M365C
Real Analysis

September 2012

Contents

Preface

These notes were written for an IBL version of Real Analysis, M365C. The gentler version is M361K and IBL notes for that course are also posted at `http://www.ma.utexas.edu/ibl`

E. Odell
September 2012

CHAPTER 1

Introduction

You do not learn how to drive a car by watching someone else drive. You do not learn how to play the flute by watching someone else play. A experienced driver or flutist may give you information or encouragement, but you still need to drive or play yourself to gain any usable skills.

The text for this course consists largely of definitions and problems, with a bit of insight and encouragement provided by the authors. You will solve problems and present your solutions in class. This will be done individually and, from time to time, in small groups. Some of the problems may not be discussed in class, just like homework problems in a standard lecture-based course.

You will be expected to maintain a notebook which contains your solutions to all of the problems. After you complete a proof, you should re-write it for the sake improvement and understanding and perhaps re-write it again after hearing the class presentation and discussion. You are writing your own textbook.

You are not to consult other notes or textbooks. This will likely be the hardest course you have taken. If you work very hard, it will also be the most rewarding.

Advice:

(a) Work hard and keep up.
(b) Never miss class except under dire circumstances.
(c) See the instructor or the GRA as soon as possible if you are having problems.
(d) Form a study group, but work the problems yourself first. Use the study group to try out your solutions and more importantly to evaluate the solutions of other individuals.
(e) Don't be afraid to make mistakes or ask "stupid" questions. We learn through our mistakes. It's better to ask a stupid question and clear up a misconception than to fail an exam.

Proofs

Hopefully all of you have seen some proofs before. The word 'proof' is the name that mathematicians give to an explanation that leaves no doubt. The level of detail in this explanation depends on the audience. In research journals and high level text books, Mathematicians often skip steps in proofs and rely on the reader to fill in the missing steps. This technique can have the advantages of improving efficiency and focusing the reader to the new ideas in the proof but it can easily lead to frustration if the reader does not have the required background in mathematics. More seriously, missing steps can conceal mistakes; many mistakes in proofs (particularly at the undergraduate level) begin with "it is obvious that…"

In this course, we will try to avoid missing any steps in our proofs: each statement should follow from a previous one by a simple property of arithmetic, by a definition, by a previous theorem, or by an axiom. The justification of each step should be clearly stated. Writing clear proofs is a skill in itself. Often the shortest proof is not the clearest.

There is no mechanical process to produce a proof but there are some basic guidelines you should follow. The most basic rule is that every object that appears should be defined. In other words, when a variable, function, or a set appears we should be able to look back and find a statement defining that object:

(a) Let $\epsilon > 0$ be arbitrary.
(b) Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = 2x + 1$.
(c) Let $A = \{x \in \mathbb{R} : x^{13} - 27x^{12} + 16x^2 - 4 = 0\}$.
(d) By the definition of continuity there exists a $\delta > 0$ such that...

It is true that mathematicians often skip defining objects if they feel the intent should be clear (just as they skip steps in proofs). In this course, however, we will avoid this practice.

Secondly, always watch out for hidden assumptions. In a proof, you may want to say "Let $x \in A$ be arbitrary," but this does not work if $A = \varnothing$ (the empty set). A common error in real analysis is to write $\lim_{n\to\infty} a_n$ or $\lim_{x\to a} f(x)$ without first checking whether the limit exists (often this is the hardest part). Furthermore, one of the purposes of this course if to prove rigorously many of the basic results of single-variable calculus. As a result, we will be proving many statements that you "know" are true. Be careful not to assume something that you have not proven if you have been taught in school that it is true.

As we mentioned above, the audience for which a proof is intended influences the style and content of the proof. Your audience is another student in the class who knows the definitions and preceding results but is clueless as to how to prove the theorem.

Logic

In this section, we review some of the basic terminology used to describe logical arguments (and proofs). The framework we are establishing is perhaps a bit dry and is certainly not the heart of mathematics, but it is a useful language for discussing the various general constructs mathematicians use.

To start with, you should be familiar with the basic logical operators: if P and Q are *propositions*, i.e., statements that are either true or false, then you should understand what is meant by:

(a) not P,
(b) P or Q: by convention, the mathematical use of "or" is inclusive (rather than not exclusive) so that P or Q is true even if both P and Q are true,
(c) P and Q,
(d) if P then Q (or P implies Q or $P \Rightarrow Q$), and
(e) P if and only if Q (sometimes written "P is equivalent to Q" or "$P \Leftrightarrow Q$").

Similarly if $P(x)$ is a *predicate*, that is a statement that becomes a proposition when an object such as a real number is inserted for x, then you should understand

(a) for all x, $P(x)$ is true and
(b) there exists an x such that $P(x)$ is true.

Simple examples of a predicate are "$x > 0$" or "x^2 is an integer."

Most of our theorems will have the form of implications: "if P then Q." P is called the hypothesis and Q the conclusion.

Definition. The *contrapositive* of the implication "if P then Q" is the implication "if not Q then not P." The contrapositive is logically equivalent to the original implication. This means that one is true if and only if the other is true.

Sometimes it is much easier to pass to the contrapositive formulation when proving a theorem.

Definition. The *converse* of the implication if P then Q is the implication if Q then P.

The converse is **not** logically equivalent to the original implication.

Definition. A statement that is always true is called a *tautology.* A statement that is always false is called a *contradiction.*

To show that an implication is false, it suffices to find one situation in which the hypotheses are true but the conclusion is false. Such a situation is called a *counterexample.*

One technique of proof is by contradiction. To prove "P implies Q" we might assume that P is true and Q is false and obtain a contradiction. This is really proving the contrapositive.

We also have symbols for *logical quantifiers.* Explicitly, "$\forall$" means "for all" and "$\exists$" means "there exists." For example, we might say $\forall\ x \in \mathbb{R}$, $x^2 \geq 0$ or $\exists\ x \in \mathbb{R}$ such that $x > 0$.

To negate "$\forall$" we change to "$\exists$." For example,"$\forall\ x \in \mathbb{R}$, $x^2 \geq 0$" has the negation "$\exists\ x \in \mathbb{R}$ such that $x^2 < 0$." Likewise, the negation of "$\exists\ x \in \mathbb{R}$, $x > 0$" is "$\forall\ x \in \mathbb{R}$, $x \leq 0$".

These symbols are useful for scratch work (especially when it comes to negations), but we should point that they are regarded as informal and one would rarely seem them in a formal math text. More seriously, their use can often disguise the meaning of a statement or lead to ambiguity or confusion. It is thus our recommendation that you avoid using them in your formal proofs (that is, in your homework solutions).

CHAPTER 2

Sets, Functions, and Numbers

More so than numbers, the most basic building block of contemporary mathematics is a set. In this text we make no attempt to formally define the term set: this is surprisingly hard to do as one is inevitably lead to very deep philosophical questions which have no place in the present course. Instead we will simply remark that, intuitively, a set is a collection of objects, typically called the *elements* or *members* of the set. If A is a set, "$x \in A$" means that x is an element of A. As one might expect, $x \notin A$ means that x is not an element of A.

For this course, we will assume all the usual basic notion of set theory, which we will describe in this section. To be precise we are assuming the set-theoretic axioms of ZFC. In mathematics, an axiom is proposition whose truth is assumed without proof. In fact, it is a fundamental result of Kurt Gödel that one cannot study mathematics without using axioms: mathematics (or logic) cannot be the ultimate source of truth. It can only build its truth on a foundation of assumption.

Nevertheless the usual axioms modern mathematics are extremely mild assumptions. Again we will not attempt to make precise formulations, but they consist of statements whose intuitive content are things like "the notion of a set is well-defined and there exists a set," "there exists and infinite set," or "given two sets, their union is also a set." Thus if one wishes to take the platonic point of view (which means that mathematics is true in some 'real' sense), one need only assume some very 'obvious' statements.

Going beyond these philosophical considerations, we remark that we will assume that the sets $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$, consisting of the natural numbers, the integers, and the rational numbers, respectively exist and have all their usual basic properties.

1. The Basic Notions of Set Theory

Individual sets are described in a number of ways. If a set is sufficiently small, we may just list its elements. For example, $A = \{1, 2, 7\}$ is a set with three elements. We might also write something like $\mathbb{N} = \{1, 2, 3, \ldots\}$ which, though not completely precise, can impart no ambiguity as to the members of the set. Frequently, we will define the elements of a set by some property. For example

$$B = \{x \in \mathbb{Q} : 6x^5 - 27x^2 + 433x + 13 = 0\}$$

is a set whose members are well defined but perhaps not readily clear to us. In this example we read "B is the set of all x in $\mathbb{Q}$ such that $6x^5 - 27x^2 + 433x + 13 = 0$." The set with no elements is called the *empty set* and is denoted by $\varnothing$.

Definition. Let A and B be sets. A is a *subset* of B (denoted $A \subset B$) if every element of A is also an element of B.

Two sets are equal if they have the same elements. Thus

$$\{1, 2, 2, 2, 3\} = \{1, 1, 2, 3\} = \{1, 2, 3\}.$$

2.1. Prove that if A and B are sets then $A = B$ if and only if $A \subset B$ and $B \subset A$.

Definition. The *union* of A and B, denoted $A \cup B$, is the set consisting of all elements which lie either in A or in B. That is,

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

Thus

$$\{1, 2, 5\} \cup \{2, 4\} = \{1, 2, 4, 5\}.$$

Definition. The *intersection* of A and B, denoted $A \cap B$, is the set consisting of all elements which lie both in A and in B. That is,

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

A and B are called *disjoint* if $A \cap B = \varnothing$.

2.2. Prove that for all sets A, $\varnothing \subset A$. Is it true that for all sets A, $\varnothing \in A$? Is it true for some sets? Is $\{\varnothing\} = \varnothing$?

2.3. Suppose that A, B, and C are sets. Prove:

(a) $A \cap B = B \cap A$
(b) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
(c) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Definition. If A is a subset of another set U, the *complement* of A in U is

$$\{x \in U : x \notin A\} .$$

Oftentimes the larger set U is clear from context. If this is the case, we refer to the above set as the complement of A and denote it by $C(A)$. Thus if we are working in $\mathbb{Q}$, and $A \subset \mathbb{Q}$, then

$$C(A) = \{x \in \mathbb{Q} : x \notin A\} .$$

The next result is often called *DeMorgan's Laws*.

2.4. Prove that for all sets A and B

(a) $C(A \cap B) = C(A) \cup C(B)$
(b) $C(A \cup B) = C(A) \cap C(B)$
(c) $C(C(A)) = A$

Definition. If A and B are sets, the *set theoretic difference* of B from A is

$$A \setminus B = \{x \in A : x \notin B\}.$$

What is the difference between a set theoretic difference and a complement?

2.5. Prove or disprove the following statements:

(a) $(A \setminus B) \setminus C = A \setminus (B \cap C)$
(b) $(A \setminus B) \setminus C = A \setminus (B \cup C)$

(Hereafter, if a problem says "prove or disprove", you should give a proof if is true and counterexample if it is false)

It is often convenient to describe a set using an 'indexing' set. We write $X = \{A_i\}_{i \in I}$ to say that X is a set whose elements are A_i for each $i \in I$ (assuming that A_i has been defined in someway). For example, every set A can be written as $A = \{a\}_{a \in A}$. Frequently we will want the indexed elements to be sets themselves. If for each i in some indexing set I, A_i is a set, we write $\bigcup_{i \in I} A_i$ to denote the union of all the A_i. That is,

$$\bigcup_{i \in I} A_i = \{x : \text{ there exists } i \in I \text{ so that } x \in A_i\}.$$

Likewise,

$$\bigcap_{i \in I} A_i = \{x : \text{ for all } i \in I,\ x \in A_i\}\ .$$

$\{A_i\}_{i \in I}$ is called disjoint if $A_i \cap A_j = \varnothing$ for each $i, j \in I$ with $i \neq j$.

Though we have not yet officially defined the real numbers $\mathbb{R}$, you may use your basic knowledge to consider the following problem (and the remaining problems that appear before we formally introduce them).

2.6. For $x \in \mathbb{R}$ let $A_x = [x-1, x+1]$. Find:

(a) $\bigcap_{x \in \mathbb{R}} A_x$
(b) $\bigcap_{x \in [0,1)} A_x$
(c) $\bigcup_{x \in \mathbb{R}} A_x$

(d) $\bigcup_{x\in[0,1)} A_x$

2.7. Let $\{A_i\}_{i\in I}$ be a collection of sets. Prove the following generalizations of DeMorgan's Laws.

(a) $C(\bigcup_{i\in I} A_i) = \bigcap_{i\in I} C(A_i)$.

(b) $C(\bigcap_{i\in I} A_i) = \bigcup_{i\in I} C(A_i)$.

The previous result is also given the name DeMorgan's Laws.

Definition. If A is a set, the *power set*, $\mathcal{P}(A)$, of A is the set of all subsets of A. Thus

$$\mathcal{P}(A) = \{B : B \subset A\}.$$

2.8. What is the power set of $\{1, 2, 3\}$?

Definition. If A and B are sets, the *Cartesian product* of A and B, denoted $A \times B$, is the collection of ordered pairs (a, b) with $a \in A$ and $b \in B$.

Likewise we define Cartesian products of more that 2 sets. If $n \in \mathbb{N}$, the nth *Cartesian power* of A, denoted A^n, is the collection of ordered n-tuples with entries in A.

2. Functions

Definition. If A and B are sets, a *function* from A to B is a subset f of the Cartesian product $A \times B$ such that for each $a \in A$, there is a unique $b \in B$, with $(a, b) \in f$. We write $f : A \to B$ to mean that f is a function from A to B. A is called the *domain* of f and B is the and whose *codomain*.

Though the precise definition of a function is defined via the Cartesian Product, we almost never think of functions in this light. Instead we think of them as "maps" or "rules" which, given an element $a \in A$, produces an element $b \in B$. We write $a \mapsto b$ or $b = f(a)$. If $f : A \to B$, we might also say that "f maps A into B.

As a particular consequence of the precise definition, two functions from A to B are equal if only they take each element of A to the same element of B. Indeed, this condition will cause them to be equal as subsets of the Cartesian product. Likewise two functions cannot be equal unless they have the same domain and the same codomain (even if they represent the same "rule"). In general, the definitions in use today have often been fine-tuned over centuries of mathematical work. As a result, many will be phrased in such a way to build in conventions that mathematicians have deemed appropriate. Thus you should always look for hidden conventions of this kind when reading a new definition.

One simple example of a function is the *identity function*. If A is a set then the identity function, $\mathrm{id}_A : A \to A$ is defined by $\mathrm{id}_A(a) = a$ for all $a \in A$. If the domain of one function is the codomain of the other, we can compose the functions. Explicitly, if $f : A \to B$ and $g : B \to C$ are functions we define the *composition*, $g \circ f$ of f and g by $g \circ f(a) = g(f(a))$ for $a \in A$.

Before we continue, a remark on language. If we are being absolutely precise the previous sentence is not well-formulated. Instead of saying "if $f : A \to B$ and $g : B \to C \ldots$" we should really say something like "if A, B, and C are sets, $f : A \to B$, and $g : B \to C \ldots$. Otherwise we have not clearly defined the meaning of the symbols. Of course it will rapidly become clear that it would be cumbersome to use this level of precision in general.

Thus we adopt the philosophy that if a sentence requires a symbol to be a certain type of object than we are assuming that it is. In other words the statement "$f : A \to B$" requires that A be a set and so when we write it, we are implicitly assuming A is a set. Likewise if we write "let $\epsilon > 0$" we really mean that $\epsilon \in \mathbb{R}$ and $\epsilon > 0$. Of course this practice could potentially lead to ambiguity and so we will look to ensure that our meaning is always clear (and you should do the same). When it doubt, it is better to be cumbersome than unclear.

Definition. Let $f : X \to Y$. If $A \subset B$, $f(A) = \{f(a) : a \in A\}$ is called the *direct image* (or simply *image*) of A under f. Likewise, if $B \subset Y$, $f^{-1}(B) = \{x \in X : f(x) \in B\}$ is the *inverse image* or *pre-image* of B under f.

Thus, if $f : X \to Y$ then we have two associated functions $f : \mathcal{P}(X) \to \mathcal{P}(Y)$ given by $A \mapsto f(A)$ $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ given by $B \mapsto f^{-1}(B)$.

Once again, we have not formally defined $\mathbb{R}$, but we will use the standard interval notation to define subset of $\mathbb{R}$. For example,

$$(0, 3] = \{x \in \mathbb{R} : 0 < x \leq 3\}$$

and

$$[0, \infty) = \{x \in \mathbb{R} : 0 \leq x\}.$$

2.9. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^2$. Find

(a) $f((0, 3])$
(b) $f^{-1}((0, 3])$
(c) $f^{-1}((-1, 2))$

2.10. Let $f : X \to Y$ and let $A \subset X$, $B \subset Y$. Prove:

(a) $x \in f^{-1}(B)$ if and only if $f(x) \in B$.
(b) $x \in f(A)$ if and only if there exists an $a \in A$ such that $f(a) = x$.

2.11. Let $f : X \to Y$ and let $A, B \subset X$. Prove or disprove:

(a) $f(A \cup B) = f(A) \cup f(B)$
(b) $f(A \cap B) \subset f(A) \cap f(B)$
(c) $f(A \cap B) = f(A) \cap f(B)$

2.12. Let $f : X \to Y$ and let $A \subset X$ and $C, D \subset Y$. Prove or disprove:

(a) $f^{-1}(C \cup D) = f^{-1}(C) \cup f^{-1}(D)$
(b) $f^{-1}(C \cap D) = f^{-1}(C) \cap f^{-1}(D)$
(c) $f^{-1}(f(A)) = A$
(d) $f(f^{-1}(C)) = C$

Definition. Let $f : X \to Y$. f is called *injective* (or *one-to-one* or *1–1*) if for all $a, b \in X$, $f(a) = f(b)$ implies $a = b$. f is *surjective* or (*onto*) if for all $y \in Y$ there exists $x \in X$ with $f(x) = y$ (equivalently, if $f(X) = Y$).

Thus injective means that each element of X is sent to a different element of Y and surjective means that every element of Y is the image of some element of B.

2.13. Let $f : X \to Y$. Prove:

(a) f is injective if and only if for all $A \subset X$, $f^{-1}(f(A)) = A$
(b) f is onto if and only if for all $C \subset Y$, $f(f^{-1}(C)) = C$.

2.14. Let $f : A \to B$ and $g : B \to C$.

(a) Prove that if f and g are injective then so is $g \circ f$.
(b) Prove that if f and g are surjective then so is $g \circ f$.
(c) Suppose $g \circ f$ is injective. Must f and g both be injective?
(d) Suppose $g \circ f$ is surjective. Must f and g both be surjective?

Definition. A function which is both injective and surjective is called a *bijection*. It is also called a *one-to-one correspondence* or simply a *correspondence*.

2.15. Let $f : X \to Y$ be a bijection. Show there exists a function $f^{-1} : Y \to X$ satisfying $f \circ f^{-1} = \mathrm{id}_Y$ and $f^{-1} \circ f\mathrm{id}_X$.

Definition. If $f : X \to Y$ is a bijection, then the function f^{-1} defined above is called the *inverse* of f.

Intuitively speaking, the inverse of f reverses the action of f.

Caution: Given $f : X \to Y$ we have defined two other functions $f : \mathcal{P}(X) \to \mathcal{P}(Y)$ and $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$. If f is a bijection, we have yet another function $f^{-1} : Y \to X$. Technically, it is improper to denote two different things by the same symbol, but it is often convenient to do so if the two objects are closely related (we refer to this practice as *abuse of notation*). Be careful when reading a problem to make certain you understand which f or f^{-1} is meant.

3. The Completeness Axiom and Other Properties of $\mathbb{R}$

The collection $\mathbb{R}$ of real numbers is arguable the most important object in all of mathematics (at least in classical mathematics). It is perhaps a bit surprising that the first rigorous definition was not given until 1871 by Georg Cantor (using a technique known as "Dedekind cuts"). In fact, it turns out that giving a rigorous definition is a bit trickier than one might expect.

In the interest of time, we will avoid giving an exact construction. Instead, we will gives a list of properties satisfied by $\mathbb{R}$, which we will deem axioms. In other words, these statements will be axioms for our purposes, but they are not actually axioms for the whole of mathematics. Using the usual axioms of set-theory, one can construct the set of real numbers and deduce our axioms as properties (we will give an indication of one way to give a precise construction, though it will not be the one originally used by Cantor).

AXIOM 1: There exists an ordered field, $\mathbb{R}$, which contains $\mathbb{Q}$ and such that the operations and order on $\mathbb{R}$ extend those on $\mathbb{Q}$.

The language of this axiom is very succinct, but it contains quite a bit of information. The reader who is unfamiliar with these terms is encouraged to consult the appendix on prerequisites. As we mention in that appendix, $\mathbb{R}$ is not the only ordered field containing $\mathbb{Q}$ ($\mathbb{Q}$ is of course already an ordered field) and so we will need to give another property of $\mathbb{R}$ before we pin it down entirely. To formulate this axiom, we must introduce some more terminology.

Definition. Let $S \subset \mathbb{R}$. Recall (from the appendix on prerequisites) that an *upper bound* for S is an element $x \in \mathbb{R}$ so that $s \leq x$ for all $s \in S$. If S has an upper bound then we say that S is *bounded above.* $x \in \mathbb{R}$ is a *supremum* (or a *least upper bound*) for S if x is a minimal element in the set of upper bounds for S.

In other words, x is a supremum for S if it is an upper bound for S with $x \leq y$ for any other upper bound y of S. You will show in the next problem that supremums are unique.

2.16. Let $S \subset \mathbb{R}$ and let x_1 and x_2 be supremums for a set S. Prove that $x_1 = x_2$.

Hence we may designate **the** supremum of a set S, if it exists, as $\sup(S)$.

2.17. Let $S \subset \mathbb{R}$. Define:

(a) x is a lower bound for S
(b) S is bounded below
(c) x is a greatest lower bound for S
(d) Show that any two greatest lower bounds for a set S must be equal.

Just as supremum is a synonym for least upper bound, *infimum* is a synonym for greatest lower bound. We denote the greatest lower bound of a set S by $\inf(S)$.

2.18. For each of the following $S \subset \mathbb{R}$, determine all upper bounds and lower bounds. Also determined $\sup(S)$ and $\inf(S)$, if they exist.

(a) $(0, 1]$
(b) $\mathbb{N}$
(c) $\varnothing$
(d) $\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots, \}$
(e) $\{\frac{9}{10}, \frac{99}{100}, \frac{999}{1000}, \ldots\}$

We now give the final axiom on $\mathbb{R}$.

Axiom 2: $\mathbb{R}$ satisfies the *completeness axiom*: if $S \subset \mathbb{R}$ is nonempty set which is bounded above, $\sup(S)$ exists.

From the completeness axiom, we may prove the corresponding statement for infimums.

2.19. Let $S \subset \mathbb{R}$ and define $-S = \{-x : x \in S\}$ (the symbol we are '$-$' we are using here is not meant to reflect any relationship with set difference).

(a) Prove that x is an upper bound for S if and only if $-x$ is a lower bound for $-S$.
(b) Prove that $x = \sup(S)$ if and only if $-x = \inf(-S)$.
(c) Prove that if S is nonempty and bounded below, then $\inf(S)$ exists.

The following results will be very important for us when we start dealing with sequences. The statement of part (a) is often called the *Archimedean Property*

2.20. Prove the following consequences of the completeness axiom.

(a) For all $x \in \mathbb{R}$ there exists $n \in \mathbb{N}$ with $n > x$.
(b) For all $\epsilon > 0$ there exists $n \in \mathbb{N}$ with $\frac{1}{n} < \epsilon$.
(c) For all $x \in \mathbb{R}$ there exists $n \in \mathbb{Z}$ with $n \leq x < n + 1$.
(d) For all $x \in \mathbb{R}$ and $N \in \mathbb{N}$ there exists $n \in \mathbb{Z}$ so that $\frac{n}{N} \leq x < \frac{n+1}{N}$.
(e) For all $x \in \mathbb{R}$ and $\epsilon > 0$ there exists $r \in \mathbb{Q}$ with $|x - r| < \epsilon$.

The following is called the *density of* $\mathbb{Q}$ *in* $\mathbb{R}$.

2.21. Let $a < b$ be real numbers and let I be the open interval (a, b). Show that $I \cap \mathbb{Q} \neq \varnothing$.

A real number which is not rational is of course called irrational. The irrational numbers are also dense in $\mathbb{R}$.

2.22. Assume that $\sqrt{2} \in \mathbb{R}$. Prove that $\sqrt{2} \notin \mathbb{Q}$. Show that every open interval in $\mathbb{R}$ contains an irrational number.

We conclude the chapter by studying the basic properties of the standard absolute value on $\mathbb{R}$.

Definition. For $x \in \mathbb{R}$ we define the *absolute value* of x to be

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0. \end{cases}$$

2.23. Let $a, b, x \in \mathbb{R}$ and $\epsilon > 0$. Prove:
(a) $|a| \geq 0$ and $|a| = 0$ if and only if $a = 0$.
(b) $|ab| = |a|\,|b|$
(c) $|a + b| \leq |a| + |b|$.
(d) $|a| \leq \epsilon$ if and only if $-\epsilon \leq a \leq \epsilon$. $|a| < \epsilon$ if and only if $-\epsilon < a < \epsilon$
(e) $|x - a| < \epsilon$ if and only if $a - \epsilon < x < a + \epsilon$
(f) $|a - b| \geq \big|\,|a| - |b|\,\big|$.

Hint: For part (c), avoid tediously checking cases by using (with proof) the fact that for $x, y \geq 0$, $x \leq y$ if and only if $x^2 \leq y^2$.

The inequality $|a + b| \leq |a| + |b|$ is extremely important in analysis and is called the *triangle inequality.*

CHAPTER 3

Metric Spaces

In a sense, the real numbers are not the realm of abstract mathematics. Indeed, those academic fields which have the most to do with the "real world," such as physics, chemistry, or engineering, often work in the real numbers and may do so even more than the mathematician. Nevertheless, we saw in the appendix on prerequisites that introducing more abstract notions from algebra allowed us to better categorize and understand some of the properties of real numbers.

In this chapter, we will continue this philosophy by meeting the beginnings of an important field called *topology*. We will then study the topological properties of $\mathbb{R}$ (as well as other related sets). Very loosely speaking, topology is the study of elements a particular set being "close" to one another. To put a topological structure on a set is to give (in a very precise way) a notion of 'closeness' to the set.

We will introduce many different concepts related to metric spaces (and topology in general). It will be our practice to give the general definition (perhaps with a few examples) and to give the general results and theorems about the concepts. We will then apply our general results to the important specific cases we are considering, namely $\mathbb{R}$ and $\mathbb{R}^n$.

1. Basic Definitions and Examples

We will not study the ideas of topology in anywhere near their full generality: just as with algebra, the ideas of topology reach across essentially every area of mathematics. Instead, we will study one particular way to define a notion of closeness: by defining an explicit distance among the elements of our set. Just as our intuition suggests, the distance between two elements should be a nonnegative real number. We now give the explicit definition.

Definition. If E is a set, a *metric* on E is a function $d : E \times E \to \mathbb{R}$ such that

(a) for $x, y \in E$, $d(x, y) \geq 0$ and $d(x, y) = 0$ if and only if $x = y$
(b) for $x, y \in E$, $d(x, y) = d(y, x)$, and
(c) for $x, y, z \in E$, $d(x, y) \leq d(x, z) + d(z, y)$.

A *metric space* is a set E together with a fixed metric on E.

As mentioned above, we think of d as the distance from x to y. We submit that all three properties mentioned should intuitively hold if d is be a distance. The third property is called the *triangle inequality.* The elements of a metric space are typically called *points.*

Technically to give a metric space is to give a set and a metric on the set. In practice, however, we will typically name our metric spaces by the set alone. If we need to denote the metric we will most often use subscripts. In other words, if E is a metric space the corresponding metric will be d_E (whether we explicitly specify it or not).

3.1. Verify that each of the following is a metric space:

(a) $E = \mathbb{R}$ and $d_E(x, y) = |x - y|$

(b) E is any set and

$$d_E(x,y) = \begin{cases} 1 & x \neq y \\ 0 & x = y. \end{cases}$$

This metric is called the *discrete metric.*

(c) E is a subset of a metric space X and d_E is the restriction of d_X to $E \times E$.

(d) $E = (0, 1]$ and $d_E(x, y) = |x - y|$.

Frequently a set comes with a usual or standard metric. For example, the absolute value metric given above is the standard metric on $\mathbb{R}$. When a set has a standard metric, we will typically use it without explicitly saying so (unless we specify another metric).

In the case that our set is a real vector space (that is, a vector space over the field $\mathbb{R}$), it is a bit easier to define a metric. Indeed, we need only define a notion of the "size" of an element. We may then take the distance between the elements to be the size of their difference. The explicit constructions follow.

Definition. Let X be a real vector space. Then a function $\|\cdot\| : X \to \mathbb{R}$ is called a *norm* on X if for $x, y, z \in X$ and $c \in \mathbb{R}$, we have:

(a) $\|x\| \geq 0$ with $\|x\| = 0$ if and only if $x = 0$,

(b) $\|cx\| = |c|\,\|x\|$, and

(c) $\|x + y\| \leq \|x\| + \|y\|$.

A (real) *normed linear space* is a real vector space together with a fixed norm.

You will probably not be surprised to learn that the third requirement above is called the *triangle inequality.* As we mentioned above, normed linear space lead to metric spaces. We point out that the absolute value on $\mathbb{R}$ makes its a normed linear space over itself. Again, whenever we discuss in properties of $\mathbb{R}$ with depend on the choice of a norm, this is the norm we use unless otherwise specified. Moreover, we will drop the adjective 'real' from the phrase 'real vector space' in that all of of our vector spaces will be vector spaces over $\mathbb{R}$.

3.2. Suppose that $\|\cdot\|$ is a norm on the vector space X and define $d : X \times X \to \mathbb{R}$ by $d(x, y) = \|x - y\|$. Show that d is a metric on X.

Definition. We say that a subset S of normed linear space $(X, \|\cdot\|)$ is *bounded* if there is a $K \in \mathbb{R}$ with $\|x\| \leq K$ for each $x \in S$. We say a function into a normed linear space is bounded if its image is bounded as set.

In particular, a subset S of $\mathbb{R}$ is bounded if there is some $K \in \mathbb{R}$ such that $|x| \leq K$ for all $x \in S$. Notice that this is equivalent to saying that S is bounded both above and below (why?). also, recall (from Section 4 of the appendix on prerequisites) that if $n \in \mathbb{N}$, $\mathbb{R}^n$ is made into a vector space by using *point-wise operations.* That is,

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$$

for $x_i, y_i \in \mathbb{R}$ and

$$c(x_1, \ldots, x_n) = (cx_1, \ldots, cx_n)$$

for $c, x_i \in \mathbb{R}$.

Likewise if A is any set, we may define operations on the collection of functions $A \to \mathbb{R}$ by putting $(f+g)(x) = f(x)+g(x)$ and $(cf)(x) = cf(x)$ for $f, g : A \to \mathbb{R}$, $c \in \mathbb{R}$ and $x \in A$. This operations are again called point-wise operations.

3.3. Prove that each of the following are normed linear spaces.

(a) $\mathbb{R}^2$ under $\|(a, b)\| = |a| + |b|$.
(b) $\mathbb{R}^2$ under $\|(a, b)\| = \max(|a|, |b|)$.
(c) $\mathbb{R}^2$ under $\|(a, b)\| = \sqrt{a^2 + b^2}$.
(d) $\{f : \mathbb{R} \to \mathbb{R} | f \text{ is bounded}\}$ under $\|f\| = \sup\{|f(x)| : x \in \mathbb{R}\}$.
(e) $\{f : A \to \mathbb{R} | f \text{is bounded}\}$ under $\|f\| = \sup\{|f(a)| : a \in A\}$, where A is any fixed set.

Just as $\mathbb{R}$ comes with a standard (or obvious) choice for a norm, so does $\mathbb{R}^n$. Once again, we will assume for now that positive real number has a unique square root.

Definition. We define the *dot product* on $\mathbb{R}^n$ by putting

$$(x_1, \ldots, x_n) \cdot (y_1, \ldots, y_n) = x_1 y_1 + \cdots x_n y_n.$$

The *standard norm* on $\mathbb{R}^n$ is then defined by $\|x\| = \sqrt{x \cdot x}$ for $x \in \mathbb{R}^n$. $\mathbb{R}^n$ together with this norm is called *n-dimensional Euclidean space* and denoted $\mathbb{E}^n$.

Of course we must prove that the standard norm is actually a norm.

3.4. Let $\|\cdot\|$ be the standard norm on $\mathbb{R}^n$ and suppose that $x, y \in \mathbb{R}^n$. Prove the following.

(a) $\|x \pm y\|^2 = \|x\|^2 \pm 2x \cdot y + \|y\|^2$

(b) $|x \cdot y| \leq \|x\|\|y\|$

Conclude that the standard norm defines a norm on $\mathbb{R}^n$.

From this point forward, if we are working in $\mathbb{R}^n$, the symbol $\|\cdot\|$ is assumed to refer to the standard norm. You will notice that the metric which results from the standard norm on $\mathbb{R}^n$ is the same is usual Euclidean distance (that is, the so-called distance formula). You will also notice that the standard norm on $\mathbb{R}$ is just the absolute value.

2. Open and Closed Sets

Having discussed some of the important ways that we define metrics, we now study some of the general properties of metric spaces.

Definition. Suppose that E is a metric space. If $p_0 \in E$ and $r > 0$, we define the *open ball of radius r centered at p_0* to be

$$B(p_0, r) = \{p \in E : d_E(p, p_0) < r\} \ .$$

Likewise the *closed ball of radius r centered at p_0* is $\{p \in E : d(p, p_0) \leq r\}$ and is denoted by $B[p_0, r]$.

We remarked at the beginning of this chapter the topology is the study of nearness. In our current setting, this nearness is expressed with open balls. An open ball (particularly on with a small radius) is to be thought of as the collection of points near the point at which the ball is centered. Of course a small radius reflects a greater degree of nearness.

In our language, we required the radius of a closed ball to be nonzero, but some authors allow the choice $r = 0$. That is, one might put $B[p_0, 0] = \{p_0\}$. It is not very useful to consider an open ball of radius zero since since a thing is always the empty set.

3.5. For each of the norms given in parts (a), (b), and (c) of Problem 3.3 and for the discrete metric on $\mathbb{R}^2$, draw the closed and open balls centered at $(0, 0)$ of radius 1. Likewise, for parts (d) and (e) of *Problem* 3.3, describe $B[0, 1]$, where 0 is the zero function (that is the function whose value at every point is zero).

Having defined open balls, we may define more general open sets. The notion of an open set is extremely important in both topology and analysis (indeed the notion of an open set is the framework which is used to define topology).

Definition. If E is a metric space, a set $U \subset E$ is called *open* (in E) if for all $p_0 \in U$ there exists an $r > 0$ so that $B(p_0, r) \subset U$. $F \subset E$ is called *closed* if $C(F) = E \setminus F$ is an open set. A set is clopen if it is both closed and open.

3.6. Classify the set A as open, closed or neither.

(a) $A = \{(x, y) : x^2 + y^2 < 2\} \subset \mathbb{E}^2$
(b) $A = \{1, \frac{1}{2}, \frac{1}{3}, \ldots\} \subset \mathbb{R}$

3.7. Let (E, d) be a metric space. Prove:

(a) $\varnothing$ and E are both *clopen* sets in E.
(b) If I is any set and for all $i \in I$, U_i is an open set in E then $\bigcup_{i \in I} U_i$ is open.
(c) If $n \in \mathbb{N}$ and $U_1, \ldots, U_N$ are open sets in E then $\bigcap_{i=1}^{n} U_i$ is open.

We often interpret parts (a) and (b) by saying that a union of open sets is open and and a finite intersection of open sets is open.

3.8. Let (E, d) be a metric space. Prove:

(a) If I is any set and for each $i \in I$, $F_i \subset E$ is closed then $\bigcap_{i \in I} F_i$ is closed.
(b) If $F_1, \ldots, F_n$ are closed in E for some $n \in \mathbb{N}$ then $\bigcup_{i=1}^{n} F_i$ is closed.

Thus we see that an intersection of closed sets is closed and a finite union of closed sets is open.

3.9. Show by counterexample that in Problem 3.7 (c), we cannot replace the intersection of finitely-many open sets by arbitrarily many and in Problem 3.8 (b) we cannot replace the union of finitely-many closed sets by infinitely many.

In other words, an arbitrary intersection of open sets need not be open and an arbitrary union of closed sets need not be closed. The next problem justifies our use of adjectives in defining open and closed balls.

3.10. Let E be a metric space. Prove:

(a) Any open ball in E is an open set.
(b) Any closed ball in E is a closed set.
(c) Any finite subset of E is closed.

Definition. Let E be a metric space and let $S \subset E$. S is *bounded* if S is contained in a ball.

We have apparently two different notions of the word bounded. For the sake of coherence we show that they coincide.

3.11. Suppose S is a subset of a normed linear space X. Show that S is bounded in the sense of normed linear spaces if and only if it is bounded in the sense of metric spaces (where X is of course made into a metric space using the norm).

3.12. Let E be a metric space and let $S \subset E$. Show that S is bounded if and only if the set $\{d_E(a,b) : a, b \in S\}$ is bounded above in $\mathbb{R}$.

This last result gives one indication of how we might measure the size of a bounded set.

Definition. If E is metric space and $S \subset E$ is bounded, we define the *diameter* of E by

$$\operatorname{diam}(E) = \sup\{d_E(a,b) : a, b \in S\}.$$

3.13. Let E be a metric space and $x \in E$. What is the diameter of $B(x,r)$. What about $B[x,r]$. In $\mathbb{R}^2$, what is the diameter of the set $\{(x,y) \in \mathbb{R}^2 | x^2 + y^2 = r^2\}$ (this set is of course called the *circle* of radius r centered at the origin).

Recall that if (E,d) is a metric space and $A \subset E$ then we can regard A to be a metric space in its own right under the metric inherited from E. When we take this perspective, we often say that A is a *subspace* of E. To some extent, this terminology is merely psychological: any subset of E can become a subspace, but the language is often useful to make our perspective clear.

3.14. Let E be a metric space and let $A \subset E$. Prove:

(a) A subset $U \subset A$ is open in (A,d) if and only if there exists an open set $\mathcal{O}$ in E with $U = A \cap \mathcal{O}$.
(b) A subset $F \subset A$ is closed in (A,d) if and only if there exists a closed set H in E with $F = A \cap H$.

3.15. Let $A = (0,1] \subset \mathbb{R}$. Let $B = (\frac{1}{2}, 1)$ and $C = (\frac{1}{2}, 1]$. Classify each of B and C as open, closed, both, or neither when regarded as

(a) subsets of A and
(b) subsets of $\mathbb{R}$.

Thus if A is a subspace of E, and $S \subset A$, we see that S being open with respect to E is a different notion that S being open with respect to A. Misconceptions along these lines often lead to errors in proofs.

A major theme throughout this chapter will be the fact that, when we apply topological notions to $\mathbb{R}$, we often get very useful and special results. The first example of this phenomenon is the following.

3.16. Let $F \subset \mathbb{R}$ be a nonempty closed set which is bounded above. Prove that $\sup(F) \in F$. Conclude that a nonempty closed set in $\mathbb{R}$ has a maximum.

Hint: Assume not and show $C(F)$ is not open.

3. Sequences

As its title suggests, this section serves to introduce the notion of sequences. Sequences are extremely important for studying the properties both of $\mathbb{R}$ and of general metric spaces.

Definition. Let E be a set. A *sequence* in E is a function $f : \mathbb{N} \to E$.

Instead of writing something like $f(n)$ for the value of a sequence at n, we typically write something like a_n. This notation is meant to reflect the intuitive notion that a sequence is an "infinite list" of elements. We denote the entire sequence by $(a_n)_{n=1}^{\infty}$ or (a_n) and call a_n the nth term.

Definition. A *subsequence* of the sequence $(a_n)_{n=1}^{\infty}$ is a sequence $(b_n)_{n=1}^{\infty}$ where for some $k_1 < k_2 < \cdots$ in $\mathbb{N}$, $b_n = a_{k_n}$.

Intuitively a subsequence is created by skipping (possibly infinitely-many) terms of the original sequence.

Definition. Let $(p_n)_{n=1}^{\infty}$ be a sequence in a set E. If S is a subset of E, we say that (p_n) is *eventually contained in* S if there is an $N \in \mathbb{N}$ such that $p_n \in S$ for $n \geq N$. If E is a metric space, and $p \in E$. We say that (p_n) *converges* to p_n if (p_n) is eventually contained in every open set of E that contains p. If (p_n) converges to p, we say that p is a *limit* of p_n and often write $p_n \to p$.

Intuitively, p_n converges to p if the terms are eventually very close to p. Not surprisingly, we have captured this intuitive notion by the use of open balls. An open ball centered at p is the collection of points which are close to p. Thus to say that the sequence eventually lies in this open ball is to say that the sequence is eventually very close to p. Regardless of the degree of closeness we require, the sequence is eventually that close.

3.17. Let p_n be a sequence in the metric space E. Show that $p_n \to p$ if and only if for every $\epsilon > 0$ there is an $N \in \mathbb{N}$ so that $d_E(p_n, p) < \epsilon$ for each $n \geq N$. Prove that $p_n \to p$ if and only if $d_E(p_n, p)$ converges to zero in $\mathbb{R}$.

Technically, a sequence in a metric space is a special type of function into the space. Since we have defined the notion of a bounded function, we get for free the notion of a bounded sequence. Explicitly, the sequence (a_n) is bounded if the set

$$\{a_n | n \in \mathbb{N}]\}$$

is contained in some ball. If our space happens to be a normed linear space, we see that (a_n) is bounded if and only if there is some $K \in \mathbb{R}$ with $\|a_n\| \leq K$ for all $n \in \mathbb{N}$.

3.18. Let $p_n \to p$ in the metric space E. Prove:

(a) If $p_n \to q$ in E then $p = q$.
(b) If (q_n) is a subsequence of $(p_n)_{n=1}^{\infty}$ then $q_n \to p$.
(c) (p_n) is bounded.

Part (a) of the previous problem justifies us in calling p **the** limit of (p_n). We often write $\lim_{n\to\infty} p_n$. If a sequence has a limit, we say that it is *convergent*. Otherwise, we say that it is *divergent*. If (a_n) is divergent, we might say that $\lim_{n\to\infty} a_n$ does not exist.

The following result is our first example of the relationship between the topological properties of a metric space and the notion of sequences.

3.19. Let S be a subset of the metric space E. Prove that S is closed if and only if for every convergent sequence $(p_n)_{n=1}^{\infty}$ with $p_n \in S$, we have $\lim_{n\to\infty} p_n \in S$.

Hint: For the reverse implication, use contradiction. More explicitly, if $C(S)$ is not open, construct a sequence in S which converges to point outside of S.

We next move to the specific case of $E = \mathbb{R}$. The following results are often called the *limit laws*.

3.20. Let $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ be sequences in $\mathbb{R}$ with $a_n \to a$ and $b_n \to b$. Prove that:

(a) $a_n + b_n \to a + b$
(b) $a_n b_n \to ab$
(c) If $b_n \neq 0$ for all $n \in \mathbb{N}$ and $b \neq 0$ then $\frac{a_n}{b_n} \to \frac{a}{b}$
(d) If $a_n \leq b_n$ for all n then $a \leq b$.

Hint: For (b), use the fact that

$$a_n b_n - ab = (a_n - a)b_n - a(b_n - b).$$

For (c), begin by proving the case that $a_n = 1$. For this case, use the fact that

$$\frac{1}{b_n} - \frac{1}{b} = \frac{b_n - b}{b_n b}.$$

Prove that you can choose N_0 so that for $n \geq N_0$, $|b_n b| > \frac{b^2}{b}$ and use this equality.

The next result is known as the *Squeeze Theorem.*

3.21. Suppose that (a_n), (b_n), and (c_n) are sequences in $\mathbb{R}$ with $a_n \leq b_n \leq c_n$ and $a_n, c_n \to a$. Then $b_n \to a$.

3.22. Which, if any, of the results in Problem 3.20 would remain true if (a_n) and (b_n) were convergent sequences in an arbitrary normed linear space?

3.23. Let $(a_n) \subset [a, b]$ and $a_n \to p$. Prove that $p \in [a, b]$.

Hint: It is not too difficult to show this directly, but with an appropriate application of some results we have already proven, the proof is trivial.

Definition. Let $(a_n)_{n=1}^{\infty}$ be a sequence in $\mathbb{R}$. $(a_n)_{n=1}^{\infty}$ is called *increasing* if $a_n \leq a_{n+1}$ for all $n \in \mathbb{N}$. Likewise it is called *decreasing* if $a_n \geq a_{n+1}$ for all $n \in \mathbb{N}$. It is called *monotone* if it is either increasing or decreasing.

3.24. Let $(a_n)_{n=1}^{\infty}$ be a bounded monotone sequence in $\mathbb{R}$. Prove:

(a) If $(a_n)_{n=1}^{\infty}$ is increasing then $a_n \to \sup\{a_n : n \in \mathbb{N}\}$.
(b) If $(a_n)_{n=1}^{\infty}$ is decreasing then $a_n \to \inf\{a_n : n \in \mathbb{N}\}$.
(c) A monotone sequence in $\mathbb{R}$ is convergent if and only if it is bounded.

The previous result is known as the *Monotone Convergence Theorem.*

3.25. Let (a_n) be a sequence in $\mathbb{R}$. Prove that (a_n) has a monotone subsequence.

Hint: Consider two cases. Begin by assuming that there exists a subsequence with no least term.

We conclude this section with a few interesting problems regarding sequences.

3.26. Let $p_n \to p$ in the metric space E. Let $S = \{p, p_1, p_2, \ldots\}$. Prove that S is a closed set.

3.27. Define a sequence by $a_n = r^n$ for $r \in \mathbb{R}$ (such a sequence is called a *geometric sequence*). For what values of r does this sequence converge? What is the limit? Give proofs.

3.28. Let $a_n \to a$ in $\mathbb{R}$. Prove that $s_n \to a$ where $s_n = \frac{1}{n}(a_1 + a_2 + \cdots + a_n)$ for $n \in \mathbb{N}$.

In other words if a sequence in $\mathbb{R}$ converges to $a \in \mathbb{R}$, the average value of the first n terms also converges to a.

4. Cauchy Sequences and Completeness

Though the examples pertinent to analysis perhaps do not demonstrate it, $\mathbb{R}$ is very special as a metric space. Of course most other spaces do not have its level of algebraic structure (that is most metric spaces do not come with a notion of addition, division, etc.), but these characteristics do not entirely describe its special nature. In this section, we formulate an extremely important property of metric spaces and show that $\mathbb{R}$ satisfies it.

Along the way, we will need the following notion. It is named after a mathematician how helped put Leibnitz's and Newton's calculus on a rigorous foundation.

Definition. Let $(a_n)_{n=1}^{\infty}$ be a sequence in the metric space E. (a_n) is called *Cauchy* if for all $\epsilon > 0$ there exists $N \in \mathbb{N}$ so that if m and n are integers with $m, n \geq N$ then $d_E(a_n, a_m) < \epsilon$.

Intuitively then a sequence is Cauchy if the terms the sequence are eventually very close together.

3.29. Let E be a metric space and let (p_n) be a sequence in E. Prove:

(a) If (p_n) is convergent then (p_n) is Cauchy
(b) If (p_n) is Cauchy then the (p_n) is bounded.
(c) If $(p_n)_{n=1}^{\infty}$ is Cauchy and contains a convergent subsequence than the sequence itself converges.

We may now formulate the property alluded to above.

Definition. A metric space is called *complete* if every Cauchy sequence is convergent.

3.30. Prove that $\mathbb{R}$ is a complete metric space.

Hint: Use the previous result, the Monotone Convergence Theorem, and 3.25.

The proof we suggested relies on the Monotone Convergence Theorem which ultimately relies on the Completeness Axiom of $\mathbb{R}$. In fact, the assumption that every Cauchy sequence converges in $\mathbb{R}$ is equivalent to the the Completeness Axiom (that is, one can prove one from the other given only the fact that $\mathbb{R}$ is an ordered field). In other words the Completeness Axiom is equivalent to the statement that $\mathbb{R}$ is complete as a metric space, thus explaining its name.

More generally, $\mathbb{E}^n$ is also complete as we now demonstrate.

3.31. Let $(p_k)_{k=1}^\infty$ be a sequence in $\mathbb{E}^n$. For each k, write

$$p_k = \left(p_k^{(1)}, p_k^{(2)}, \ldots, p_k^{(n)}\right).$$

with $p_k^{(i)} \in \mathbb{R}$. Prove:

(a) For each $1 \leq i \leq m$, $k, \ell \in \mathbb{N}$, $\left|p_k^{(i)} - p_\ell^{(i)}\right| \leq \|p_k - p_\ell\|$.

(b) For each $k, \ell \in \mathbb{N}$ $\|p_k - p_\ell\| \leq \sum_{i=1}^{n} \left|p_k^{(i)} - p_\ell^{(i)}\right|$.

(c) $(p_k)_{k=1}^\infty$ is Cauchy in $\mathbb{E}^n$ if and only if for all $1 \leq i \leq m$ the sequence $(p_k^{(i)})_{k=1}^\infty$ is Cauchy in $\mathbb{R}$.

(d) $(p_k)_{k=1}^\infty$ is convergent in $\mathbb{E}^m$ if and only if for all $1 \leq i \leq m$ the sequence $(p_k^{(i)})_{k=1}^\infty$ is convergent in $\mathbb{R}$. Moreover, if $p_k^{(i)} \to p^{(i)}$ for $1 \leq i \leq m$ then

$$p_k \to \left(p^{(1)}, \ldots, p^{(n)}\right).$$

3.32. Prove that $\mathbb{E}^m$ is complete for all $m \in \mathbb{N}$.

3.33. Let A be a closed subset of a complete metric space E. Prove that the subspace A is also a complete metric space (as a subspace of E).

3.34. Show that $(0, 1)$ is not complete.

5. Interior and Closure

We continue to study the topological properties of general metric spaces. Of course not every set in a metric space is open, but given a set we do have a natural way of finding an open set.

Definition. Let A be a subset of the metric space (E, d). $p \in E$ is called an *interior point* of A if there is an open ball centered p which is contained in A. The collection of interior points of A is called the *interior* of A and is denoted $\text{int}(A)$.

3.35. Let A be a subset of the metric space E. Prove:

(a) $\text{int}(A) \subset A$.
(b) $\text{int}(A)$ is an open set.
(c) If $U \subset A$ and U is open then $U \subset \text{int}(A)$.
(d) $\text{int}(A)$ is the union of all the open sets contained in A.
(e) A is open if and only if $A = \text{int}(A)$.

The part (d) of the previous problem tells us that the interior of A is the largest open set contained in A. This interpretation of the interior leads to a corresponding notion for closed sets.

Definition. Let A be a subset of the metric space E. The *closure* of A, denoted by $\bar{A}$, is the intersection of all the closed sets containing A.

3.36. Let A be a subset of the metric space E. Prove:

(a) $A \subset \bar{A}$.
(b) $\bar{A}$ is closed.
(c) A is closed if and only if $A = \bar{A}$
(d) A point $p \in E$ lies in A if and only if every open ball centered at p intersects A.
(e) A point of E lies in $\bar{A}$ if and only if there is a sequence in A that converges to.

Definition. Let A be a subset of the metric space E. The *boundary* of A, denoted $\text{bd}(A)$ is defined by $\bar{A} \setminus \text{int}(A)$.

3.37. Let A be a subset of the metric space E. Prove:

(a) The point $p \in E$ lies in $\text{bd}\, A$ if and only if every every ball centered at p intersects both A and its complement.
(b) A point of E lies in $\text{bd}\, A$ if only if there is a sequence in A converging to it and a sequence in $C(A)$ converging to it.
(c) E is the disjoint union of $\text{int}(A)$, $\text{int}(C(A))$, and $\text{bd}(A)$.
(d) A is closed if and only if $A \supset \text{bd}\, A$.
(e) A is open if and only if $A \cap \text{bd}\, A = \varnothing$.

3.38. For each of the following "$A \subset E$" find $\text{int}(A)$, $\bar{A}$, $\text{bd}\, A$ and determine if A is open, closed, both, or neither.

(a) $\mathbb{Q} \subset \mathbb{R}$
(b) $\mathbb{Q} \subset \mathbb{R}$, where $\mathbb{R}$ is given the discrete metric
(c) $(0, 1] \subset (0, 1] \cup (2, 3)$

(d) $(2,3) \subset (0,1] \cup (2,3)$
(e) $(0,1) \subset [0,1]$, where $[0,1]$ is given the discrete metric
(f) $\{(x,y) : x^2 + y^2 \leq 1\} \subset \mathbb{E}^2$
(g) $\{(x,y) \in \mathbb{R}^2 : x \in \mathbb{Q}\} \subset \mathbb{E}^2$

Definition. If E is a metric space, $A \subset E$ is called dense in E if $\bar{A} = E$.

The proof of the next result will explain why the density of $\mathbb{R}$ is called is such.

3.39. Show that $\mathbb{Q}$ is dense in $\mathbb{R}$.

6. Compactness

Though its formulation may seem strange at first, the topological notion of compactness is extremely important and has consequences which are far-reaching in many branches of mathematics.

Definition. Let E be a metric space and let $K \subset E$. An *open cover* of K is a collection of open sets in E, $(\mathcal{O}_\alpha)_{\alpha \in I}$, such that $K \subset \bigcup_{\alpha \in I} \mathcal{O}_\alpha$. K is *compact* if for every such open cover there exists a finite set $F \subset I$ with $K \subset \bigcup_{\alpha \in F} \mathcal{O}_\alpha$ (the collection $(\mathcal{O}_\alpha)_{\alpha \in F}$ is said to be a *finite subcover*). E is called a *compact metric space* if E is a compact subset of itself.

We have seen previously that whether a set is open (or closed) depends on the space of which it is considered a part. For example, if E is a metric space and $A \subset E$ is a set then A may or may not be open as a subset of E. It is however always open as a subset of itself (considered as a subspace). Thus the notion of openness depends on the larger set in which one is working. The next result shows that this is not the case when it compactness.

3.40. Let $K \subset (E,d)$. Prove that K is a compact subset of E if and only if K is a compact metric space when it viewed as metric space itself.

Hence compactness is more intrinsic to a metric space then is openness. The observant reader will note that this is also the case with completeness, but completeness is a less important notion than compactness as it is only a notion of metric spaces and not of topological spaces in general (we remarked at the beginning of this chapter that metric spaces are a special type of topological space but that they do not tell the whole story).

3.41. Prove that $K = \{0, 1, \frac{1}{2}, \frac{1}{3}, \ldots\} \subset \mathbb{R}$. is compact and $A = (0,1] \subset \mathbb{R}$ is not.

3.42. Suppose that F is closed subset of a compact metric space E. Show that F is compact.

3.43. Let K be a compact subset of a metric space. Show that K is bounded.

The following statement is very useful in proving results about compact spaces.

3.44. Let K be a compact metric space and suppose $(F_n)_{n\in\mathbb{N}}$ is a decreasing sequence of nonempty closed subsets of K (to say that (F_n) is decreasing is to say that $F_1 \supset F_2 \supset \cdots$). Prove that $\bigcap_{n=1}^{\infty} F_n \neq \varnothing$. Suppose in addition that $\lim_{n\to\infty} \operatorname{diam}(F_n) = 0$ show that $\bigcap_{n=1}^{\infty} F_n$ consists of exactly one point.

Hint: Suppose that the intersection is empty and consider the collection $\{C(F_n)\}$.

We will next give several different characterizations of compact metric spaces.

3.45. Prove that a compact metric space is complete.

3.46. Show that a compact subset of a metric space is closed.

Compact subsets of $\mathbb{E}^n$ have a very easy characterization. The following result is known as the *Heine-Borel Theorem.*

3.47. Let $K \subset \mathbb{E}^n$. Prove that K is compact if and only if K is closed and bounded.

Hint: We give the outline of the reverse implication. You should fill in the details. Let K be a set which is closed and bounded and suppose that it is not compact. Then there exists a collection $(U_i)_{i\in I}$ of open sets covering K without a finite subcover. We may put $K \subset \bigcup_{j=1}^{n_1} B_j^1$ where B_j^1 is a closed ball of radius 1. Then there exists $j_1 \leq n_1$ so that $K \cap B_{j_1}^1$ cannot be covered by finitely many of the U_i. Repeat the argument for

$$K \cap B_{j_1}^1 \subset \bigcup_{j=1}^{n_2} B_j^2$$

where the B_j^2 is a closed ball of radius $1/2$. Continuing in this manner, we conclude that for all n, $K \cap B_{j_1}^1 \cap \cdots \cap B_{j_n}^n$ cannot be covered by finitely many of the U_i. For each n choose a point

$$p_n \in K \cap B_{j_1}^1 \cap \cdots \cap B_{j_n}^n.$$

Show the sequence (p_n) thus defined is Cauchy. Let p be the limit. Show that $p \in K$ and conclude that $p \in U_r$ for some $r \in I$. Deduce that $K \cap B^1_{j_1} \cap \cdots \cap B^1_{j_n} \subset U_r$ for some n, a contradiction.

Unfortunately the previous result is a very special property of $\mathbb{E}^n$.

3.48. Give an example of a metric space E and set K which is closed and bounded and yet not compact.

Nevertheless, if we modify the formulation a bit, we can prove a result that is still somewhat nice. Along the way we will prove another extremely useful characterization of compact metric spaces.

Definition. A metric space, E, is called *totally bounded* if, for all $\epsilon > 0$, E may covered by finitely many balls of radius ϵ.

3.49. Prove:

(a) The metric space E is totally bounded if and only if for all $\epsilon > 0$ there exist a finite number of balls of radius less than ϵ whose union is E.
(b) A subset of $\mathbb{R}$ is bounded if and only if it is totally bounded
(c) A subset of $\mathbb{E}^n$ is bounded if and only if it is totally bounded.

3.50. Give an example of a metric space and a subset A that is bounded but not totally bounded.

Definition. Let A be a subset of the metric space E and let $p \in E$. p is called a *cluster point* of A if every ball centered at p intersects A in infinitely-many points. The collection of cluster points of A is denoted A'. Points in the set $A \setminus A'$ are called *isolated* points of A.

3.51. Let A be a subset of the metric space E and suppose $p \in E$. Show that $x \in A'$ if and only every ball centered at p intersects A at a point other than p.

3.52. Let E be a compact metric space. Show that every infinite set in E has a cluster point (which may or may not lie in the set).

Hint: If A has no cluster point, then for every $p \in E$ there is an open ball centered at p which intersects A at finitely-many points.

3.53. Let $A = \{\frac{1}{n} + \frac{1}{m} : n, m \in \mathbb{N}\} \subset \mathbb{R}$. Find A'.

3.54. Let A be a subset of the metric space E. Prove:

(a) The point $p \in E$ lies in A' if and only if there is a sequence (p_n) in A with $p_n \neq p$ for all n and $p_n \to p$.
(b) $A' \subset \bar{A}$.
(c) A is closed if and only if $A' \subset A$.

3.55. Give an example, if possible, of each of the following. Otherwise, prove no example exists.

(a) $A \subset \mathbb{R}$, A infinite, $A' = \varnothing$.
(b) A complete bounded noncompact metric space E.
(c) A metric space E so that for every closed ball B in E, B is not complete.

Definition. A metric space is called *sequentially compact* if every sequence contains a convergent subsequence.

We can now state the characterizations of compact spaces that we have been working towards.

3.56. Prove that the following are equivalent for a metric space E.

(a) E is compact,
(b) E is sequentially compact, and
(c) E is complete and totally bounded.

Note similarity between the equivalence of the first and the third conditions on the one hand and the Heine-Borel Theorem on the other. The analogy is even more clear when we remark that a subset of $\mathbb{E}^n$ is complete if and only if it is closed (why?).

7. Connectedness

Like compactness, connectedness is an extremely important topological concept. It is also a bit more intuitively motivated. Essentially a metric space if it cannot be "broken into separate parts." For example, we will show that the real line $\mathbb{R}$ is connected: it is one continuous line. On the other hand the set $[0, 1] \cup [1, 2]$ is not connected.

Definition. The metric space E is called *connected* if the only clopen subsets of E are E and $\varnothing$. $A \subset E$ is connected if it is connected as a metric space (when given the subspace metric).

By definition, connectedness enjoys the same intrinsic nature as does compactness.

3.57. Let A be a subset of the metric space E. Prove:

(a) A is connected if and only if there do not exist disjoint nonempty open sets U_1 and U_2 in A with $U_1 \cup U_2 = A$.
(b) A is connected if and only if there do not exist open sets $\mathcal{O}_1$ and $\mathcal{O}_2$ of E with $\mathcal{O}_1 \cap A \neq \varnothing$, $\mathcal{O}_2 \cap A \neq \varnothing$, $A \cap \mathcal{O}_1 \cap \mathcal{O}_2 = \varnothing$, and $A \subset \mathcal{O}_1 \cup \mathcal{O}_2$.
(c) A is connected if and only if there do not exist closed sets F_1 and F_2 in E with $F_1 \cap A \neq \varnothing$, $F_2 \cap A \neq \varnothing$, $A \cap F_1 \cap F_2 = \varnothing$, and $A \subset F_1 \cup F_2$.

The condition given in part (a) demonstrates why the definition of connectedness matches the intuitive notion described above: A is divided into the two "separate parts" U_1 and U_2. We often say that the sets $\mathcal{O}_1$ and $\mathcal{O}_2$ above *disconnect* A.

Definition. A set, S, in $\mathbb{E}^n$ is called convex if whenever $a, b \in S$, we have $\lambda a + (1 - \lambda)b \in S$ for any $\lambda \in [0, 1]$.

In other words, S is convex if given any two points of S, the line segment between them lies in S. This definition can of course be generalized to any vector space.

3.58. Let $A \subset \mathbb{R}$. Prove that A is convex if only if for every pair of points $a, b \in A$ the interval $[a, b]$ is contained in A. Prove that a set in $\mathbb{R}$ is convex if and only it if is an interval.

3.59. Let $A \subset \mathbb{R}$. Prove that A is connected if and only if A is an interval.

Hint: If A is not an interval, find a point which is not in A and yet lies between two points of A. Use this point to construct two sets that disconnect A. Conversely suppose $\mathcal{O}_1$ and $\mathcal{O}_2$ disconnect A. Let $a \in \mathcal{O}_1 \cap A$ and $b \in \mathcal{O}_2 \cap A$ and assume $a < b$. Let $x = \sup\{c \in \mathbb{R} : [a, c) \subset \mathcal{O}_1 \cap A)$. Show $x \notin \mathcal{O}_1 \cup \mathcal{O}_2$.

3.60. Show that a convex subset of $\mathbb{E}^n$ is connected.

8. Series and Decimals

You probably already know of the importance of series in regards to $\mathbb{R}$. In fact, the concept generalizes to any normed linear space (but we will still use it mostly to study the properties of $\mathbb{R}$).

3.61. Suppose that (a_n) is a sequence in a normed linear space X. The *sequence of partial sums* of (a_n) is the sequence, (s_n) given by $s_n = \sum_{i=1}^{n} a_n$ for $n \in \mathbb{N}$. If $s_n \to L$ we put $L = \sum_{n=1}^{\infty} a_n$ and we say that (a_n) *sums* to L.

A bit less formally, we call $\sum_{n=1}^{\infty} a_n$ a series. As one might expect, we say that $\sum_{n=1}^{\infty} a_n$ does not exist if the sequence of partial sums of (a_n) is divergent.

3.62. For what values of $x > 0$ does $\sum_{n=1}^{\infty} x^n$ converge? Prove it rigorously and give the limit if it exists.

Hint: If $x \neq 1$ prove that for $n \in \mathbb{N}$

$$1 + x + x^2 + \cdots + x^n = \frac{1 - x^{n+1}}{1 - x}.$$

3.63. Let (a_n) is a sequence and $b \in \mathbb{R}$. Suppose that $0 \leq a_n \leq b$. Show that $\sum_{n=1}^{\infty} a_n b^{-n}$ converges and the the limit lies in $[0, 1]$.

As we have mentioned many times, our definition of $\mathbb{R}$ is purely axiomatic, but using series we can give systematic ways of describing them. For notation simplicity, we will want to allow sequences to be indexed by $\mathbb{Z}$ rather than just $\mathbb{N}$. Technically an *integrally indexed sequence* is a function $\mathbb{Z} \to \mathbb{R}$, but we will write $(a_n)_{n\in\mathbb{Z}}$ and give this notation in its obvious meaning.

Definition. Fix a natural number $b \geq 2$. A *base b decimal expansion* is a integrally indexed sequence $(x_n)_{n\in\mathbb{Z}}$ with $x_n \in \mathbb{Z}$ and $0 \leq x_n < b$. Moreover, we require that there is some $N \in \mathbb{Z}$ such that $x_n = 0$ for $n < N$.

You have probably already guessed that given a base b decimal expansion $(x_n)_{n\in\mathbb{Z}}$, we get a real number. Firstly we get an integer $z = x_0 + x_{-1}b + x_{-2}b^2 + \cdots$. This sum is immediately well-defined: since only finitely many x_n for negative n are nonzero, it is actually just a finite sum in $\mathbb{Z}$. Likewise, by Problem 3.63, we get a real number d in $[0, 1]$ by putting $d = \sum_{n=1}^{\infty} d_n b^{-n}$. The real number $z + d$ is called the *real number coming from the decimal expansion* $(x_n)_{n\in\mathbb{Z}}$.

Needless to say, if $(x_n)_{n\in\mathbb{Z}}$ is a decimal expansion with $x_n = 0$ for $n > N$ (with $N < 0$), we usually denote the associated real number x by

$$x = x_N x_{N+1} \ldots x_0.x_1 x_2 x_3 x_4 \ldots$$

(usually of course cutting on extra zeros on the left in the obvious sense).

3.64. Suppose that $x \in \mathbb{R}$ show that there is a base b decimal expansion $(x_n)_{n\in\mathbb{Z}}$ which gives x.

Unfortunately, decimal expansions are not unique. For example, for $b = 10$, (that is the usual base), we may represent the number 1 with the decimal expansion, $(x_n)_{n\in\mathbb{Z}}$, where $x_0 = 1$ and $x_n = 0$ for all $n \in \mathbb{Z}\setminus 0$ or we can represent it with $(y_n)_{n\in\mathbb{Z}}$ where $y_n = 0$ for $n \leq 0$ and $y_n = 9$ for $n > 0$.

3.65. Show that $x \in \mathbb{R}$ has a unique base b decimal expansion if and only if x does not have the form a/b^n for $a \in \mathbb{Z}$ and $n \in \mathbb{N}$. If x does not have a unique decimal expansion, show that it has exactly 2, one which ends with a constant sequence of zero and one which ends with a constant sequence of $b - 1$.

3.66. Suppose that $x, y \in \mathbb{R}$ have decimal expansions $(x_n)_{n\in\mathbb{Z}}$ and $(y_n)_{n\in\mathbb{Z}}$, respectively. Describe (with) proof an algorithm for getting:

(a) a decimal expansion for $x + y$,
(b) a decimal expansion for $x - y$,
(c) a decimal expansion for xy, and
(d) a decimal expansion for x/y (assuming $y =\neq 0$).

Describe also (with proof) a method for determining if $x = y$, $x > y$, or $x < y$.

The previous problem hints at a way of constructing the real numbers rigorously. Indeed, we fix a b and let $\mathbb{R}$ the collection of integrally defined sequences, $(x_n)_{n\in\mathbb{Z}}$ with $0 \leq x_n \leq b - 1$, which are zero for n sufficiently small (in the obvious sense) and which do not end with a constant sequence of $b - 1$. We define algebraic operations on the set using the algorithms given in the previous problem with care be taken to avoid an expansion ending with a constant sequence of $b-1$. Likewise we define an order on the collection. Finally we have to verify that all the axioms of $\mathbb{R}$ are satisfied.

Showing in this way that $\mathbb{R}$ exists is very tedious and certainly long, but it can be (and has been) done. The ambitious reader is encouraged to attempt the outlined program. Of course an arbitrary integer $b \geq 2$ will work, but it is perhaps easier to make a specific choice of base (such as $b = 10$).

CHAPTER 4

Continuous Functions

The idea of assigning a metric to a set is to have a notion of closeness between the points. When one considers a function between two metric spaces then, one is naturally led to ask if the function respects this closeness. We are led to the notion that we will study in this chapter.

1. Basic Definitions and Characterizations

Definition. Let $f : E \to G$ be a function of metric spaces. Suppose $p_0 \in E$. We say that f *is continuous at* p_0 if for every ball B_G, centered at $f(p_0)$, there is a ball, B_E centered at p_0 such that $f(B_E) \subset B_G$. If $A \subset E$, f is said to be *continuous on* A if for all $p_0 \in A$, f is continuous at p_0. f is called *continuous* if f is continuous on E.

Hence, loosely speaking, the function f is continuous at p_0 if points near p_0 are taken by f to points close to $f(p_0)$.

4.1. Suppose that $f : E \to G$ is a function of metric spaces and let $p_0 \in E$. Prove:

(a) f is continuous at p_0 if and only if for all $\epsilon > 0$ there exists a $\delta > 0$ so that for all $p \in E$, if $d_E(p, p_0) < \delta$ then $d_G(f(p), f(p_0)) < \epsilon$.
(b) f is not continuous at p_0 if and only if there exists an $\epsilon > 0$ so that for all $\delta > 0$ there exists $p_\delta \in E$ with $d_E(p_\delta, p_0) < \delta$ and $d_G(f(p_\delta), f(p_0)) \geq \epsilon$.

4.2. Let $f : \mathbb{R} \to \mathbb{R}$ be given as below. Prove that f is continuous at $p = 2$ and $p = 5$.

(a) $f(x) = -6x + 7$
(b) $f(x) = x^2 + x + 1$

Hint: For (b) with $p = 2$, let $\epsilon > 0$. We must find $\delta > 0$ so that if $|x-2| < \delta$, then $|f(x) - f(2)| = |x - 2||x + 3| < \epsilon$. Make sure the eventually choice of δ is less than 1 so that $|x - 2| < \delta \leq 1$ implies that $x \in (1, 3)$ and thus $|x + 3| < 6$. Thus for an appropriate choice of δ, the size of $|x + 3|$ can be controlled and that of $|x - 2|$ can be made very small.

In general the choice of δ is allowed to depend both on ϵ and on p_0. Is this dependence necessary for the two functions of the previous problem?

4.3. Prove that each of the following functions, f, is continuous.

(a) Let E be any metric space and $q \in E$. Define $f : E \to \mathbb{R}$ by given by $f(p) = d_E(p, q)$.
(b) Define a function $f : E \to G$ between to metric spaces by choosing a point $q_0 \in G$ and putting $f(p) = q_0$ for all $p \in E$ (such a function is of course called a *constant function*).
(c) $f = \mathrm{id}_E$ where E is any metric space.
(d) Define $f : [0, \infty) \to \mathbb{R}$ by $f(x) = \sqrt{x}$ (once again you may assume that each nonnegative real number has a unique square root).

4.4. Define $f : [0, 1] \to \mathbb{R}$ by

$$f(x) = \begin{cases} 0 & \text{if } x \text{ is irrational} \\ 1 & \text{if } x \text{ is rational.} \end{cases}$$

Prove that f is not continuous at any point in $[0, 1]$.

4.5. Let $f : [0, 1] \to \mathbb{R}$ be given by $f(0) = 0$ and

$$f(x) = \begin{cases} 0 & \text{if } x \text{ is irrational} \\ 1/m & \text{if } x = \frac{n}{m} \text{ in lowest terms, } n, m \text{ nonnegative integers} \end{cases}$$

Prove that:

(a) f is discontinuous at each rational in $(0, 1]$.
(b) f is continuous at each irrational in $[0, 1]$.

When considering continuity, it is very important to take account the domain of the function. For instance, consider the following example.

4.6. Define a function $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1 & \text{if } x \in [0, 1] \\ 0 & \text{otherwise.} \end{cases}$$

Show that f is **not** continuous on $[0, 1]$. However, show that the restriction of f to $[0, 1]$ is continuous.

The next result is sometimes called the *Global Continuity Theorem.*

4.7. Let $f : E \to G$ be a function of metric spaces. Prove:

(a) f is continuous if and only if for all open sets $U \subset G$, $f^{-1}(U)$ is open in E.

(b) f is continuous if and only if for all closed sets $F \subset G$, $f^{-1}(F)$ is closed in E.

4.8. Let $f : E \to G$ and $g : G \to H$ be functions of metric spaces. Let f be continuous at $p_0 \in E$ and let g be continuous at $f(p_0) \in G$. Prove that $g \circ f$ is continuous at p_0.

In a calculus course, continuity is typically defined in terms of limits of functions and we will give this characterization now.

Definition. Suppose that E and G are metric spaces, $A \subset E$ and p_0 is a cluster point of A. Suppose that $f : A \to G$. Let $q \in G$. We say that $\lim_{p \to p_0} f(p) = q$ if for every open ball B_G centered at q there is an open ball of E B_E centered at p_0 with $f(B_E \cap (A \setminus \{p_0\})) \subset B_G$.

Note that in the previous definition we do not insist that f is defined on all of E nor do we insist that is even defined at p_0 (only at points 'close' to p_0). Furthermore if f does happen to be defined at p_0, its value there is irrelevant to its limit at p_0.

Notice that our definition is not completely correct because writing an equality like $\lim_{p \to p_0} f(p) = q$ probably implies the uniqueness of limits of functions, which we have yet to demonstrate. Fortunately, as we will now show, limits of functions are unique and so our notation is justified.

4.9. Let E be a subset of a metric space, $A \subset E$ and p_0 a cluster point of A. Suppose G is another metric space and let $f : A \to G$. Prove that if $q, q_0 \in G$, $\lim_{p \to p_0} f(p) = q$, and $\lim_{p \to p_0} f(p) = q_0$ then $q = q_0$.

Thus we are justified in saying **the** limit and use the notation $\lim_{p \to p_0} f(p) = q_0$. One might also write "$f(p) \to q$ as $p \to p_0$." The following result is a characterization of continuity that one might see in a calculus course. Of course in such a course, the text often does not worry about technicalities like cluster points or isolated points.

4.10. Let $f : E \to G$ be a function of metric spaces and let $p_0 \in E$. Prove:

(a) If p_0 is an isolated point of E then f is automatically continuous at p_0.
(b) If p_0 is a cluster point of E then f is continuous at p_0 if and only if $\lim_{p \to p_0} f(p) = f(p_0)$.

The next result says that a continuous function is determined by its values on any dense subset.

4.11. Suppose that E and G are metric spaces and A is a dense subset of E. Suppose that $f, g : E \to G$ are continuous. Prove that if $f(x) = g(x)$ for all $x \in A$ then $f = g$.

The next result is called the *Sequential Characterization of Continuity.* It is very useful in showing results about continuity of functions because it allows us to use the results that we have already proven regarding sequences.

4.12. Let $f : E \to G$ be a function of metric spaces and let $p_0 \in E$. Prove that f is continuous at p_0 if and only if the following condition holds: For all sequences $(p_n) \subset E$ such that $p_n \to p_0$ we have $f(p_n) \to f(p_0)$.

4.13. Let $f : E \to \mathbb{R}$ and $g : E \to \mathbb{R}$. Suppose that f and g are both continuous at $p_0 \in E$. Prove:

(a) $f + g$ and fg are both continuous at p_0.
(b) If $g(p_0) \neq 0$ then $\frac{f}{g} : E \setminus \{p \in E : g(p) = 0\} \to \mathbb{R}$ is continuous at p_0.

4.14. Which of the above results hold if $\mathbb{R}$ is replaced by an arbitrary normed linear space?

Recall that a function $f : \mathbb{R} \to \mathbb{R}$ is a *polynomial function* if it has the form

$$f(x) = a_n x^n + \cdots + a_1 x + a_0$$

for $n \in \mathbb{Z}$, $n \geq 0$ and $a_i \in \mathbb{R}$.

4.15. Let $p : \mathbb{R} \to \mathbb{R}$ be a polynomial. Prove that p is continuous.

Definition. Let S be any set and let $f : S \to \mathbb{E}^n$. The *component functions* (or *coordinate functions*) of f are given by

$$f(x) = (f_1(x), f_2(x), \ldots, f_n(x)).$$

Thus $f_i : E \to \mathbb{R}$ for $1 \leq i \leq n$.

4.16. Let E be a metric space, let $f : E \to \mathbb{E}^n$, and let $p_0 \in E$. Prove that f is continuous at p_0 if and only if for each $1 \leq i \leq n$, the ith component function, f_i, of f is continuous at p_0.

2. Uniform Continuity and Lipschitz Functions

In some circumstances the condition of continuity is actually not a stronger enough condition to put on a function. In this section we introduce and study to stronger notions that are often very useful in real analysis.

Definition. A function $f : E \to G$ of metric spaces is *uniformly continuous* on E if for all $\epsilon > 0$ there exists $\delta > 0$ so that for all $p, q \in E$ if $d_E(p, q) < \delta$ then $d_G(f(p), f(q)) < \epsilon$.

Note that in the above definition, δ is allowed to depend on ϵ but not on the point.

4.17. Let $f : E \to G$ be a function of metric spaces. Prove:

(a) If f is uniformly continuous on E then f is continuous on E.
(b) f is not uniformly continuous on E if and only if there exists $\epsilon_0 > 0$ so that for all $\delta > 0$ there exist $p_\delta, q_\delta \in E$ with $d_E(p_\delta, q_\delta) < \delta$ and $d_G(f(p_\delta), f(q_\delta)) \geq \epsilon_0$.
(c) Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^2$. Then f is not uniformly continuous.

Definition. A function $f : E \to G$ is said to be *Lipschitz* if there exists $K \in \mathbb{R}$ so that for all $p, q \in E$, $d_G(f(p), f(q)) \leq K d_E(p, q)$.

4.18. If a function $f : E \to G$ of metric spaces Lipschitz then it is uniformly continuous.

4.19. Let $f : [0, \infty) \to \mathbb{R}$ be given by $f(x) = \sqrt{x}$. Prove that f is uniformly continuous but not Lipschitz (as usual, you may assume that every positive real number has a unique positive square root).

4.20. Let $f : E \to G$ be a function of metric spaces and suppose that E is compact. Prove that f is uniformly continuous if and only if it is continuous.

Hint: Let f be continuous. Given $\epsilon > 0$ for $x \in E$ choose $\delta_x > 0$ so that $f(B(x, \delta_x)) \subset B(f(x), \frac{\epsilon}{2})$ and then take a finite subcover of $\{B(x, \frac{\delta_x}{2})\}_{x \in E}$. Let $\delta = \min \frac{\delta_x}{2}$ where the minimum is taken over the finite subcover.

Uniform continuity is incredibly useful when it comes to extending functions. Specifically, we have the following result.

4.21. Let S be a subspace of the metric space E and let $f : S \to G$ be a uniformly continuous function into a complete metric space. Prove that there exists a unique continuous extension $g : \bar{S} \to G$ of f and that it is uniformly continuous (to say that g is an extension of f is to say that $g(s) = f(s)$ for $s \in S$).

The previous result does not hold if we only us continuity.

4.22. Give an example of a metric space E with a subset $S \subset E$ and a function $f : S \to G$, where G is a complete metric space such that f has no continuous extension to $\bar{S}$.

3. Theorems about Continuity

One property of continuous functions is that the inverse image of a closed set is always itself closed. This is not the case for direct images.

4.23. Give an example of a continuous function $f : E \to G$ of metric spaces and a closed set $F \subset E$ such that $f(F)$ is not closed.

Nevertheless if we require the stronger condition of compactness, the result we would want does hold.

4.24. Let $f : E \to G$ be continuous. Prove that if $K \subset E$ is compact then $f(K)$ is compact.

Hint: Use the Global Continuity Theorem.

A similar result holds for connected spaces.

4.25. Let $f : E \to G$ be continuous. Prove that if E is connected then $f(E)$ is connected.

4.26. Let $f : E \to G$ be a continuous bijection. Prove:

(a) If E is compact then $f^{-1} : G \to E$ is also continuous.
(b) If E is not compact, f^{-1} need not be continuous.

As usual, these theorems have particularly nice interpretations when applied to $\mathbb{R}$.

4.27. Let $f : [a, b] \to \mathbb{R}$ be continuous. Prove that $f([a, b]) = [c, d]$ for some $c \leq d$.

4.28. Let $f : [a, b] \to \mathbb{R}$ be continuous. Prove:

(a) There exist $x_0, x_1 \in [a, b]$ so that $f(x_0) \leq f(x) \leq f(x_1)$ for all $x \in [a, b]$.
(b) Let $x_0, x_1 \in [a, b]$ and let $y \in \mathbb{R}$ so that $f(x_0) < y < f(x_1)$. Then there exists $x \in [a, b]$, lying between x_0 and x_1 with $f(x) = y$.

The statement of part (a) above is often called *Extreme Value Theorem.* Reworded, its conclusion is the f attains a maximum and a minimum on $[a, b]$. Likewise. statement (b) is often called the *Intermediate Value Theorem*. They can actually be generalized.

4.29. Suppose that E is a metric space and $f : E \to \mathbb{R}$ is continuous.

(a) If E is compact then f attains a maximum and a minimum on E.
(b) If E is connected then for each $y_0, y_1 \in f(C)$, and $y \in \mathbb{R}$ with $y_0 < y < y_1$, there is a $p \in E$ with $f(p) = y$.

As an application of the Intermediate Value Theorem, we may finally prove the existence of nth roots.

Definition. Suppose that $a \in \mathbb{R}$ and $n \in \mathbb{N}$. We say that a number b is an *nth root of a* if $b^n = a$.

4.30. Use the Intermediate Value Theorem to show that every positive real number has an nth root for all $n \in \mathbb{N}$. Show that if n is odd than this root is unique and that it is positive. Likewise, show that if n is even then there are two roots exactly one of which is positive.

4.31. Show that if $a < 0$, a has an nth root if and only if n is odd. Show this root is unique.

If $n \in \mathbb{N}$ and $a \in \mathbb{R}$, we use the symbols $\sqrt[n]{a}$ and $a^{1/n}$ to denote the unique positive nth root of a. Likewise we set $0^{1/n} = \sqrt[n]{0} = 0$. If $m/n \in \mathbb{Q}$ and $a \geq 0$, we define $a^{m/n} = (\sqrt[n]{a})^m$.

4.32. Show that if $m/n = p/q$ then $a^{m/n} = a^{p/q}$. Show that when exponentiating positive real numbers by rational numbers the usual rules of exponents apply.

4. Convergence of Functions

Thus far all of our sequences have been sequences of points. In this section, we modify our perspective a bit to allow for sequences of functions.

Definition. Let E and G be metric spaces and let (f_n) be a sequence of functions, $f_n : E \to G$. Let $f : E \to G$ be another function. We say that f_n *converges to f pointwise* on E if for all $p \in E$, $f_n(p) \to f(p)$. We say that f_n *converges to f uniformly* on E if for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ so that for all $p \in E$ and $n \geq N$, $d_G(f_n(p), f(p)) < \epsilon$.

4.33. Let $f_n : E \to G$ be a sequence of functions between metric spaces and let $f : E \to G$. Prove that if $f_n \to f$ uniformly on E, then $f_n \to f$ pointwise on E.

4.34. Let $f_n : [0,1] \to \mathbb{R}$ be given by $f_n(x) = x^n$ for $x \in [0,1]$ and $n \in \mathbb{N}$. Show that f_n converges pointwise but not uniformly on $[0,1]$.

You will notice that the limit function of the previous problem is not continuous.

4.35. Let $f_n : E \to G$ be continuous for each $n \in \mathbb{N}$. Suppose that $f_n \to f$ uniformly on E for some $f : E \to G$. Prove that f is continuous on E.

Definition. Let $f_n : E \to G$ be a sequence of functions between metric spaces. (f_n) is called *uniformly Cauchy* if for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ so that for all $m, n \geq N$ and $p \in E$, we have $d_G(f_n(p), f_m(p)) < \epsilon$.

4.36. Let $f_n : E \to G$ be a sequence of functions between metric spaces. Prove:

(a) If $f_n \to f$ uniformly on E for some $f : E \to G$ then (f_n) is uniformly Cauchy on E.
(b) If G is complete and (f_n) is uniformly Cauchy then (f_n) is uniformly convergent on E.

In the case that $G = \mathbb{R}$ and E is compact, we can actually realize uniform convergence of functions as convergence of points in a metric space. To be precise, if E is any compact metric space, we denote the set of continuous real-valued functions on E by $C(E)$. Notice that there is some conflict of notation here as $C(E)$ could also denote a set-theoretic complement of E. The intent of the notation should be clear from the context.

Using Problem 4.13, we see that $C(E)$ is a vector space (since we have seen that a constant function is continuous). We define a norm on $C(E)$ by putting $\|f\| = \sup\{|f(x)| : x \in E\}$ (though from the Extreme Value Theorem, we see that this supremum is actually a maximum).

4.37. Let E be a compact metric space. Show that $C(E)$ is a normed linear space.

4.38. Let E be a compact metric space, let (f_n) be a sequence of functions in $C(E)$, and let $f \in C(E)$. Show that $f_n \to f$ uniformly if and only if (f_n) converges to f as points in the metric space $C(E)$.

4.39. Show that $C(E)$ is complete as a metric space.

Of course there are many variations on this idea. We can make some of the same notions work if we drop the assumption that E is compact and consider only bounded functions (an assumption that holds automatically if E is compact). Likewise we might consider the domain to be a more general normed linear space and use some of same ideas. Spaces of these types are prototypical examples of Banach Spaces. Banach Spaces are are an extremely important class of object in more advanced Real Analysis.

CHAPTER 5

Single Variable Calculus

You probably know that Calculus was first created independently by Sir Isaac Newton and Gottfried Wilhelm Leibnitz in the 17th Century. It was major step forward in the study of mathematics and of science. The applications of the theory are endless.

In this chapter we formulate the basic notions of calculus on $\mathbb{R}$ in rigorous modern language. Calculus of course has two branches: differential calculus and integral calculus. The Fundamental Theorem of Calculus (which we will prove) unifies the two branches.

Needless to say the ideas of calculus can also be applied to $\mathbb{E}^n$ to create a subject often called vector calculus. We will not study vector calculus in this text.

1. Differentiation

We probably do not need to give motivation or explanation behind the concept of a derivative. Intuitively, it is a way to measure the way a function is changing or to measure the instantaneous slope of the graph of the function.

Definition. Let $A \subset \mathbb{R}$, $f : A \to \mathbb{R}$ and $x_0 \in \text{int}(A)$. We say that f is *differentiable* at x_0 if

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists. In the case that f is differentiable at x_0, we denote the corresponding limit by $f'(x_0)$. If $D \subset \mathbb{R}$, we say that f is differentiable on D if $D \subset \text{int}(A)$ and f is differentiable at every point of D.

If f is differentiable on all on its domain than we say that it is a differentiable function. Notice that a necessary condition for this to hold is that the domain of f must be open.

5.1. Let $A \subset \mathbb{R}$, $f : A \to \mathbb{R}$, and $x_0 \in \text{int}(A)$. Prove:

(a) If f is differentiable at x_0 then

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

(b) f is differentiable at x_0 if and only if there exists $L \in \mathbb{R}$ so that for all $\epsilon > 0$ there exists $\delta > 0$ such that if $|x - x_0| < \delta$ then

$$|f(x) - f(x_0) - L(x - x_0)| \leq \epsilon |x - x_0| \ .$$

5.2. Suppose $A \subset \mathbb{R}$, $f : A \to \mathbb{R}$, and $x_0 \in A$ Prove that if f is differentiable at x_0 then f is continuous at x_0.

5.3. Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = x^2$. Show $f'(0) = 0$. Let $g(x) = -x^2$. Show $g'(0) = 0$. Suppose $h : \mathbb{R} \to \mathbb{R}$ is a function with $g(x) \leq h(x) \leq f(x)$ for all $x \in \mathbb{R}$. Show $h'(0) = 0$.

5.4. Construct a function $f : \mathbb{R} \to \mathbb{R}$ so that $f'(x)$ exists for all $x \in \mathbb{R}$ but $f'(x)$ is not continuous at $x = 0$.

Hint: Use the previous problem

.

5.5. Let A be a subset of $\mathbb{R}$, $f : A \to \mathbb{R}$ and $x_0 \in \text{int}(A)$. Prove that f is differentiable at x_0 if and only if there is a function $\phi : A \to \mathbb{R}$ that is continuous at x_0, such that for $x \in A$,

$$f(x) = f(x_0) + (x - x_0)\phi(x).$$

Moreover, prove that if such a ϕ exists, then $f'(x_0) = \phi(x_0)$.

The next several problems can be solved using either the definitions directly or by using the previous problem.

5.6. Let A be a subset of $\mathbb{R}$ and suppose that $f : A \to \mathbb{R}$ and $g : A \to \mathbb{R}$ are both differentiable at $x_0 \in \text{int}(A)$. Let $c \in \mathbb{R}$. Prove:

(a) cf is differentiable at x_0 and $(cf)'(x_0) = cf'(x_0)$.
(b) $f + g$ is differentiable at x_0 and $(f + g)'(x_0) = f'(x_0) + g'(x_0)$.
(c) fg is differentiable at x_0 and $(fg)'(x_0) = f(x_0)g'(x_0) + f'(x_0)g(x_0)$.
(d) If $g(x_0) \neq 0$ then f/g is differentiable at x_0 and

$$\left(\frac{f}{g}\right)'(x_0) = \frac{g(x_0)f'(x_0) - f(x_0)g'(x_0)}{[g(x_0)]^2}.$$

Hint: For (d), start with the case that $f(x) = 1$ and then use (c).

Of course part (c) is typically called the *Product Rule* and part (d) is called the *Quotient Rule.*

5.7. Let $p : \mathbb{R} \to \mathbb{R}$ be given by a polynomial $p(x) = a_0 + a_1 x + \cdots + a_n x^n$. Prove that p is differentiable on $\mathbb{R}$ and that $p'(x) = a_1 + 2a_2 x + \cdots + n a_n x^{n-1}$ for all $x \in \mathbb{R}$.

Next we prove the *Chain Rule.*

5.8. Let A and B be subsets of $\mathbb{R}$. Let $f : A \to B$ and $g : B \to \mathbb{R}$ and assume that $f(A) \subset B$. Suppose that $x_0 \in \text{int}(A)$ and $f(x_0) \in \text{int}(B)$. If f is differentiable at x_0 and g is differentiable at $f(x_0)$, then $g \circ f$ is differentiable at x_0 and $(g \circ f)'(x_0) = g'(f(x_0))f'(x_0)$.

Hint: Write $g(x) = g(f(x_0)) + (x - f(x_0))\psi(x)$ for $x \in B$. Replace x by $f(x)$ for $x \in A$ and use $f(x) - f(x_0) = (x - x_0)\phi(x)$.

The next result is of course extremely important in optimizing functions (and in doing routine calculus problems).

5.9. Suppose that $f : [a, b] \to \mathbb{R}$ is continuous. Suppose that $x_0 \in (a, b)$ is a maximum for f. Then either f is not differentiable at x_0 or $f(x_0') = 0$. State and prove a corresponding result for minimums.

In other words, any maximum or minimum of f occurs either at an endpoint or at a "critical point" (that is a point at which the derivative is zero or does not exist). The next result is an extremely important result in the theory of differentiation. It is called the *Mean Value Theorem.*

5.10. Let $a < b$ and let $f : [a, b] \to \mathbb{R}$. Assume that f is differentiable on (a, b) and continuous at a and b. Then there exists a $c \in (a, b)$ with

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Hint: Prove this by following these steps.

(1) Prove the mean value theorem under the additional assumption that $f(a) = f(b) = 0$. To be precise: prove that if $f : [a, b] \to \mathbb{R}$ is differentiable on (a, b) and continuous at a and b with $f(a) = f(b) = 0$ then there is a $c \in (a, b)$ with $f'(c) = 0$.
(2) To prove the Mean Value Theorem its full generality, construct a linear function L which agrees with $f(x)$ at a and b and set $g(x) = f(x) - L(x)$. Then use (b).

The special case of the Mean Value Theorem where $f(a) = f(b) = 0$ which we have used to prove the full Mean Value Theorem is call *Rolle's Theorem.*

5.11. Prove the following corollaries of the Mean Value Theorem:

(a) Let $f : (a,b) \to \mathbb{R}$ satisfy $f'(x) = 0$ for all $x \in (a,b)$. Then f is constant on (a,b).
(b) Let $f, g : (a,b) \to \mathbb{R}$ be differentiable and satisfy $f'(x) = g'(x)$ for all $x \in (a,b)$. Then there exists a constant c so that $f(x) = g(x)+c$ for $x \in (a,b)$.
(c) Let $f, g : \mathbb{R} \to \mathbb{R}$ be differentiable with $f'(x) = g'(x)$ for all $x \in \mathbb{R}$. Then there exists a constant c so that $f(x) = g(x)+c$ for all $x \in \mathbb{R}$.

Definition. Let $A \subset \mathbb{R}$ and $f : A \to \mathbb{R}$. f is *increasing* if for all $x, y \in A$ with $x < y$ we have $f(x) \le f(y)$. f is *strictly increasing* if for all $x, y \in A$ with $x < y$ we have $f(x) < f(y)$.

Similarly we can define *decreasing* and *strictly decreasing.*

5.12. Let $f : (a,b) \to \mathbb{R}$ and assume that f is differentiable. Prove:

(a) If $f'(x) > 0$ for all $x \in (a,b)$ then f is strictly increasing on (a,b).
(b) If $f'(x) \ge 0$ for all $x \in (a,b)$ then f is increasing on (a,b).
(c) Prove the corresponding results for $f'(x) < 0$ and $f'(x) \le 0$.

5.13. Let $f : [a,b] \to \mathbb{R}$ be continuous. Assume that f is differentiable and f' is bounded on $[a,b]$. Prove that f is Lipschitz.

The next result shows that derivative functions satisfy a sort of Intermediate Value Theorem automatically.

5.14. Let $A \subset \mathbb{R}$ and $f : A \to \mathbb{R}$. Assume that $[a,b] \subset \text{int}(A)$ and that f' exists on $[a,b]$. Let γ lie between $f'(a)$ and $f'(b)$. Show that there exists $c \in (a,b)$ with $f'(c) = \gamma$.

Hint: Consider $F(x,y) = \frac{f(x)-f(y)}{x-y}$ on the set $\{(x,y) : a \le x < y \le b\}$. F is a continuous function on a connected set so its range is an interval.

5.15. Let f be the function of Problem 4.5. Is f differentiable at any point of $[0,1]$?

2. The Riemann Integral

Once again, we all are familiar with the purpose of the integral: its gives a way to measure the area under the graph of a function on an interval. The idea behind this is relatively straight forward: we divide the interval into pieces and make a rectangle out of each piece by using a value of the function to determine the height. To approximate the area under the graph, we add up the areas of these rectangles. Using more and more pieces gives better and better approximations. Nevertheless, the rigorous formulation is a bit technical so you should keep the intuitive ideas in mine when reading the following definition.

Definition. Let $a < b$. A *partition* P of $[a, b]$ is an ordered finite set $\{a = x_0 < x_1 < \cdots < x_n = b\}$. The *norm* or *width* of P is

$$\|P\| = \max\{\Delta x_i : 1 \leq i \leq n\}$$

where $\Delta x_i = x_i - x_{i-1}$. Let $f : [a, b] \to \mathbb{R}$. A *Riemann sum for f with respect to P* is a number

$$R = \sum_{i=1}^{n} f(x_i')\Delta x_i$$

where $x_i' \in [x_{i-1}, x_i]$ for each $1 \leq i \leq n$. We denote the collection of Riemann sums of f (with respect to P) by $R(f; P)$. The function f is said to be *Riemann integrable* or just *integrable* on $[a, b]$ if there exists an $I \in \mathbb{R}$ such that for all $\epsilon > 0$ there exists a $\delta > 0$ so that if P is any partition of $[a, b]$ with $\|P\| < \delta$ then for all $R \in R(f; P)$, $|I - R| < \epsilon$.

5.16. Show that if I and J both satisfy the definition above for f then $I = J$.

Thus I is unique and we call it the *Riemann Integral of f on* $[a, b]$ (or just the integral of f on $[a, b]$). We write $I = \int_a^b f = \int_a^b f(x)dx$.

5.17. Let $a < \alpha < \beta < b$ and define $f : [a, b] \to \mathbb{R}$ by

$$f(x) = \begin{cases} c & x \in [\alpha, \beta] \\ 0 & x \notin [\alpha, \beta] \end{cases}$$

for $c \in \mathbb{R}$. Prove that f is integrable on $[a, b]$ and $\int_a^b f = c(\beta - \alpha)$.

5.18. Let $f : [0, 1] \to \mathbb{R}$ be given by

$$f(x) = \begin{cases} 1 & x \text{ rational} \\ 0 & x \text{ irrational} \end{cases}$$

Prove that f is not integrable on $[0, 1]$.

5.19. Let f and g be integrable functions on $[a, b]$. Let $c \in \mathbb{R}$. Prove:

a) $f + g$ is integrable on $[a, b]$ and $\int_a^b (f + g) = \int_a^b f + \int_a^b g$

b) cf is integrable on $[a, b]$ and $\int_a^b cf = c\int_a^b f$

5.20. Let f and g be integrable on $[a, b]$. Prove:

(a) If $f \geq 0$ on $[a,b]$ then $\int_a^b f \geq 0$.
(b) If $f \geq g$ on $[a,b]$ then $\int_a^b f \geq \int_a^b g$.

5.21. Let f be integrable on $[a,b]$.

(a) Let $(P_n)_{n=1}^{\infty}$ be a sequence of partitions of $[a,b]$ with $\lim_{n\to\infty} \|P_n\| = 0$. Let $R_n \in R(f;P_n)$ for $n \in \mathbb{N}$. Prove that $\lim_{n\to\infty} R_n = \int_a^b f$.
(b) Prove

$$\int_a^b f = \lim_{n\to\infty} \sum_{i=1}^{n} f\left(a + \tfrac{i}{b-a}\right)\left(\tfrac{b-a}{n}\right).$$

5.22. Show that $\int_0^1 x dx = 1/2$ and $\int_0^1 x^2 dx = 1/3$, under the assumption that x and x^2 are integrable on [0,1].

5.23. Let f be integrable on $[a,b]$. Prove that f is bounded on $[a,b]$.

The following result gives a sort of Cauchy criteria for integrability.

5.24. Let $f : [a,b] \to \mathbb{R}$. Prove that f is integrable on $[a,b]$ if and only if for all $\epsilon > 0$ there exists a $\delta > 0$ with the following property: for all partitions P_1 and P_2 of $[a,b]$ with $\|P_1\|, \|P_2\| < \delta$ and for all $R_1 \in R(f;P_1)$, $R_2 \in R(f;P_2)$ we have $|R_1 - R_2| < \epsilon$.

Hint: For the forward implication, let (P_n) be partitions with $\|P_n\| \to 0$ and let $R_n \in R(f;P_n)$. Show that (R_n) is Cauchy. Let $I = \lim R_n$. Prove that $\int_a^b f = I$.

Definition. A *step function* on $[a,b]$ is a function $s : [a,b] \to \mathbb{R}$ such that for some partition $P = (x_1 < \cdots < x_n)$ of $[a,b]$, s is constant on each open subinterval (x_{i-1}, x_i) for $1 \leq i \leq n$. In other words, there exist $c_i \in \mathbb{R}$ so that $s(x) = c_i$ if $x_{i-1} < x < x_i$. $s(x_i)$ can be either c_i or c_{i+1}.

5.25. Let $s : [a,b] \to \mathbb{R}$ be a step function with $s(x) = c_i$ on (x_{i-1}, x_i) for some partition $(x_0 < \cdots < x_n)$ of $[a,b]$. Prove that s is integrable on $[a,b]$ and

$$\int_a^b s = \sum_{i=1}^{n} c_i \Delta x_i.$$

5.26. Let $f : [a,b] \to \mathbb{R}$. Prove that f is integrable on $[a,b]$ if and only if for all $\epsilon > 0$ there exist step functions s_1 and s_2 with $s_1 \leq f \leq s_2$ and $\int_a^b (s_2 - s_1) < \epsilon$.

Hint: For the forward implication, let $I = \int_a^b f$ and $\epsilon > 0$. Choose $\delta > 0$ so that $\|P\| < \delta$ implies $|R - I| < \epsilon/3$ for $R \in R(f;P)$. Fix such a P with $P = (x_0 < \cdots < x_n)$ and for $1 \leq i \leq n$, set

$$m_i = \inf f[x_{i-1}, x_i] \text{ and } M_i = \sup f[x_{i-1}, x_i].$$

Let $s_1(x) = m_i$ and $s_2(x) = M_i$ if $x \in (x_{i-1}, x_i)$.

For the reverse implication, let $s_1 \leq f \leq s_2$ with $\int_a^b (s_2 - s_1) < \epsilon/3$. Find $\delta > 0$ so that if $\|P\| < \delta$ and $R \in R(s_i;P)$ then $|R - \int_a^b s_i| < \epsilon/3$ for $i = 1, 2$. Fix $P = (x_0 < \cdots < x_n)$ with $\|P\| < \delta$. Let $R \in R(f;P)$, with $R = \sum_{i=1}^n f(x_i')\Delta x_i$. Let $R_j = \sum_{i=1}^n s_j(x_i')\Delta x_i$ for $j = 1, 2$. Show $R \in (\int_a^b s_1 - \frac{\epsilon}{3}, \int_a^b s_2 + \frac{\epsilon}{3})$.

Definition. Let $A \subset \mathbb{R}$. We say that A has *measure zero*, denoted $m(A) = 0$, if for all $\epsilon > 0$ there exists a sequence of intervals $(I_n)_{n=1}^\infty$ so that $A \subset \bigcup_{n=1}^\infty I_n$ and $\sum_{i=1}^\infty m(I_n) < \epsilon$ (where by $m(I_n)$ we mean the length of the interval I_n; when we say that the infinite sum is less than ϵ, we of course mean that it converges and that the limit is less than ϵ).

The previous definition is part of a very important branch of real analysis called measure theory. Intuitively it gives a precise way of measuring the 'size' of a set (in terms of length, areas, and volumes rather then cardinality; we give a discussion on the latter in the appendices). Some sets cannot be measured, but all reasonable ones can be.

As our language suggests, sets of the type above are exactly those those which measure to zero. Measure theory also leads to a more robust notion of the integral, known as the Lebesgue integral. It is more robust in the sense that any function which is Riemann integrable will also be Lebesgue integrable (and give the same integral), but some functions are Lebesgue integrable without being Riemann integrable. Thus Lebesgue integration is a strictly superior theory of integration.

Returning to our present setting, an observation that might help for the problems below is that

$$\frac{\epsilon}{2} + \frac{\epsilon}{4} + \frac{\epsilon}{8} + \cdots = \epsilon.$$

Make sure to prove this if you use it.

5.27. Let $A \subset \mathbb{R}$. Prove:

(a) A has measure zero if and only if for all $\epsilon > 0$ there exists a sequence of open intervals $(I_n)_{n=1}^\infty$ with $A \subset \bigcup_{n=1}^\infty I_n$ and $\sum_{n=1}^\infty m(I_n) < \epsilon$.

(b) A has measure zero if and only if for all $\epsilon > 0$ there exists a sequence of closed intervals $(J_n)_{n=1}^\infty$ with $A \subset \bigcup_{n=1}^\infty J_n$ and $\sum_{n=1}^\infty m(J_n) < \epsilon$.

5.28. Prove:

(a) Finite sets have measure zero.
(b) $m(\mathbb{Q}) = 0$.
(c) $[0, 1]$ does not have measure zero.
(d) If A is contained in a set of measure zero than A has measure zero.
(e) Suppose that (A_n) is a sequence of sets of measure zero. Then $\bigcup_{n=1}^{\infty} A_n$ also has measure zero.
(f) Countable sets have measure zero (the definition of countable can be found in the appendices).

5.29. Let $f : [a, b] \to \mathbb{R}$ and let $D(f)$ be the collection of points at which f is discontinuous. Prove that f is integrable if and only if f is bounded and $m(D(f)) = 0$.

Hint: For the forward implication, first show $D(f) = \bigcup_{n=1}^{\infty} D(f, \frac{1}{n})$, where $D(f, \frac{1}{n})$ is the collection of $x \in [a, b]$ such that for all $\delta > 0$ there exist $y, z \in [a, b]$ with $|y - x| < \delta$, $|z - x| < \delta$ and $f(y) - f(z) \geq 1/n$. Show $m(D(f, \frac{1}{n_0})) = 0$ by using Problem 5.26 for ϵ replaced by $\epsilon/(2n_0)$: let $J = \{1 \leq i \leq n : (x_{i-1}, x_i) \cap D(f, \frac{1}{n_0}) \neq \varnothing\}$ and note that

$$\sum_{i \in J} \frac{1}{n_0} \Delta x_i \leq \int_a^b (s_2 - s_1) < \frac{\epsilon}{2n_0}.$$

For the reverse implication again use Problem 5.26. To produce $s_1 \leq f \leq s_2$ with $\int_a^b s_2 - s_1 < \epsilon$, note that $D(f, \frac{\epsilon}{2(b-a)})$ is a closed set of measure zero so it can be covered by finitely many open intervals whose lengths sum to less than $\epsilon/(4K)$ where $|f| \leq K$. This defines a partition of $[a, b]$, $Q = (y_0 < \cdots < y_m)$. Moreover if we sum the lengths Δy_i of those intervals which have non-empty intersection with $D(f, \frac{\epsilon}{2(b-a)})$ we get a number less than $\epsilon/4K$. If $[y_{i-1}, y_i] \cap D(f, \frac{\epsilon}{2(b-a)}) = \varnothing$ then each point in the interval is the center of an open interval on which f varies less than $\epsilon/(2(b - a))$. So using compactness we can partition each such interval thusly.

5.30. Let f be the function of 4.5. Prove that f is integrable on $[0, 1]$. What is $\int_0^1 f$?

5.31. Let f be integrable on $[a, b]$ and let $a < c < b$. Prove that the restrictions $f : [a, c] \to \mathbb{R}$ and $f : [c, b] \to \mathbb{R}$ are both integrable and

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

If $f : [a, b] \to \mathbb{R}$ is integrable we define $\int_b^a f = -\int_a^b f$. We also define $\int_a^a f = 0$ for any function f (defined at a).

5.32. Let f be integrable on $[a, b]$ and let $c, d, e \in [a, b]$. Prove that

$$\int_c^e f = \int_c^d f + \int_d^e f,$$

regardless of the order of c, d, and e.

5.33. Let $f : [a, b] \to \mathbb{R}$ be increasing. Prove that f is integrable on $\mathbb{R}$.

Hint: Show $D(f)$ is countable. Note that if $x_0 \in D(f)$ then the "right and left hand limits" exist at x_0 but are not equal. Thus there exists a nonempty open interval $I_{x_0} = (\lim_{x\to x_0^-} f(x), \lim_{x\to x_0^+} f(x))$. If $x_1 \in D(f)$ with $x_1 \neq x_0$ then $I_{x_0} \cap I_{x_1} = \varnothing$.

5.34. Let $f : [a, b] \to \mathbb{R}$ be integrable. Prove that $|f|$ is integrable.

Hint: How does $D(|f|)$ relate to $D(f)$?

5.35. Let f and g be integrable on $[a, b]$. Prove that fg is integrable.

3. The Fundamental Theorem of Calculus

Those of us who were taught calculus in the modern era of mathematics have trouble appreciating the power and beauty of the Fundamental Theorem of Calculus. Being fairly intuitive, the idea of the integral was around long before calculus. Certainly mathematicians and scientists found it advantageous and necessary to know the volumes and areas of certain shapes. Nevertheless computing the area of any shape which was not extremely simple proves to be a very difficult task.

For example, even under the assumption that $f(x) = x^2$ is integrable on the interval $[0, 1]$, our computation of the area under the this segment of the parabola was necessarily very elaborate. The power of the Fundamental Theorem of Calculus is the it transforms the computationally difficult problem of the integral to the computationally simple (at least in certain situations) problem of an antiderivative. The beauty of the theorem is that it unifies to seemingly unrelated notions of mathematics. To us the derivative and the integral have always been intertwined, but in the 17th century it was quite a remarkable discovery.

5.36. Let $f : [a, b] \to \mathbb{R}$ be continuous. Define a function $F : [a, b] \to \mathbb{R}$ by $F(x) = \int_a^x f$.

(a) Prove F is Lipschitz on $[a, b]$. Moreover if $x \in (a, b)$ then $F'(x) = f(x)$.

(b) Let $G : [a, b] \to \mathbb{R}$ be continuous with $G'(x) = f(x)$ if $a < x < b$. Prove $\int_a^b f = G(b) - G(a)$.

5.37. Let $f : [a, b] \to [0, \infty)$ be continuous and not identically 0. Prove that $\int_a^b f > 0$.

5.38. Let $f : [a, b] \to \mathbb{R}$ be continuous. Prove that there exists $c \in [a, b]$ with $\int_a^b f = f(c)(b - a)$.

5.39. Let $f : [a, b] \to \mathbb{R}$ be continuous, $f \geq 0$. Let $M = \max f[a, b]$. Prove that $\lim_{n\to\infty}(\int_a^b f^n)^{1/n} = M$.

5.40. Define $F : C[a, b] \to \mathbb{R}$ by $F(f) = \int_a^b f$. Prove that F is uniformly continuous.

APPENDIX A

Prerequisites

1. Algebra

Though Real Analysis and Algebra are regarded as separate branches of mathematics, Real Analysis does rely on some of the basic fundamental constructs of Algebra. Indeed, calculus is the precursor to Real Analysis and one could not fathom a robust calculus course which avoided the use of the algebraic properties of numbers. In this section, we will describe briefly a bit of the more abstract algebraic language. Our primary purpose for introducing this language is that it will allow us to easily encapsulate many of the important properties of numbers. The language itself is not necessary for an understanding of real analysis, but the organization it permits can be quite convenient.

Definition. If A is a set, an *operation* on A is a map $A \times A \to A$.

Thus an operation is a rule which takes a pair of elements of A and returns a third element. There is an unlimited number of important examples of operations in mathematics. For our purposes, however, one should focus on addition and multiplication within various systems of numbers (like $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, or $\mathbb{R}$). In this vein, we often denote operations using symbols like '+' or '$\cdot$'. In other words, if our operation is '+,' we say the image of (a, b) is denoted $a + b$. If our operation is dnoted by '$\cdot$', we will often write ab rather than $a \cdot b$.

The abstract notion of an operation is far too general to be studied with any seriousness. Hence we typically place extra conditions on our operations. We will describe some of them.

Definition. Suppose $\cdot$ is an operation on a set A. We say that $\cdot$ is *associative* if for $a, b, c \in A$, we have $(a \cdot b) \cdot c = a \cdot (b \cdot c)$. Likewise, we say the operation is *commutative* if $a \cdot b = b \cdot a$ for $a, b \in A$. An element $e \in A$ is an *identity* for the operation if for each $a \in A$, we have $e \cdot a = a \cdot e = a$.

A.1. Suppose that $A \times A \to A$ is an operation. Show that if there is an identity for this operation then there is only one.

Since identities are unique, we see that the following definition makes since.

Definition. Suppose $\cdot$ is an operation on A and suppose that $e \in A$ is an identity. If $a \in A$, we say that $b \in A$ is an *inverse* of a with respect to $\cdot$ if $a \cdot b = b \cdot a = e$.

Definition. An *abelian group* is a set G, together with an associative commutative operation such that G contains an identity and such the every element has an inverse.

Thus we see that the collection of integers $\mathbb{Z}$ is an abelian group under the operation of addition. The same can be said of the set of rational numbers $\mathbb{Q}$. What is the identity and the inverses in these cases? Notice that the collection, $\mathbb{Q}^{\times}$, of nonzero rational numbers is also an abelian group under multiplication. Again what is the identity and the inverses? Why must we exclude zero?

Unlike $\mathbb{Q}^{\times}$, the collection of nonzero integers fails to be a abelian group under multiplication. Indeed, it satisfies all the requirements except for the existence of inverses. We thus say it is a *commutative monoid*. Explicitly, a commutative monoid is a set together with an associative commutative operation which has an identity. Thus $\mathbb{Q}$ and $\mathbb{Z}$ are both commutative monoids under multiplication.

Of course with numbers, it is important to consider the interaction of multiplication and addition, rather than just considering the two operations separately. We are lead to the following definition.

Definition. A *field* is a set F together with two operations $+$ and $\cdot$, called addition and subtraction respectively, subject to the following conditions:

(a) F is an abelian group under addition,
(b) the collection of nonzero elements of F is an abelian group under multiplication,
(c) and multiplication distributes over addition, which is to say that for $x, y, z \in F$, we have $x \cdot (y+z) = x \cdot y + x \cdot z$ and $(x+y) \cdot z = x \cdot y + y \cdot z$.

We typically denote the additive identity (that is, the identity with respect to addition) by 0_F or just 0 and the multiplicative identity by 1_F or 1. The additive inverse of x is denoted by $-x$ and (assuming $x \neq 0$) the multiplicative inverse is denoted by $1/x$ or x^{-1}. We define for x^n if $n \in \mathbb{Z}$ in the obvious manner.

A.2. If F is a field and $x \in F$, show that $0 \cdot x = x \cdot 0 = 0$. Show that multiplication in F is associative and commutative.

There are many basic properties of fields that may be derived from the above field axioms and we will not attempt to give even a partial list here. It is however, important to realize that any statement you wish to use in this course must be proven from the axioms. In general, however, it can be a bit tedious to prove all the 'obvious' statements we will be using (such as $(-1)x = -x$ or $\frac{ac}{bc} = \frac{a}{b}$) and so we will allow you to assume these. Nevertheless, if you have any doubt as to whether a statement is clear, you should prove it rigorously. If you think that something is 'obvious' and yet you are having difficulty showing it, it is probably something that you need to prove.

2. Orders

Thus far, we have used defined some important algebraic notions. In this section we discuss the notion of ordering a set. In addition to allowing us to formulate more properties of $\mathbb{R}$, the language of orders will lead us to two very important proof techniques.

Definition. A *relation* R on a set S is a subset of $S \times S$. The relation R on S is a *partial order* if

(a) R is *reflexive*, meaning $(a, a) \in R$ for each $a \in S$;
(b) R is *antisymmetric*, meaning that for $a, b \in S$, if $(a, b) \in R$ and $(b, a) \in R$ then $a = b$;
(c) and R is *transitive*, meaning that for $a, b, c \in S$, if $(a, b) \in R$ and $(b, c) \in R$ then $(a, c) \in R$.

We usually denote partial orderings by symbols like $\leq$. In other words, if $(a, b) \in R$, we write $a \leq b$. The related symbols $<$, $>$, and $\geq$ then have their obvious meanings. If we say $(S, \leq)$ is a partially ordered set we mean that $\leq$ is a partial ordering on the set S.

A.3. In each case, decide if the relation R on the set S is a partial order.

(a) $S = \mathbb{N}$ where $(m, n) \in R$ if m divides n.
(b) $S = \mathbb{N}$ where $(m, n) \in R$ if n divides m.
(c) $\mathbb{N}$ where $(m, n)R$ if m is an integer greater than or equal to n
(d) $S = \mathbb{Z}$ under $(m, n) \in R$ if $m - n$ is a multiple of 17.
(e) $S = \mathcal{P}(U)$ under $(A, B) \in R$ if $A \subset B$.

We point out that the word 'partial' is used to reflect that the fact that, given two elements $a, b \in S$, we do not necessarily have one of $a \leq b$ or $b \leq a$. In other words, some pairs of elements may not compare to each other. If we do have either $a \leq b$ or $b \leq a$, we say that a and b are *comparable*.

Definition. An order under which each pair of elements is comparable is called a *total order* or more simply an *order*.

We note that the usual ordering gives a total order on $\mathbb{N}$, $\mathbb{Z}$, and on $\mathbb{Q}$.

Definition. Let $(S, \leq)$ be a partially ordered set. If T is a subset, an *upper bound* for T is an $b \in S$ so that $a \leq b$ for all $a \in T$. A *maximal element* of T is an element $b \in T$ so that for $a \in B$, if $a \leq b$ then $a = b$. A *chain* in S is a subset consisting of elements each pair of which are comparable.

A.4. Let $(S, \leq)$ be a partially ordered set and let T be a subset Show by example that a maximal element of T need not be an upper bound for T. Show, however, that if S is totally ordered a maximal element of T is an upper bound for T and T can have at most one maximal element.

A.5. Use your basic knowledge of $\mathbb{R}$ to find find all upper bounds in $\mathbb{R}$ (if any) and all maximal elements (if any) for each of the following subsets.

(a) $\mathbb{N}$
(b) $(-1, 2)$
(c) $(-\infty, 2]$

A.6. Let $(S, \leq)$ be given by $S = \mathcal{P}(\mathbb{N})$ and $A \leq B$ if $A \subset B$.

(a) Give an example of an infinite chain $C \subset S$.
(b) Does every subset of S have a maximal element? An upper bound?

The following is an important technique for proofs in many fields of mathematics.

Zorn's Lemma: Let $(S, \leq)$ be a partially ordered set with $S \neq \varnothing$. Assume every chain in S has an upper bound. Then S has a maximal element.

Zorn's lemma can be shown to be logically equivalent to the so-called "axiom of choice" which, roughly speaking, states that given a set of sets, one can select precisely one element from each set and form a new set. Nearly all mathematicians consider this axiom to be " obviously true" and so consider it a valid assumption in mathematics (there is a slight controversy in that the axiom of choice leads to strange and counter-intuitive results in certain exotic settings).

We mentioned that the standard collection of set-theoretic axioms is called ZFC. We remark here that the 'C' stands for 'choice,' reflecting the fact that the axiom of choice is assumed. If you are curious, the 'Z' and the 'F' stand for Zermelo and Fraenkel, respectively. These two mathematicians were foundational in formulating some of the basic notions of set theory.

A partially ordered set with the property that every chain has an upper bound is called *inductively ordered*. Thus Zorn's Lemma states that every inductively ordered set has a maximal element.

Even if you are not familiar with Zorn's Lemma, you are probably familiar with the second proof technique of this section: mathematical induction. In fact this technique is based on a very special property of $\mathbb{N}$.

Definition. An ordered set is called *well-ordered* if every nonempty set contains a minimum element.

The quintessential example of a well-ordered set is $\mathbb{N}$ and we add this fact to our list of axioms. Once again, the axioms of set theory allow one to construct $\mathbb{N}$ precisely and then prove that it is well-ordered.

THE THEOREM OF MATHEMATICAL INDUCTION: Let $P(1), P(2), \ldots$ be a sequence of statements. Suppose that

1) $P(1)$ is true, and
2) for all $n \in \mathbb{N}$ if $P(n)$ is true then $P(n+1)$ is true.

Then for all n, $P(n)$ is true.

A.7. Prove the Theorem of Mathematical Induction.

A.8. Use mathematical induction to prove that for all $n \in \mathbb{N}$

$$1 + 2 + 3 + \cdots + (n-1) + n = \frac{n(n+1)}{2}$$

and

$$1^2 + 2^2 + \cdots + (n-1)^2 + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

A.9. Consider the following statement. Let $n \in \mathbb{N}$. In any classroom of n students if at least one has red hair then all have red hair. Find the error in the 'proof' below.

PROOF. Let $P(n)$ be the statement "in any classroom of n students if at least one has red hair then all have red hair." Clearly $P(1)$ is true. Assume $P(n)$ is true. We will prove $P(n+1)$ is true. To illustrate our argument in full generality we show $P(6)$ implies $P(7)$. Suppose a classroom has 7 students in it and at least one has red hair. Keep the red head seated and ask one of the other students to step outside the room. Now we have a classroom of 6 students and at least one has red hair. Since $P(6)$ is true, by assumption, all 6 have red hair. Invite the 7th student back in and send any of the other students outside. Again we have 6 students and at least one (in fact at least 5) have red hair so all have red hair. Invite the external student back inside the classroom and gaze upon the 7 red heads. □

Needless to say the above false proof demonstrates how even a seemingly small error in logic can lead to ridiculous consequences.

In the following problem, you may assume that every positive real number has a unique positive square root in $\mathbb{R}$. This is of course the case, but does not follow from the axioms we have introduce so far. We will prove it later.

A.10. Let $a_1 = \sqrt{2}$ and for $n \in \mathbb{N}$, $a_{n+1} = \sqrt{2 + a_n}$. Show that for all n, $a_n < 2$.

3. Ordered Fields

Thus far we have introduce two types of properties: algebraic properties and ordering properties. In the case of $\mathbb{R}$, these two type properties interact favorable in a manner which we now describe.

Definition. Suppose that F is a field and $\leq$ is an order on F. Then F is an *ordered field* with respect to $\leq$ if

(a) $\leq$ is preserved by addition, meaning that if $x, y, z \in F$ and $x \leq y$ then $x + z \leq y + z$ and
(b) if $x, y \in F$ and $x, y \geq 0$, then $xy \geq 0$.

If F is an ordered field, we get a notion of positive. That is, we define an element x to be *positive* if $x > 0$. We likewise define negative, nonnegative, and nonpositive (though the latter term is not used very often).

A.11. Suppose that F is an ordered field and $a \in F$. Show that $a^2 \geq 0$ and that $a^2 = 0$ if and only if $a = 0$. Show also that $-1 < 0$.

In particular, since -1 is a square in $\mathbb{C}$, the field of complex numbers, we conclude that there is no order on $\mathbb{C}$ which makes it an ordered field. Nevertheless, there are many examples of ordered fields. It does turn out to be the case that every ordered field must contain $\mathbb{Q}$ (in an appropriate sense).

4. Vector Spaces

In this section we recall the basic definitions of Linear Algebra.

Definition. Suppose that F is a field. A *vector space* over F is an abelian group V equipped with an action $F \times V \to V$, called *scalar multiplication* and denoted $(x, v) \mapsto xv$ such that

(a) $x(v_1 + v_2) = xv_1 + xv_2$ for $x \in F$ and $v_1, v_2 \in V$ and
(b) $(xy)v = x(yv)$ for $x, y \in F$ and $v \in V$.

Since V is an abelian group, it has an identity element. We typically call this element the zero vector and denote it with the symbol 0_V or 0.

The basic example of a vector space is F^n. Addition is performed pointwise, meaning

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots x_n + y_n)$$

for $x_i, y_i \in F$ and scalar multiplication is defined by $x(y_1, \dots, y_n) = (xy_1, \dots, xy_n)$ for $x, y_i \in F$. You should verify for yourself that F^n satisfies all the requirements on a vector space.

Definition. Let S be a subset of the F-vector space V. An F-linear combination of S is an expression of the form $x_1 s_1 + \cdots + x_r s_r = 0$ with $s_i \in S$ and $x_i \in F$. The x_i in such a sum are called the *coefficients*. S is said to be*linearly independent* (over F) if the only way to write zero as a linear combination of S is to make all the coefficients zero. Likewise, S is said to *span* V (over F) if each element of V may be written as a linear combination of S. S is called a *basis* for V (over F) if it is linearly independent and spans V.

The following characterization of bases demonstrates their value.

A.12. Show that a subset S of a F-vector space V is basis if and only if each element $v \in V$ may be written uniquely as linear combination of S.

A.13. Suppose V is a vector space over the field F. Suppose that S_1 is a linearly independent set and S_2 is set containing S_1 which spans V. Show there is a basis B with $S_1 \subset B \subset S_2$. Show that every vector space contains a basis.

Hint: This problem can be solved with an application of Zorn's Lemma. Find an appropriate ordered set.

A.14. Find a basis for F^n.

APPENDIX B

Supplementary Material and Problems

1. Cardinality

Like many of the great ideas in mathematics, the notion of cardinality is a way to make an intuitive idea precise. We all know that finite sets carry with them a notion of 'size,' namely the number of ideas in the set. Cardinality is a way to extend this notion even to infinite sets. We begin by defining what it means for two sets to have the same 'size.' The mathematician Georg Cantor (who also gave the first rigorous definition of the real numbers) is one of the fathers of set theory and is widely credited as the creator of cardinality.

Definition. Let A and B be sets. We shall say that A and B have the *same cardinality*, denoted $|A| = |B|$ if there exists a bijection $f : A \to B$. If $f : A \to B$ and $g : B \to C$ we define $g \circ f : A \to C$ by $g \circ f(a) = g(f(a))$.

B.1. Let A, B and C be sets. Prove:

(a) $|A| = |A|$
(b) $|A| = |B|$ implies $|B| = |A|$
(c) $|A| = |B|$ and $|B| = |C|$ implies $|A| = |C|$

B.2. Use Zorn's lemma to prove that for all sets A and B either $|A| \leq |B|$ or $|B| \leq |A|$.

Hint: Let S consist of all ordered triples (C, D, f) where $C \subset A$, $D \subset B$ and $f : C \to D$ is 1–1 and onto. We order S by $(C, D, f) \leq (C', D', g)$ if $C \subset C'$, $D \subset D'$ and g extends f. Thus if $x \in C$ then $g(x) = f(x)$. Show that S contains a maximal element.

Definition. A set A is *finite* if $A = \varnothing$ or if there exists $n \in \mathbb{N}$ so that $|A| = |\{1, 2, \ldots, n\}|$. In this last case, we write $|A| = n$. A is *infinite* if A is not finite. A is *countably infinite* if $|A| = |\mathbb{N}|$. A is *countable* if A is finite or countably infinite. A is *uncountable* if A is infinite but not countably infinite.

The next result should demonstrate the motivation behind the definition of having the same cardinality.

B.3. Let A and B be finite sets. Show that $|A| = |B|$ if and only if A and B have the same number of elements.

B.4. Let A be a set. Prove that A is countably infinite if and only if we can write A in the form

$$A = \{a_1, a_2, \ldots\} = \{a_n\}_{n=1}^{\infty},$$

where $a_n \neq a_m$ if $n \neq m$.

B.5. Prove that A is countable if and only if $A = \varnothing$ or we can write $A = \{a_1, a_2, a_3, \ldots\}$. Here the a_i's need not be distinct as in the previous problem.

B.6. Let A and B be sets. Prove

(a) If $A \subset B$ and B is countable then A is countable.
(b) If $A \subset B$ and A is uncountable then B is uncountable.
(c) If A and B are countable then $A \times B$ is countable.
(d) $\mathbb{Q}$ is countably infinite.
(e) If I is countable and for all $i \in I$, A_i is countable, then $\bigcup_{i \in I} A_i$ is countable (we say that a countable union of countable sets is countable).
(f) $[0, 1]$ is uncountable.
(g) $\mathbb{R}$ is uncountable.
(h) $\mathbb{R} \setminus \mathbb{Q}$ is uncountable.

If A and B are sets we say that the cardinality of A is less than or equal to the cardinality of B, denoted "$|A| \leq |B|$" if there exists an injection $f : A \to B$.

B.7. For all sets A, B, C, prove:

(a) $|A| = |B|$ implies that $|A| \leq |B|$.
(b) $|A| \leq |A|$.
(c) $|A| \leq |B|$ and $|B| \leq |C|$ implies $|A| \leq |C|$.
(d) $|A| \leq |B|$ and $|B| \leq |A|$ implies that $|A| = |B|$.

Hint: Part (d) is a bit trickier than one might expect. In fact, even the Cantor himself gave a flawed proof. By hypothesis there exist injective functions $f : A \to B$ and $g : B \to A$. We must produce a bijection $h : A \to B$. Let $b \in B$. If there exists $a \in A$ with $f(a) = b$ we call a a *first ancestor* of b. If there exists $c \in B$ with $g(c) = a$ we call c a *second ancestor* of b (and a first ancestor of a) and so on. Let

$$A_0 = \{a \in A : a \text{ has an odd number of ancestors}\}$$
$$A_e = \{a \in A : a \text{ has an even number of ancestors}\}$$
$$A_i = \{a \in A : a \text{ has an infinite number of ancestors}\}.$$

Similarly we may define B_0, B_e and B_i. Show that

$A = A_0 \cup A_e \cup A_i$ and
$A_0 \cap A_e = \varnothing$, $A_0 \cap A_i = \varnothing$ and
$A_e \cap A_i = \varnothing$ (and similarly for B).

In this case we call (A_e, A_0, A_i) a *partition* of A. Show that f maps A_e bijectively onto B_0 and A_i 1–1 onto B_i. Show g maps B_e bijectively onto A_0. Then define $h : A \to B$ by $h(a) = f(a)$ if $a \in A_e \cup A_i$ and if $a \in A_0$ let $h(a)$ be the first ancestor of A.

Although our language has alluded to it, we have not yet proved that there exists a set which is not countable. In some sense it is rather counter-intuitive that this is the case. Essentially we are saying that some infinite sets are "strictly bigger" than others. One might say that there is more than one kind of infinity.

B.8. Prove that for all sets A, $|\mathcal{P}(A)| > |A|$.

Hint: First show there exists an injection $f : A \to \mathcal{P}(A)$. Then show there does not exist a surjective $g : A \to \mathcal{P}(A)$ as follows. Suppose for a contradiction that g is such. Let $C = \{a \in A : a \notin g(a)\}$. Suppose $g(c) = C$. Is $c \in C$? Is $c \notin C$?

Thus, for example, the cardinality of $\mathcal{P}(\mathbb{N})$ is strictly larger than that of $\mathbb{N}$. In particular, $\mathcal{P}(\mathbb{N})$ is not countable. The same result demonstrates even more: given **any** set, we may find a set with strictly larger cardinality. A consequence of this fact is there are infinitely-many different cardinalities (of infinitely-many different types of infinite).

One is naturally led to ask: which infinity describes the number of different infinities? In fact this a very deep question and is essentially unanswerable: the collection of all cardinalities is too 'large' to be a set itself and thus we cannot give it a cardinality (if one tries to consider the set of all cardinalities, one is led inevitable to certain logical paradoxes).

We will show below that $|\mathbb{R}| = |\mathcal{P}(\mathbb{N})|$. In practice, mathematicians only need to deal with certain small cardinalities: finite ones, the cardinality of $\mathbb{N}$, the cardinality of $\mathbb{R}$, and the cardinality of sets like the set of all functions from $\mathbb{R} \to \mathbb{R}$.

B.9. Prove:

(a) $|\mathbb{R}| = |\mathbb{R}^2|$
(b) $|[0,1]| = |\mathbb{R}|$
(c) $|S| = |[0,1]| = |(0,1)|$ where S is the set of all sequences of 0's and 1's.
(d) Let $A = \{p(x) : p(x)$ is a polynomial with integer coefficients$\}$. Then
$$|A| = |\mathbb{N}| \ .$$

Cantor himself was shocked to discover that the cardinality of $\mathbb{R}$ was the the same as that of $\mathbb{R}^2$ because he expected that the space $\mathbb{R}^2$, which has larger dimension than $\mathbb{R}$, should be comprise a larger set than did $\mathbb{R}^2$. He famously said "I proved it, but I don't believe it."

B.10. Prove that $|\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$.

Hint: Use binary decimal expansions to give a bijection between $[0, 1]$ and the collection of infinite subsets of $\mathbb{N}$. Show that the latter had the same cardinality as $\mathcal{P}(\mathbb{N})$.

For historical reasons, the cardinality of $\mathbb{R}$ is called the *cardinality of the continuum.* We have shown that the cardinality of the continuum is strictly larger than the cardinality of $\mathbb{N}$. One is naturally led to wonder if there are any cardinalities in between. The *continumm hypothesis* is the statement that there is this is false: there is no cardinality between that of $\mathbb{R}$ and that of $\mathbb{N}$.

In fact it was demonstrated in the 20th Century that the truth of truth of the continuum hypothesis is independent of the usual axioms of set theory (even when when one assumes the axiom of choice). In other words, one cannot prove that the continuum hypothesis is true and one cannot prove that the continuum hypothesis is false.

2. Challenge Problems: A Space Filling Curve

In the previous section, we saw that there exists a bijection between $\mathbb{R}$ and $\mathbb{R}^2$, a fact that surprise Georg Cantor. Perhaps even stranger is the result we will show in this section: one can construct a **continuous** surjective function $f : [0, 1] \to [0, 1]^2$. We typically think of the plane and the line as two different worlds, but this result shows that the continuous line segment $[0, 1]$ is "big enough" to fill the square $[0, 1]^2$. Results like this forced mathematicians to rethink many of the basic notions of dimension.

The following steps guarantee the existence of an onto continuous $f : [0, 1] \to [0, 1]^2$. Let $f_0 : [0, 1] \to [0, 1]^2$ be the diagonal map. That is, put $f_0(t) = (t, t)$ for $t \in [0, 1]$.

Define f_1 as follows: divide the square $[0, 1]^2$ into nine equal sub-squares (like a tic-tac-toe board). Map $[0, 1/9]$ to the 'northeast' diagonal

$$\{(t, t) : t \in [0, 1/3]\}$$

of the lower-left sub-square in the obvious way, namely linearly in the 'north-east'direction: $x \mapsto (3x, 3x)$. Likewise map $[1/9, 2/9]$ to the 'northwest' diagonal in the left-center subsquare. Map $[2/9, 3/9]$ to the appropriate diagonal of the upper-left sub-square. Proceed in this way to the upper-center, center, lower-center, lower-right, center-right, and upper center sub-squares (in that order).

We define f_2 by further sub-dividing. Indeed, we divide each of the nine sub-squares into nine individual pieces and divided each of the intervals $[i/9, (i+1)/9]$ into subintervals and map accordingly. We continue this process obtaining $f_3, f_4, \ldots$. As an aside, this construction is an example of a 'fractal.'

B.11. Show that each f_n is continuous and that (f_n) is uniformly Cauchy. Let f be the uniform limit of the f_n and show that $f : [0,1] \to [0,1]$ is continuous and surjective.

3. Challenge Problems: Convergence of Functions

B.12. Prove or disprove: Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. There exists a sequence $(f_n)_{n=1}^{\infty}$ of functions from $\mathbb{R} \to \mathbb{R}$ so that for all n and all $x_0 \in \mathbb{R}$, f_n is discontinuous at x_0 and $f_n \to f$ uniformly.

B.13. For $n \in \mathbb{N}$ let $f_n(x) = x^n(1-x)$ for $0 \leq x \leq 1$. Show that (f_n) is uniformly convergent on $[0,1]$.

B.14. For $n \in \mathbb{N}$ define $f_n : [0,1] to \mathbb{R}$ by

$$f_n(x) = \frac{nx}{1+nx^2}.$$

Is (f_n) uniformly convergent on $[0,1]$? Prove or disprove.

4. Challenge Problems: The Cantor Set

The *Cantor set* is a set with very interesting set-theoretic, measure-theoretic, and topological properties (which is to say that is very useful for constructing counter-examples to would be theorems). We will denote it by Δ and define it as follows.

Let

$$\begin{aligned}
K_0 &= [0,1] \ , \\
K_1 &= \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right] \ , \\
K_2 &= \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right] \ , \\
&\cdots
\end{aligned}$$

In general, K_n is a union of 2^n closed sub-intervals of $[0,1]$ and to get K_{n+1} from K_n we remove the middle third of each sub-interval. We define

$$\Delta = \bigcap_{n=0}^{\infty} K_n \ .$$

B.15. Prove:

(a) Δ is uncountable so that $|\Delta| = |[0,1]|$
(b) $m(\Delta) = 0$
(c) Δ is compact and $\text{int}(\Delta) = \varnothing$.
(d) $\Delta' = \Delta$.
(e) $[0,1] \setminus \Delta$ is dense in $[0,1]$.

From the Cantor set, we can define the *Cantor function*. It also has very interesting properties. Define the Cantor function $c : [0,1] \setminus \Delta \to [0,1]$ as follows. On the interval $K_0 - K_1$, c is defined to be $\frac{1}{2}$. $K_1 - K_2$ is two subintervals, c is define to be $\frac{1}{4}$ on the first and $\frac{3}{4}$ on the other. We proceed in this manner.

B.16. Show that c is uniformly continuous. Thus c has a unique uniformly continuous extension to $[0,1]$ which we also denote by c. Show that $c : [0,1] \to [0,1]$ is onto and increasing and that $c'(x) = 0$ for all $x \notin \Delta$.

5. Challenge Problems: Separability

Definition. A metric space (E, d) is called *separable* if contains a countable dense subset.

B.17. Prove:

(a) $\mathbb{E}^n$ is separable for all $n \in \mathbb{N}$.
(b) $\mathbb{R}$, when given the discrete metric, is not separable.
(c) $C[0,1]$ is separable.
(d) Letℓ_∞ denote the normed linear space of all bounded sequences $x = (x_i)_{i=1}^\infty$ of reals under the norm

$$\|x\| = \sup\{|x_i| : i \in \mathbb{N}\} .$$

Show that ℓ_∞ is not separable.

B.18. Consider the following statement: if E is a separable metric space and $A \subset E$ then A is separable. Find an error in the following proof.

Proof. Let $D \subset E$ with $\bar{D} = E$ and D countable. Let $D_0 = A \cap D$. Then D_0 is countable and $\bar{D}_0 = A$. □

B.19. Give a correct proof of the previous statement.

B.20. Let E be a metric space and let $A \subset E$ be nonempty. If A is bounded, we have previously defined the diameter. If A is unbounded, we define its diameter to be ∞. If $A, B \subset E$ are both nonempty we let

$$d(A, B) = \inf\{d_E(a, b) : a \in A,\, b \in B\}.$$

(a) Prove that if $A \subset E$ is compact then there exists $x, y \in A$ with $d_E(x, y) = \operatorname{diam}(A)$.
(b) Show by example that the conclusion of the previous part need not be true if A is not compact.
(c) Prove that if $A, B \subset E$ are compact and both sets are nonempty then there exists an $a \in A$ and a $b \in B$ with $d(A, B) = d_E(a, b)$.
(d) Show by example that the conclusion of the previous part need not be if A is not compact.

B.21. Let $f_n : [0, 1] \to \mathbb{R}$ be continuous for all $n \in \mathbb{N}$. Assume that for all $x \in [0, 1]$, the sequence $(f_n(x))_{n=1}^{\infty}$ is decreasing and converges to zero. Show that f_n converges uniformly to the zero function.

6. Challenge Problems: The Baire Category Theorem

The next result which seeming very technical nonetheless has many important consequences. It is called the Baire Category Theorem.

B.22. Let (E, d) be a complete metric space. If $E = \bigcup_{n=1}^{\infty} A_n$ and each A_n is closed then for some n, $\operatorname{int}(A_n) \neq \varnothing$.

Hint: Use contradiction. Find a closed ball B_1 of radius less than 1 with $B_1 \cap A_1 = \varnothing$. Find a closed ball $B_2 \subset B_1$ of radius less than 1/2 with $B_2 \cap A_2 = \varnothing$. Proceed in this manner and derive a contradiction.

B.23. Let E be a complete metric space. For $n \in \mathbb{N}$, let $A_n \subset E$ be open and dense. Then $\bigcap_{n=1}^{\infty} A_n$ is also dense in E.

Hint: Use the Baire Category Theorem.

As we mentioned, the Baire Cateogory Theorem has many consequences. Here is an interesting one.

B.24. For $K \in \mathbb{N}$, let F_K be the collection of $f \in C[0, 1]$ such that there is an $x_0 \in [0, 1 - 1/K]$ with $|f(x) - f(x_0)| \leq K(x - x_0)$ for all $x_0 \leq x \leq 1$. Prove:

(a) F_K is a closed subset of $C[0, 1]$.
(b) $C[0, 1] \setminus F_K$ is dense in $C[0, 1]$.
(c) Let $f \in C[0, 1]$ be differentiable at $x_0 \in (0, 1)$. Then $f \in F_K$ for some $K < \infty$.
(d) There exists a continuous function $f : [0, 1] \to \mathbb{R}$ which is not differentiable at a single point.

In fact, one can construct such a function more directly, but the Baire category theorem greatly simplifies the proof.

B.25. Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and not bounded. Then there exists $x \in \mathbb{R}$ with $(f(nx))_{n\in\mathbb{N}}$ unbounded.

Hint: Let $A_K = \{x \in [0, 1] : |f(nx)| \leq K \text{ for all } n \in \mathbb{N}\}$.

7. Challenge Problems: Miscellaneous

B.26. Let K be a compact metric space and let $K \supset K_1 \supset K_2 \supset \cdots$ where each K_i is nonempty and closed. Let G be an open set with $\bigcap_{n=1}^{\infty} K_n \subset G$. Then there exists n with $K_n \subset G$.

Definition. Let E be a metric space. A set $A \subset E$ is called *path connected* if for each $a, b \in A$ there is a map $f : [0, 1] \to A$ with $f(0) = a$ and $f(1) = b$.

B.27. Let $G \subset E^n$ be open and connected. Prove that G is *path connected*.

Definition. A function $f : E \to G$ of metric spaces is called an *isometry* if for all $p, q \in E$ we have

$$d_G(f(p), f(q)) = d_E(p, q).$$

In other words an isometry is a function which preserves distances (whereas a continuous function only preserves closeness).

B.28. Let E be a compact metric space and let $T : E \to E$ be an isometry. Show that T is surjective. Show by example that this is not necessarily true if E is not compact.

B.29. Let $T : [a, b] \to [a, b]$ be continuous. Let $a \leq x_1 < x_2 < x_3 < x_4 \leq b$ with $Tx_1 = x_2$, $Tx_2 = x_3$,$Tx_3 = x_4$ and $Tx_4 \leq x_1$. Show that there exists an $x \in [a, b]$ with $T \circ T \circ T(x) = x$ but $T \circ T(x) \neq x$.