## Contents

Preface

### 1 Some Basic Tools

1.1 Sets

1.1.1 Operations upon sets

1.1.2 More concerning set operations

1.1.3 The Cartesian Product of two sets

---

Preface

1 Some Basic Tools

1.1 Sets

1.1.1 Operations upon sets

1.1.2 More concerning set operations

1.1.3 The Cartesian Product of two sets
CONTENTS

1.1.4 Relations and Functions ................................................................. 1-13
1.2 Logic ................................................................................................. 1-17
  1.2.1 Propositions, operations, and truth tables .............................. 1-17
  1.2.2 Quantifiers ................................................................................. 1-21

2 Integers and Rational Numbers ......................................................... 2-1
  2.1 Introductory Remarks ...................................................................... 2-2
  2.2 Basic properties of the integers ...................................................... 2-3
    2.2.1 Algebraic properties of the integers ...................................... 2-4
    2.2.2 Order properties of the integers .......................................... 2-7
    2.2.3 The Well Ordering Principle and Mathematical Induction ... 2-10
    2.2.4 Other properties of the integers .......................................... 2-13
  2.3 A small excursion ........................................................................... 2-14
  2.4 Construction of the rational numbers ........................................... 2-18
  2.5 Inadequacy of the rational numbers ............................................. 2-22

3 Building the real numbers ............................................................... 3-1
  3.1 Introduction ...................................................................................... 3-2
  3.2 Sequences ......................................................................................... 3-6
3.3 Sequences of rational numbers .................................................. 3-7
3.4 The real numbers ................................................................. 3-10
3.5 Completeness ........................................................................ 3-18

4 Series ......................................................................................... 4-1
4.1 Finite series and sigma notation ............................................... 4-2
4.2 Tools – the binomial theorem .................................................. 4-7
4.3 Infinite series ........................................................................ 4-8
4.4 Tools – the geometric series .................................................... 4-11
4.5 A small excursion – the definition of \( e \) ................................ 4-12
4.6 Some discussion of the exponential function ......................... 4-15
4.7 More on convergence of series ............................................... 4-17
  4.7.1 The Root Test and the Ratio Test ....................................... 4-17
  4.7.2 Conditional Convergence and the Alternating Series Test .... 4-20

5 Cardinality .................................................................................. 5-1
5.1 Finite and countable sets ......................................................... 5-1
5.2 Uncountable sets .................................................................... 5-6
6 Representations of the real numbers
  6.1 Introduction ......................................................... 6-2
  6.2 Decimal representation ............................................... 6-3
  6.3 Binary representation ............................................... 6-5
  6.4 Other representations ............................................... 6-6
  6.5 The Cantor Set ....................................................... 6-8

7 Conclusion of MATH 3100  .................................................. 7-1
Preface

This document will serve several purposes. First, it will serve as a textbook for the course MATH 3100, Spring 2013. The material which is presented will be organized around various topics, which introduce basic concepts in mathematics and, even more important, present the internal logic of the subject, some of the basic questions which need to be answered in mathematics, and why those questions need to be answered. No other textbook is used. Therefore, this document is presented to the students and to their future instructors so that everyone interested in the matter can know what has been done here.

The material presented here to the students is subject to revision during the semester, which may be done a chapter or a section at a time. The students are encouraged to keep the book in a binder, or, perhaps even better, to keep it in
electronic form. In either event, it should be kept in a form which makes it easy to add pages as they are handed out, or to replace portions which are revised or expanded. Taking into consideration this possibility of revision during the semester, the page numbering will be done chapter by chapter instead of consecutively from beginning to end. This will minimize the inevitable disruption of the page numbering if revisions need to be done.

These materials will be presented in two formats. One of them will contain standard-sized pages which are ready for printing, and the other will use short and wide pages which are suitable for viewing in landscape mode on a laptop, netbook, or handheld electronic device. Both versions will contain identical text, and both versions will be made available at <www.auburn.edu/~kilgota>.

The choice of the topics has been guided by the instructor’s own sense of what the students need for further mathematical development. The choice is also determined in part by the instructor’s sense of what the place of the introductory course in theoretical mathematics is, in the current course structure in the mathematics program at Auburn University. What the student most needs from this course is an introduction to mathematics as a logical discipline. Systematic and step-by-step development have been neglected in the previous mathematics training of the students, or even sometimes avoided or postponed to the indefinite future. Thus, the understanding of mathematics as a subject which is developed step by step from basic assumptions, called axioms, has languished.

Typically, the students have taken the calculus sequence and perhaps both an introductory differential equations course and an introductory linear algebra course. To a too great extent, these courses present a very rushed introduction to huge areas of mathematics, a bewildering multitude of topics, which are in fact related at a deeper level but that may not be obvious to the students. The contents and the emphases in these courses are also to a too great extent determined
by the nature of these courses as service courses, in which the students are expected to learn that bewildering multitude of
topics in a huge rush for future use in various other subjects, such as in the engineering fields or in the sciences. Inevitably,
the logical and methodical development which is the very heart of mathematics has been sacrificed to institutional and
curricular needs, dictated by pressures which are external to mathematics and which work against actually learning
and comprehending the subject. That is a very unfortunate thing, precisely because mathematics cannot be learned
well, some say cannot be learned at all, without due attention to the logical and methodical development. Logical and
methodical development have been practiced in mathematics literally for millenia. And this has been true, after all, for
a reason.

Students who have seen through all of these institutional constraints, who have nevertheless tried to pay attention to
logical and methodical development, and who have tried to integrate their knowledge by relating it to logical development
and unifying principles instead of trying to learn by memorizing seemingly unconnected formulas have almost without
exception been rewarded for doing that immediately. Such students get better grades and tend to retain what they
have learned instead of just having “studied for the test” which beyond a certain point is a sure prescription for failure
to understand the subject. Even if they have taken those mathematics courses only as service courses, they are still
rewarded. For, having understood that mathematics is the key to much of the content of their chosen discipline, they
have taken the time and put forth the effort to learn the mathematics. However, those students have even so had to do
a great deal of that effort on their own time and resources. Here, we take the approach that the logical development and
unifying principles are the things which are truly important. We revisit some of the basics of the subject which have
been so cavalierly glossed over in those previous courses.
The main topic of this course will be the step-by-step logical construction of the real number system, starting out with what we know of the integers. The logical and set-theoretic abilities which are needed for doing that will be introduced and developed as required. Also, some concepts which are studied in other areas of mathematics such as abstract algebra will be introduced along the way, too. For, mathematics is viewed best as a whole.

There are several reasons for the choice of the real number system as a main topic and priority. One of these reasons is that the topic is very basic and important and thus can serve as a unifying theme for the entire course. Another very good reason is that the student may possibly not see the detailed construction of the real number system later on. For, in the instructor’s experience this systematic development seems to be omitted from the later parts of the curriculum, in particular from the introductory course in real analysis. Rather, in higher-level courses it is assumed that the student has a working knowledge and an intuitive understanding already about just what the set of real numbers is. For example, in the “standard” introductory analysis course, the material usually begins in medias res with an axiomatic description of the real numbers rather than with an actual construction. By all means, the analysis course can do just that – if the student has actually been through a systematic construction. But those students who have never seen such a systematic construction are at that point shortchanged.

Here, after a brief introductory chapter dealing with sets and logic, the integers are taken as the basic building blocks of the number system. As a first step, then, some basic properties of the integers are described and explored. Then the rational numbers are developed as a quotient field over the integers. Some care is given to describing the logical and practical need, after this, to expand the rational numbers in order to obtain what we call completeness. There are several equivalent ways to carry out this construction of the real number system. Here, one of those methods
will be used. Other methods will be mentioned, too, sometimes in homework problems.

As to the internal organization of the course, a heavy emphasis will be placed upon classroom participation, problem solving, and classroom presentations by the students of the results of their efforts. Classroom discussions and presentations will comprise a large portion of the grade, too.

The problems interspersed in these notes will be directly related to the fundamental theme of the course. Indeed, much of the work which will be done will use the problems as building blocks. By the time that the course is completed, most of the problems given here will have actually been presented by one or another student during class time. Some of them, in the interests of democracy and fairness, will have been presented by the instructor. Some of them may be assigned as homework to be written, or may be used as test problems. But, it is intended that the students should know how to complete these problems at the end. Indeed, one might go so far as to say that much of the material is presented in the problems.

It is the hope of the instructor that the students, on completing this course, will have had an experience which is positive, if at times challenging, and that the students will have assimilated and mastered enough of the subject for future success.
Chapter 1

Sets and logic

1.1 Sets

Mathematics is built up as an axiomatic structure. This means, that we start out with some basic assumptions which either look reasonable to us, or they are assumptions which we would like to take on in order to explore what kind of
system comes out as a consequence. One of the basic tools for doing our work is the concept of a set. But nobody knows what a set is. Funny, that. Well, actually, what this means is, a definition of what a set is has no place inside our formal system, precisely because it is something which we must presuppose and therefore logically precedes the formal system. If one perceives that we have already descended into thinking which is not serious, then one might try to play the “dictionary game” in order to see what the problem is. The dictionary game is played by looking up a word in the dictionary. Then one looks up all of the words used in the definition of the word which was looked up. This procedure is repeated until one comes back to one of the words previously defined. Usually, it does not take very many steps before this happens.

Well, then, what is a set? We can make an informal definition easily enough. A set is a collection of objects, things, concepts, or whatever. Why can we not make this formal? Well, perhaps we could. But then we would require a rigorous definition of “collection,” “objects,” “things,” “concepts,” and so on. As already stated, we have to start somewhere, and we might as well start here.

Things are either in a given set, or they are not. The things which are in the set are called elements of the set. If we have decided that the name of the set is $S$, and $x$ is some object, then we write $x \in S$ to mean that $x$ is an element of $S$. If it is not true that $x \in S$, then we write $x \notin S$. Now that we are getting all formalistic, we can restate the first sentence of this paragraph as saying that an element is either in a given set, or it is not. A set can be defined either by explicitly listing the elements in it, or it can be defined by giving some kind of characterization which completely describes the elements in it, in other words defining the set by some defining property of the elements. Here are some rather stock examples:
1. \( \{1, 2, 3\} \) describes the set which contains the three listed integers (it is customary to use the curly brackets when giving a definition of a set). Also notice that the order in which the elements are listed does not matter. For example, the set \( \{3, 1, 2\} \) is the same set.

2. \( \{x \mid x \text{ is an even integer} \} \)

3. \( \{x : x \text{ is an even integer} \} \) (same as previous)

4. \( \{x \mid x = 2k, \text{ for some integer } k \} \) (equivalent to previous two – how do we define an even integer, anyway?)

The astute, or those who have already seen things like this before, may be aware that there can already be problems at this basic level. We have to be careful, else we cannot even pretend to build a logical structure on what we have already done. Clearly, it is possible to put words together to come up with a description which stumbles over something basic. The problem which arises from that is, how do we know that we are doing such a thing on a given occasion, or not? Ideally, we would like to be able to put some kind of an airtight system in place which automatically disallows such an act, even if it is committed by the well-intentioned who would prefer always to stick to common sense. A rather obvious problem arises, now that we have agreed we know all about sets, is the following:

\[
S = \{x \mid x \notin S\}
\]

For, if an \( x \) is given to us to test and decide whether it is in \( S \) or not, then, if it is, then it is not, and if it is not, then it is.
Exercises

1.1.1 This exercise relates to the example just above. One might suppose that the difficulties which it presents
need somehow to be avoided. Perhaps the way out is to agree that the example somehow does not give a proper
definition of a set? Has anything already been stated as a property of sets, which this “definition” violates? If so,
find it in the previous text.

In the alternative, it may be that in fact nothing at all has been violated, but the situation is nonetheless strange.
Is there something which can be done about the matter? Or, have we successfully defined a set after all, but just a
very strange one?

Another, more serious matter is implicit here. If you are looking for a hint about how to solve the previous problem,
then do not look further. No hint will be given in what follows. But we do have yet another problem, independent of
the previously posed one. Namely, we have to be careful about where our elements come from as well as where our sets
come from. Consider the following definition, that

\[ S = \{ Y | Y \notin Y \}, \]

in which the \( Y \) in the definition is, by its usage, a set. This would seem to make sense, until one looks more closely
and asks that just how the definition can be applied to \( S \) itself. For, if \( S \) is not an element of itself, then \( S \in S \). But,
alternatively, if \( S \in S \), then \( S \) must satisfy the definition of \( S \) which says then that \( S \notin S \). This is the famous Russell’s
paradox, named after Bertrand Russell, who stumbled upon it approximately 100 years ago.
There are several solutions to Russell’s paradox and to similar logical conundrums. One of the most practical, which we will adopt here, is to realize something which goes all the way back to Aristotle’s logic. Aristotle recognized the relevance of a “universe of discourse” which, in terms of set theory is a “universal set” which serves as the context for the discussion of everything inside of it. That is, the universal set is not the set which includes everything. Rather, the universal set represents the result of prior agreement on what is relevant to the discussion. We are on quite safe ground if we declare (for example) that on a given occasion we are speaking of integers, cards in a standard deck, or whatever. Then, if we want to speak of sets of integers, those sets are not integers. If we wish to speak of poker or bridge hands, then those hands are not individual cards and cannot be confused with such. There are other, more sophisticated approaches, but this is the one that we will adopt.

Another reason that we use a universal set is that it quite often makes a difference what the universal set is, what is the correct answer to a given question. For example, is 11 divisible by 4? Yes? Or, no? The answer to that question depends very much upon context, in other words, upon the universal set which is presumed to underlie the question.

Homework or (not unduly long because this is not a class in metaphysics or epistemology, but in mathematics) class discussion:

Exercises

1.1.2 Other logical paradoxes? Think of some examples. Discuss.

1.1.3 Does the employment of a universal set really solve all our problems? Or are we still glossing something over?
Is it possible to have a poorly defined universal set and to land right back in the soup again?

1.1.4 There once was a town famous for the fact that it had a barber who shaved every man in town who did not shave himself. Who shaved the barber? Is this another example of a paradox? Or is there some other kind of logical problem?

1.1.1 Operations upon sets

Sets, of course, are used as a tool in mathematics. Therefore, we need to pay attention to some of the basic operations used upon them and to familiarize ourselves with some of the notation. Let us first agree that we have chosen a universal set, $U$, for the purposes of the ensuing discussion.

If $A$ and $B$ are any two subsets of $U$, then we can define three operations on each of them. First, it is possible to consider the complement of each of these sets:

We agree that, given the set $A$, then concerning each element $x \in U$ it is the case that either $x \in A$, or $x \notin A$. We have already agreed that 

$$\{x|x \in A\}$$

is merely an alternative and equivalent definition of $A$. But, then, what about the set

$$\{x|x \notin A\}?$$
This set is the complement of $A$. Several notations are in use for this set. It can be denoted with

$$U \setminus A$$

or sometimes as

$$U - A$$

$$U \sim A$$

or, sometimes, as just $\sim A$, or $\tilde{A}$ or $A^c$. Similarly we can construct the complement of $B$ or any other subset of $U$.

Now, the above discussion raises an obvious question, that what is $U \setminus U$? Does this make sense? Well, certainly it does if we want to agree that every subset of $U$ has a complement. The alternative would be to say that every subset except $U$ has a complement. However, we shall see that it becomes very inconvenient if one tries to formulate general statements and rules and then to start listing exceptions. Thus, perhaps we ought to say that we have just defined the **empty set** which we will denote by $\Phi$. There is the obvious problem that there might be more than one empty set, and that problem needs to be resolved. We will defer the discussion of that until we have presented the other set operations.

Given our two sets $A$ and $B$ it is clear that we can consider the set

$$\{x | x \in A \text{ and } x \in B\}$$
which is called the **intersection** of $A$ and $B$. Notice that the word “and” here means that the statements to the left and right of it are both simultaneously true. Notice that the intersection of the two given subsets $A$ and $B$ may not contain any elements at all. There is nothing, after all, which says that it has to. In such an eventuality, we could say that the intersection is empty, and we do. We would be tempted to define that empty intersection as the set $\emptyset$ which has already been discussed. But there is that little problem which needs to be cleared up. Observe that if we cannot do so, then our set theory is already a kind of a mess because we can’t really say that any two sets *have* an intersection, even if there is a universal set lurking in the background.

Also there is the set

$$\{x|x \in A \text{ or } x \in B\}$$

which is called the **union** of $A$ and $B$. Notice that here the word “or” signifies that at least one of the statements to its left and right is true. In particular, we do not intend to exclude the possibility that both are true. Naturally, a situation could arise in which we intend to say “one or the other but not both” but then we would need to write that down with some other words, which exactly specify what we mean to say.

We now need to consider possible relationships between our two sets $A$ and $B$. It might be, for example, that every element of $A$ is also an element of $B$. In that case, we say that $A$ is a **subset** of $B$. The notation for this is

$$A \subseteq B$$

Of course, we could also write

$$B \supseteq A$$
which we might express as “$B$ contains $A$” which obviously means the same thing.

There is also the possibility, of course, that the two sets $A$ and $B$ are in fact identical, and then we write

$$A = B.$$  

Exercises

1.1.5 Given two subsets $A$ and $B$ of some universal set $U$, show that

i. If $A = B$, then $A \subseteq B$ and $B \subseteq A$

ii. The converse of the previous statement is true, too.

1.1.6 If $\Phi_1$ and $\Phi_2$ are two subsets of a universal set $U$ and neither of them has any elements, then in fact $\Phi_1 = \Phi_2$. (Hint: Use the previous problem).

1.1.2 More concerning set operations

The set operations obey certain rules. The following are left as exercises for the students.
Exercises

1.1.7 Associative Laws. Given any three subsets $A$, $B$, and $C$ of a universal set, $U$, we have

$$A \cap (B \cap C) = (A \cap B) \cap C \quad \text{and} \quad A \cup (B \cup C) = (A \cup B) \cup C$$

1.1.8 Distributive Laws

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \quad \text{and} \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

1.1.9 Commutative Laws

$$A \cap B = B \cap A \quad \text{and} \quad A \cup B = B \cup A.$$ 

There are also laws which relate specifically to $U$ and $\Phi$. Continuing our convention that $A$, $B$, $C$ are arbitrarily chosen subsets of $U$, we have

$$A \cup \Phi = A \quad \text{and} \quad A \cap \Phi = \Phi$$

and

$$A \cup U = U \quad \text{and} \quad A \cap U = A.$$
There are also some rules relating to complements. First of all, we extend the notion of the complement to include the **relative complement**. That is, we define the notation $B \setminus A$ to mean the set

$$B \cap (U \setminus A).$$

De Morgan’s Laws state that

$$\sim (A \cap B) = (\sim A) \cup (\sim B) \quad \text{and} \quad (\sim A \cup B) = (\sim A) \cap (\sim B).$$

Now, in view of the fact that union and intersection operations each obey the associative law, we can in fact define the union and intersection of more than two sets. First, suppose that we have some sets $A_1, \ldots, A_n$. Then we can define

$$\bigcup_{i=1}^{n} A_i = \left( \bigcup_{i=1}^{n-1} A_i \right) \cup A_n$$

and similarly define

$$\bigcap_{i=1}^{n} A_i = \left( \bigcap_{i=1}^{n-1} A_i \right) \cap A_n.$$

These are examples of **inductive** definitions. The way a definition like this works is, we know what $A_1 \cup A_2$ is, and we can work up from there to any value of $n$ (similar remarks if the operation was “∩” instead).
We can also replace the “\( n \)” by “\( \infty \)” if we want to. How? Well, we cannot use an inductive definition there, so we had better introduce a concept which gets us away from counting entirely, the \textbf{index set}.

Let us consider that we have a collection of subsets of \( U \) called \( \mathcal{C} \). Then, \( \mathcal{C} \) itself is a set. It is not a subset of \( U \), obviously, but we can say that it lives in some other universal set, for example the set of all subsets of \( U \), which is called the power set of \( U \). Well, then, the elements of \( \mathcal{C} \) are subsets of \( U \). We can define

\[
\bigcup_{C \in \mathcal{C}} C = \{x \mid x \in C \text{ for at least one } C \in \mathcal{C} \}
\]

and similarly define

\[
\bigcap_{C \in \mathcal{C}} C = \{x \mid x \in C \text{ for all } C \in \mathcal{C} \}.
\]

Observe that we have gotten away completely from the need to \textit{count} the sets when thus defining the union and the intersection of the sets in the collection \( \mathcal{C} \).

\textbf{Exercises}

1.1.10 Show that these general definitions of the union and the intersection are compatible with the definitions previously given, for two sets.
1.1.11 What form is taken by De Morgan’s laws, as applied to arbitrary collections of sets? Prove that these generalized De Morgan’s laws are valid.

Now, if we need more of general set theory, we will try to introduce it in the course of doing the rest of our business. We move to a brief discussion of some of the fundamentals of logic.

1.1.3 The Cartesian Product of two sets

Let two sets $A$ and $B$ be given. The Cartesian product of these two sets is then denoted by $A \times B$ and consists of all ordered pairs $(a, b)$ in which $a \in A$ and $b \in B$. The students are presumably quite familiar with the original Cartesian product which is represented by two crossed axes and the thus generated plane used for graphing.

1.1.4 Relations and Functions

Any subset of a Cartesian product $A \times B$ is considered to be a relation, since it describes a relationship or a linkage between certain elements of either of the sets $A$ or $B$ to elements in the other set.

A function from $A$ to $B$ may be defined as a relation on the two sets which has the two additional properties that

i. Every element of $A$ may be found in some element of the relation.

ii. If $a \in A$ and $b_1, b_2 \in B$, then $(a, b_1)$ and $(a, b_2)$ are not both present in the function unless $b_1 = b_2$. 
Some would say that the above represents the “graph” of a function, however, and would say instead that the function consists of a rule which associates to each element of \( A \) one and only one element of \( B \). In view of the fact that a function is often given by such a rule, procedure, or computation, this definition of function is perhaps more intuitive. One often denotes the rule by a symbol such as \( f \) and writes \( f : A \rightarrow B \).

In certain respects, more precise terminology is needed here. Perhaps unfortunately, this more precise terminology is not universally agreed upon. If we have two sets \( A \) and \( B \) and a function \( f : A \rightarrow B \), it is generally agreed that the function \( f \) is defined upon all of \( A \) and that \( A \) is called the domain of \( f \). That is, the function \( f \) “uses all of \( A \.” The set \( B \) is usually called the range of \( f \). Careful examination of the definition above, however, will show that nothing in it requires all of \( B \) to be used. Some, such as many algebra and calculus books, may add the additional restriction that all of the set \( B \) needs to be in use, or, worse, tend to require that in practice without ever actually saying so. However, others would say that there is nothing wrong at all. The definition of a function is perfectly good as stated above. But, if one wishes to refer by name to the set of those elements in \( B \) which are actually of importance here, then one needs to speak of the image of \( f \), not the range of \( f \). The image is then the set

\[
\text{Im}(f) = \{ b \in B \mid b = f(a) \text{ for some } a \in A \}.
\]

We will follow this usage here. The consequence is that, in our usage, the range (the set \( B \)) can be any set containing the image. If the image of \( f \) and range of \( f \) happen to be the same set, then the function \( f \) is said to be onto by some mathematicians, and others, more concerned about grammatical correctness than linguistic purity, say that \( f \) is surjective, or is a surjection.
A further source of possible confusion is the frequently used expression “a real-valued function of a real variable” to denote a function whose domain is any subset of the real numbers (aside comment: we officially do not know what those are, as yet!) and whose image is any subset of the real numbers. For, often this fact is important all by itself. And then it is possible to give all those exercises, just like in the calculus book, where the students are supposed to find the domain of the function and the “range” of the function (that is, what we have just defined as the image of the function).

Now, let us notice that indeed the prescription of the domain is part of the definition of a function. Consider the function given by $f(x) = x^2$. If we take the domain to be the largest set possible, then the domain consists of all real numbers. The range can be any set containing the image, which comprises all non-negative real numbers. This function is not invertible, which means that there is no function going back from the image to the domain which reverses the action of $f$. However, if we took the domain to be the set of all non-negative numbers, instead, then the resulting function is invertible. Thus, the two functions are not identical even though given by the same computational procedure. It can, however, be said that the second function can be obtained from the first by restriction of the domain. This is a term of the art, which is why it has been put in boldface.

A function $f : A \to B$ for which $f(a_1) = f(a_2)$ cannot happen if $a_1 \neq a_2$ is called one to one, or, by those who like fancy words, injective, or to be an injection. Note that in the example mentioned in the previous paragraph are seen two functions. One of them is one to one, and the other is not, even though the same computational formula is used for both. To complete the introduction of necessary terminology, a function which is both one to one and onto is called a one to one correspondence or is said to be bijective. Finally, the word mapping is often used as a synonym for the word “function” on occasions when it seems more descriptive of what is happening, but the formal definition is
identical.

One might ask why, exactly, have we not agreed that a function could have several output values instead of only one? The answer to this question is that the cumulative experience of mathematicians leads to the conclusion that the definition is most useful the way it is, and not some other way. The definition of function is, after all, a tool for making mathematics easier to do, not an end in itself. An example might serve to illustrate what can happen if we allow more than one value to “come out” when just one value is “put in”:

Let the function $f$ be defined by $f(x) = \sqrt{x}$ and $g$ defined by $g(x) = \sqrt{x} + 1$. It is advantageous to be able to say that $f(x) + g(x)$ will define a “new function” which we could (and do) call $f + g$. But if (contrary to standard usage!!!) we allowed the square root sign to denote both the positive and the negative square root of a positive number and were to call the result a function, then how many possible answers are there when we add the two together? It is easy to see that the result would be a rapid descent into nonsense and chaos, One of the important reasons why we developed the concept of function in the first place was to make our lives simpler. But this kind of thing would violate the clarity which we are seeking. Such problems are easily avoided by insisting that a function has to have unambiguous outcomes.
1.2 Logic

1.2.1 Propositions, operations, and truth tables

A proposition is a statement which is either true or false. This means, among other things, that whatever statements can not classified as either true or false might not be fit very easily into the scheme and are thus excluded. We can denote a proposition either by issuing the related statement, or we can also label propositions with letters, such as $p, q, r$.

Now, let $p, q, r$ stand for propositions, as described. We can connect propositions by symbols representing the connectors “and” and “or” and “not.” We will use the following notation, which is fairly standard:

- $p \lor q$ means $p$ or $q$
- $p \land q$ means $p$ and $q$
- $\sim p$ means not $p$

A compound statement or proposition is a statement in which several substatements are given, linked by such connectors as are described here, or the one more which is not listed,

- $p \Rightarrow q$ which means $p$ implies $q$

Now that we have our notation in place, we need to look for ways to analyse compound statements to see whether they have an intended logical meaning, or not. One of the most basic ways to do this is with a truth table, which schematically lists all possible combinations of true and false for the component propositions, and for the compound statement, too. Here are some examples:
• The statement $\sim p$ is of course easiest. The table for it is

<table>
<thead>
<tr>
<th>$p$</th>
<th>$\sim p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>$F$</td>
</tr>
<tr>
<td>$F$</td>
<td>$T$</td>
</tr>
</tbody>
</table>

• The table for $p \lor q$ is

<table>
<thead>
<tr>
<th>$p$</th>
<th>$q$</th>
<th>$p \lor q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
</tr>
<tr>
<td>$T$</td>
<td>$F$</td>
<td>$T$</td>
</tr>
<tr>
<td>$F$</td>
<td>$T$</td>
<td>$T$</td>
</tr>
<tr>
<td>$F$</td>
<td>$F$</td>
<td>$F$</td>
</tr>
</tbody>
</table>

• The table for $p \land q$ is

<table>
<thead>
<tr>
<th>$p$</th>
<th>$q$</th>
<th>$p \land q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
</tr>
<tr>
<td>$T$</td>
<td>$F$</td>
<td>$F$</td>
</tr>
<tr>
<td>$F$</td>
<td>$T$</td>
<td>$F$</td>
</tr>
<tr>
<td>$F$</td>
<td>$F$</td>
<td>$F$</td>
</tr>
</tbody>
</table>

• The table for $p \Rightarrow q$ is (discuss)
We say that two statements are **equivalent** if they have the same truth table.

### Exercises

1.2.1 Construct the truth tables for $q \lor p$ and for $q \land p$. The idea is, of course, to see that these come out equivalent to two of those given above, respectively.

1.2.2 Set up a truth table for three statements, $p, q, r$ and then construct the entries for $(p \lor q) \lor r$ and for $p \lor (q \lor r)$. The two should look identical, of course.

1.2.3 Is $p \lor (q \land r)$ equivalent to $(p \lor q) \land (p \lor r)$?

1.2.4 Construct the truth table for $(\sim p) \lor q$. Is it identical to any of those above?

1.2.5 What do De Morgan’s laws say in the context of propositions? Construct a truth table for each of the two De Morgan’s laws.

<table>
<thead>
<tr>
<th>$p$</th>
<th>$q$</th>
<th>$p \Rightarrow q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>$T$</td>
<td>$T$</td>
</tr>
<tr>
<td>$T$</td>
<td>$F$</td>
<td>$F$</td>
</tr>
<tr>
<td>$F$</td>
<td>$T$</td>
<td>$T$</td>
</tr>
<tr>
<td>$F$</td>
<td>$F$</td>
<td>$T$</td>
</tr>
</tbody>
</table>
1.2.6 How might one go about writing $\sim (p \Rightarrow q)$ in terms of the other listed operations?

1.2.7 We have two connectors, namely $\lor$ and $\land$, as well as the “not” operation, $\sim$. Is this more than is actually needed? Could we get away with just two of them? For example, is it possible using a cunning sequence of $\land$ and $\sim$ operations to write something equivalent to $p \lor q$ by using only those other two operations?

1.2.8 Show using a truth table that $p \lor (\sim p)$ always has to come out as true, and that $p \land (\sim p)$ always comes out as false. A statement which is always true, no matter what truth values of its respective inputs are, is called a tautology. A statement which is always false, no matter what the truth values of its respective inputs are, is called a contradiction or a self-contradiction.

Now, in principle we could do all logic relating to propositions or statements by appeal to truth tables. However, notice that when we put three statements into a truth table, things already got a bit tedious. We had to have eight rows in order to take all possibilities into account. If we have four statements to glue together, then sixteen rows are required, and if we had five then thirty-two rows are required. This kind of thing gets old really fast, which is one very good reason why we don’t just quit using logical arguments and bring in truth tables instead. Nevertheless, truth tables certainly have their uses for showing some basic things.
1.2.2 Quantifiers

There is an important class of statement-like things which we use all the time, but which do not exactly fit into the scheme above, of statements which are either true or false. In dealing with sets, we have in fact already touched upon such things. A construction such as \( x \in A \) may not be such a proposition or sentence as we have considered, at all. If it is asserted about some particular, specific \( x \), then in principle it can be decidable as true or false. But if it is an element of the universal set which is chosen quite at random, or, better, which will be chosen at random after we make the statement, then we do not know and cannot decide whether \( x \in A \) is true, or not. An example of this might be that we pick an integer at random, and we want to investigate whether it is even, or not. We can inspect the integer to see if it is even or odd, but unless and until we do inspect the number, the statement that it is even is undecidable. For, the truth or falsity of the statement depends upon the number. A statement of the sort described, the truth or falsity of which depends upon some other input or inputs, is an open statement.

In connection with open statements, it is usually possible to save the situation by asserting the statement to be true for all \( x \), some \( x \) (taken to be equivalent to at least one \( x \)), or for no \( x \) at all. In symbolic language, we have the quantifiers

\[
\forall x \quad \text{For all } x \\
\exists x \quad \text{There exists } x
\]

Now, let \( p(x) \) be an open statement. We can say things such as \( \forall x p(x) \) to mean that \( p(x) \) is true for all \( x \). We can say that \( \exists x p(x) \) if we mean that there is at least one \( x \) for which \( p(x) \) is true.

A reasonable question is, how do we formulate the negations of statements when quantifiers are present? We should
notice the following are true (whether read from left to right or from right to left)

\[ \sim (\forall x p(x)) \] means the same as \[ \exists x (\sim p(x)) \]

\[ \sim (\exists x p(x)) \] means the same as \[ \forall x (\sim p(x)) \]

and that, in fact, nothing else would work. Also, once understood, the procedure for constructing the negation of a statement with quantifiers in it can be made completely mechanical, which is nice when things get complicated.

The principles just outlined work also if there is an open statement which requires more than one input variable, and thus more than one quantifier. One can work through one quantifier at a time. To negate such a statement, first change the outermost quantifier and negate everything which comes after. Then go to the next quantifier, and continue in like fashion.

Also, in constructing statements which involve more than one input and more than one quantifier, the order in which the quantifiers are stated can very much affect the meaning. For an example that every reader probably knows, consider the difference between “Everybody gets a Ring of Power” (\( \forall \) precedes \( \exists \)) and “One Ring to rule them all” (\( \exists \) precedes \( \forall \)).

More relevant to our mathematical interests is the meaning of statements involving quantifiers in mathematics. A very important example of a statement with several quantifiers is the definition of continuity of a real-valued function of a real variable. Let us assume that the function is called \( f \) and is defined upon a set \( D \), its domain. Then, \( f \) is said to be continuous at the point \( x \) in \( D \) provided that the following condition is met:

For all \( \varepsilon > 0 \) there is \( \delta > 0 \) such that for all \( t \in D, |t - x| < \delta \) implies \( |f(t) - f(x)| < \varepsilon \). This may be slightly
shortened in our new symbolic language as

\[ \forall \epsilon > 0 \ \exists \delta > 0 \ \forall t \in D, \ (|t - x| < \delta \Rightarrow |f(t) - f(x)| < \epsilon). \]

We also say that \( f \) is continuous on \( D \) if it is continuous at each individual \( x \in D \). That is,

\[ \forall x \in D \ \forall \epsilon > 0 \ \exists \delta > 0 \ \forall t \in D, \ (|t - x| < \delta \Rightarrow |f(t) - f(x)| < \epsilon). \]

If \( D \) is a closed interval \([a, b]\), then we can say, of course,

\[ \forall x \in [a, b] \ \forall \epsilon > 0 \ \exists \delta > 0 \ \forall t \in [a, b], \ (|t - x| < \delta \Rightarrow |f(t) - f(x)| < \epsilon). \]

**Exercises**

1.2.9 Negate the statement which expresses what is meant by the continuity of \( f \) at \( x \).

1.2.10 Let us suppose that \( D \) is an interval \([a, b]\). Then, why does the statement below mean something other than the statement that \( f \) is continuous at each \( x \in [a, b] \)? What is actually said here?

\[ \forall \epsilon > 0 \exists \delta > 0 \forall t \in [a, b] \text{ and } \forall x \in [a, b], \ (|t - x| < \delta \Rightarrow |f(t) - f(x)| < \epsilon). \]

What if \( D \) is an open or a half-open interval? Can you think of a function which is continuous on \( D \) but does not satisfy this new property?
1.2.11 Why is the statement below different from all of the previous statements? What does it actually say about $f$?

$$\exists \delta > 0 \ \forall \epsilon > 0 \ \forall t \in [a, b] \text{ and } \forall x \in [a, b], \ (|t - x| < \delta \Rightarrow |f(t) - f(x)| < \epsilon).$$

To conclude this section, we should notice that the definition of continuity at a single point $x$ given above, which is the definition actually used by mathematicians in their own work, is not the same as the definition which is given in the typical calculus book. For, if the domain $D$ is for example the set of positive integers, then the statement of the definition given here implies automatically that $f$ is continuous upon $D$, whereas according to the calculus book that is not so. Mercifully, the definition used in the typical calculus book is equivalent to the above if the domain $D$ is an open interval, which is what the calculus book is mostly interested in. The discrepancy between the two definitions of continuity is pointed out here to avoid future confusion, and also to point out but one example of the shortcuts which have to be made in order to get the job done in the time allotted for teaching calculus – we actually feel forced to engage in “white lies.”

We should also notice that a statement with a “for all” quantifier in it does not imply the existence of anything in particular, for it may in fact be true precisely upon all the elements in the empty set. This may seem strange, forced, and a little bit tricky. Nevertheless, it does go along with ordinary understanding. A sign in a park which says “All dogs in this park must be kept on a leash” certainly does not imply that there is always at least one dog in the park.

But the observation above about the “for all” quantifier does have another consequence. Namely, the negation of a statement with “for all” in it is a statement with “there exists” in it, which is a bald assertion that something exists.
What do we mean by saying that something exists? That’s a good question. To take what ought to be a rather everyday example, we all believe that irrational numbers exist. Otherwise, we probably would not be meeting together in a classroom to study mathematics. But have you, the reader, ever really met an irrational number? Probably not. What, then, ought it to mean if one says that irrational numbers exist?
Chapter 2

Construction of the rational numbers

2.1 Introductory Remarks

The real number system is basic (obviously) in the study of real analysis, which deals with properties of the number system and the properties of real-valued functions of a real variable, especially continuous functions. In order to begin
the subject of mathematics from its foundations, one of the things which we should do is to perform the steps of the construction of the real number system. To start with what we know, one should start by the formal construction of the rational numbers as what is called a quotient field over the integers and then construct the real numbers from the rational numbers. Unfortunately, we will not be able to carry out all aspects of this program due to obvious constraints of time. Thus, rigor and step-by-step development will not universally be followed at the expense of all other goals. Nevertheless, enough of the construction will be given or indicated that the difficulties and peculiar problems involved should be made clear to the student. That is, if we do not have time systematically to explore all of the answers, which in fact we will not (and, one should realize, some questions may not have answers!), then at least we should be aware of the questions. One reason for that is, some of the questions that can or could be posed have in fact given rise to various important areas of mathematics. Of course, we continue the pattern already used in the previous chapter. There will be problems interspersed in the material for the student to solve, and further progress may use or even depend upon the solution of some of the problems.

2.2 Basic properties of the integers

The set of integers will be denoted by the standard symbol \( \mathbb{Z} \). Two algebraic operations are defined upon \( \mathbb{Z} \), addition and multiplication. However, the set of integers also satisfies several other properties as well. We discuss these properties in more detail in the following sections.
2.2.1 Algebraic properties of the integers

The addition satisfies the associative and commutative laws, and there is a zero, and there are additive inverses. The multiplication is associative and commutative, and there is a multiplicative unit. The two operations are linked in the distributive law. More specifically, where $x, y, z$ signify arbitrary integers, then we can list the above properties more systematically. It remains to observe that “equals” means that what on either side of it may freely be replaced by what is on the other. As a couple of examples of what this means, consider that item A4 below says among other things, that the expression “0” may be freely replaced by “$x + (−x)$” where $x$ is any particular integer. The entire list of properties now follows:

A1. $x + (y + z) = (x + y) + z$

A2. $x + y = y + x$

A3. There is an integer called zero, denoted by 0, satisfying $x + 0 = x$

A4. To each integer $x$ there corresponds an integer $(−x)$, such that $x + (−x) = 0$.

M1. $x(yz) = (xy)z$

M2. $xy = yx$
2.2. BASIC PROPERTIES OF THE INTEGERS

M3. There is an integer 1, not equal to 0, which satisfies $1x = x$

M4. $x(y + z) = xy + xz$

Before stating the exercises which follow, it should be emphasized very much that in doing them nothing should be assumed except what is stated in A2 through A4 and M1 through M4, above, and the immediately preceding discussion of what is meant by “equals.” The exercises can be done, indeed must be done, based upon these properties and upon nothing else. If the student has not done similar exercises previously, they may not all be that easy. For, one of the hardest things to do is to put aside the “knowledge” based upon previous experience and to meet the problems with nothing in hand except for the above-listed properties. Nobody can learn to play a guitar without passing through a stage of sore muscles and fingertips. Nobody can learn mathematics without learning to concentrate and to put forth mental effort and struggle, nor, even more to the point, without learning to prove the next step in a chain of logical development by appeal to nothing whatsoever except for that which came before in that same chain.

Exercises

2.2.1 Based upon the properties A1 through A3, the zero is unique. That is, if A1 and A2 hold true, there can be no other integer besides 0 for which the property A3 holds true.

2.2.2 Given $x$, the integer $(-x)$ is unique, based upon the properties A1 through A4.
2.2.3 If \( x + z = y + z \), then \( x = y \). This is often referred to as the Law of Cancellation.

2.2.4 Based upon M1 through M3, the integer 1 is unique.

2.2.5 Based upon all of the properties and in particular upon M4, \( 0x = 0 \) for every integer \( x \). Note that this is not the same thing as to say, “If \( x \) and \( y \) are any two integers such that \( xy = 0 \), then either \( x = 0 \) or \( y = 0 \).” This second statement does also happen to be true, but as we shall see from later discussion it does not follow from the list of algebraic axioms given above this problem set.

2.2.6 For every integer \( x \) it is the case that \( (-x) = (-1)x \). Note that, once this matter has been settled, we can simply write \( -x \) instead of \( (-x) \).

2.2.7 \((-xy)) = (-x)y = x(-y)\)

2.2.8 If \( x = y \), then \( zx = zy \), where \( z \) can be any integer (still true, but not so interesting, if \( z = 0 \)). The statement that if \( zx = zy \) with \( z \neq 0 \), then \( x = y \) would obviously follow if we know that, whenever the product of two integers is zero, then one of the two has to be zero. But please see the remark in Exercise 2.2.5 above.

Now, certainly we agree that the integers satisfy the properties listed above. However, these properties do not fully describe the integers. The properties above are considered by mathematicians in the category of algebraic properties. The same properties listed above also hold in many other situations, and in still other situations they almost hold, with
only one or two of them failing. Let us try to think of some examples of this. If we do, then perhaps we can add some
problems to those just above.

2.2.2 Order properties of the integers

One of the big differences that the integers have from some other systems which satisfy the above algebraic axioms is
the fact that the integers also satisfy order relations “<” and “≤.” For, any two integers can be compared using either
of these. An efficient way to describe and to establish these relations is to postulate the existence of two sets, $P$ (set of
positive integers) and $N$ (set of negative integers). The properties of these sets are that

I1. The union of $N$ and $P$ and $\{0\}$ is all of the integers, and the three sets are pairwise disjoint.

I2. An integer $n$ is in $P$ if and only if $-n \in N$

I3. The sum of two elements of $P$ is in $P$

I4. The product of two elements of $P$ is in $P$

We will say that $x < y$ (or $y > x$) if and only if $y - x \in P$. We will show in exercises that “<” forms a strict order
relation on the integers, that is,
O1. \( x < x \) is never true.

O2. If \( x < y \), then \( y < x \) is false.

O3. If \( x < y \) and \( y < z \), then \( x < z \).

Also very important is the fact that this order relation pertains to any two integers. The Law of Trichotomy, states that

O4. Given integers \( x \) and \( y \), either \( x < y \) or \( x = y \) or \( y < x \).

**Remark:** It is the property the 2.2.2 which guarantees that the ordering of the integers allows us to compare any two integers. An order relation which allows one to compare any two elements of a given set is a total ordering. One may not have this, and still have an order relation, that is a relation satisfying 2.2.2, or perhaps satisfying the properties 2.2.2 which are listed below. The most obvious examples of such **partial orderings** are based on properties of set inclusion.

As far as the set of integers is concerned, the properties O1 through O4 will be seen in the exercises to follow from the properties I1 through I4.
Exercises

2.2.9 1 ∈ P, and more generally \( z^2 ∈ P \) for all \( z \neq 0 \).

2.2.10 If \( x \) and \( y \) are integers satisfying \( xy = 0 \), then either \( x = 0 \) or \( y = 0 \).

2.2.11 Show the properties O1 through O4, based upon I1 through I4.

2.2.12 If \( x > y \), then

(a) \( x + z > y + z \)

(b) \( xz > yz \) if \( z > 0 \)

(c) \( xz < yz \) if \( z < 0 \) (order of the inequality is reversed)

2.2.13 Using the order properties listed in 2.2.2 or, in the alternative, the order properties O1 through O4, it should now follow that if \( zx = zy \) with \( z \neq 0 \), then \( x = y \) (See Exercise 2.2.5 for a previous discussion of this topic).

It should be obvious that “≤” is also an order relation of a slightly different kind. It is clearly related to the previous order relation, with only small changes in what the describing axioms say. They would now have to read something like

- \( x ≤ x \) is always true.
• $x \leq y$ and $y \leq x$ are both true if and only if $y = x$

• If $x \leq y$ and $y \leq z$, then $x \leq z$.

In fact, it might even be possible to combine the first two properties in one. However, observe that the property O4 is now harder to state.

2.2.3 The Well Ordering Principle and Mathematical Induction

Now, since we have listed the order properties of the integers, we note one other important property, related specifically to the positive integers:

The **Well Ordering Principle** states that any non-empty subset of the positive integers must have a smallest element in it.

It should be clear to the reader that the well ordering property is distinct from the order properties listed in the previous subsection, and is one of the things which make the set of integers quite different and distinct from some other sets with otherwise similar properties. Among the first consequences of the Well Ordering Principle, we list the following, in the form of exercises.

Exercises
2.2.14 As I am sure we are all aware, it is generally agreed that 1 is the smallest positive integer, there being no positive integer which is less than 1. Based upon the properties of the integers previously described in this section, and not upon what “everybody knows” can you prove this? (Hint: Suppose that the set of integers $n$ such that $0 < n < 1$ is not empty. Then by the Well Ordering Principle it would have a least element. Call that least element by the name $N$. Then, based upon the laws of inequalities previously established, can you show that $N^2 < N$ would follow?)

2.2.15 Let $n$ be any integer. Show that it follows from the result stated in the previous problem that there are no integers between $n$ and $n + 1$.

Further to illustrate the use of the Well Ordering Principle, we show the following:

**Proposition:** Given any two positive integers $m$ and $n$, there exist two unique integers $q \geq 0$ and $r$, satisfying $m = nq + r$, and furthermore that $0 \leq r < n$.

**Proof:** First, we notice that if $n = 1$, the proof is obvious. In that case, we have $m = m \cdot n$ exactly, with $r = 0 < n$. We therefore can proceed on the assumption that $n > 1$.

Now, with $n > 1$, let us define

$$K = \{ x | x = ny, \text{ for some positive integer } y \text{ and } x > m \}.$$ 

To see that $K \neq \Phi$, note that $mn > m$ and hence $mn \in K$. This can be seen to follow from the fact that $n > 1$, in the following steps:
n > 1 implies that n - 1 > 0 (definition of what is meant by n > 1).
In turn, mn - m > 0 because m(n - 1) > 0 (being the product of two positive integers), and we have established that mn ∈ K, showing that K is not empty.

Now, to continue the proof we invoke the Well Ordering Principle in order to notice that K has a smallest element. Since every element of K is a multiple of n by some other integer, we can assume that this smallest element of k may be represented as x₀ = nk₀. By the definition of k₀, we see that nk₀ > m, whereas n(k₀ - 1) ≤ m. It is further clear that k₀ > 0. For, nk₀ ∈ K implies that nk₀ > m and it is not true that 0 = n · 0 > m > 0. Consequently, since k₀ > 0 it follows by Exercise 2.2.14 that k₀ - 1 ≥ 0.

If we now define q = k₀ - 1, then we can also define r = m - nq, and it follows that r ≥ 0. Also, since nq + r = m < nk₀, it is seen that r < nk₀ - nq, whence r < n.

Finally, we must address the question of unicity. Let us suppose that there were to exist q' ≥ 0 and r' which satisfy 0 ≤ r' < n and that m = nq' + r', possibly in addition to the q and r already found, above. We need to show that the only way this can happen is that, in fact, q' = q and r' = r.

First, we show that q' = q. If q' ≠ q, then either q' > q, or q' < q. That q' > q is not possible. For, then by Exercise 2.2.14 we have q' ≥ q + 1, and then nq' ∈ K and hence nq' > m. If on the other hand it were the case that q' < q, then we would have m = nq + r = nq' + n(q - q') + r = nq' + r', and we see that then r' = n(q' - q) + r would have to hold true. But then it would also follow from Exercise 2.2.14 that q - q' ≥ 1, whence q' - q ≤ -1, and n(q' - q) < -n. And then r' = n(q' - q) + r < -n + r < 0, which violates the restriction that r' > 0. Thus, q' = q.

Now, since q' = q, we have both m = nq + r and m = nq' + r'. Hence, both r' = m - nq and r = m - nq, and the
equality $r' = r$ must follow. The proof is completed.

Logically equivalent to the Well Ordering Principle for the positive integers is the Principle of Mathematical Induction:

Let $P(n)$ be a statement which depends upon the positive integer $n$. If the following two conditions hold, then $P(n)$ is true for all positive integers $n$

(i) $P(1)$ is true.

(ii) For any $n \geq 1$ the implication $P(n) \Rightarrow P(n + 1)$ is true.

Exercises

2.2.16 The Principle of Mathematical Induction and the Well Ordering Principle for the positive integers are logically equivalent.

2.2.4 Other properties of the integers

Finally, we mention that certain other properties of the integers will be taken here as known, without further discussion. Some of these properties are in themselves quite important in their own right, but they are more appropriately learned elsewhere.

Examples of such properties include the existence of prime numbers and results related to them, such as
• There list of prime numbers is not finite; there is no largest prime.

• If $p$ is a prime number and $p$ is a divisor of the product $ab$, where $a$ and $b$ are both integers, then either $p$ is a divisor of $a$, or $p$ is a divisor of $b$. The previous statement may or may not be true if $p$ is not prime.

• The Prime Factorization Theorem, which states that every positive integer can be factored as a product of primes, and further states that if the primes are listed in their natural order then the factorization is unique.

• The Euclidean Algorithm, and its uses for finding the greatest common divisor and the least common multiple of any two positive integers.

We leave this section with the observation that the listed properties, among some others, are very important for day-to-day use in the manipulation of integers. Most particularly, they are important in handling a sum of fractions with different denominators and writing the result in lowest terms. For, in doing such problems it is necessary to find the least common multiple of the denominators or else the problem can easily become quite unmanageable.

Exercises

2.2.17 If you have taught the arithmetic of fractions or have reason to believe that you might ever have to teach such material, what approach would you take?
2.3 More about the algebraic properties of the integers: a small excursion

It is stated in the previous section that the set of integers has two algebraic operations and that these two algebraic operations obey certain laws. It is further stated that there are properties relating to order. Here, we play a little bit of theme and variations. One purpose of this section is to show the power of abstraction in mathematics, which consists of the distillation of general principles from a number of specific examples or manifestations, followed by the exploration of the consequences of those general principles taken in isolation from other factors. Abstraction has been one of the driving principles behind the development of mathematics in the last one hundred fifty years or so, and it has been very successful in keeping things manageable while the field has grown so greatly in size and complexity.

Consider the following examples:

Exercises

2.3.1 We construct a set containing precisely two elements, which we will call 0 and 1. We construct rules for addition and multiplication, as follows

\[
\begin{array}{c|cc}
+ & 0 & 1 \\
0 & 0 & 1 \\
1 & 1 & 0 \\
\end{array}
\quad
\begin{array}{c|cc}
\cdot & 0 & 1 \\
0 & 0 & 0 \\
1 & 0 & 1 \\
\end{array}
\]
Show that with the indicated algebraic operations the set \( \{0, 1\} \) is an algebraic system which obeys all of the laws of algebra which have been prescribed for the integers. It goes without saying, of course, that the order properties fail to work. Nevertheless, the result shown in Exercise 2.2.13 is valid in this system.

2.3.2 In a manner similar to the previous example, we can construct an algebraic system defined upon the set \( \{0, 1, \ldots, n - 1\} \), in which the sum of any two numbers is their usual sum, unless that sum exceeds \( n - 1 \), in which case we take the remainder upon division by \( n \). Parenthetically, one might note that the effect is to identify \( n \) with 0. We define the multiplication similarly. As an example, of this, we take \( n \) to be 12. Then an everyday specimen of this system is seen on a clock face. Show that the algebraic operations for the integers all hold here. Show also that it is sometimes possible for two numbers to be non-zero, but their product in this 12-based system can come out zero nevertheless. Can you account for the fact that the result in Exercise 2.2.13 does not hold here, but it does hold in Exercise 2.3.1? Are there values of \( n \) (other than 12, obviously, and in addition to 2) for which the product of two non-zero numbers is never zero?

2.3.3 Consider the set of all subsets of a given universal set, \( X \). We call this set of subsets of \( X \) the power set of \( X \), and it is usually denoted by \( \mathcal{P}(X) \). We define two operations upon \( \mathcal{P}(X) \). One of them is intersection, which we already know. That one will play the role of the multiplication. The second operation, which will serve as the addition, is the symmetric difference defined by

\[
A \triangle B = (A \cup B) \setminus (A \cap B).
\]
Parenthetically, this operation corresponds to an exclusive “or” operation in logic and to the XOR bitmap operation on computers. Here, however, the problem is to show that \( P(\mathcal{X}) \) satisfies the listed algebraic properties of the integers, using these two indicated operations.

The axioms which define the basic algebraic properties of the integers are those which, in the modern field of algebra, are taken to define an algebraic structure called a **commutative ring with unit**, the unit being, of course, 1. It is obviously possible to come up with structures which have some but not all of these properties, but inside of which it is still meaningful and relevant to do algebra-like activities. Consider the following further examples:

**Exercises**

2.3.4 The set of all functions which are continuous on \([0, 1]\) but which all satisfy \( f(0) = 0 \) (nothing here corresponds to the number 1, so it seems to be a ring without unit). Also, again it is possible to have two functions in this set whose product is zero, even though neither of the two functions are zero.

2.3.5 If we restrict the functions in the previous example to be polynomials, then it is no longer possible to have two non-zero polynomials whose product is identically zero (why?).

2.3.6 The set of all 2 by 2 matrices with integer coefficients, with the usual rules for addition and multiplication of them by one another. Here, the commutative law for multiplication clearly fails, which in turn requires that the distributive law for multiplication, M4, has to be amended and written as a “left distributive law” and a “right
distributive law" in order fully to describe the situation. Also, the product of two non-zero two by two matrices with integer coefficients can come out to be zero.

We can devote a little bit of time to presentations and proofs of some of the properties of some of the previous examples. Any volunteers?

2.4 The construction of the rational numbers as a quotient field

The construction of the rational numbers from the integers is the classic example of what is called in abstract algebra the construction of a quotient field over the integers. To perform the construction, one begins with the Cartesian product \( \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) \). Upon this set, an equivalence relation is then imposed, namely we will say that \((m, n) \sim (p, q)\) if and only if \(mq = np\).

Note 1: The Cartesian product of two sets was defined in the previous chapter, in the subsection 1.1.3.
Note 2: What is an equivalence relation? The definition is seen in the problem immediately below.

Exercises

2.4.1 The advertised equivalence relation is indeed an equivalence relation. We say that \( \sim \) is an equivalence relation on a set \( S \), provided that
2.4. CONSTRUCTION OF THE RATIONAL NUMBERS

E1. $s \sim s$ for all $s \in S$
E2. $s \sim t$ if and only if $t \sim s$
E3. $s \sim t$ and $t \sim u$ imply together that $s \sim u$

Now, we define the algebraic operations of addition and multiplication upon our set $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$. The addition is defined by

$$(m, n) + (p, q) = (mq + np, nq),$$

and the multiplication is defined by

$$(m, n)(p, q) = (mp, nq).$$

We will now define the set of rational numbers $\mathbb{Q}$ to be the quotient set $(\mathbb{Z} \times (\mathbb{Z} \setminus \{0\}))/\sim$. That is, the set of rational numbers $\mathbb{Q}$ is the set of equivalence classes which is generated by the above equivalence relation. In $\mathbb{Q}$, of course, it is customary to write $(m, n)$ as $\frac{m}{n}$, and we will start to do that as soon as possible, after showing that it all works. The immediate problem, of course, is that we have to show that this all actually works. As a trivial, concrete example of what we have to do, we have to show not only how to add two fractions such as $\frac{1}{2}$ and $\frac{1}{3}$, but we have to show that the procedure still works if we take any other fractions which are respectively equivalent to the two given ones, by providing a sum which is equivalent to the sum of the two given fractions. We also have to do the same thing for multiplication. Hence, the following exercises.
Exercises

2.4.2 Show that the addition is well-defined. That is, show that if

\((m_1, n_1) \sim (m_2, n_2)\) and \((p_1, q_1) \sim (p_2, q_2)\),
then \((m_1, n_1) + (p_1, q_1) \sim (m_2, n_2) + (p_2, q_2)\).

The purpose of this exercise is to show that the addition defined in the entire set \(Z \times (Z \setminus \{0\})\) is compatible with the definition of \(Q\), that is, most particularly, the proposed method for addition does not do violence to the equivalence relation that underlies our work.

2.4.3 Show that the multiplication is well-defined.

2.4.4 Show that the algebraic properties A1 through A4 and M1 through M4 which held for the integers also hold for the set \(Z \times (Z \setminus \{0\})\) and hence hold in \(Q\). Show that in addition there are multiplicative inverses for all elements \((m, n)\) for which \(m \neq 0\). The result of this exercise is to show that \(Q\) has the algebraic properties of a field (see just below these exercises for the definition of a field).

2.4.5 Establishing an order relation on the rational numbers can be done by at least two methods which both lead in the end to the same result. In this problem, we explore one of these, and in the next problem we look at the other.

We can define sets of rational numbers which are similar in their functionality to those sets \(P\) and \(N\) of integers which we have already defined. Let us call them \(P\) and \(N\). We will say that the ordered pair \((m, n)\) which represents
2.4. CONSTRUCTION OF THE RATIONAL NUMBERS

A rational number is in \( P \) if and only if \( mn > 0 \) and is in \( N \) if and only if \( mn < 0 \). Show that these sets \( P \) and \( N \) are compatible with the equivalence class structure of the rational numbers. Show that the properties I1 through I4 which are listed in Section 2.2.2 can be successfully extended to the new sets. We can then define an order relation on all of \( Q \) by agreeing that \((m,n) < (p,q)\) if \(((p,q) - (m,n)) \in P\).

2.4.6 In this problem, we extend the order relation without bothering to construct first the sets \( P \) and \( N \) of the previous problem. We say that \((m,n) < (p,q)\) if both \( n > 0 \) and \( q > 0 \) and it is true that \( mq < np \). Show without using Exercise 2.4.5 but only based upon the definition in the previous sentence, that the order relation just now defined makes sense on all of \( Q \) and in fact defines an order relation between any two elements of \( Q \) which are not equal to each other. For, every pair \((m,n)\) in \( \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) \) is equivalent to another pair in which the second number is positive (that is, if \( n < 0 \) we can use \((-m,-n)\) in any such comparison instead of \((m,n)\) ). Show that the results in this problem are also compatible with the result of Exercise 2.4.5 and could have been derived from Exercise 2.4.5 instead of having been defined in isolation from Exercise 2.4.5.

2.4.7 Show that the integers \( \mathbb{Z} \) can be embedded into \( Q \), by the rule that \( z \mapsto (z,1) \) (or, equivalently at this point, \( z \mapsto \frac{z}{1} \)). Show further that this mapping preserves the algebraic operations on \( \mathbb{Z} \).

2.4.8 Show that the definition for the order relation on \( Q \) which was given in Exercise 2.4.5 above is compatible with the order relation defined previously upon \( \mathbb{Z} \), under the natural embedding defined in problem 2.4.7.

We have given a construction of the set of rational numbers, starting from the set of integers, and we have shown that
the set of rational numbers satisfies the axioms $A_1$, $A_2$, $A_3$, and $A_4$ and $M_1$, $M_2$, $M_3$, $M_4$, as does the set of integers, and in addition we have a new property of multiplication exists which was not previously present. It was proved in Problem 2.4.4. We will call it $M_5$, and it says

M5. For each non-zero rational number $q$ there is a corresponding rational number $q^{-1}$ such that $(q)(q^{-1}) = 1$.

The axioms $A_1$, $A_2$, $A_3$, $A_4$, and $M_1$, $M_2$, $M_3$, $M_4$, $M_5$ are the axioms of an algebraic structure called a field.

Exercises

2.4.9 Prove that the system described in Problem 2.3.1 is also an algebraic field.

2.4.10 Prove that the properties listed in Problem 2.2.12 remain true in the set of rational numbers.

When the proof of this problem is completed, we have shown that the set of rational numbers consists of an ordered field, a property which the system discussed in Problem 2.3.1 obviously can not share.

2.4.11 In addition to the properties shown in the previous problem, we have the following property:

If $x \in \mathbb{Q}$ and $y \in \mathbb{Q}$ with $x > y$ and $y > 0$, then $0 < 1/x < 1/y$.

We have also shown that the integers may be naturally embedded into the rational numbers (Problem 2.4.7), and that the order properties of the rational numbers (introduced in Problems 2.4.5 and 2.4.8) naturally and compatibly extend the order properties of the integers which were given in $I_1$, $I_2$, $I_3$, $I_4$ and in $O_1$, $O_2$, $O_3$, $O_4$. 
2.5 Inadequacy of the rational numbers

This goes way back.

When the theorem of geometry known nowadays as the Pythagorean Theorem was proved, Pythagoras and his disciples are said to have sacrificed a hundred oxen in celebration of the great discovery. Unfortunately for their world view, which had as a doctrine the idea that all of nature was built on ratios of numbers, it was seen very quickly that there is no rational number which is equal to the square root of two. That number, of course, would express the length of the hypotenuse of an isosceles right triangle with both sides having the length one unit. In the language used by the Greek geometers, the legs and the hypotenuse of said triangle were seen to be “incommensurable” which is an equivalent statement. The historical accounts agree that this was discovered the very next day after the big result was proved and the oxen were slaughtered. After agreeing on this, the historical accounts also agree that the entire school of Pythagoras was sworn to secrecy about the second discovery. For, as said above, this discovery struck at the heart of their world view and if widely publicized would have made them look foolish. Accounts differ about what happened to the disciple who made the second discovery. Some of the accounts say that he was executed. Needless to say and whether the hapless discoverer was or was not executed, the truth of the matter could not be concealed for very long. It was rediscovered fairly soon.

Exercises

2.5.1 Show that there can be no rational number which is the square root of 2. (Hint: Start by assuming that
there is such a number, and observe that if it does exist it can be expressed as a fraction in lowest terms. From this starting point, show that both the numerator and the denominator of the fraction are necessarily divisible by two, and thus the fraction, which was assumed to be in lowest terms, is not in lowest terms and can not be put in lowest terms. Therefore, the initial assumption, that the fraction exists, must be false.

In view of the previous account and the previous exercise, there is at least one quantity which we would think ought to be expressible as some kind of a number, but it is seen that one must step outside the set of rational numbers to do so. As so often happens in the development of mathematics, one little example might very well uncover a multitude of similar problems. One single counterexample to a general statement does not cause much discomfort if it is also possible to show that it is the only counterexample. The so-called general statement has just been shown false, but it is then easy to revise it by excluding the counterexample. The really worrisome situations are those in which the counterexample points in many directions. If as in this case it is quite unclear where the problem begins and where it ends, one needs to put in a lot more work to lay the potential problems to rest. As far as the deficiencies of the rational number system is concerned, it is quite clear that the only way out of such problems is, if possible, to construct a number system in which such things can not happen. Most ideally, that new number system needs to incorporate the rational numbers inside it, too, in similar fashion to the way in which the integers can be seen as embedded in the rational numbers. In the next chapter, we carry out this project.
Chapter 3

Construction of the real numbers from the rational numbers

3.1 Introduction

Informally, one might think of the set of real numbers as a set containing the rational numbers, and any other “number” which one can compute or estimate within any previously given error tolerance, starting with some kind of description
3.1. INTRODUCTION

of a property which defines the new “number.” Let us take as an example of this a very simple way to get an estimate for $\sqrt{2}$, which at the end of the previous chapter was seen (whatever else it might be) not to be something which is expressible as a rational number. Even so, it is possible to construct rational numbers whose squares approximate 2 within any previously prescribed accuracy.

One might start by noticing that 1 is too small because $1^2 = 1 < 2$, and 2 is too big because $2^2 = 4 > 2$. Thus, the logical thing to do is to check what happens at the value of $3/2$, which is halfway between.

We get ourselves ready to repeat the above procedure in the following manner:

First, we must decide whether $(3/2)^2$ is more than two, or less. That is done of course by actually computation. If the result is more, then we know that 1 is too small and $3/2$ is too large. Therefore, we can use the point halfway between them as the next approximation, and the second step is completed. If on the other hand the result is less then $3/2$ is too small and 2 is too large. Again, we can consider the average of these two numbers as the next approximation, and the second step is completed.

In the actual event, $(3/2)^2 = 9/4 > 2$, and so we now can narrow our search to the left half-interval, which has 1 on its left end and $3/2$ on the right.

We can repeat the logical steps described above in order to continue the procedure as many times as we want. At each step, we have a left-hand number which is less than $\sqrt{2}$ and a right-hand number which is greater, and the resulting approximation is the average of the two. We only need to notice, now, that what we have done can be repeated as many times as one wishes, and moreover that after $n$ steps we have an interval which is of length $2^{-n}$. Moreover, the square of the number at the left endpoint of that interval is less than 2 and the square of the number at the right endpoint is
greater than 2, and the difference between the squares of these numbers tends to zero, too. Under the circumstances, it is both convenient and natural for us to believe that inside of all of these progressively smaller intervals there is a number which is “exactly equal” to $\sqrt{2}$.

Again, the above argument does not compute a rational number which is “exactly equal” to $\sqrt{2}$, because, as we have already seen, there can be no such rational number. But, the above procedure will certainly produce any desired degree of accuracy even though at every step we are staying inside the set of rational numbers. In a way, the question of whether or not there is any number of any kind which is “exactly equal” to $\sqrt{2}$ can now be answered by “Well, perhaps that depends upon what you mean by saying that such a number exists. If you are willing to believe that a construction of a process which gives us arbitrary accuracy is a proof for existence, then the question is answered.”

Of course, the above method is not the only method for computing $\sqrt{2}$ within arbitrary, pre-prescribed accuracy. There are lots of other methods, too, and the successive approximation steps in those other methods might not at all be identical to the ones outlined above, and the successive steps might well result in numerical outputs which are different from those above. Therefore, if we would adopt the idea that the real number which we wish to define is to be identified, somehow, with a procedure for approximating it, we would have to ensure (just for example) that all procedures for approximating $\sqrt{2}$ would actually have to approximate the same thing. This and other difficulties need to be solved, and not only for $\sqrt{2}$ but in general and for all occasions where questions might arise.

The rest of this chapter will deal with laying out the necessary groundwork for the construction of the real number system. In view of the sort of thing we did in the actual example we just investigated, which prescribed a sequential approximation method for a certain, specific real number, the first step is to start with some discussion of sequences,
which we will do in the next section of this chapter.

Before we get started with the rest of the chapter, though, we need to realize a few things. What we are doing is to construct a number system which expands the system of rational numbers in certain appropriate ways. We do not adopt the attitude that we “have” this new number system or that we “know” what it is until we have finished this construction. As a couple of prime examples of what this means, consider the following:

- “The real numbers consist of all the numbers between minus infinity and infinity.” is quite unsuitable as a definition. It leaves completely open the important question of what kind of numbers these are. The rational numbers, too, are between these two extremes. Are there any other numbers? Yes, or no? Why, or why not? In other words such a “definition” may be all very fine for helping some people sort of understand what a real number is, or some might think so. But it is a statement which actually has very little content and is totally unhelpful for our purposes.

- There are other ways to construct the system of real numbers besides the approach we are taking here. All of these methods are quite legitimate. Each of them also seems easy and natural from some standpoint, but complications arise when one gets into the details. One of these approaches is to describe the set of real numbers as equal to the set of all (possibly infinite) decimal expansions. This method has some natural features to recommend it. But one of the basic problems with it is that some numbers have two decimal expansions (Which ones? Do you happen to know?). Another problem is that, in fact, decimal expansions are sequences of a certain particular form, so why not start out by learning about sequences? How to do arithmetic with nonterminating decimals, anyway? How does it actually work? What happens when one needs to do carrying during addition, or borrowing during
CHAPTER 3. BUILDING THE REAL NUMBERS

3.2 Sequences

Formally, a sequence of elements from any set $S$ is a function from the natural numbers with range in $S$. The sequence is often represented by $\{s_n\}_{n=1}^\infty$, or more briefly, if the context makes the matter clear, as just $\{s_n\}$. Of course, more informally, a sequence may “start” at 0 or at some other integer, even a negative one. A finite sequence is a similar function defined on the set $\{1, \ldots, n\}$ for some natural number $n$. A subsequence of a sequence defined upon the set $S$ may be formally viewed as a composition of a function from the natural numbers to $S$ (a sequence) with a (strictly) increasing function from the natural numbers to the natural numbers. The effect is to “skip” some of the terms of the original sequence Given a sequence $\{s_n\}$, we write a subsequence as $\{s_{n_k}\}_{k=1}^\infty$ or simply as $\{s_{n_k}\}$.
3.3 Sequences of rational numbers

Now, after the above generalities about sequences, let us assume that we are working in particular with sequences of rational numbers.

We define the limit (if it exists) of a sequence \( \{s_n\} \) of rational numbers as a number \( s \) for which:

Given \( \epsilon > 0 \), there exists a number \( N \), such that, if \( n \geq N \), then \( |s_n - s| < \epsilon \).

If \( s \) is the limit of the sequence \( \{s_n\} \), we also write

\[
\lim_{n \to \infty} s_n = s.
\]

It should be kept in mind that, as we have given no definition to any set of numbers containing the rational numbers, this definition lives in a restricted context. The number \( \epsilon \) must perforce be rational, and the numbers \( s_n \) must be rational, and the number \( s \) must for the moment be rational, too. For we have given no formal definition for more than this.

A closely related concept to that of the limit of a sequence is that of a Cauchy sequence. The sequence \( \{s_n\} \) has the Cauchy property, provided that:

Given \( \epsilon > 0 \), there exists a number \( N \) such that if \( m \geq N \) and \( n \geq N \), then \( |s_m - s_n| < \epsilon \).

Note that again we are operating under restrictions. The numbers \( s_n \) and the number \( \epsilon \) must still be rational. We are working completely within the context of the rational numbers. We have not gone outside that context. Indeed, after we have completed the construction of the real numbers we will need to give these definitions again in that context.
The definitions will look the same, but they will not be the same because all of the quantities in them which have to be rational numbers here will be allowed to take on irrational (that is, other than rational) values there.

Exercises

3.3.1 If a sequence of rational numbers has a limit, then it does not have another, distinct limit.

3.3.2 Given two sequences \( \{a_n\} \) and \( \{b_n\} \) which have limits respectively equal to \( a \) and to \( b \), we can define
- their sum as the sequence \( \{s_n\} \), where \( s_n = a_n + b_n \).
- their product as the sequence \( \{p_n\} \), where \( p_n = a_n b_n \).
- their quotient as the sequence \( \{q_n\} \), where \( q_n = a_n / b_n \), provided that none of the \( b_n \) are zero.

Given a rational number \( q \) and a sequence \( \{a_n\} \), we can also construct the sequence \( \{qa_n\} \).
Show that, in all of these cases, the limit of the new sequence is what should be expected, if the original sequences \( \{a_n\} \) and \( \{b_n\} \) have limits. However, what additional condition, if any, must hold before the quotient sequence can be said to have a limit?

3.3.3 Who was the mathematician named Cauchy, and why does the concept of a Cauchy sequence carry his name?
Do some reading, and report to the class about this topic.

3.3.4 Show that every Cauchy sequence of rational numbers is bounded.
3.3.5 Taking into consideration the definitions in problem 3.3.2, show that, if the sequences \( \{a_n\} \) and \( \{b_n\} \) are assumed to be Cauchy sequences, then so are the sequences \( \{s_n\} \), \( \{p_n\} \), and (under reasonable conditions which you should determine) \( \{q_n\} \) (Problem 3.3.4 is useful here).

3.3.6 Show that every convergent (that is, possessing a limit) sequence of rational numbers is a Cauchy sequence.

3.3.7 If \( \{s_n\} \) is a Cauchy sequence of rational numbers, then every subsequence of it is a Cauchy sequence. Furthermore, \( \{s_n\} \) has a limit \( s \) if and only if every subsequence of \( \{s_n\} \) has limit \( s \), too.

3.3.8 If \( \{s_n\} \) is a Cauchy sequence of rational numbers and \( \{s_n\} \) does not have 0 as its limit, then there is \( b > 0 \) and there is some \( N \) (depending upon \( b \)) such that for all \( n \geq N \) we have either \( s_n > b \), or \( s_n < -b \). In either of these two possible cases there can be no more than a finite number of terms whose value is zero.

What this problem is saying is, either the values \( s_n \) are “ultimately positive” or they are “ultimately negative” and in either situation the terms are bounded away from zero. Note that the requirements here are only that \( \{s_n\} \) is a Cauchy sequence, and that the sequence does not have 0 as its limit. This language does not say that the sequence has a limit, only that if it does have a limit then that limit is not 0. A good start toward doing the problem is to construct the negation of the statement that the limit of the sequence is zero, and proceed from there, by combining that with the statement that the sequence is a Cauchy sequence.
3.4 The real numbers as Cauchy sequences of rational numbers

We now introduce the set of real numbers, which will often be denoted by the symbol \( \mathbb{R} \). As we did while constructing the rational numbers, we begin by defining an equivalence relation, this time upon the set of all Cauchy sequences of rational numbers. This equivalence relation could be defined by saying that, for any two sequences \( \{a_n\} \) and \( \{b_n\} \), we have \( \{a_n\} \sim \{b_n\} \) if and only if

\[
\lim_{n \to \infty} (a_n - b_n) = 0.
\]

Exercises

3.4.1 The advertised equivalence relation is indeed an equivalence relation.

We now can define the set of real numbers as the set of all equivalence classes of Cauchy sequences of rational numbers, using the equivalence relation described immediately above.

To extend the algebraic and order relations of the rational numbers to the set of real numbers, let us suppose that \( r \) and \( s \) are two real numbers. This means that, just as a rational number actually was an equivalence class containing many individual specimens, the same is true about \( r \) and \( s \). For each of \( r \) and \( s \) there corresponds a sequence of rational numbers, \( \{r_n\} \) a member of the equivalence class for \( r \) and \( \{s_n\} \) a member of the equivalence class for \( s \). Then:
3.4. THE REAL NUMBERS

- The equivalence class of 0 is any sequence tending to 0
- The equivalence class of 1 is any sequence tending to 1.
- The sum \( r + s \) may be taken to be the equivalence class of the sum of the two sequences \( \{r_n\} \) and \( \{s_n\} \)
- The product of \( r \) and \( s \) is the equivalence class of the product of the sequences \( \{r_n\} \) and \( \{s_n\} \)
- If \( s \) is not zero, then no sequence \( \{s_n\} \) which represents it has limit 0. If the sequence which has been chosen to represent \( s \) also has no terms in it which are equal to 0, then it is not a problem to construct the quotient sequence whose respective terms are \( \frac{1}{s_n} \). If, however, some of the terms in \( \{s_n\} \) do happen to be zero, then there are at most a finite number of such terms (see Exercise 3.3.8, above). For those terms, we can replace \( \frac{1}{s_n} \) (which otherwise does not exist) by anything we like. For example, we can handle this by declaring the corresponding term in the quotient sequence to be equal to zero – or anything else, for that matter, even a random pattern of values for those particular, offending \( n \), For, in no event are there more than a finite number of such terms. By this means, we can define \( s^{-1} \) to be the equivalence class of any sequence such that \( s(s^{-1}) = 1 \), where 1 also signifies an equivalence class of sequences, too. Namely, it is the equivalence class of all those sequences of rational numbers which have the limit 1.

Exercises
3.4.2 Addition of real numbers is well-defined.

3.4.3 Multiplication is well-defined.

3.4.4 The results in this problem have already been foreshadowed above. Please work out any details which need to be covered. If the real number $s$ is not equal to zero, then there is a representative $\{s_n\}$ from the equivalence class which defines $s$ satisfying the condition that $s_n \neq 0$ for all $n$. More than this, any sequence $\{s_n\}$ representing $s$ has the property that $\{n : s_n = 0\}$ is a finite set, and therefore the sequence $\{s_n\}$ is always equivalent to another sequence in which there are no zero terms.

3.4.5 Subject to the condition that a sequence representing the number $s \neq 0$ has had all of its zero terms suitably replaced, or that we replace it with any subsequence of itself which starts somewhere beyond any zero terms, the real number $s^{-1}$ is well-defined.

3.4.6 Show that the set of real numbers, with the algebraic operations which we have defined upon it, is an algebraic field (cf. Problem 2.4.4).

3.4.7 Show that the rational numbers can be naturally embedded into the real numbers, with algebraic operations preserved.

We now wish to extend the order relations defined upon the rational numbers to the real numbers. To this end, we declare a real number $r$ to be positive, provided that there is a (Cauchy) sequence $\{r_n\}$ of rational numbers representing...
it which is “ultimately positive.” That is, there is $N$ such that, if $n \geq N$, then $r_n > \delta$, for some rational number $\delta > 0$. We declare $r$ to be negative if $-r$ is positive.

**Remark:** The following exercises are to be done using the construction of the real numbers which we have performed here. Most particularly, they are not to be done by appeal to the often-used description of the real numbers as all of those numbers which can be represented by a decimal expansion. Although that statement happens to be true, we have **not** used it, nor have we at this point shown that it is correct. Most particularly, a decimal expansion of a real number comprises a representation of that number by a certain, specific kind of sequence generated by an infinite series. Infinite series will be discussed in the next chapter of this text. To establish the equivalence of the real numbers to the set of all decimal expansions will thus depend both upon the present chapter and the next chapter after this, and will be presented later on. This present material precedes that discussion. Please stick to using things which we have already done when working the problems currently presented here, not things which we will do in the future. That is the way that mathematics is developed.

**Exercises**

3.4.8 Between zero and any positive real number, there is a rational number. (Note that what is actually meant here is, between zero and any positive real number, there is a rational number, placed in its location by the natural embedding of the rational numbers into the real numbers. But we don’t need to go through saying all of that all the time, because we all know at this point what is really meant!)
3.4.9 Between zero and any positive rational number, there is an irrational (meaning, other than rational) real number. Indeed, between any two rational numbers there is an irrational number and also, for that matter, a rational number.

3.4.10 There is not a largest real number.

3.4.11 Between any two irrational numbers there is a rational number.

3.4.12 Every real number is either positive, negative, or zero.

3.4.13 (cf. Problem 3.4.6.) There is a natural embedding of the rational numbers into the real numbers which preserves the ordering of the rational numbers, as well as the algebraic operations upon them.

The real numbers are said to be the completion of the rational numbers. What this exactly means will be discussed in more detail in the next section. Here, let us only note here that to give the full justification of this statement, we must first of all extend the definition of the limit of a sequence and the definition of a Cauchy sequence to real numbers. To do this, we need only to drop the previous restrictions that \( \epsilon \) had to be rational, that that the terms of the sequence itself had to be rational, and that the limit had to be rational. Other than this, the definition remains at this point identical to the one which we have already been using. Here follow the formal statements.

We define the limit (if it exists) of a sequence \( \{s_n\} \) of real numbers as a number \( s \) for which:

Given \( \epsilon > 0 \), there exists a number \( N \), such that, if \( n \geq N \), then \( |s_n - s| < \epsilon \).
3.4. THE REAL NUMBERS

If \( s \) is the limit of the sequence \( \{s_n\} \), we also write

\[
\lim_{n \to \infty} s_n = s.
\]

To re-emphasize, the above “new” definition for the limit of a sequence of real numbers uses the same words and symbols as before, with the only change being that all the symbols which previously had to be rational numbers are now permitted to be real numbers. In other words, we now allow \( \epsilon \) to be a real number, the terms in the sequence to be real numbers, and the limit itself to be a real number. It should also be clear that, in this new context, every Cauchy sequence of rational numbers now has a limit, equal to that real number which the sequence of rational numbers represents.

The restatement of the definition of a Cauchy sequence of real numbers is, after this, perhaps not surprising. The sequence \( \{s_n\} \) of real numbers has the Cauchy property, provided that:

Given \( \epsilon > 0 \), there exists a number \( N \) such that if \( m \geq N \) and \( n \geq N \), then \( |s_m - s_n| < \epsilon \).

In view of what has come before, we can essentially repeat the contents of a previous set of exercises in this new context, too. The proofs of these results will not differ in any essential details, either, from those for the very similar results for rational numbers.

Exercises

3.4.14 If a sequence of real numbers has a limit, then it does not have another, distinct limit.
3.4.15 Given two sequences \( \{a_n\} \) and \( \{b_n\} \) which have limits respectively equal to \( a \) and to \( b \), we can define

- their sum as the sequence \( \{s_n\} \), where \( s_n = a_n + b_n \).
- their product as the sequence \( \{p_n\} \), where \( p_n = a_n b_n \).
- their quotient as the sequence \( \{q_n\} \), where \( q_n = a_n / b_n \), provided that none of the \( b_n \) are zero.

Given a real number \( r \) and a sequence \( \{a_n\} \), we can also construct the sequence \( \{ra_n\} \).

Show that, in all of these cases, the limit of the new sequence is what should be expected, if the original sequences \( \{a_n\} \) and \( \{b_n\} \) have limits. However, what additional condition, if any, must hold before the quotient sequence can be said to have a limit?

3.4.16 Show that every Cauchy sequence of real numbers is bounded.

3.4.17 Taking into consideration the definitions in problem 3.4.15, show that, if the sequences \( \{a_n\} \) and \( \{b_n\} \) are assumed to be Cauchy sequences, then so are the sequences \( \{s_n\} \), \( \{p_n\} \), and (under reasonable conditions which you should determine) \( \{q_n\} \) (Problem 3.4.16 is useful here).

3.4.18 Show that every convergent (that is, possessing a limit) sequence of real numbers is a Cauchy sequence.

3.4.19 If \( \{s_n\} \) is a Cauchy sequence of real numbers, then every subsequence of it is a Cauchy sequence; if \( \{s_n\} \) has a limit \( s \), then every subsequence of \( \{s_n\} \) has limit \( s \), too.

Now, as the final step of the construction of the real numbers, we show that the set of real numbers is indeed complete. We use the word “complete” in this context because we confronted a situation with the rational numbers in which things
were missing which we believe ought to be present. We have provided logical steps by which we were able to supply those things which seemed to be missing. So, what we have done is a completion of the rational numbers and have obtained a bigger set of numbers. So now the situation needs to be investigated, whether there is anything which seems to be missing in this new set of real numbers, or not? If we do what we already did, are we going to get some still bigger new set, or will we get the real numbers back again, with nothing new added? If we follow the same procedures as before, then in fact we do not get anything new. That is what we mean by saying that the real number system is complete.

Therefore, the final result which needs to be shown is that every Cauchy sequence of real numbers converges to a real number and not to some new entity. The following problem is therefore essential to all that comes after it.
3.4.20 Let \( \{r_n\} \) be a Cauchy sequence of real numbers. Then \( \{r_n\} \) has a limit which is a real number. (Note: The proof of this statement has been preceded by a certain method for the construction of the real numbers. Namely, every real number can be represented by a Cauchy sequence of rational numbers, and thus each of the real numbers in the sequence \( \{r_n\} \) has such a representation. Thus, to complete this problem it is a requirement to use that construction of the real numbers. Construct a Cauchy sequence \( \{q_n\} \) of rational numbers which is equivalent to the given sequence, \( \{r_n\} \), in the sense that \( \lim_{n \to \infty} (q_n - r_n) = 0 \). Then from the fact that it is possible to do this, explain why it follows that to find limit of the given Cauchy sequence \( \{r_n\} \) of real numbers, one does not need to go outside of the newly constructed system of real numbers to do another job of completion).

3.5 Equivalent formulations of completeness in R

We have taken the approach of constructing the real numbers, by starting with the rational numbers and studying Cauchy sequences of rational numbers. Then what we have shown in Problem 3.4.20 is the following:

**Statement of Completeness of R** Every Cauchy sequence of real numbers converges to a real number.

It is precisely this statement, that every Cauchy sequence of real numbers converges to a real number, which is what we mean here by saying that the real numbers are complete. The statement that the real numbers are the completion of the rational numbers means that the rational numbers are embedded in the real numbers, that the real numbers are
complete, and that any proper subset of the real numbers which contains the rational numbers will fail to be complete. It should be fairly obvious at this point to see that this must be true, since what we have exactly done is to use the set of Cauchy sequences of rational numbers to construct the real numbers.

Now, the main point of this section is to take notice that the Statement of Completeness, as stated in the form above, implies several other statements, which are in fact equivalent to it. The following problems explore some of these equivalent formulations. However, before stating them we give the definitions of supremum and infimum:

The supremum (if any) of a set $A$ is written $\sup A$ and is defined to be that (real) number $s$ satisfying the two conditions

(i) For all $a \in A$, $a \leq s$

(ii) For each $t < s$, there exists $a \in A$ such that $t < a$.

Similarly the infimum (if any) of a set $A$ is written $\inf A$ and is defined to be that (real) number $i$ satisfying the two conditions

(i) For all $a \in A$, $i \leq a$

(ii) For each $j > i$, there exists $a \in A$ such that $j > a$.

It should be noticed that, by these definitions, a set which is empty can have neither supremum nor infimum. It should also be noticed that no non-empty set can have two suprema, nor two infima (Remark about language: the words supremum and infimum are in fact Latin nouns, and their properly constructed plurals are indeed as given here).
In the following set of exercises, please use the Statement of Completeness in the form given here, that is, assume that every Cauchy sequence of real numbers converges.

Exercises

3.5.1 Every bounded monotone sequence has a limit. Show this based upon our Statement of Completeness (above) and also show that the result in this problem follows from the conclusion of Problem 3.5.2.

3.5.2 Every non-empty subset of \( \mathbb{R} \) which is bounded above has a supremum.

Note: You are asked to show that this statement follows logically from the Statement of Completeness given above. Past experience with teaching this material indicates that students tend to have psychological barriers which severely impede the ability to find a solution. If for any reason you think that what this problem asks you to do is simply impossible because it asks you to prove what you think you have heard or read somewhere is some kind of an “axiom” and everybody knows that you can’t prove an axiom so this problem must be mis-stated or must simply be wrong, then you are mistaken and for more than one reason. If you are having such an experience, then it might be helpful to re-read the title of this section, and the paragraphs in this section which precede this set of exercises, and to digest what those words are saying. Also it might help if you read the Remark which is below this problem set, which says some of the same things again.

3.5.3 Every non-empty subset of \( \mathbb{R} \) which is bounded below has an infimum.
3.5.4 Every nested sequence of bounded closed intervals has a non-empty intersection. More exactly:
Let \( \{I_n\} \) be a sequence of bounded closed intervals, such that \( I_{n+1} \subseteq I_n \) for each \( n \). Then

\[
\bigcap_{i=1}^{\infty} I_n \neq \Phi.
\]

The above statement need not be true if the word “bounded” is omitted. Find an example of a nested sequence of unbounded closed intervals for which the intersection is, in fact, empty.

And there are even more equivalent formulations of the Statement of Completeness, not listed here.

Remark: Quite often, mathematics courses whose content depends upon the properties of the real number system do not contain a construction of the real numbers. Rather, it is quite common that the real numbers are taken as a “given” and the course starts off with a list of their properties which is purely descriptive in nature. These listed properties are then called by the name “Axioms.” This has not been our approach in regard to the completeness property. Instead, what has been done in the preceding pages is to perform an actual construction of the real numbers by performing the completion of the set of rational numbers. This is something other than a mere description or listing of whatever properties the real numbers have, taken in isolation from an actual construction. Usually, when the merely descriptive approach has been taken, the Statement of Completeness is listed as one of the descriptive axioms. Usually, it is stated in the form given in Problem 3.5.2, and dignified with the name “The Completeness Axiom.” No proof is then given for that statement precisely because it has been taken as an axiom. There are three things which the student
should note well at this point:

i. The Statement of Completeness, given above as the statement that every Cauchy sequence of real numbers converges to a real number is not an axiom. It is a theorem. A proof for it has been given. The proof depended upon the construction of the real numbers which preceded the statement. We did not need to take it as an axiom because we actually constructed the set in which it is true and while doing the construction we showed that the statement is true for what we constructed. Treatments of the real numbers which do not do a construction but merely describe, have to state completeness as an axiom.

ii. Quite independently of whether or not the real numbers are constructed or merely described, all of the statements in the previous set of exercises are logically equivalent to one another, and also to our Statement of Completeness. That means, it is possible to assume any one of them as a hypothesis and prove any other of them as a conclusion. Whether or not any particular one of them is deemed to be an axiom has no relationship to their logical equivalence, which is a fact in itself.

iii. Do not make the mistake of assuming that the statement in any one of the above problems is true while attempting to prove it. Especially, the temptation to refer to the statement of Problem 3.5.2 as “The Completeness Axiom” is very great. If you succumb to this temptation, it will severely impede your understanding. It will probably vitiate all of your efforts to complete Problem 3.5.2. Why is that? Because you will believe that what you are supposed to prove in Problem 3.5.2 is an “axiom” in spite of everything that is being said right here. If you have
fallen into this trap, you will somehow think that Problem 3.5.2 is requesting you to use an axiom to prove itself, which it most definitely is not, and which would obviously be an absurd thing to ask anyone to do. Again, the reason for this confusion: many books, including even the typical calculus text, do indeed refer to a “Completeness Axiom” and state it in the same way as Problem 3.5.2. Those books intend merely to describe the number system and then to get on with some other business. Please do understand that our Statement of Completeness was a theorem which follows from an actual construction of the real number system. In our context, then the name “Completeness Axiom” is a misnomer. The real number system which has been constructed in the preceding pages certainly does have the property of completeness, but that property does not come in the form of an axiom.

Now, it is also true that other constructions of the real number system are feasible (for example, see Problem 3.5.8). If the set of real numbers has been constructed by some other method of construction, then the Statement of Completeness might very well have a natural expression which naturally derives from that construction, just as our Statement of Completeness naturally derives from our construction. Moreover, even if the real numbers are merely presented as some pre-existing entity which one wishes to describe with a set of “axioms” there is the necessity of showing that all of the different possible expressions of the “Completeness Axiom” are logically equivalent. For example, if a “Completeness Axiom” has been stated in the form which is given in Problem 3.5.2, it is then clearly necessary to prove as a theorem that every Cauchy sequence converges, that every bounded monotone sequence has a limit, and any or all of the other statements which are equivalent to the stated “Completeness Axiom” so that these other statements are ready for use when needed. The following sequence of problems explores this “reverse” implication and some other related topics:
CHAPTER 3. BUILDING THE REAL NUMBERS

Exercises

3.5.5 Every non-empty subset of $\mathbb{R}$ which is bounded above has a supremum if and only if every non-empty subset which is bounded below has an infimum.

3.5.6 If every non-empty subset of $\mathbb{R}$ which is bounded above is presumed to have a supremum, then every bounded nondecreasing sequence has a limit. A similar statement is true about sets which are bounded below and nonincreasing sequences. Note that, in this problem, one shows that the statements about bounded monotone sequences can be derived from the conclusion stated in Problem 3.5.2.

3.5.7 If every bounded monotone sequence is presumed to converge, it follows that every Cauchy sequence also converges (Hints: Show that every sequence contains a monotone subsequence, that a Cauchy sequence is bounded and therefore that any subsequence of a Cauchy sequence must also be bounded, and that, if a Cauchy sequence has any convergent subsequence, then it must itself be convergent to the same limit as the subsequence).

3.5.8 As there are several characterizations of completeness, so there are alternative methods for the construction of the real numbers. One of those methods is the method of Dedekind cuts. Do a class presentation outlining this method.

Finally, it should be noticed that there is a characterization of the real numbers which is often stated in textbooks used in more introductory classes in mathematics than this one. Namely, the real numbers are often described as the
set of all numbers which can be described by a decimal expansion. The justification of this statement is a project in itself. For, to prove that this is indeed true requires a lot of machinery which we have not developed. Nevertheless, it is also a true and accurate characterization of the real number system. Some aspects of this characterization will be described or delved into, later on. For the here and now, let us only note that the decimal expansion of any real number can be viewed as a sequence of a special type which converges to the real number in question. To see this, let us assume without loss of any actual generality that the real number in question is somewhere in the interval $(0, 1]$ Then, we have a decimal expansion for the number $r$ which looks like

\[ r = .a_1a_2a_3a_4 \ldots \]

in which each of the numbers $a_k$ is one of \{0, 1, \ldots, 9\}. Then, what we have in terms of sequences and limits is the statement that

\[ r = \lim_{n \to \infty} A_n \]

in which \( \{A_n\} \) is the sequence of fractions defined by \( A_n = \frac{a_1 \ldots a_n}{10^n} \) where the numerator \( a_1 \ldots a_n \) is the integer which has the given digits in its (standard) decimal expansion. Consequently, the representation of a real number as a decimal expansion is but a special case of the representation of a real number as a Cauchy sequence. Indeed:

Exercises
3.5.9 Let \( \{A_n\} \) any sequence whose terms are of the form

\[
A_1 = \frac{a_1}{10}
\]

and for \( n > 1 \)

\[
A_n = A_{n-1} + \cdots + \frac{a_n}{10^n}
\]

in which for each \( n \) the integer \( a_n \) satisfies \( 0 \leq a_n \leq 9 \). Then \( \{A_n\} \) is a Cauchy sequence.
Chapter 4

Series finite and infinite

4.1 Finite series and sigma notation

Let us start with a finite sequence of numbers \( \{a_1, \ldots, a_n\} \). Then, the notation

\[
\sum_{k=1}^{n} a_k
\]
denotes the sum of the numbers in the sequence. The number denoted by $k$ is called the *index of summation* or, sometimes, the “counter” or the “pointer.” Notice that its presence here is important because of its role, not its name or its particular symbol. Instead of $k$, one could have called it $i$ or $j$ or $m$ upon whim, without changing the meaning at all. It should also be clear that one can systematically change the upper and lower limits for the index which are given in the sum and at the same time can systematically relabel the terms in the sum, and the sum comes out to be the same thing. Moreover, though the description above considers a list of consecutive numbers as indices or counters which starts with 1, the list of the counters could obviously start with any number and proceed from there. Indeed, in many applications, especially those involving successive powers of a number, it is quite usual to start the indicated sum with the initial value 0 instead of 1 (see for example Problem 4.1.4 below).

Because of its flexibility in expression, the sigma notation is quite powerful and makes a lot of things easy which are otherwise not at all obvious, but it is often introduced so suddenly (usually in the middle of a calculus course which is already moving too fast) that students come away with phobias. Let us try to overcome those phobias in case that any remain. First, let us note that the following two properties are obvious:

\[
\sum_{k=1}^{n} a_k = a_1 + \sum_{k=2}^{n} a_k = a_n + \sum_{k=1}^{n-1} a_k
\]
• More generally, when \( 1 \leq m < n \) we have

\[ \sum_{k=1}^{m} a_k + \sum_{k=m+1}^{n} a_k = \sum_{k=1}^{n} a_k \]

Here follow some introductory exercises. Note that, for really rigorous proofs of some of them, you may wish to use Mathematical Induction or the Well Ordering Principle.

**Exercises**

4.1.1 Show that \( \sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k = \sum_{k=1}^{n} (a_k + b_k) \).

4.1.2 Further, show that the distributive law holds, namely that if \( c \) is any number, we have \( c \sum_{k=1}^{n} a_k = \sum_{k=1}^{n} ca_k \).

4.1.3 Exercise on “the gentle art of index sliding”:

\[ \sum_{k=1}^{n} a_k = \sum_{k=0}^{n-1} a_{k+1} = \sum_{k=2}^{n+2} a_{k-1} \]
4.1.4 (Sum of the finite geometric series) Use the properties given in the previous exercises to show that, when \( r \) is any number, we have

\[
(1 - r) \sum_{k=0}^{n} r^k = \sum_{k=0}^{n} r^k - \sum_{k=0}^{n} r^{k+1} = 1 + \sum_{k=1}^{n} r^k - \sum_{k=1}^{n} r^k - r^{n+1} = 1 - r^{n+1}
\]

From this, find the formula which gives \( \sum_{k=0}^{n} r^k \) whenever \( r \neq 1 \).

Note: An agreement on a small point regarding notation is relevant here. The value of \( r^0 \) is equal to 1 for all values of \( r \) which are not zero. We agree in the present context that \( r^0 = 1 \) even when \( r = 0 \). Without any such carefully restricted context, the expression "0^0" can not be assigned any precise meaning.

4.1.5 Note that \( \sum_{k=1}^{n} 1 = n \), and note that \( \sum_{k=1}^{n+1} k^2 - \sum_{k=1}^{n} k^2 = (n+1)^2 \). The first series on the left can now be rewritten using “index sliding,” obtaining

\[
\sum_{k=1}^{n+1} k^2 = \sum_{k=0}^{n} (k + 1)^2 = \sum_{k=0}^{n} k^2 + \sum_{k=0}^{n} 2k + \sum_{k=0}^{n} 1
\]
Therefore, from

\[ \sum_{k=1}^{n+1} k^2 - \sum_{k=1}^{n} k^2 = (n + 1)^2 \]

one obtains (noticing that all positive powers of zero are zero),

\[ \sum_{k=1}^{n} k^2 + \sum_{k=1}^{n} 2k + n + 1 - \sum_{k=1}^{n} k^2 = \sum_{k=1}^{n} 2k + n + 1 = (n + 1)^2 \]

Verify the above steps for yourself, and complete the problem by finding a formula for \( \sum_{k=1}^{n} k \).

Note that Problem 4.1.5 actually gives a constructive, step by step derivation of a summation formula. The student may well have been given the resulting formula or a similar one and asked to prove by mathematical induction that the formula, as given, is correct. But mathematical induction is not very helpful, at all, if one does not know the formula and needs to find it.

Exercises
4.1.6 The previous problem can be generalized to a method for finding the summation formula for \( \sum_{k=1}^{n} k^p \) in which \( p \) is any positive integer. Show in detail how this can be done when \( p = 2 \), and explain what is needed in order to obtain the formula corresponding to any higher value of \( p \).

4.2 Tools – the binomial theorem

The binomial theorem states that, for a given positive integer \( n \)

\[
(a + b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k},
\]

in which

\[
\binom{n}{k} = \frac{n!}{k!(n-k)!}
\]

is often read as “n choose k” from another application of the binomial theorem, in probability.
Exercises

4.2.1 Prove the Binomial Theorem (Hint: try using induction).

4.3 Infinite series

An infinite series is written formally as

$$\sum_{n=1}^{\infty} a_n,$$

where \(\{a_n\}\) is some sequence of numbers, which in this connection are called terms of the series. Closely related to the series is the sequence of partial sums \(\{A_N\}_{N=1}^{\infty}\), where \(A_N\) is defined by

$$A_N = \sum_{n=1}^{N} a_n.$$

We say that the series converges or is convergent provided that

$$\lim_{N \to \infty} A_N$$
exists and is finite (this statement is for clarity only, as we have not at this time defined limits which are other than finite). A series which is not convergent is often called *divergent*, or it is said that the series *diverges*.

Students in calculus classes often find the difference between “has a limit” and “converges” to be confusing. There really ought to be no confusion, provided only that the student realizes the reason we have these two slightly different concepts in the first place. Namely, the use of the term “converges” instead of “has a limit” harks back to the use of infinite series and infinite sequences as computational procedures. For, when one is actually using a procedure for computation it is neither interesting nor useful if the results of the procedure increase without bound and “go to infinity.” In fact, such a result usually indicates that something needs to be done over again because nothing good came out of it.

**Exercises**

4.3.1 If $\sum_{n=1}^{\infty} |a_n|$ converges, then $\sum_{n=1}^{\infty} a_n$ also converges. A given series $\sum_{n=1}^{\infty} a_n$ for which $\sum_{n=1}^{\infty} |a_n|$ is known to converge is said to **converge absolutely**. Thus, the result shown in this problem is often restated as “A series which converges absolutely must itself converge.” The converse of this statement is obviously not universally valid, of course. It is quite possible for a given series $\sum_{n=1}^{\infty} a_n$ containing some mixture of positive and negative terms to converge while
the series $\sum_{n=1}^{\infty} |a_n|$ diverges. In such a situation, the series $\sum_{n=1}^{\infty} a_n$ is said to converge conditionally. There are many examples of this. Hint for the proof: We know that every Cauchy sequence converges. Therefore show that the sequence of partial sums is a Cauchy sequence.

4.3.2 Consider two series $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$. If there is $N$ such that $0 \leq a_n \leq b_n$ holds for all $n \geq N$, then

(i) if the second series converges, then the first also converges
(ii) if the first series diverges, then the second also diverges.

4.3.3 Given two series $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$. If both converge, then $\sum_{n=1}^{\infty} (a_n + b_n)$ also converges, and moreover

$$\sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} (a_n + b_n)$$

Similarly, if $k$ is any non zero number, then $\sum_{n=1}^{\infty} a_n$ converges if and only if $\sum_{n=1}^{\infty} ka_n$ converges. Moreover, when both
converge, we have
\[
\sum_{n=1}^{\infty} k a_n = k \sum_{n=1}^{\infty} a_n
\]
The above formula is also true when \( k = 0 \) if \( \sum_{n=1}^{\infty} a_n \) converges, of course, but is hard to interpret if \( \sum_{n=1}^{\infty} a_n \) diverges (The left side is obviously zero then, but just what exactly does the right side then mean?).

4.4 Tools – the geometric series

A particularly interesting and basic infinite series is the geometric series
\[
\sum_{k=0}^{\infty} r^k.
\]

Exercises

4.4.1 The geometric series diverges for \(|r| \geq 1\), while for \(|r| < 1\) its sum is equal to \(\frac{1}{1-r}\). Note: An integral part of this problem is to show rigorously, by appeal to things done previously in this text only, that
\[
\lim_{n \to \infty} r^n = 0 \text{ whenever } |r| < 1.
\]

4.4.2 Any repeating decimal defines a rational number, and any rational number can be represented as a repeating decimal.

4.5 A small excursion – the definition of \( e \)

Taking a historical view of mathematics and its applications, the number \( e \) arose from the study of compound interest. If the period of compounding is decreased to zero, then one is presumed to have “continuous compounding,” and the question naturally arose as to what happens then. As a very simple case, suppose one has annual compounding at 100% interest, starting with $1.00 principal. Then, after one year one has $2.00. If one compounds quarterly, then one has four periods per year, and the interest rate is therefore 25% per period. The amount in dollars accumulated after one year (ignoring such inconveniences as roundoff to the nearest cent, which we intend to ignore here for the purposes of higher mathematics) is

\[
(1 + \frac{1}{4})^4.
\]
More generally, if the compounding is with $n$ periods per year, then the amount accumulated after one year at 100% interest, starting with one dollar, is

$$(1 + \frac{1}{n})^n.$$ 

At some point, some clever money lender noticed that he could get a lot of word-of-mouth publicity among potential investors and steal a march on his competition by shortening the period of compounding, and he also did some calculations which seemed to indicate that there seems to be some kind of ceiling on what comes out of this sequence, so he could see that the buzz thus generated was not really going to cost him very much money. Based upon these origins and taking into account what we know now, it is natural to try to define a number, which is usually denoted by the symbol $e$, as

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n.$$ 

This number is called $e$ in honor of the eighteenth-century mathematician Euler. To put the matter into modern terms, he gave a careful description of it and a proof that it actually exists, after being given a royal grant by Catherine the Great, Tsarina of Russia, to explore the topic of compound interest.

Of course, the above definition of $e$ is only an attempt at a definition unless and until it can be shown actually to define something. To this end, it is necessary to show that the sequence in the definition actually does have a limit and that this limit actually is a finite number. The details of proving these two things are left to the student, in the first two of the following three problems.
Exercises

4.5.1 Let $s_n = (1 + \frac{1}{n})^n$. Show that $\{s_n\}$ is an increasing sequence. Hint: Use the Binomial Theorem to expand $(1 + \frac{1}{n})^n$ and $(1 + \frac{1}{n+1})^{n+1}$ and carefully compare the indexed terms from 0 to $n$.

4.5.2 Show that the same sequence $\{s_n\}$ is bounded above (it is fairly easy to show that the sequence is bounded above by something like 10, a bit harder to show it is bounded above by 4, harder yet to show it is bounded above by 3).

4.5.3 If you followed a fairly obvious line of reasoning in order to establish the previous problem, you ought to be able to generalize your argument to show (using only the tools of this chapter) that in fact

$$e = \sum_{k=0}^{\infty} \frac{1}{k!}$$
4.6 Some discussion of the exponential function

Without going into all the details at this time, let us note before moving on to other topics that it is possible to generalize what was done in the previous section to the discussion of a function of $x$ defined by

$$\lim_{n \to \infty} (1 + \frac{x}{n})^n$$

and similarly it is possible to see that the value of this function can actually be calculated by the infinite series

$$\sum_{k=0}^{\infty} \frac{x^k}{k!}.$$ 

The argument by which this can be shown is similar to what has been done in the previous section, with a few refinements.

What is not clear at first, of course, is that the series just given is also equal to $e^x$, where $e$ is the number described in the previous section. For, to establish this would require the proof of the identity

$$\lim_{n \to \infty} (1 + \frac{x}{n})^n = \lim_{n \to \infty} (1 + \frac{1}{n})^{nx}$$

This identity can indeed be shown without using any of the tools of “calculus” other than the concept of the limit of a sequence or sum of a series, as has been the practice in this chapter. Probably, the most effective way to carry out
the proof is to do it in stages, first assuming that $x$ is an integer, then that $x$ is a fraction of the form $\frac{1}{m}$, then proving it for any rational number, then showing that the result can be “completed” in a manner which, essentially, uses the constructions of the previous chapter in order to get the job done.

Note, however, that the above identity and trivial themes and variations on it lie behind many of the standard exercises in the calculus book on L’hôpital’s Rule. We can not use L’hôpital’s Rule here, of course, to prove anything at all about such limits. One reason for our inhibitions is that we have not at this point discussed derivatives, not even limits of functions, and so we are certainly not free to use such things. But there is a much more fundamental reason, too. Namely, the typical calculus book has not really defined the exponential function at all and indeed has not even taken the trouble to give a rigorous definition of $e$. Instead, the book has just done some vague mumble and hand-waving about such matters. In most of the books currently in use, this vague mumble and hand-waving about $e$ takes place in some very early chapter. Quite often, that very early chapter is even a chapter which is skipped during the course because there is too much other stuff which needs to get done, and the contents of that chapter are thus ignored both by the teacher and the students. Taking a somewhat cynical point of view, it seems it is apparently hoped (and, alas, not unreasonably so in practice) that by the time the students come to learn about L’Hôpital’s Rule they have forgotten just how vague the mumble and hand-waving were concerning that funny number $e$ – if indeed they were actually asked in the first place to look at that introductory chapter. Therefore, the arguments found in the Calculus book which, based upon mumble and hand-waving, purport to establish results via L’hôpital’s Rule about limits involving the exponential and logarithm functions are circular arguments by their fundamental nature. It is merely hoped that the students are too inattentive to catch on to what is being done to them.
Also note that such is not what we have done here.

4.7 More on Convergence of Series

4.7.1 The Root Test and the Ratio Test

There exist two more tests which are often used for showing the convergence or divergence of an infinite series. These tests are called the Root Test and the Ratio test. The most general of the two is the Root Test. A complete statement of it depends upon the concept of the “limit superior” of a sequence, which we have not introduced. Because of this, we present here the Root Test in the more restricted version which is often found in calculus books. We also give the Ratio Test.

These two tests are specifically related to series with positive terms. It goes without saying, therefore, that they can be used to show the absolute convergence of any series \( \sum_{n=1}^{\infty} a_n \) in which the terms are not all of the same sign. In that case, the Root Test or the Ratio Test would be applied to the related series \( \sum_{n=1}^{\infty} |a_n| \) and then, if the test shows convergence of the second series then the first one converges, too, and converges absolutely. See Exercise 4.3.1.
Exercises

4.7.1 (a weak version of the Root Test, what is presented in most calculus texts) Consider an infinite series \( \sum_{n=1}^{\infty} a_n \) in which all of the terms (or all of the terms for \( n \geq N \), some \( N \)) are non-negative. Then, let \( r = \lim_{n \to \infty} (a_n)^{\frac{1}{n}} \), provided that the limit exists. If \( r < 1 \) the series converges. If \( r > 1 \) the series diverges. If \( r = 1 \) then this test tells absolutely nothing about the convergence of the series.

4.7.2 (the Ratio Test) Consider an infinite series \( \sum_{n=1}^{\infty} a_n \) in which all of the terms (or all of the terms for \( n \geq N \), some \( N \)) are positive. Then, let \( r = \lim_{n \to \infty} \frac{a_{n+1}}{a_n} \), provided that the limit exists. If \( r < 1 \) the series converges. If \( r > 1 \) the series diverges. If \( r = 1 \) then this test tells absolutely nothing about the convergence of the series.

A hint for the proof of both of the preceding exercises is to do a comparison test (see Problem 4.3.2) of the given series with a geometric series \( \sum_{n=0}^{\infty} \rho^n \) (or, better, with the terms in the “tail” of the geometric series consisting of all the terms beyond a sufficiently large index). If \( r < 1 \), then \( \rho \) can be chosen as any number satisfying \( r < \rho < 1 \), and if \( r > 1 \) then choose \( \rho \) to be any number such that \( 1 < \rho < r \). And note that when \( r = 1 \) it is not possible to do either of these.
The above hint shows that the Root Test and the Ratio Test depend, both of them, upon a comparison with a geometric series. In the previous discussion of $e$ we did in fact use an argument based upon the geometric series to establish error estimates and thereby to prove convergence.

Also, a few words about the use of these two theorems. The Ratio Test is often easier to set up, but it is less general. Even in the form seen in this section the Root Test is more general. The Ratio Test is completely destroyed if, for example, a lot of the terms $a_n$ could be zero. The problem becomes especially bad if the occurrence of zero terms happens in some apparently irregular fashion. The Root Test is also a lot quicker to do, if one happens to be able to compute the limit involved, which is sometimes not easy. To this end, the following problems are relevant. Note that it is forbidden while working these problems to use L'Hôpital's rule or any other tool involving derivatives. We have not even defined derivatives in this course, and so we certainly can not use them for anything at all. We also have not developed the exponential function, nor the logarithm function. Please stick to using machinery which is developed in the preceding pages, only.

Exercises

4.7.3 (this one is a loose end which needs to be tied up) Formulate an appropriate definition for what we ought to mean by the statement $\lim_{n \to \infty} a_n = \infty$.

4.7.4 If $a > 0$, then $\lim_{n \to \infty} a_{\frac{1}{n}} = 1$. Hint: Show first that if $a > 1$ we have a decreasing sequence which is bounded
below by 1. Show that, from this, it follows that when $a > 1$ the limit must exist and must be greater than or equal to 1. Find a way to complete the argument by showing that the limit (which is already shown to exist) can not, in fact, be any number strictly greater than 1.

Similarly, if $a < 1$ show that we have an increasing sequence which is bounded above by 1. The rest of the argument is quite similar to one for the case that $a > 1$.

4.7.5 $\lim_{n \to \infty} \frac{1}{n} = 1$. Hint: As the first step, show that this sequence is decreasing for $n \geq N$. This means, find that $N$, for one thing. Also notice that $\frac{1}{n} \geq 1$ for all $n$. Therefore, the series has a limit, and the limit is not less than 1. Complete the problem by showing that indeed the limit is 1.

4.7.6 $\lim_{n \to \infty} (\frac{n}{n!})^{\frac{1}{n}} = \infty$. Hint: Show that if $n > K$, where $K$ is any fixed but arbitrary integer (think of it as large) then $n! > K! \cdot K^{n-K} = (K-K!K!K)$. Now notice that the factor in parentheses is a constant, and so by Problem 4.7.4 we have $\lim_{n \to \infty} (\frac{n}{K-K!})^{\frac{1}{n}} = 1$. One might infer from this that $\lim_{n \to \infty} (\frac{n}{n!})^{\frac{1}{n}}$ is bounded below by $K$. Fill in the details, and complete the proof.

4.7.2 Conditional Convergence and the Alternating Series Test

Note: What is in this section is very interesting, but it is not very important for us to do in MATH 3100. We may skip this section or pass over it very lightly, depending on how the calendar looks.
In a previous section, in Exercise 4.3.1, the concepts of absolute and conditional convergence were introduced. We have seen just above that there are several extensive and detailed results about absolute convergence. What about results for conditional convergence? Well, that is a bit harder to prove, except in some very special circumstances. Without doubt, there are many series for which it is easy to show that they do not converge absolutely, and thus if they converge at all then they must converge conditionally, but no clear path toward such a proof seems to open itself for that, either. In exactly one circumstance, though, we do have a result which shows convergence of a series in which not all the terms are of the same sign, and it is rather nice. The result is called the Alternating Series Test. It also comes with an error estimate for the difference between the sum of the series and its $N$th partial sum.

**Theorem:** Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of non-negative numbers. Then the series

$$\sum_{n=0}^{\infty} (-1)^{n-1} a_n$$

will converge provided that the following two conditions are met:

(i) $a_n \to 0$ as $n \to \infty$

(ii) $a_n \geq a_{n+1}$ for all $n$.

Moreover, if the above two conditions are met, then the difference between the sum of the series and the $N$th partial
sum of the series is not greater than $a_{N+1}$ and the furthermore $N$th partial sum exceeds the sum of the series if $N$ is odd and is less than the sum of the series if $N$ is even.

Exercises

4.7.7 Prove the above result. As a hint, consider that the sequence of partial sums can be rewritten by grouping the terms two at a time in two different ways.

4.7.8 The series

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

is called the harmonic series. Show that for every $k \geq 1$ that

$$\sum_{j=2^k+1}^{2^{k+1}} \frac{1}{j} > \sum_{j=2^k+1}^{2^{k+1}} \frac{1}{2^{k+1}} = \frac{1}{2}.$$

Based upon this result, show that the harmonic series diverges.
4.7.9 Show that the alternating harmonic series defined by
\[
\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}
\]
converges, as it meets the criteria of the Alternating Series Test.

4.7.10 Consider the sequence \( \{a_n\} \) defined by \( a_n = \frac{2}{n} \) when \( n \) is odd and \( a_n = \frac{1}{n} \) when \( n \) is even, and the series
\[
\sum_{n=0}^{\infty} (-1)^{n-1} a_n
\]
based upon this sequence \( \{a_n\} \). Does this series meet all of the conditions which are required in the Alternating Series Test? Does it converge?

**Concluding Remarks:** The Alternating Series Test is a test for convergence. It applies to any series in which the terms of the series satisfy the hypotheses. All by itself, the Alternating Series Test says nothing at all, neither one way or the other, about whether a series converges conditionally or absolutely. Thus:

1. If a given series can be seen to converge by appeal to the Alternating Series Test but fails to converge absolutely, then that series does converge conditionally.
2. If a given series can be seen to converge by appeal to the Alternating Series Test and also can be seen (by some other test) to converge absolutely, then the series does converge absolutely. The fact that the Alternating Series Test applies to the given series does not violate the concept of absolute convergence.

3. If a series is seen to converge by the Alternating Series Test then the error estimate which is part of the Alternating Series Test also pertains to that series. Again, the matter of whether the Alternating Series Test pertains to the series has nothing whatsoever to do with the question of whether the given series converges absolutely, or does not converge absolutely.
Chapter 5

Cardinality

5.1 Finite and countable sets

A set is said to be finite if it is the empty set (which contains no elements at all), or if it can be put into a one-to-one correspondance with a set of the form \{1, \ldots, n\}, in which case the given set contains \(n\) elements.
A set is called **infinite** if it is not finite.

A set is called **countable** if it is finite, or if it can be put into one-to-one correspondence with the set of natural numbers, \( \mathbb{N} \). It should be obvious from this definition that \( \mathbb{N} \) itself is countable, and is not finite. This will be the first of a small group of exercises, none of which appear to be difficult, at least to the instructor. Of course, since the comparison of two sets almost invariably involves the construction of a function with one of them as the domain and the other as the range (or as the image), a revisiting of that portion of Chapter 1 which deals with functions and related terminology and properties can be helpful.

### Exercises

5.1.1 The set \( \mathbb{N} \) itself is countable and infinite.

5.1.2 Any subset of \( \mathbb{N} \) is countable, even if not finite.

5.1.3 The set \( \mathbb{Z} \) of all integers is countable.

5.1.4 The set of even integers in \( \mathbb{N} \) is countable and infinite.

5.1.5 The set of square integers is countable and infinite (this was explicitly noticed by Galileo, who was probably also not the first to notice it).
5.1.6 Since there is no largest prime number, the set of prime numbers is infinite and is also obviously countable (The proof that there is no largest prime number seems to date back at least to the ancient Greek mathematician Eratosthenes).

5.1.7 Any subset of \( \mathbf{N} \) which is not finite can be put into one-to-one correspondence with all of \( \mathbf{N} \).

5.1.8 Any subset of any countable set is countable.

The above exercises are, as stated above, fairly trivial. Less obvious is the fact that the set of rational numbers, \( \mathbf{Q} \), is countable. For the sake of making the argument a little bit simpler, let us confine our proof to the set of rational numbers which are non-negative:

Let the rational number \( q > 0 \), which can be given in lowest terms in the form \( q = m/n \), be mapped to the integer \( 2^m \cdot 3^n \), and let the rational number \( 0 \) be mapped to \( 1 \). Than the mapping thus defined is one-to-one, in that no two rational numbers can be mapped to the same integer. The output of this mapping can then be ordered by appeal to the natural ordering of the integers and thereby be represented as an increasing sequence.

Exercises

5.1.9 Can you find a simple way to adapt the above argument in order to show that the entire set \( \mathbf{Q} \) is countable?
An alternative method of proving that the set of positive rational numbers is countable is to prove that all of $\mathbb{N} \times \mathbb{N}$ is countable, in which we do not even consider the fact that rational number $q$ is identified with a pair $(m, n)$ that gives $q$ in lowest terms. To see that $\mathbb{N} \times \mathbb{N}$ is countable, visualize the set as graphed using a vertical and horizontal axis. Then one sees the set as represented by a grid of dots in the first quadrant (think of all the places on a piece of standard graph paper where a vertical and a horizontal line intersect). Then, observe that for each positive integer $k$ there are exactly $k$ pairs $(m, n)$ such that $m + n = k + 1$. That is, $(1, 1)$ is the only pair in which $m + n = 2$, and then $(2, 1)$ and $(1, 2)$ are the only two pairs in which $m + n = 3$, and then $(3, 1)$ and $(2, 2)$ and $(1, 3)$ are the only three pairs in which $m + n = 4$, and the same will be true for every value of $k$.

**Important Remark:** Notice that any of the schemes described above for counting the rational numbers will be highly incompatible with the standard order relation of the rational numbers. But there is no contradiction. We are not describing the rational numbers here as an ordered set, in which the order is compatible with the algebraic operations on $\mathbb{Q}$. We are only showing how to count them, which is something entirely different. Indeed, it is surely impossible to perform a counting operation which preserves the order relation previously defined, for the simple reason that between any two rational numbers there is another one. In that sense, there cannot be a rational number and then a “next” one. Obviously not. But, again, that is not what we are doing here.

**Exercises**

5.1.10 (The “salt and pepper” function) Let any method be established which counts the rational numbers in the
interval $[0, 1]$. That is, we can form a sequence $\{q_n\}$ in which every rational number in $[0, 1]$ is equal to $q_n$ for appropriate $n$. Let the function $f_n$ be defined by $f_n(q_n) = 1$ and $f_n(x) = 0$ for all other values of $x$ in $[0, 1]$. Then the function $f$ defined by $f(x) = \sum_{n=1}^{\infty} f_n(x)$ has the property that $f(x) = 1$ whenever $x$ is rational, and $f(x) = 0$ whenever $x$ is irrational.

5.1.11 Show that the function $f$ defined in the previous problem is continuous at no point in its domain. (Students in MATH 3100 are excused from doing this problem because we have done no formal treatment of continuity.)

Remark: No formal treatment of continuity has been done in MATH 3100. We have been sufficiently busy with doing an actual construction of the real number system. Nevertheless, I leave the above problem in the text even though it obviously cannot be assigned, on the grounds that it is something to think about. Another thing to think about is, how would one define the integral of this “salt and pepper” function if we were to wish to define it? Should it have an integral? Ought we to care? Why or why not? These are very good questions, by the way, and some of you might face these questions more seriously in the (not very near) future.
5.2 Uncountable sets, and the real numbers are uncountable

In Chapter 1, there was brief mention of the power set of a given set. It was stated there that the power set of the given set is the collection of all subsets of the given set. Let us describe the size of the power set, relative to the given set.

Exercises

5.2.1 Suppose that the set $S$ is finite and contains exactly $n$ elements. Then the power set of $S$, which we will denote by $\mathcal{P}(S)$, contains exactly $2^n$ elements.

Here, we can extend this result, showing that the power set of $\mathbb{N}$ itself contains somehow “more” elements than does $\mathbb{N}$, that is, that $\mathcal{P}(\mathbb{N})$ is uncountable. The way that we do this, of course, is to show that any such counting scheme which is alleged to work, in fact does not work. To show that any such scheme cannot work, it furthermore suffices to show that no such scheme can even complete the counting of the infinite subsets of $\mathbb{N}$, on the grounds that if those cannot be counted, then it is certainly not possible to count the finite subsets in addition to the infinite ones.

We prove that the collection of infinite subsets of $\mathbb{N}$ is uncountable:

Let $S_1, S_2, \ldots$ be any alleged counting of the infinite subsets of $\mathbb{N}$. For convenience, let us write for each $k$ the set $S_k$ in double subscript notation as a strictly increasing sequence $\{s_{kn}\}_{n=1}^{\infty}$. Now consider the “diagonal” set $D$ described by the sequence $\{s_{nn}\}$. From $D$, construct a new set, $D_{\text{new}}$, given by the sequence $\{d_n\}$, in which $d_{11} = s_{11} + 1$ and for each $n > 1$,
\[ d_n = \max \{ s_{nn} + 1, d_{n-1} + 1 \} \]. Then \( D_{\text{new}} \), by its construction, cannot be the same set as the set \( S_k \). For, \( D_{\text{new}} \) is also given in the form of a strictly increasing sequence of elements \( d_k \), and \( d_k \) cannot be equal to \( s_{kk} \) for any value of \( k \).

Now, we end with the main result of this chapter of these notes, by showing that the set \( \mathbb{R} \) is uncountable. To prove this, it clearly suffices to show that the interval \((0, 1)\) is uncountable. To do this, it is necessary to understand the ternary expansion of a real number. It is part of the final examination in this class to show that every real number \( r \) in this interval can be represented in the form of an infinite series, as

\[ r = \sum_{n=1}^{\infty} \frac{c_n}{3^n}, \]

in which \( c_n \) is either 0 or 1 or 2. This representation is unique for “most” values of \( r \). It is part of the final exam to determine exactly when it is not unique, but the two possibilities are that a number can either repeat \( c_n = 0 \) after a certain point, or it can repeat \( c_n = 2 \). Let us agree here that the infinite repetition of 2 is disallowed, and then the representation is unique. Once we have done that, we can agree that the real number \( r \) can be identified with the sequence of coefficients, \( \{c_n\} \), used in the above sum.

The proof now proceeds along lines similar to the proof of the previous result. Namely, we pretend that we have been presented with some attempt to count the real numbers in \((0, 1)\). That is, we have a listing of real numbers in the interval which could be represented as a sequence \( \{r_k\} \). Then, in turn each of the numbers \( r_k \) can be identified with the sequence of the coefficients, \( \{c_{kn}\} \), in its ternary expansion. Then, we define \( a_n = 0 \) if \( c_{nn} = 1 \), \( a_n = 1 \) if \( c_{nn} = 0 \),
and \( a_n = 1 \) if \( c_{nn} = 2 \). Having done this, it is clear that the sequence \( \{a_n\} \) is not identical to any of the sequences \( \{c_1n\}, \{c_2n\}, \ldots \). Moreover, the sequence \( \{a_n\} \) can not repeat 2. Therefore, the number \( r \), which is given by

\[
    r = \sum_{n=1}^{\infty} \frac{a_n}{3^n}
\]

can not equal any of the numbers \( r_1, r_2, \ldots \). Therefore, \( r \) is not in the alleged listing. It is clear that \( 0 \leq r < 1 \). If it is also true that \( r > 0 \), we are finished with the proof. But if \( r = 0 \) then we see that this outcome has required in particular that \( c_{11} = 1 \) to begin with. In that case we re-define \( r \) by requiring \( a_1 = 2 \) instead of requiring \( a_1 = 0 \), as above. and we keep all of the other coefficients \( a_n \) as defined already. In this eventuality, it is still the case that \( r \) can not be equal to any of the listed numbers \( r_1, r_2, \ldots \). And it is clear that \( 0 < r < 1 \).

Therefore, we have shown that any alleged counting of the real numbers in the interval \((0, 1)\) has fallen short of counting all of them, by constructing from the given act of counting itself a number which is in the interval \((0, 1)\) and which was omitted from the count. It follows that the interval \((0, 1)\) is an uncountable set.

The above proof contained an unexpected technicality at the end. Namely, we had to look out for the case that a number falls outside the interval. Interestingly, a direct proof of the same result can be constructed by invoking the expansion of the numbers in base 2 instead of base 3. It is comparatively easy to show that the numbers in the interval \((0, 1)\) can be expanded in base 2 instead of in base 3. This is given as an exercise on the final exam. But if this method of representation is used and a proof similar to the one given above for the uncountability of \((0, 1)\) is attempted, then the technicalities seem to increase.
Chapter 6

Representations of the real numbers by expansions

6.1 Introduction

The purpose of this chapter is to present some other standard ways of representing the real numbers. The results contained here will be presented in problem form, as their proofs would depend upon the concepts developed previously.
during the course.

6.2 The decimal representation of the real numbers

It is often stated in more elementary courses in mathematics, even at the pre-university level, that the real numbers consist of every number which can be represented as a decimal expansion, whether terminating or non-terminating, and if non-terminating then either repeating or non-repeating. This statement is true, of course. But there are some drawbacks to its use as a basic characterization of the number system. One of those drawbacks is that some numbers have more than one representation. Thus, these numbers would need to be characterized. Another, more serious problem is that true comprehension of the concept of a non-terminating decimal expansion clearly involves at a very basic level the understanding of the concept of a sequence and of the limit for a sequence. It is even better if the student has an understanding of series representations. The results, however, can be summarized in the following problem. The problem is stated only for the interval [0, 1) because the extension of its results to real numbers outside of that interval is presumed obvious.

Exercises

6.2.1 Every real number in the interval [0, 1) can be represented using a decimal expansion, or, more easily visualized,
in series form as

\[ \sum_{n=1}^{\infty} \frac{a_n}{10^n} \]

using appropriate choices for the coefficients \( \{a_n\} \), with the restrictions that \( a_n \) will be an integer, and \( 0 \leq a_n \leq 9 \). Moreover, every infinite series of this form represents a real number in \([0, 1)\), with the exception that if \( a_n = 9 \) for all \( n \) the sum of the series is exactly 1.

The representation is said to be **terminating** if there is some \( N \) such that for all \( n > N \) we have \( a_n = 0 \). Show further that the representation of a number is unique if and only if the number has no terminating expansion. But if a number has a terminating expansion then it also has a non-terminating expansion. Give a description or characterization of all those numbers which have a terminating expansion. Also explain how the non-terminating expansion for such numbers is constructed.

**Note:** A proper proof of all of the statements in the above problem will require the systematic use of the material in the previous chapters. Naturally, since the description in the problem involves a particular kind of series it is good to know about series. Also useful here is the Statement of Completeness in one or perhaps in more than one of its several equivalent formulations.
6.3 The binary representation of the real numbers

In the above discussion of the representation of the real numbers as decimals, nothing was particularly of interest about the number 10, except that we use it for such purposes by force of habit. Indeed, in some circumstances it is not even advantageous to use 10 as a base at all. Inside a computer, for example, every number is represented by a string of ones and zeroes. This has pretty much seeped into everyday culture at this point, and most people are quite aware of this, at least as far as integers are concerned. Computer scientists, computer engineers, and software engineers are quite accustomed to thinking in base 2 or in base 16 (hexadecimal). But what about numbers between 0 and 1? Can they also be represented in a form similar to the above? Yes, they can. We can’t call the results “decimals” of course, because “decimal” means that we are using base 10.

Exercises

6.3.1 Every real number in the interval [0, 1) can be represented using a binary expansion (the decimal point now becomes a “binary point”). Of course, for the purposes of this problem we can more easily visualize a “binary decimal” in series form. We write the number as

\[ \sum_{n=1}^{\infty} \frac{b_n}{2^n} \]

using appropriate choices for the coefficients \( \{b_n\} \), with the restriction that either \( b_n = 0 \) or \( b_n = 1 \) is true for every
n. Moreover, every series of this form represents a number in the interval [0, 1), with the sole exception that the series sums exactly to 1 if \( b_n = 1 \) for all \( n \).

The representation is said to be **terminating** if there is some \( N \) such that for \( n > N \) we have \( b_n = 0 \). Show further that the representation of a number is unique if and only if the number has no terminating expansion. But if a number has a terminating expansion then it also has a non-terminating expansion. Give a description or characterization of all those numbers which have a terminating expansion. Also give the form of the non-terminating expansion for such numbers.

**6.3.2** Represent the fractions 1/3 and 2/3 and 1/5 in their binary expansions. Show work, and explain why your work is valid.

### 6.4 Representation of the real numbers using any base

The above results clearly generalize. Let \( p \) be any integer greater than 1. You, the student, probably already know how to write any integer in its expansion base \( p \), and I am assuming that you do. But also the following is true.

In the above discussion of the representation of the real numbers as decimals, nothing was particularly of interest about the number 10, except that we use it for such purposes by force of habit. Indeed, in some circumstances it is not even advantageous to use 10 as a base at all. Inside a computer, for example, every number is represented by a string of ones and zeroes. This has pretty much seeped into everyday culture at this point, and most people are quite
aware of this, at least as far as integers are concerned. But what about the numbers between 0 and 1? Can they also be represented in a form similar to the above? Yes, they can.

Exercises

6.4.1 Let $p$ be any integer greater than 1. Then every real number in the interval $[0, 1)$ can be represented using an expansion in base $p$ (the decimal point now becomes a “$p$-ary” point), or, more easily visualized, in series form as

$$\sum_{n=1}^{\infty} \frac{c_n}{p^n}$$

using appropriate choices for the coefficients $\{c_n\}$, with the restriction that $b_n \in \{0, \ldots, p-1\}$ is true for every $n$. Moreover, every series of this form represents a number in the interval $[0, 1)$, with the sole exception that the series sums exactly to 1 if $c_n = p - 1$ for all $n$.

The representation is said to be terminating if there is some $N$ such that for $n > N$ we have $b_n = 0$. Show further that the representation of a number is unique if and only if the number has no terminating expansion. But if a number has a terminating expansion then it also has a non-terminating expansion. Give a description or characterization of all those numbers which have a terminating expansion. Also give The form of the non-terminating expansion for such numbers.
The reader should note that the above result has been applied already, using the value $p = 3$, in the previous chapter. But the proof that we can actually do that was deferred. In fact, the proof that it really can be done is found in this problem.

6.5 The Cantor set

The Cantor set is a somewhat infamous construction in mathematics. Frequently, it has been used as a counterexample to very natural conjectures which turn out instead to be based upon naive intuitions.

For example, now that we have seen that the real numbers in the interval $[0, 1]$ are uncountable, it is clear that that in fact many subsets of $\mathbb{R}$ as an interval can not be represented as a sequence. The reason is that the set of numbers used in a sequence is, by construction, a countable set. Again, this statement is clear enough and at this point should be obvious to all of us.

But the natural, human tendency after learning that an interval is not countable is to incline toward the belief that perhaps an uncountable set has to contain something like an interval somewhere inside it. This natural and naive impression turns out to be quite false, and the standard counterexample is the Cantor set.

The Cantor set may be defined in the following way:

Let us start with the representation of the real numbers in the interval $[0, 1]$ using base three. That is, we can write
each number $x$ which is in this interval as

$$x = \sum_{n=1}^{\infty} \frac{a_n}{3^n}$$

The usual rules, of course, restrict $a_n$ to be either 0 or 1 or 2, depending upon what is needed in order to express $x$ correctly. But we also notice that some values of $x$ have two expansions, one of them terminating and the other having $a_n = 2$ for all $n$ beyond some specific $N$. For some examples of this, we notice that we can write

$$1 = \sum_{n=1}^{\infty} \frac{2}{3^n}$$

and

$$\frac{1}{3} = \sum_{n=2}^{\infty} \frac{2}{3^n}$$

or as

$$\frac{1}{3} = \frac{1}{3} + \sum_{n=2}^{\infty} \frac{0}{3^n}.$$ 

Moreover, we can write the fraction $2/3$ in two different ways. One of them is the obvious, terminating one, and the
other represents it as

\[ \frac{2}{3} = \frac{1}{3} + \sum_{n=2}^{\infty} \frac{2}{3^n}. \]

Now, the Cantor set is defined as

\[ C = \{ x \mid x = \sum_{n=1}^{\infty} a_n \frac{3^n}{3^n} \text{ and } x \text{ has a representation in which } a_n \neq 1 \forall n \} \]

From the above construction of the Cantor set, it ought to be clear that it contains 0 and 1, but lots of “holes” are left in the interval. For example, just as a start it contains 1/3 and 2/3, but the open interval \((1/3, 2/3)\) is removed, as it were, in the first step of constructing \(C\). Indeed, another way to construct \(C\) (which has more intuitive visual content) is to remove that interval, and what remains is the union of the two intervals \([0, 1/3]\) and \([2/3, 1]\). In the second step, remove the middle third of each of these two intervals. In the third step, remove the middle third of each of the intervals which remained after the second step, and keep on in like fashion. However, the presence of some of the numbers which do remain in the Cantor set can be somewhat counterintuitive. It is easily verified, for example, that the fraction \(1/4\) is in the set because it most definitely has a representation in which the coefficients \(a_n\) are either 0 or 2.

The Cantor set \(C\) can be shown to contain no interval whatsoever, and it has many other properties which we have not defined here. Indeed, it has been the shipwreck of many plausible but false hypotheses during the development of
the modern concepts used in mathematics. But a most fundamental and important property of the Cantor set is that it is clearly not countable.

To see that the Cantor set is uncountable, let $x$ be an element of it. Recalling that in the above definition we used coefficients $a_n$ which are either 0 or 2, let $b_n = a_n/2$ for each $n$. Then we let $x$ correspond to the number

$$
\sum_{n=1}^{\infty} \frac{b_n}{2^n}.
$$

It is not difficult to see that this correspondence is one-to-one and that it maps the Cantor set $C$ onto the entire interval $[0, 1]$. We have established a one-to-one correspondence between the Cantor set and another set which we already know is uncountable.

Most particularly, this construction demonstrates that an uncountable set need not be a set which resembles an interval, nor even a set which contains an interval as a subset.

Finally, as a historical note it should be mentioned that the Cantor set was described not only by Georg Cantor in 1882, but also by at least three other mathematicians who actually published the description of it before he did. Those were Henry J. S. Smith (1874), Paul du Bois-Reymond (1880), and Vito Volterra (1881).
Chapter 7

Conclusion of MATH 3100

A major portion of the final examination was in take-home format. The main topic of the take-home portion of the final examination dealt with yet another equivalence of the completeness axiom for the real numbers, namely that every real number can be represented as a decimal, or analogously in a binary or a ternary expansion.

In addition to the above topic, the students sat for a final examination which dealt with some basic results in the
development of the rational numbers, and with certain “review” problems relating to the concepts of logic presented in Chapter 1 of this text.

Mathematics, Auburn University, Auburn AL 36849-5307