has an upper bound, then X contains a maximal element.

Definition. [4, 53]. An element x in X is a **least upper bound** or **supremum** for A , denoted $\sup A$, if and only if x is an upper bound for A and $x \leq y$ for any upper bound y for A . Hence, a supremum is an upper bound that is a lower bound for the set of all upper bounds.

Definition. [4, 53]. An element x in X is a **greatest lower bound** or **infimum** for A , denoted $\inf A$, if and only if x is a lower bound for A and $x \geq w$ for any lower bound w for A . Hence, an infimum is a lower bound that is an upper bound for the set of all lower bounds.

Chapter 2: The Natural Number Sequence and Its Generalizations

2.1: The Natural Number Sequence

The rigorous study of the natural numbers began with Peano. The central importance of his postulates is evident from the fact that Russell begins his *Introduction to Mathematical Philosophy* by discussing them.

Stoll's naive approach to the natural numbers system \mathbb{N} is as follows.

Definition. [4, 57]. The **successor function**, denoted $'$ (prime), is a primitive function, used in generating the natural numbers. We start with an initial object 0 (zero) and say that the *successor* of any object n already generated is another uniquely determined object n' . Therefore, the natural numbers \mathbb{N} appear as a set of objects

$$0, 0', (0')', ((0')')', (((0')')')' \dots$$

or, more simply,

$$0, 0', 0'', 0''' \dots$$

The transition to the usual notation is made upon introducing

$$0, 1, 2, 3 \dots$$

to stand for

$$0, 0', 0'', 0''' \dots,$$

and then employing decimal notation. The remainder of the description of this function can be expressed in two properties:

1. ' is a one-to-one mapping on \mathbb{N} into $\mathbb{N} \setminus \{0\}$.
2. if M is a subset of \mathbb{N} such that $0 \in M$ and $m' \in M$ whenever $m \in M$, then $M = \mathbb{N}$.

Property (2) is the basis for the **principle of induction** (or the **principle of weak induction**).

This approach is cleaned up a bit by the following formulation.

Modern Formulation of Peano's Postulates:

There exists a non-empty set \mathbb{N} , called the set of natural numbers.

- **Postulate 1:** There exists a natural number $0 \in \mathbb{N}$.[†]
- **Postulate 2:** There exists a function S , called the successor function, from \mathbb{N} to $\mathbb{N} \setminus \{0\}$.
- **Postulate 3:** The function S is one-to-one, that is, $S(m) = S(n) \implies m = n$.
- **Postulate 4:** There is no $m \in \mathbb{N}$ such that $S(m) = 0$.
- **Postulate 5:** The Principle of Mathematical Induction: If $M \subseteq \mathbb{N}$, $0 \in M$, and $n \in M \implies S(n) \in M$, then $M = \mathbb{N}$.

[†]Although Peano actually used the natural number 1 as the unique non-successor, we will, for the sake of consistency, use the natural number 0.

Postulates (1) through (4) give us a series of successors, resulting in an endless series of continually new numbers. And by the mathematical induction of (5), every number belongs to the series.

From these postulates it was customary to define the addition and multiplication of natural numbers, as follows.

Definition. Addition, $+$, is the function from $\mathbb{N} \times \mathbb{N}$ into \mathbb{N} defined by

$$\forall m, n \in \mathbb{N} : \begin{cases} m + 0 & = m \\ m + S(n) & = S(m + n). \end{cases}$$

Definition. Multiplication, \cdot , is the function from $\mathbb{N} \times \mathbb{N}$ into \mathbb{N} defined by

$$\forall m, n \in \mathbb{N} : \begin{cases} m \cdot 0 & = 0 \\ m \cdot S(n) & = m + (m \cdot n). \end{cases}$$

Russell saw various problems with these definitions. For example, are addition and multiplication both well-defined? That is, does a given number actually define the number we want it to define? How do we know, for instance, that $5 + 12 = 10 + 7$?

As a result of Russell's criticism of the Peano Postulate method of deriving the natural numbers, the following approach was developed.

Definition. [4, 58]. A triple (X, g, x_o) , where X is a set, g is a function on X into X (in other words, a **unary operation in X**), and x_o is an element of X , is a **unary system**.

Definition. [4, 58]. An **integral system** is a unary system (X, g, x_o) such that

1. g is a one-to-one mapping on X into $X - \{x_o\}$
2. if Y is a subset of X such that $x_o \in Y$ and $yg \in Y$ whenever $y \in Y$, then $Y = X$.

The existence of such an integral system is guaranteed by the **Axiom of Infinity**, which Zermelo was the first to recognize as necessary. This axiom states that there exists a set containing 0 and containing the successor of each of its elements [2, 44].

Hence, the natural number system of Peano's Postulates listed above may be summarized by the assertion that $(\mathbb{N}, S, 0)$ is an integral system.

Definition. [4, 59]. Two integral systems (X, s, x_0) and (Y, t, y_0) are **isomorphic** if there exists a one-to-one correspondence f between X and Y with $f(x_0) = y_0$ and $f(xs) = (fx)t$ for all x in X . This means that the elements of X can be paired with those of Y in such a way that successors of corresponding elements correspond

Definition. [4, 59]. Let (X, g, x_0) be a unary system. The **set of descendents of x_0 under g** (in symbols, $D_g x_0$) is the intersection of all subsets A of X , such that $x_0 \in A$ and $xg \in A$, whenever $x \in A$.

Lemma (1.1). [4, 59]. Let (X, g, x_0) be a unary system. Then $D_g x_0$ is the smallest subset of X which contains x_0 and which is closed under g . Alternatively, $x \in D_g x_0$ if and only if $x = x_0$ or there exists a y in $x \in D_g x_0$ such that $x = yg$.

Lemma (1.2). [4, 59]. Let (X, s, x_0) be an integral system and (Y, t, y_0) be a unary system. Define

$$s \nabla t : X \times Y \rightarrow X \times Y \text{ with } (x, y)s \nabla t = (xs, yt).$$

Then $(X \times Y, s \nabla t, (x_0, y_0))$ is a unary system. If f is the set of descendents of (x_0, y_0) under $s \nabla t$, then

1. f is a function on X into Y .
2. $fx_0 = y_0$ and $f(xs) = (fx)t$ for all x in X , and
3. f is uniquely determined by the properties in (2).

This result is called by Halmos the **Recursion Theorem**. It enables us to proceed to the non-naive definitions of the addition and multiplication of natural numbers.

Theorem (1.1). [4, 61]. Any two integral systems are isomorphic.

We now consider the particular integral system $(\mathbb{N}, S, 0)$ of Peano's Postulates.

Theorem (1.2). [4, 61]. Let B be a nonempty set, c be an element of B , and g be a function on $\mathbb{N} \times B$ into B . Then there exists exactly one function $k : \mathbb{N} \rightarrow B$ such that

$$k(0) = c \text{ and } k(S(n)) = g(n, k(n)).$$

Theorem (1.3). [4, 63]. The relation \leq well-orders \mathbb{N} .

Theorem (1.4). [4, 63]. For the integral system $(\mathbb{N}, S, 0)$ there exists exactly one function $\alpha : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ such that

1. for each n in \mathbb{N} , $\alpha(0, n) = n$, and
2. for all m and n in \mathbb{N} , $\alpha(S(m), n) = S(\alpha(m, n))$.

This function is **addition** in \mathbb{N} . We shall henceforth write $m + n$ instead of $\alpha(m, n)$.

Theorem (1.5). [4, 64]. Addition in \mathbb{N} has the following properties.

1. Associativity. For all m, n , and p in \mathbb{N} ,

$$m + (n + p) = (m + n) + p$$

2. Commutativity. For all m and n in \mathbb{N} , $m + n = n + m$.
3. Cancellation laws. For all m, n , and p in \mathbb{N} , $p + m = p + n$ implies $m = n$ and $m + p = n + p$ implies $m = n$.
4. For all m, n , and p in \mathbb{N} , $m \leq n$ if and only if there exists p in \mathbb{N} such that $p + m = n$.
5. For all m, n , and p in \mathbb{N} , $m < n$ if and only if $p + m < p + n$.
6. For all m and n in \mathbb{N} , $m + n = 0$ implies $m = 0$ and $n = 0$.

Theorem (1.6). [4, 66]. For the integral system $(\mathbb{N}, S, 0)$ there is exactly one function $\mu : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ such that

1. for each n in \mathbb{N} , $\mu(0, n) = 0$, and

2. for all m and n in \mathbb{N} , $\mu(S(m), n) = \mu(m, n) + n$.

This function is **multiplication** in \mathbb{N} . We shall henceforth write mn instead of $\mu(m, n)$.

Theorem (1.7). [4, 66]. Multiplication in \mathbb{N} has the following properties.

1. Associativity. For all $m, n,$ and p in \mathbb{N} ,

$$m(np) = (mn)p$$

2. Commutativity. For all m and n in \mathbb{N} , $mn = nm$.
3. Cancellation laws. For all $m, n,$ and p in \mathbb{N} , $p \neq 0$ and $pm = pn$ or $mp = np$ imply $m = n$.
4. Distributivity over addition. For all $m, n,$ and p in \mathbb{N} , $m(n + p) = mn + mp$ and $(n + p)m = nm + pm$.
5. For all $m, n,$ and p in \mathbb{N} , $p \neq 0$ implies that $m < n$ if and only if $pm < pn$.
6. For all m and n in \mathbb{N} , $mn = 0$ implies $m = 0$ or $n = 0$ or, what is equivalent, if $m \neq 0$ and $n \neq 0$, then $mn \neq 0$.

Theorem. [4, 68]. Let (X, s, x_0) and (X^*, s^*, x_0^*) be integral systems. Let $+, \cdot,$ and \leq be the addition, the multiplication, and the ordering relation, respectively, in X which satisfy the earlier definitions. Let $+^*, \cdot^*,$ and \leq^* be the corresponding relations in X^* . Then there exists a one-to-one mapping f on X onto X^* which preserves each of these relation in the following sense:

1. $f(x + y) = f(x) +^* f(y)$,
2. $f(x \cdot y) = f(x) \cdot^* f(y)$,
3. $x \leq y$ if and only if $f(x) \leq^* f(y)$.

2.3: Cardinal Numbers

Definition. [4, 79]. Two sets are **similar** or **equinumerous**, symbolized

$$A \sim B,$$

if and only if there exists a one-to-one correspondence between A and B .

Definition. [4, 80]. Recall that the set of all subsets of a set A is the **power set** of A , symbolized by $\mathcal{P}(A)$. Similarity is an equivalence relation on $\mathcal{P}(U)$ and a **cardinal number** is a similarity set. If $A \in \mathcal{P}(U)$, then the cardinal number of A , symbolized

$$\overline{\overline{A}} \text{ or card } A,$$

is the cardinal number having A as an element. For example, let $U = \{1, 2, 3\}$, so that

$$\mathcal{P}(U) = \{\{\}, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, \{1, 2, 3\}\}.$$

Let $A = \{1\}$. The cardinal numbers of U are

$$\{\{\}, \{\{1\}, \{2\}, \{3\}\}, \{\{1, 2\}, \{2, 3\}, \{1, 3\}\}, \text{ and } \{\{1, 2, 3\}\}.$$

Since the cardinal number $\{\{1\}, \{2\}, \{3\}\}$ has A as its element, it follows that $\{\{1\}, \{2\}, \{3\}\}$ is the cardinal number of A .

This definition is slightly different from that of Frege and Russell, who identified the cardinal number $\overline{\overline{M}}$ with the set of all sets similar to M . Both definitions are satisfactory, however, since it can be successfully argued that it is irrelevant to know in mathematics what cardinal numbers *are*, so long as cardinal numbers have the property $\overline{\overline{A}} = \overline{\overline{B}}$ if and only if $A \sim B$.

Definition. [4, 81]. To compare cardinals, we define the notion of “domination” for sets. If A and B are sets such that A is similar to a subset of B , we write $A \lesssim B$, and say that A is **dominated by** B or that B **dominates** A .

Theorem (3.1). [4, 81]. If $A \lesssim B$ and $B \lesssim A$, then $A \sim B$. This is known as the **Schröder–Bernstein Theorem**.

Definition. [4, 82]. We define $A < B$ for sets A and B to mean that $A \lesssim B$ and not $B \lesssim A$ (abbreviating “it is not the case that $B \lesssim A$ ” to “not $B \lesssim A$ ”).

Lemma (3.1). [4, 82]. For sets A and B , $A \lesssim B$ if and only if either $A \sim B$ or $A < B$.

Lemma (3.2). [4, 82]. For cardinal numbers a and b , $a < b$ if and only if there exist respective representatives A and B such that $A < B$.

Definition. [4, 83]. The natural numbers in the role of cardinal numbers are the **finite cardinals**, and sets which have these cardinals are **finite sets**.

Theorem (3.2). [4, 84]. For each natural number n , the finite cardinal n is the cardinal of the set of natural numbers which precede n in the natural ordering.

Theorem (3.3). [4, 84]. For each natural number n , if $\overline{A} = n$, then A is not similar to a proper subset of itself.

Theorem (3.4). [4, 85]. In their new role as cardinals, the natural numbers are subject to the ordering of cardinal numbers generally, as defined above, following Cantor; this ordering we write temporarily as $<_c$. In their original role as members of \mathbb{N} , the natural numbers possess the familiar ordering, which we write as $<_{\mathbb{N}}$. The natural ordering and the cardinal ordering agree on \mathbb{N} . That is, for all natural numbers p and q ,

$$q <_{\mathbb{N}} p \text{ if and only if } q <_c p.$$

Definition. [4, 85]. A non-finite cardinal is an **infinite** or **transfinite cardinal**.

Definition. [4, 85]. If the cardinal number of a set is infinite, then the set is called **infinite**.

Definition. [4, 85]. The cardinal number of the set of natural numbers is symbolized by

$$\aleph_0.$$

Theorem (3.5). [4, 85]. If n is a finite cardinal, then $n < \aleph_0$.

Theorem (3.6). [4, 86]. For every set A , $A < \mathcal{P}(A)$ or, in other words, $\overline{\overline{A}} < \overline{\overline{\mathcal{P}(A)}}$.

The Cantor Paradox. [4, 128]. This paradox is derived from the set defined by the formula

$$x \text{ is a set.}$$

Let C be the set defined by this formula. Then C is the set of *all* sets. By Theorem 3.6, $\overline{\overline{\mathcal{P}(C)}} > \overline{\overline{C}}$. Also, since C is the set of all sets and $\mathcal{P}(C)$ is a set (the set whose members are the subsets of C), $\mathcal{P}(C) \subseteq C$. Hence, $\overline{\overline{\mathcal{P}(C)}} \leq \overline{\overline{C}}$ or, in other words, it is false that $\overline{\overline{\mathcal{P}(C)}} > \overline{\overline{C}}$. Thus, it follows that both “ $\overline{\overline{\mathcal{P}(C)}} > \overline{\overline{C}}$ ” and the negation of the statement are valid. This is a contradiction.

2.4: Countable Sets

Definition. [4, 87]. A set is **denumerable** if and only if it has cardinal number \aleph_0 .

Definition. [4, 87]. A set is **countable** if and only if it is either finite or denumerable.

Definition. [4, 87]. An **enumeration** of a denumerable set A is a specific one-to-one correspondence between \mathbb{N} and A .

Theorem (4.1). [4, 89]. A subset of a countable set is countable.

Theorem (4.2). [4, 89]. If the domain of a function is countable, then its range is also countable.

Theorem (4.3). [4, 90] $\mathbb{N} \times \mathbb{N}$ is denumerable.

Corollary. If X is a denumerable set, then so is $X \times X$. More generally, if n is a natural number, then X^{n+1} is denumerable.

Theorem (4.4). [4, 91] If \mathcal{A} is a nonempty finite collection of denumerable sets, then $\bigcup \mathcal{A}$ is denumerable. If \mathcal{A} is a nonempty finite collection of countable sets, then $\bigcup \mathcal{A}$ is countable.

Theorem (4.5). [4, 92] If \mathcal{A} is a countable collection of countable sets, then $\bigcup \mathcal{A}$ is countable.

Definition. [4, 92]. We denote $\text{card } \mathcal{P}(\mathbb{N})$ by

\aleph .

Definition. [4, 92]. To say that a set is **uncountable** means that it is infinite and non-denumerable.

Theorem (4.6). [4, 93] 2^{\aleph} is an uncountable set.

Proof. If $a_1, a_2, \dots, a_n, \dots$ is any enumeration of elements from 2^{\aleph} , then an element a of 2^{\aleph} can be constructed that does not correspond to any a_n in the enumeration. Consider the following enumeration of elements from 2^{\aleph} :

$$\begin{aligned} a_1 &= (0, 0, 0, 0, 0, 0, 0, \dots) \\ a_2 &= (1, 1, 1, 1, 1, 1, 1, \dots) \\ a_3 &= (0, 1, 0, 1, 0, 1, 0, \dots) \\ a_4 &= (1, 0, 1, 0, 1, 0, 1, \dots) \\ a_5 &= (1, 1, 0, 1, 0, 1, 1, \dots) \\ a_6 &= (0, 0, 1, 1, 0, 1, 1, \dots) \\ a_7 &= (1, 0, 0, 0, 1, 0, 0, \dots) \\ &\vdots \end{aligned}$$

By definition, the **complementary** of 0 is 1 and the complementary of 1 is 0. To construct a sequence a , we choose its first digit to be the complementary of the first digit of a_1 . Similarly, we choose the second digit of a to be the complementary of the second digit of a_2 , the third digit as complementary to the third digit of a_3 , and so on. In general, for every n , the n th digit of a is chosen to be the complementary of the n th digit of a_n . For the example

given above, this procedure yields:

$$\begin{aligned}
 a_1 &= (\underline{0}, 0, 0, 0, 0, 0, 0, \dots) \\
 a_2 &= (1, \underline{1}, 1, 1, 1, 1, 1, \dots) \\
 a_3 &= (0, 1, \underline{0}, 1, 0, 1, 0, \dots) \\
 a_4 &= (1, 0, 1, \underline{0}, 1, 0, 1, \dots) \\
 a_5 &= (1, 1, 0, 1, \underline{0}, 1, 1, \dots) \\
 a_6 &= (0, 0, 1, 1, 0, \underline{1}, 1, \dots) \\
 a_7 &= (1, 0, 0, 0, 1, 0, \underline{0}, \dots) \\
 &\vdots \\
 a &= (\underline{1}, \underline{0}, \underline{1}, \underline{1}, \underline{1}, \underline{0}, \underline{1}, \dots)
 \end{aligned}$$

The procedure guarantees that a is an element of $2^{\mathbb{N}}$, but it differs from every a_n , since their n -th digits differ (as highlighted in the example). Since a cannot be included in the enumeration, $2^{\mathbb{N}}$ is an uncountable set. \square

Definition. [4, 93-94]. The question of whether \aleph is the smallest cardinal greater than \aleph_0 is known as the **continuum problem**.

Definition. [4, 94]. It has been discovered that a number of theorems, some of them important, can be based on the hypothesis that the answer to the continuum problem is in the affirmative. This conjecture is known as the **continuum hypothesis**.

2.5: Cardinal Arithmetic

Definition. [4, 95]. The **sum**, $u + v$, of the cardinal numbers u and v is $\overline{A \cup B}$, where A and B are disjoint representatives of u and v , respectively.

Theorem (5.1). [4, 95]. For cardinal numbers u , v , and w ,

1. $u + v = v + u$,
2. $u + (v + w) = (u + v) + w$,
3. $u \leq v$ implies $u + w \leq v + w$.

Definition. [4, 95]. The **product**, uv , of the cardinal numbers u and v is $\overline{A \times B}$, where A and B are representatives of u and v , respectively

Theorem (5.2). [4, 96]. For cardinal numbers u , v , and w ,

1. $uv = vu$,
2. $u(vw) = (uv)w$,
3. $u \leq v$ implies $uw \leq vw$,
4. $(u + v)w = uw + vw$.

Definition. [4, 97]. If u and v are cardinals, the v th **power** of u , in symbols u^v , is $\text{card } A^B$, where A and B are representatives of u and v , respectively.

Theorem (5.3). [4, 97] For cardinal numbers u , v , and w ,

1. $u^v u^w = u^{v+w}$,
2. $(uv)^w = u^w v^w$,
3. $(u^v)^w = u^v w$,
4. $u^1 = u$ and $1^u = 1$,
5. $u \leq v$ implies $w^u \leq w^v$,
6. $u \leq v$ implies $u^w \leq v^w$.

2.6: Order Types

Definition. [4, 98]. Two chains X and Y are called **ordinally similar**, symbolized

$$X \approx Y,$$

if and only if they are isomorphic ordered sets.

Definition. [4, 99]. An equivalence class under ordinal similarity is called an **order type**.

Definition. [4, 100]. Let A and B be disjoint sets of order types α and β . Then the **sum**, $\alpha + \beta$, of α and β is the order type of $A \cup B$, totally ordered as follows. Pairs in A and pairs in B are ordered according to the total orderings of A and B , respectively, and each a in A precedes each b in B .

Definition. [4, 100]. The **product**, $\alpha\beta$, of α and β is the order type of $A \times B$ ordered by

$$(a, b) < (a', b') \text{ if and only if } b < b', \text{ or } b = b' \text{ and } a < a'.$$

2.7: Well-ordered Sets and Ordinal Numbers

Definition. [4, 103]. The principle of **proof by transfinite induction** is as follows, where, as earlier, $P(x)$ stands for “the element x has the property P .” If $P(x_0)$, where x_0 is the first element of the well-ordered set X , and if for all z in X , $P(y)$ for all $y < z$ implies $P(z)$, then $P(x)$ for all x in X .

Definition. [4, 103]. If A is a well-ordered set and if $x \in A$, then $\{a \in A \mid a < x\}$ is called the **initial segment** determined by x ; this is denoted by A_x .

Definition. [4, 103]. If B is an arbitrary nonempty set, then by a **sequence of type x in B** we shall mean a function on A_x into B .

Definition. [4, 103]. The principle of **definition by transfinite induction** may be stated as follows. Let A be a well-ordered set having a_0 as its least element, let B be a set, and let c be a member of B . If h is a function whose range is included in B and whose domain is the set J of all sequences j of type x in B for some $x \neq a_0$, then there exists exactly one function $k : A \rightarrow B$ such that

$$k(a_0) = c \text{ and } k(x) = h(k|A_x),$$

for each x in A other than a_0 .

Theorem (7.1). [4, 103]. If A is a well-ordered set and f is an isomorphism of A into itself, then $a \leq f(a)$ for each a in A .

Theorem (7.2). [4, 104]. A well-ordered set is not ordinally similar to any of its initial segments.

Corollary (4, 104). If A is a well-ordered set and if $A_x \approx A_y$, then $x = y$.

Theorem (7.3). [4, 104]. If A and B are ordinally similar well-ordered sets, then there exists exactly one isomorphism between them.

Theorem (7.4). [4, 104]. If A and B are well-ordered sets, then exactly one of the following hold: A is ordinally similar to B , A is ordinally similar to an initial segment of B , or B is ordinally similar to an initial segment of A .

Corollary (4, 105). For well-ordered sets A and B , exactly one of $\overline{\overline{A}} = \overline{\overline{B}}$, $\overline{\overline{A}} < \overline{\overline{B}}$, $\overline{\overline{B}} < \overline{\overline{A}}$ holds. In other words, any two cardinal numbers which have well-ordered sets as representatives are comparable.

Definition. [4, 105]. The order type of a well-ordered set is called an **ordinal number**, or simply an **ordinal**.

Definition. [4, 105]. The ordinals which are not natural numbers are called **transfinite ordinals**.

Definition. [4, 105]. If α and β are ordinals, we shall say that α is **less than** β , symbolized

$$\alpha < \beta,$$

if and only if there exists a representative of α which is ordinally similar to an initial segment of β .

Theorem (7.5). [4, 106]. The set $s(\alpha)$ of all ordinals less than the ordinal α is a well-ordered set of ordinal number α .

Theorem (7.6). [4, 106]. Any set of ordinals is well-ordered.

Theorem (7.7). [4, 106]. If Δ is any set of ordinals, then there exists ordinals greater than any ordinal of Δ . Indeed, there exists a smallest such ordinal.

Theorem (7.8). [4, 107]. If α and β are ordinals and $\beta > 0$, then $\alpha + \beta > \alpha$.

Definition. [4, 107]. Ordinals having a predecessor are **ordinal numbers of the first kind**.

Definition. [4, 107]. Ordinals having no predecessor are **ordinal numbers of the second kind**.

Theorem (7.9). [4, 107]. Let α and β be ordinals with $\alpha < \beta$. Then there exists exactly one ordinal $\gamma > 0$ such that $\alpha + \gamma = \beta$.

Theorem (7.10). [4, 109]. For ordinal numbers α , β , and γ ,

1. $\alpha < \beta$ implies $\gamma + \alpha < \gamma + \beta$ and conversely;
2. $\alpha < \beta$ implies $\alpha + \gamma \leq \beta + \gamma$; conversely, $\alpha + \gamma < \beta + \gamma$ implies $\alpha < \beta$.
3. $\alpha < \beta$ and $\gamma > 0$ imply $\gamma\alpha < \gamma\beta$; conversely, $\gamma\alpha < \gamma\beta$ implies $\alpha < \beta$;
4. $\alpha < \beta$ implies $\alpha\gamma \leq \beta\gamma$; conversely, $\alpha\gamma < \beta\gamma$ implies $\alpha < \beta$;
5. $\gamma + \alpha = \gamma + \beta$ implies $\alpha = \beta$;

6. $\gamma\alpha = \gamma\beta$ and $\gamma > 0$ imply $\alpha = \beta$.

Theorem (7.11). [4, 109]. If α and β are ordinals and $\beta > 0$, then α has a unique representation in the form

$$\alpha = \beta\xi + \gamma \text{ where } 0 \leq \gamma < \beta.$$

2.10: Some Theorems Equivalent to the Axiom of Choice

Theorem (10.3). [4, 125]. (Hartog). The axiom of choice is equivalent to the assertion that any two cardinal numbers are comparable.

Chapter 3: The Extension of the Natural Numbers to the Real Numbers

3.1: The System of Natural Numbers

Definition. [4, 132] A binary operation \star has the **cancellation property** if and only if each of $x \star z = y \star z$ and $z \star x = z \star y$ implies that $x = y$.

3.2: Differences

Definition. [4, 133]. By a **difference** we shall mean an ordered pair (m, n) . In the set $\mathbb{N} \times \mathbb{N}$ of all differences we introduced the relation \sim_d (the subscript is for “difference”) by defining

$$(m, n) \sim_d (p, q) \text{ if and only if } m + q = p + n.$$

Lemma (2.1). [4, 133]. \sim_d is an equivalence relation on $\mathbb{N} \times \mathbb{N}$.

Definition. [4, 133]. We shall call a difference (m, n) **positive** if and only if $m > n$.

Lemma (2.3). [4, 133]. If x, y, u , and v are differences and $x \sim_d u$ and $y \sim_d v$, then $x + y \sim_d u + v$.

Lemma (2.4). [4, 133]. Addition of differences is associative and commutative. The sum of two positive differences is a positive difference. Further, addition is cancellable with respect to \sim_d .

Lemma (2.5). [4, 133]. If x and y are differences, then there exists a difference z such that $z + x \sim_d y$.

Definition. [4, 133]. Another binary operation in $\mathbb{N} \times \mathbb{N}$, which we call **multiplication** and symbolize by \cdot , is defined for differences by

$$(m, n) \cdot (p, q) = (mp + nq, mq + np).$$

Usually we shall write “ xy ” or “ $x \cdot y$ ” for a product of differences.

Lemma (2.6). [4, 134]. If x, y, u , and v are differences and $x \sim_d u$ and $y \sim_d v$, then $xy \sim_d uv$.

Lemma (2.7). [4, 134]. Multiplication of differences is associative and commutative, and distributes over addition. The product of two positive differences is a positive difference. Further, multiplication is cancellable with respect to \sim_d for differences other than those of the form (m, m) .

3.3: Integers

Definition. [4, 134]. Recalling Lemma 2.1, we define an **integer** to be a \sim_d -equivalence class. We shall write

$$[x]_i$$

for the equivalence class determined by the difference x (The new subscript is for “integer”). The set of integers will be symbolized by \mathbb{Z} .

Definition. [4, 134]. We shall call an integer **positive** if and only if one of its members is a positive difference. The set of positive integers will be symbolized by \mathbb{Z}^+ .

Definition. [4, 134]. Consider the relation $\mathbb{Z} \times \mathbb{Z}$ into \mathbb{Z} :

$$\{([x]_i, [y]_i), [x + y]_i \mid x \text{ and } y \text{ are differences}\}.$$

We call this operation **addition** and symbolize it by $+$. Hence,

$$[x]_i + [y]_i = [x + y]_i.$$

Lemma (3.1). [4, 134]. Addition of integers is associative and commutative, and has the cancellation property. Further, the sum of two positive integers is a positive integer.

Lemma (3.2). [4, 135]. If x and y are integers, then there exists exactly one integer z such that $z + x = y$.

Definition. [4, 135]. Let x be an integer. The **negative** of x , denoted $-x$, is an integer such that

$$(-x) + x = x + (-x) = [(0, 0)]_i.$$

Definition. [4,]. Let \mathbb{Z} be the set of positive integers. Consider the relation $\mathbb{Z} \times \mathbb{Z}$ into \mathbb{Z} :

$$\{([x]_i, [y]_i), [xy]_i \mid x \text{ and } y \text{ are differences}\}.$$

We call this operation **multiplication** and symbolize it by \cdot . Hence,

$$[x]_i \cdot [y]_i = [xy]_i.$$

Lemma (3.3). [4, 135]. Multiplication is associative and commutative, distributes over addition, and has the cancellation property if $(0, 0)$ is not a member of the factor to be cancelled. Further, the product of two positive integers is a positive integer.

Definition. \mathbb{Z}^0 is the set of integers of the form $[(n, 0)]_i$.

Definition. [4, 135]. Theorem 2.1.8 implies that the mapping f in \mathbb{N} into \mathbb{Z} such that $f(n) = [(n, 0)]_i$ is one-to-one, onto \mathbb{Z}^0 , and preserves addition, multiplication, and less than. We summarize these properties of f by calling it an **order-isomorphism** of \mathbb{N} **onto** \mathbb{Z}^0 and indicate the relationship of \mathbb{Z}^0 to \mathbb{N} by referring to \mathbb{Z}^0 as an **order-isomorphic image** of \mathbb{N} (or, saying that \mathbb{Z}^0 is **order-isomorphic** to \mathbb{N}).

Definition. [4, 136]. The order-isomorphism of \mathbb{N} onto \mathbb{Z}^0 suggests that we call the members of \mathbb{Z}^0 the integers which *correspond* to the natural numbers and adopt “ 0_i ,” “ 1_i ,” “ 2_i ” \cdots as names for them. Since the remaining integers (that is, the members of $\mathbb{Z} - \mathbb{Z}^0$) have the form $[(0, m)]_i$ with $m \in \mathbb{N} - \{0\}$, and since

$$[(0, m)]_i = -[(m, 0)]_i = -m_i,$$

we acquire “ -1_i ,” “ -2_i ,” \cdots as names for the so-called **negative integers**.

Theorem (3.1). [4, 136]. The operations of addition and multiplication for integers, together with $0_i, 1_i$, and the set \mathbb{Z}^+ of positive integers, have the following properties for all integers x, y , and z .

1. $x + (y + z) = (x + y) + z$.
2. $x + y = y + x$.
3. $0_i + x = x$.
4. There exists an integer z such that $z + x = 0_i$.
5. $x(yz) = (xy)z$.
6. $xy = yx$
7. $1_i x = x$.
8. $x(y + z) = xy + xz$.
9. $xz = yz$ and $z \neq 0_i$ imply that $x = y$.
10. $0_i \neq 1_i$.
11. $x, y \in \mathbb{Z}^+$ imply that $x + y \in \mathbb{Z}^+$.
12. $x, y \in \mathbb{Z}^+$ imply that $xy \in \mathbb{Z}^+$.
13. Exactly one of $x \in \mathbb{Z}^+, x = 0_i, -x \in \mathbb{Z}^+$ holds.
14. If $<_i$ is defined by $x <_i y$ if and only if $y - x \in \mathbb{Z}^+$, then $<_i$ totally orders \mathbb{Z} and well-orders $\{0_i\} \cup \mathbb{Z}^+$.

3.4: Rational Numbers

Definition. [4, 138]. An ordered pair (a, b) with $b \neq 0_i$ is called a **quotient**. The quotient (a, b) will be written as

$$\frac{a}{b}.$$

Definition. [4, 138]. The relation \sim_q is introduced into the set of all quotients by defining

$$\frac{a}{b} \sim_q \frac{c}{d} \text{ if and only if } ad = bc.$$

This is an equivalence relation on the set of all quotients and has the further property that

$$\frac{ac}{bc} \sim_q \frac{a}{b} \text{ if } c \neq 0_i.$$

Definition. [4, 138]. A quotient is **positive** if and only if ab is a positive integer.

Definition. [4, 138]. We introduce **addition** and **multiplication** into the set of quotients by way of the following definitions:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$$

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

Since $b \neq 0_i$ and $d \neq 0_i$ imply that $bd \neq 0_i$, these are operations in the set of quotients.

Lemma (4.1). [4, 138]. If x, y, u , and v are quotients and $x \sim_q u$ and $y \sim_q v$, then $x + y \sim_q u + v$, $xy \sim_q uv$ and, if x is positive, then u is positive.

Definition. [4, 138]. We define a **rational number** to be a \sim_q -equivalence class. The rational number having the quotient x as a representative we write as

$$[x]_s.$$

The letter s stands for “rational”—the letter r is reserved for the real numbers. The set of rational numbers is symbolized by \mathbb{Q} .

Definition. [4, 139]. We say $[x]_s$ is **positive** if and only if it contains a quotient y such that y is positive. The set of positive rationals we symbolize by \mathbb{Q}^+ .

Theorem (4.1). [4, 139]. The operations of addition and multiplication for rational numbers, together with $0_s, 1_s$, and the set \mathbb{Q}^+ of positive rationals, have the following properties for all rationals x, y , and z .

1. $x + (y + z) = (x + y) + z$.
2. $x + y = y + x$.
3. $0_s + x = x$.
4. There exists a z such that $z + x = 0_s$.
5. $x(yz) = (xy)z$.

6. $xy = yx$
7. $1_s x = x$.
8. If $x \neq 0_s$, there exists a z such that $zx = 1_s$.
9. $x(y + z) = xy + xz$.
10. $1_s \neq 0_s$
11. $x, y \in \mathbb{Q}^+$ imply that $x + y \in \mathbb{Q}^+$.
12. $x, y \in \mathbb{Q}^+$ imply that $xy \in \mathbb{Q}^+$.
13. Exactly one of $x \in \mathbb{Q}^+, x = 0_s, -x \in \mathbb{Q}^+$ holds.
14. If P is the intersection of all subsets of \mathbb{Q}^+ which contain 1_s and are closed under addition, then, for each $x \in \mathbb{Q}^+$, there exist $a, b \in P$ such that $xb = a$.

Definition. [4, 140]. For each $x \neq 0_s$ the solution of $zx = 1_s$ is unique. This solution is called the **inverse** of x and is symbolized by x^{-1} .

Theorem (4.2). [4, 141]. Between any two distinct rational numbers there is another rational number.

Theorem (4.3). [4, 141]. (**Archimedean property**). If r and s are positive rational numbers, then there exists a positive integer n (properly, a positive integral rational number n) such that $nr > s$.

3.5: Cauchy Sequences of Rational Numbers

Definition. [4, 143]. A **Cauchy sequence of rational numbers** is a sequence x of rational numbers such that for every positive rational number ϵ there exists a positive integer N such that for every $m, n > N$

$$|x_n - x_m| < \epsilon.$$

Definition. [4, 144]. The operations of **addition** and **multiplication** for sequences of rational numbers is defined in the following way:

$$\begin{aligned} x + y = u \text{ where } u_n &= x_n + y_n \\ xy = v \text{ where } v_n &= x_n y_n. \end{aligned}$$

Lemma (5.1). [4, 145]. If x is a Cauchy sequence of rational numbers, then there exists a positive rational number δ such that for every n

$$|x_n| < \delta.$$

Lemma (5.2). [4, 145]. If x and y are Cauchy sequences of rational numbers, then $x + y$ and xy are Cauchy sequences of rational numbers.

Definition. [4, 146]. If x and y are Cauchy sequences of rational numbers, then

$$x \sim_c y$$

if and only if for every positive rational number ϵ there is an integer N such that for every $n > N$,

$$|x_n - y_n| < \epsilon.$$

Lemma (5.3). [4, 146]. The relation \sim_c is an equivalence relation on the set of all Cauchy sequences of rational numbers.

Definition. [4, 146]. If x is a Cauchy sequence of rational numbers, then x is called **positive** if and only if there is a positive rational number ϵ and an integer N such that for every $n > N$

$$x_n > \epsilon.$$

Lemma (5.4). [4, 146]. If x, y, u , and v are Cauchy sequences of rational numbers and $x \sim_c u$ and $y \sim_c v$, then $x + y \sim_c u + v, xy \sim_c uv$ and, if x is positive, then u is positive.

Lemma (5.5). [4, 147]. The sum and product of two positive Cauchy sequences are positive Cauchy sequences. Further, if x is any Cauchy sequence, then exactly one of the following hold: x is positive, $x \sim_c 0_c$, $-x$ is positive.

Lemma (5.6). [4, 148]. If the Cauchy sequence x is not equivalent to 0_c , then there is a Cauchy sequence z such that $zx \sim_c 1_c$.

3.6: Real Numbers

Definition. [4, 149]. We define a **real number** as a \sim_c -equivalence class of Cauchy sequences of rational numbers. The real number having the Cauchy sequence x as a representative we write as

$$[x]_r.$$

The letter r stands for the real numbers. The set of real numbers is symbolized by \mathbb{R} .

Definition. [4, 149]. A real number is **positive** if and only if it contains a positive Cauchy sequence.

Theorem (6.1). [4, 150]. The operations of addition and multiplication for real numbers, together with 0_r , 1_r , and the set of positive reals, have properties (1) – (13) listed in Theorem 4.1.

Theorem (6.2). [4, 151]. Between any two distinct real numbers there is a rational real number. Precisely, if x and y are distinct real numbers, then there exists a rational real number z such that if $x < y$, then $x < z < y$ while if $y < x$, then $y < z < x$.

Theorem (6.3). [4, 152]. (**Archimedean property**). If x and y are positive real numbers, then there exists a positive integer n (properly, a real number n which corresponds to a rational which, in turn, corresponds to a positive integer) such that $nx > y$.

Theorem (6.4). [4, 152]. A nonempty set of real numbers which has an upper bound has a least upper bound.

3.7: Further Properties of the Real Number System

Definition. [4, 154]. A **Cauchy sequence of real numbers** is a sequence x of real numbers such that for every positive real number ϵ there exists a positive integer N such that for every $m, n > N$

$$|x_n - x_m| < \epsilon.$$

Definition. [4, 155]. The real number y is a **limit** of the sequence x of real numbers if and only if for every positive real number ϵ there exists a positive integer N such that for every $n > N$

$$|x_n - y| < \epsilon.$$

Lemma (7.1). [4, 155]. A sequence of real numbers has at most one limit.

Lemma (7.2). [4, 155]. Let a be a sequence of rational numbers and let x be the sequence of real numbers such that for every n , $x_n = (a_n)_r$, the real number corresponding to a_n . Then x is a Cauchy sequence if and only if a is a Cauchy sequence. Further, if a is a Cauchy sequence and y is the real number which it defines, then $\lim x_n = y$.

Theorem (7.1). [4, 156]. (Cauchy convergence principle). A sequence of real numbers has a limit if and only if it is a Cauchy sequence.

Definition. [5, 160]. The famous method of defining the real numbers, by **Dedekind cuts**, is outlined as follows. A **cut** of the rational numbers is an ordered pair (A, B) of sets such that

1. A and B are both non-empty,
2. $A \cup B =$ the set of rationals,
3. if $x \in A$ and $y \in B$, then $x < y$.

A is called the **lower class** and B the **upper class**, since every element of A precedes every element of B . A **real number** is then simply a cut of the rationals.

4 Conclusion

We conclude by discussing the material in Chapters 1 to 7 of Russell's *Introduction to Mathematical Philosophy* (Part I) that has since been adopted by modern mathematics (Part II).

Russell introduces the concept of similarity (a “one-to-one relation”) and uses it to define the notion of a finite set. Specifically, if there exists a one-to-one relation between two sets, then they are said to have the same size or “cardinality.” These concepts are also adopted in Robert Stoll's “Set Theory and Logic.” In Chapter 2 of Stoll's book, the concept of a one-one correspondence is used to define the notion of equinumerosity between sets [4, 79]. Moreover, Stoll uses a one-one correspondence to compare the sizes of infinite sets. In particular, he introduces the concept of countability, which is a way to classify sets with the same cardinality as the set of natural numbers [3, 87].

Russell also focuses on the concept of finitude and its relation to mathematical induction. Specifically, Russell says that mathematical induction is, above all else, the essential characteristic that distinguishes the finite from the infinite [3, 27]. It is thus unsurprising that the principle of mathematical induction is treated by Stoll as a way to define the natural numbers, which in the role of cardinal numbers are the finite cardinals [4, 83].

Russell also discusses the concept of order and how it is defined. Keeping in mind that a relation is aliorelative (and hence asymmetric) if and only if it is both antisymmetric and irreflexive, and that Russell's “connected” and “aliorelative” are synonymous with Stoll's “comparable” and “irreflexive,” respectively, we may say that Russell's “serial” relation (which is aliorelative, transitive, and connected) [3, 34] is, for all intents and purposes, equivalent to Stoll's “totally ordered” relation (which is reflexive, antisymmetric, transitive, and comparable) [4, 50]. (The true equivalent of Russell's serial relation, not treated directly by Stoll, is a “strict totally ordered” relation, which is irreflexive, antisymmetric, transitive, and comparable).

It is worth noting that Stoll's definition of a “totally ordered” relation is more flexible than that of Russell's “serial” relation, in two ways. For one,

we could remove the “comparable” property from Stoll’s definition and be left with the definition of a partially-ordered relation. Second, we could add the property that “each non-empty subset of the set has a least element” to Stoll’s definition and have a well-ordered relation. Such additions and deletions are difficult with Russell’s definition, owing to the fact that the “aliorelative” relation encompasses both antisymmetry and irreflexivity.

Russell also touches on the concept of “similarity between relations,” defining it as the existence of a one-to-one relation between the elements of one relation and the elements of the other. Specifically, if there is a one-to-one correspondence between the domain of the first relation and the domain of the second relation, and a one-to-one correspondence between the co-domain of the first relation and the co-domain of the second relation, and so on for all the terms of the relations, then the two relations are similar. In such a case the two relation “do not depend upon the actual terms in their fields” [3, 54]. This idea closely resembles that of an isomorphism [4, 52].

Russell then talks about the development of the different number systems.

Russell and Stoll both use the natural numbers as a starting point for the construction of the real numbers. Russell defines the “natural numbers” as the posterity of 0 with respect to the relation ”immediate predecessor” (which is the converse of ”successor”) [3, 27]. But then generalizes this definition, defining the natural numbers to be those “to which proofs by mathematical induction can be applied, i.e., as those that possess all inductive properties” [3, 27]. Stoll (as noted above) incorporates the principle of mathematical induction to define the natural numbers, but asserts more formally that $(\mathbb{N}, S, 0)$ is an integral system [4, 58].

Russell and Stoll then move to positive and negative integers. Russell says that, if m is any natural number, $+m$ will be the relation of $n + m$ to n (for any n), and $-m$ will be the relation of n to $n + m$ [4, 63]. On the other hand, Stoll calls an integer positive if and only if one of its elements is a positive difference. And if x is any integer, then there exists exactly one integer, symbolized by $-x$ (the “negative of x ”), such that

$$(-x) + x = x + (-x) = [(0, 0)]_i \text{ [4, 134-135].}$$

Russell defines a fraction as being that relation which holds between two

inductive numbers x, y , when $xn = ym$. Similarly, Stoll introduces the relation \sim_q into the set of quotients by defining

$$\frac{a}{b} \sim_q \frac{c}{d} \text{ if and only if } ad = bc.$$

Finally, both Russell and Stoll arrive at the real numbers. Russell uses Dedekind cuts, which involves dividing all the terms of a series into two sets, of which the one “wholly precedes” the other. Russell confines himself to cuts in which the lower section has no maximum; these no-maximum sections are called “segments.” This allows him to define a “real number” as a segment of the series of ratios in order of magnitude. Stoll introduces a relation, symbolized by \sim_c , in the set of all Cauchy sequences of rational numbers. If x and y are Cauchy sequences of rational numbers, then

$$x \sim_c y$$

if and only if for every positive rational number ϵ there is an integers N such that for every $n > N$,

$$|x_n - y_n| < \epsilon.$$

This allows Stoll to define a real number as a \sim_c -equivalence class of Cauchy sequences of rational numbers.

References

- [1] Frege, Gottlob, *The Foundations of Arithmetic*, translated by J. L. Austin, second revised edition, Harper Torchbooks/The Science Library, Harper and Brothers, New York, 1960.
- [2] Halmos, Paul R., *Naive Set Theory*, The University Series in Undergraduate Mathematics, D. Van Nostrand Co., Inc. Princeton, 1960.
- [3] Russell, Bertrand A. W., *Introduction to Mathematical Philosophy*, Second Ed., Dover-Hill, 1993.
- [4] Stoll, Robert R., *Set Theory and Logic*, First Ed., W. H. Freeman and Company, 1963.
- [5] Suppes, Patrick, *Axiomatic Set Theory*, D. Van Nostrand Co., Inc. Princeton, 1965.