

MATH 131A: REAL ANALYSIS

NICKOLAS ANDERSEN

The textbook for the course is Ross, *Elementary Analysis* [2], but in these notes I have also borrowed from Tao, *Analysis I* [3], and Abbott, *Understanding Analysis* [1].

1. INTRODUCTION

Real analysis is a rigorous study of the (hopefully familiar) objects from calculus: the real numbers, functions, limits, continuity, derivatives, sequences, series, and integrals. While you are likely quite familiar with computing with these objects, this course will focus on developing the theoretical foundation for the definitions, theorems, formulas, and algorithms that you are used to using. We will start by building up the real numbers from scratch, i.e., from just a few basic axioms, then we will focus our attention on proving many of the things you already believe about functions, sequences, and series. Along the way we will encounter several “pathological” objects which will hopefully convince you that our careful approach is necessary and worthwhile.

To get an idea of how subtle some questions in analysis can be, ask yourself: what is a real number? We will answer this question in due time, but for now let's focus on one specific real number that got a lot of attention from the ancient Greeks: $\sqrt{2}$. Prior to the discovery that $\sqrt{2}$ is an irrational number, it was assumed that: given any two line segments $\overline{AB}$ and $\overline{CD}$, there is a rational number r so that the length of $\overline{CD}$ equals r times the length of $\overline{AB}$. However, the length of the diagonal of a square of side length 1 (using the Pythagorean Theorem) equals $\sqrt{2}$, so by the previous assumption, we must have that $\sqrt{2}$ is a rational number. The proof of the following theorem is one of the most classic proofs in mathematics.

Theorem 1.1. *There is no rational number whose square is 2.*

Proof. Suppose that $x^2 = 2$ and that x is a rational number. Recall that a rational number is one which can be expressed as p/q for integers p and q . To prove that there are no integers p, q for which $x = p/q$ we will employ an important proof technique called *proof by contradiction*. That is, we will assume that $x = p/q$ for some integers p, q and we will carefully follow logical steps until we end up with something absurd. Thus, our original assumption must have been faulty. Here we go.

Suppose that there are integers p, q for which

$$\left(\frac{p}{q}\right)^2 = 2. \tag{1.1}$$

We may assume that p and q have no common factors, since we could just cancel the common factors and write p/q in lowest terms. Equation (1.1) implies that

$$p^2 = 2q^2. \tag{1.2}$$

It follows that p^2 is an even number (it's 2 times an integer), and therefore p is an even number (you can't square an odd number and get an even number). So we can write $p = 2r$

for some integer r . Then equation (1.2) becomes (after cancelling $2s$)

$$2r^2 = q^2.$$

By the previous discussion, this implies that q is even. But that's ridiculous! We assumed that p and q had no factors in common, but we just showed that p and q were both even. Since we have reached a *contradiction*, it must be that our initial assumption (1.1) was false. Thus 2 is not the square of any rational number. $\square$

The previous theorem shows that there is a “hole” at $\sqrt{2}$ in the rational numbers. The importance of this fact cannot be overstated. Later it will lead us to the *Axiom of Completeness* which is an essential property that the real numbers enjoy which basically states that there are no holes in the set of real numbers. This will lead us to limits, derivatives, continuity, and eventually integrals. But first we need to take a few steps back and start at the beginning.

2. THE NATURAL NUMBERS

If our aim is to construct the real numbers and do calculus on the set of real numbers, then we must start with the simplest numbers first and build our way up. Thus our story begins with the *natural numbers* (a.k.a whole numbers or counting numbers) which we can informally define as the elements of the set

$$\mathbb{N} := \{1, 2, 3, 4, 5, \dots\}.$$

We are no longer in the business of informal definitions, so we'll need to build $\mathbb{N}$ from scratch. In excruciating detail.

Let's think about what we want in a set of numbers. In mathematics, it is often desirable not to think too carefully about the actual elements in a set, but more about what you want those elements to do (i.e. what operations or functions do you want to apply to those elements?). A few moments of thought might lead you to say that the most important thing we do with the natural numbers is *counting* (you might have said addition or multiplication, but addition is just repeated counting, and multiplication is just repeated addition). So it stands to reason that we should construct the natural numbers so that we can count with them.

We will begin with two concepts: the number 1, and the successor $n + 1$. Note that we haven't defined addition yet (we don't even know what the numbers are!) so $n + 1$ doesn't mean n “plus” 1. Yet. It's just an expression that we use to denote the successor of n . Informally (but less informally than before), we will define the natural numbers as the set containing 1, the successor $1 + 1$, the successor $(1 + 1) + 1$, the successor $((1 + 1) + 1) + 1$, etc. This leads to our first two axioms.

Axiom 1: 1 is a natural number.

Axiom 2: If n is a natural number, then $n + 1$ is a natural number.

By Axioms 1 and 2, we see that

$$((((((1 + 1) + 1) + 1) + 1) + 1) + 1) + 1$$

is a natural number. Don't worry, we won't write numbers like this; instead we'll use the notation we're all familiar with. So the number above is called 8. But for now, the symbol 8 means nothing other than a shorthand notation for the successor of the successor of the successor of the successor of the successor of the successor of 1.

It may seem like this is enough to define the natural numbers, but consider the set consisting of all natural numbers from 1 to 12, where the successor of 12 is 1 (this is not some crazy thing, it's how clocks work!). Even though this number system obeys Axioms 1 and 2, it doesn't even allow us to count how many fingers and toes we have, so it must not be right. Let's add another axiom.

Axiom 3: 1 is not the successor of any natural number; i.e. $n + 1 \neq 1$ for all n .

Now we can prove statements like the following.

Lemma 2.1. $4 \neq 1$.

Proof. By definition $4 = 3 + 1$. By Axioms 1 and 2, 3 is a natural number (since $3 = (1 + 1) + 1$). Thus by Axiom 3, $3 + 1 \neq 1$. Therefore $4 \neq 1$. $\square$

At this rate we'll never get to derivatives! (Don't worry, we're going to go through the construction of the natural numbers in painful detail so that you can see what goes into a rigorous mathematical foundation of analysis. Then we will move a bit faster so that we can cover other things.)

Have we constructed $\mathbb{N}$ yet? Unfortunately, there are still weird pathological number systems which satisfy the first three axioms, but which are not the natural numbers (as we would like them to be). Consider the number system

$$1, 1 + 1 = 2, 2 + 1 = 3, 3 + 1 = 4, 4 + 1 = 4, 4 + 1 = 4, \dots$$

You can check that this doesn't break our first three axioms, but it's still definitely not right. Let's add another axiom.

Axiom 4: If n and m are natural numbers and $n \neq m$ then $n + 1 \neq m + 1$.

Equivalently, if $n + 1 = m + 1$ then $n = m$.

Now we can't have the above pathology.

Lemma 2.2. $4 \neq 2$.

Proof. Suppose, by way of contradiction, that $4 = 2$. Then $3 + 1 = 1 + 1$, so by Axiom 4 we have $3 = 1$. But that contradicts Axiom 3, so our original assumption must have been wrong. Thus $4 \neq 2$. $\square$

We're not out of the woods yet. We have constructed a set of axioms which confirms that all of the numbers that we think should be natural numbers (i.e. $1, 2, 3, \dots$) are elements of $\mathbb{N}$. But we can't rule out the existence of other numbers masquerading as natural numbers. For example, the set

$$\{.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, \dots\}$$

satisfies all of the axioms so far. So we'll need one final axiom. This one is so important that it gets its own name. (You'll want to chew on this one for a bit.)

Axiom 5 (The principle of mathematical induction):

Let P_n be any statement or proposition that may or may not be true.

Suppose that P_1 is true, and that whenever P_n is true, P_{n+1} is also true.

Then P_n is true for every natural number n .

The principle of mathematical induction allows us to prove that a statement is true by simply checking two things: first we check that the statement is true for $n = 1$, then,

assuming it is true for n , we check that it is true for $n + 1$. Here is an example (note: this example really belongs later in the course, after we've defined addition and multiplication and division, but I'm happy to time travel for a few seconds if you are).

Proposition 2.3. *For all natural numbers n we have*

$$1 + 2 + \dots + n = \frac{n(n+1)}{2}.$$

Proof. For each n , the statement we want to prove is

$$P_n : "1 + 2 + \dots + n = \frac{n(n+1)}{2}."$$

We proceed by induction. We begin with P_1 , which states that $1 = \frac{1(1+1)}{2}$. This is certainly true. Suppose that P_n is true, i.e. suppose that

$$1 + 2 + \dots + n = \frac{n(n+1)}{2}$$

is a true statement. We wish to show that P_{n+1} is true. Add $n + 1$ to both sides to obtain

$$\begin{aligned} 1 + 2 + \dots + n + n + 1 &= \frac{n(n+1)}{2} + n + 1 \\ &= \frac{(n+1)(n+1+1)}{2}. \end{aligned}$$

Thus P_{n+1} holds if P_n holds. By the principle of mathematical induction, P_n holds for all natural numbers n . $\square$

Note that we didn't prove P_n directly for any n except for $n = 1$. We just proved P_1 , and we proved that if P_1 is true, so is P_2 (thus P_2 is true), and we proved that if P_2 is true, so is P_3 (thus P_3 is true), and we proved that if P_3 is true, so is P_4 (thus P_4 is true), etc. It's like stacking up an infinite line of dominoes and knocking over the first one.

Since we will use induction regularly to prove things in lecture, homework, and exams, it might be useful to have a template for such proofs.

Proposition 2.4. *A property P_n is true for all natural numbers n .*

Proof. We proceed by induction on n (it's good to specify the variable if there are several variables in the statement you want to prove). We first verify the *base case* $n = 1$, i.e. we prove that P_1 is true. [Insert proof of P_1 here]. Now suppose that P_n has already been proven. We show that P_{n+1} is true. [Insert proof of P_{n+1} , assuming P_n]. It follows that P_n is true for all natural numbers n . $\square$

3. THE INTEGERS AND THE RATIONALS

At this point, if we had all the time in the world, we would maintain the glacial pace of the last section and carefully develop the theory of addition and multiplication on the natural numbers. But for the sake of time (and to maintain our sanity) we will take some things for granted as we build up the integers and the rational numbers. This means that the discussion in the section will be quite informal. I hope our careful approach to the natural numbers was enough to convince you that there is a careful way to construct the integers and the rational numbers from the natural numbers. If you would like to learn about this in more detail, see [3].

Informally, the set of *integers* is made up of the positive integers, the negative integers and 0. We know the positive integers well; these are just the natural numbers. We can add ($2 + 3 = 5$) and multiply ($2 \cdot 3 = 6$) natural numbers as usual. (Formally, $m + n$ is defined as applying the successor to m , n times; multiplication is then defined as repeated addition.) These operations satisfy the rules

$$\begin{aligned} a + b &= b + a, & (\text{commutative law for addition}) \\ (a + b) + c &= a + (b + c), & (\text{associative law for addition}) \\ a \cdot b &= b \cdot a, & (\text{commutative law for multiplication}) \\ (a \cdot b) \cdot c &= a \cdot (b \cdot c), & (\text{associative law for multiplication}) \\ a \cdot 1 &= a, & (\text{multiplicative identity}) \\ (a + b) \cdot c &= a \cdot c + b \cdot c. & (\text{distributive law}) \end{aligned}$$

In order to define subtraction, we introduce the *additive inverse* of a number and the *additive identity element*. The additive identity element, 0, is defined by its behavior under addition via

$$n + 0 = 0 + n = n. \quad (\text{additive identity})$$

(Note: we could have defined the natural numbers to include 0, and some authors do this. But Ross [2] doesn't, so we won't.) For each natural number n , the additive inverse of n , which we will call $-n$, is defined by the property

$$n + (-n) = (-n) + n = 0. \quad (\text{additive inverse})$$

Subtraction is then defined via

$$a - b := a + (-b).$$

(Note: here I'm being quite informal and sweeping some subtleties under the rug. For example, we would like to say that the integers $1 - 5$ and $2 - 6$ and $3 - 7$, etc., are all the same, but the definition I've given you above doesn't account for that. But to do this properly, we would have to introduce the notion of *equivalence classes*, which we won't do here. See Tao's book [3] if you are interested in learning more.)

We define the integers as the elements of the set

$$\mathbb{Z} := \mathbb{N} \cup \{0\} \cup (-\mathbb{N}).$$

Here $-\mathbb{N}$ is the set consisting of the additive inverses of all the natural numbers, i.e., the negative integers. The notation $\cup$ means *union*. For any two sets A and B , the set $A \cup B$ consists of all the elements that are in A and/or B . As you are already aware, the laws listed above extend to the entire set $\mathbb{Z}$ of integers. Using these laws, we can prove some familiar properties of integers.

Proposition 3.1. *Suppose that $a, b, c \in \mathbb{Z}$. Then*

- (1) $a + c = b + c$ implies that $a = b$,
- (2) $a \cdot 0 = 0$,
- (3) $(-a) \cdot b = -(a \cdot b)$,
- (4) $(-a) \cdot (-b) = a \cdot b$,
- (5) $a \cdot b = 0$ implies that $a = 0$ or $b = 0$ (or both),
- (6) $a \cdot c = b \cdot c$ and $c \neq 0$ together imply that $a = b$.

Proof. (1) If $a + c = b + c$ then $(a + c) + (-c) = (b + c) + (-c)$. So by associativity of addition we have $a + (c + (-c)) = b + (c + (-c))$. Then by the additive inverse property, we have $a + 0 = b + 0$, and by the additive identity property, we conclude that $a = b$.

(2) By the additive identity property and the distributive law we have

$$0 + a \cdot 0 = a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0.$$

Using (1) we find that $0 = a \cdot 0$.

(3) Starting with $a + (-a) = 0$, we multiply both sides by b and use the distributive law to see that

$$a \cdot b + (-a) \cdot b = (a + (-a)) \cdot b = 0 \cdot b = 0,$$

where we have used (2) in the last equality. Thus we see that $(-a) \cdot b = -(a \cdot b)$.

(4) Exercise.

(5) We will prove that $a, b \neq 0$ implies that $a \cdot b \neq 0$ (convince yourself that this is enough! This is an example of proof using *contrapositive*).

First assume that $a, b \in \mathbb{N}$. Then since multiplication is repeated addition, which is the same as taking the successor repeatedly, by Axiom 3 we cannot have $a \cdot b = 0$ (this is slightly informal, but since we never formally defined addition or multiplication, I'm going to let it slide).

Now suppose that $a \in (-\mathbb{N})$ and $b \in \mathbb{N}$. Then $a = -n$ for some $n \in \mathbb{N}$, and by (3) we have $a \cdot b = (-n) \cdot b = -(n \cdot b)$. Since $n, b \in \mathbb{N}$, the previous case shows that $n \cdot b \neq 0$, so $-(n \cdot b) \neq 0$.

Now suppose that $a, b \in (-\mathbb{N})$. Then $a = -m$ and $b = -n$ for some $m, n \in \mathbb{N}$, and by (4) we have $a \cdot b = (-m) \cdot (-n) = m \cdot n$. Since $m, n \in \mathbb{N}$, the first case applies again and we have $m \cdot n \neq 0$.

(6) Exercise. □

Just as we defined subtraction by creating the negative integers and using addition, we can define division by creating reciprocals and using multiplication. For each *nonzero* integer n , we define the *multiplicative inverse* n^{-1} by the relation

$$n \cdot n^{-1} = n^{-1} \cdot n = 1. \quad (\text{multiplicative inverse})$$

Division is then defined by (assuming that $b \neq 0$)

$$a/b := a \cdot b^{-1},$$

and the *rational numbers* are all of the numbers a/b for $a, b \in \mathbb{Z}$ with $b \neq 0$. We use $\mathbb{Q}$ to denote the set of rational numbers. In set-builder notation,

$$\mathbb{Q} := \left\{ \frac{a}{b} : a, b \in \mathbb{Z} \text{ and } b \neq 0 \right\}.$$

(As with subtraction, we have some subtlety here: we would like to think of the rational numbers $1/2$ and $2/4$ and $3/6$ as being equal. To do this properly requires an equivalence relation; see [3] for more details if you are interested.)

As you know already, the rational numbers $\mathbb{Q}$ inherit the same properties as the integers, along with the properties listed in Proposition 3.1. For completeness, we list them all together

here. If $a, b, c \in \mathbb{Q}$ then

$$\begin{aligned}
 a + b &= b + a, & (\text{commutative law for addition}) \\
 (a + b) + c &= a + (b + c), & (\text{associative law for addition}) \\
 n + 0 &= n, & (\text{additive identity}) \\
 n + (-n) &= 0, & (\text{additive inverse}) \\
 a \cdot b &= b \cdot a, & (\text{commutative law for multiplication}) \\
 (a \cdot b) \cdot c &= a \cdot (b \cdot c), & (\text{associative law for multiplication}) \\
 a \cdot 1 &= a, & (\text{multiplicative identity}) \\
 n \cdot n^{-1} &= 1, & (\text{multiplicative inverse}) \\
 (a + b) \cdot c &= a \cdot c + b \cdot c. & (\text{distributive law})
 \end{aligned}$$

A number system which satisfies all of these properties (and for which $0 \neq 1$; that would be weird) is called a *field*.

The set field $\mathbb{Q}$ is actually an *ordered field*, which means that it has an order structure $\leq$ which obeys the following rules (which you are already familiar with):

- (O1) For all $a, b \in \mathbb{Q}$, either $a \leq b$ or $b \leq a$.
- (O2) If $a \leq b$ and $b \leq a$ then $a = b$.
- (O3) If $a \leq b$ and $b \leq c$ then $a \leq c$.
- (O4) If $a \leq b$ then $a + c \leq b + c$.
- (O5) If $a \leq b$ and $c \geq 0$ then $ac \leq bc$.

Using these rules, we can prove the following properties. Note that $a > b$ means that $a \geq b$ and $a \neq b$.

Proposition 3.2. *If $a, b, c \in \mathbb{Q}$ then the following properties hold.*

- (1) *If $a \leq b$ then $-a \geq -b$.*
- (2) *If $a \leq b$ and $c \leq 0$ then $ac \geq bc$.*
- (3) *If $a \geq 0$ and $b \geq 0$ then $ab \geq 0$.*
- (4) *$a^2 \geq 0$ for all a .*
- (5) *$0 < 1$.*
- (6) *If $a > 0$ then $a^{-1} > 0$.*
- (7) *If $0 < a < b$ then $0 < b^{-1} < a^{-1}$.*

Proof. (1) We apply (O4) with $c = (-a) + (-b)$. If $a \leq b$ then

$$-b = -b + a + (-a) = a + (-a) + (-b) \leq b + (-a) + (-b) = (-a) + b + (-b) = -a,$$

as desired.

- (2) Suppose that $c \leq 0$. Then by (1) we have $-c \geq 0$. If $a \leq b$ then by (O5) we have $(-c)a \leq (-c)b$ which implies that $-(ac) \leq -(bc)$. Applying (1) again we find that $ac \geq bc$.
- (3) Follows immediately from (O5) with $a = 0$.
- (4) By (O1), either $a \geq 0$ or $a \leq 0$. If $a \geq 0$ then we apply (3) with $b = a$. If $a \leq 0$ then by (1) we have $-a \geq 0$. By (3) again we have $(-a)(-a) \geq 0$, i.e. $a^2 \geq 0$.
- (5) Exercise.

- (6) Suppose, by way of contradiction, that $a > 0$ but $a^{-1} \leq 0$. Applying (2) to $0 \leq a$ we find that

$$0 = 0 \cdot a^{-1} \geq aa^{-1} = 1,$$

which contradicts (5).

- (7) Exercise. □

4. ABSOLUTE VALUE

One of the most important basic concepts in analysis is that of *absolute value*. The purpose of absolute value is to measure distance between two numbers; for example (we will see this over and over again in this course) if we want to say that two numbers x and y are very close, we would write

$$|x - y| < \varepsilon,$$

where $\varepsilon > 0$ is some very small positive number. The absolute value is defined by

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

From this definition, we immediately see that

$$|x| \geq 0 \quad \text{for all } x.$$

The absolute value plays nicely with multiplication:

$$|xy| = |x| \cdot |y|,$$

which you can prove easily by splitting into the four cases $x, y \geq 0$, $x, -y \geq 0$, $-x, y \geq 0$, and $x, y \leq 0$.

The following theorem is ~~probably~~ the most important and most used inequality in analysis. It's so important that it has its own name.

Theorem 4.1 (The Triangle Inequality). *For all $x, y \in \mathbb{Q}$ we have*

$$|x + y| \leq |x| + |y|.$$

To prove the triangle inequality, we require the following lemmas which you will prove as exercises.

Lemma 4.2. *For all $x \in \mathbb{Q}$ we have $-|x| \leq x \leq |x|$.*

Lemma 4.3. *For all $x, y \in \mathbb{Q}$ we have $|x| \leq |y|$ if and only if $-|y| \leq x \leq |y|$.*

Proof of Theorem 4.1. Using Lemma 4.3, it is enough to prove that

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

But this follows by adding together the inequalities

$$\begin{aligned} -|x| &\leq x \leq |x| \\ -|y| &\leq y \leq |y| \end{aligned}$$

from Lemma 4.2. □

5. THE REAL NUMBERS

We are now ready to construct the real numbers $\mathbb{R}$. Recall that the property which sets the real numbers apart from the rational numbers is that, in the reals, all of the “holes” are filled (i.e. $\sqrt{2}$ is not a rational number, but it is a real number). The purpose of this section is to make this property precise and then to construct the real numbers to satisfy it. We will freely use all the properties of the rational numbers, ordering, and the absolute value without explicitly stating them.

Let’s start with $\sqrt{2}$. It may help to time travel for a bit; in the following discussion it will be helpful to know that

$$\sqrt{2} = 1.414\,213\,562\,373\,095\,048\ldots$$

Although we haven’t defined the decimal expansion of a number, you already know that $1.414213 = \frac{1414213}{1000000}$, etc. While we can’t find a rational number whose square is 2, we can find rational numbers whose square is *arbitrarily close to 2*, as follows:

$$\begin{aligned} 1^2 &= 1, \\ 1.4^2 &= 1.96, \\ 1.41^2 &= 1.9881, \\ 1.414^2 &= 1.999396, \\ 1.4142^2 &= 1.99996164, \\ 1.41421^2 &= 1.9999899241. \end{aligned}$$

Assuming we can continue this process indefinitely (and we can; there will be a homework problem about it later) this defines a *sequence* of rational numbers that *converges* to $\sqrt{2}$ (if we can make sense of what those words mean). Using this example as a template, we will “fill in the gaps” in the rational numbers by asserting that every sequence that *should* converge *does* converge. Let’s try to make this precise.

First, a *sequence* is just a list of numbers

$$a_1, a_2, a_3, a_4, \dots$$

indexed by the natural numbers $1, 2, 3, \dots$. For now, all of the $a_i \in \mathbb{Q}$ (because we haven’t defined $\mathbb{R}$) but later they can be in $\mathbb{R}$. We will usually start the index of our lists at 1, but sometimes it is convenient to start them at some other integer. We usually think of sequences as being infinitely long lists, but often it is useful consider finite sequences; sometimes we will think of a finite sequence as a finite list, and sometimes we will just implicitly repeat the last element of the list indefinitely, i.e. $a_1, a_2, a_3, \dots, a_n, a_n, a_n, \dots$. Usually this distinction will not make too big a difference.

We would like our sequences to be convergent; roughly speaking, that means that a_n and a_{n+1} should be getting closer together as n gets larger. Actually, this is not quite enough; instead we will use the following definition.

Definition 5.1. A *Cauchy sequence* (of rational numbers) is a sequence $(a_1, a_2, a_3, a_4, \dots)$ such that for every rational $\varepsilon > 0$ there exists a positive integer N (which is allowed to depend on ε) such that

$$|a_m - a_n| < \varepsilon \quad \text{whenever} \quad m, n \geq N.$$

Note: the restriction that ε needs to be a rational number is there purely because we don't know what a real number is yet. Later we will consider Cauchy sequences of real numbers and we will think of ε as being any positive real number. You should not think about this distinction too much, as it will not be important in the long run.

We should unpack Definition 5.1; it's quite important to understand this well. It's saying that a Cauchy sequence is a sequence whose terms *eventually* (i.e. there exists a positive integer N) *get really close together* (i.e. $|a_m - a_n| < \varepsilon$). Note that this needs to hold for all $m, n \geq N$, not just n and $n + 1$, say. You should think of ε as being a microscopically tiny number; so for $\varepsilon = \frac{1}{10}$ or $\frac{1}{100}$ or $\frac{1}{10^{10}}$ or $\frac{1}{10^{100}}$, etc. the terms of our sequence are eventually at most a distance of ε apart.

Proposition 5.2. *The sequence $(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \dots)$ is a Cauchy sequence.*

Proof. We would like to come up with a natural number N so that

$$\left| \frac{1}{m} - \frac{1}{n} \right| < \varepsilon \quad \text{whenever} \quad m, n \geq N.$$

Using the fact that $m, n \geq N$, together with the triangle inequality, we find that

$$\left| \frac{1}{m} - \frac{1}{n} \right| \leq \frac{1}{m} + \frac{1}{n} \leq \frac{2}{N}.$$

So it is enough to come up with a number N for which $2/N < \varepsilon$ or, in other words, $N > 2/\varepsilon$. But this is possible by the Archimedean property (the following proposition). Now that we've done some scratch work, let's write out a formal proof.

Given $\varepsilon > 0$, choose a natural number $N > 2/\varepsilon$. Then if $m, n \geq N$ we have

$$|a_m - a_n| = \left| \frac{1}{m} - \frac{1}{n} \right| \leq \frac{1}{m} + \frac{1}{n} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Therefore (a_i) is a Cauchy sequence. □

To complete the proof we need the Archimedean property.

Proposition 5.3 (The Archimedean property). *For each $x \in \mathbb{Q}$ there exists an $n \in \mathbb{N}$ such that $n > x$.*

Proof. Since $x \in \mathbb{Q}$ there exist integers p, q with $q \geq 1$ such that $x = p/q$. If $p \leq 0$, we can just take $n = 1$. So we may assume that $p \geq 1$. Then we can take $n = p$ since $q \geq 1$ implies that $p/q \leq p < p + 1$. □

Here's a non-example.

Proposition 5.4. *For each $n \geq 1$ let $H_n := 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n}$. Then the sequence $(H_1, H_2, H_3, \dots)$ is not a Cauchy sequence.*

Proof. In order to show that a sequence is not Cauchy, it suffices to find a specific $\varepsilon > 0$ such that for all $N \in \mathbb{N}$ we have

$$|a_m - a_n| \geq \varepsilon \quad \text{for some} \quad m, n \geq N.$$

That is, no matter how big you take N , we will always be able to find a pair m, n for which a_m and a_n are greater than ε apart.

For this particular sequence we can take $\varepsilon = \frac{1}{2}$. For each $n \in \mathbb{N}$, consider the difference

$$|H_{2n} - H_n| = \frac{1}{n+1} + \frac{1}{n+2} + \frac{1}{n+3} + \dots + \frac{1}{2n}.$$

Notice that there are exactly n terms in the sum, and each term is larger than (or equal to) the smallest one $\frac{1}{2n}$. Thus

$$|H_{2n} - H_n| \geq \frac{1}{2n} + \cdots + \frac{1}{2n} = \frac{1}{2}.$$

Therefore, no matter how big you choose N , we will always be able to find a pair $n, m (= 2n)$ such that $|a_m - a_n| \geq \frac{1}{2}$, which shows that the sequence (H_n) is not Cauchy. $\square$

Right now, proving that a sequence is Cauchy requires a bit of work, even for a sequence as simple as $(\frac{1}{n})$. Later we will prove some limit laws that allow us to determine when a sequence is Cauchy more easily.

We would like to do arithmetic on Cauchy sequences like addition and multiplication, but first we need to develop the notion of boundedness.

Definition 5.5. A sequence (a_i) is *bounded* if there exists a (rational) number M such that $|a_i| \leq M$ for all $i \geq 1$. In this case, we say that (a_i) is *bounded by M* .

Lemma 5.6. *Every finite sequence $(a_1, \dots, a_n)$ is bounded.*

Proof. We proceed by induction on the length n . If $n = 1$ then the statement of the lemma is clear by taking $M = |a_1|$. We want to show that the sequence $(a_1, \dots, a_n, a_{n+1})$ is bounded. By the induction hypothesis, every list of length n is bounded; in particular $(a_1, \dots, a_n)$ is bounded. Then there exists some M such that $|a_i| \leq M$ for $1 \leq i \leq n$. Since the list $(a_1, \dots, a_{n+1})$ is bounded by $M + |a_{n+1}|$, we are done. $\square$

Proposition 5.7. *Every Cauchy sequence is bounded.*

Proof. Suppose that (a_i) is a Cauchy sequence. Taking $\varepsilon = 1$, this means that there is some $N \in \mathbb{N}$ for which $|a_m - a_n| \leq 1$ for all $m, n \geq N$. This splits our sequence into a finite piece (indices 1 through N) and an infinite piece (index N and after). (Note that we are using a_N in both pieces.) By Lemma 5.6, the finite piece is bounded, say by M . Using the Triangle Inequality we find that

$$|a_n| = |a_n - a_N + a_N| \leq |a_n - a_N| + |a_N| < 1 + M.$$

It follows that (a_n) is bounded by $M + 1$. $\square$

The trick we used in the preceding proof of adding and subtracting a_N (clever addition of 0) comes up all the time in analysis (and throughout mathematics). You should keep this trick in your back pocket for quick use at all times.

We can now add and multiply Cauchy sequences.

Proposition 5.8. *If (a_i) and (b_i) are Cauchy sequences then so are $(a_i + b_i)$ and $(a_i b_i)$.*

Proof. Given $\varepsilon > 0$ there are natural numbers N_a and N_b such that

$$|a_m - a_n| < \frac{\varepsilon}{2} \quad \text{and} \quad |b_m - b_n| < \frac{\varepsilon}{2}$$

if $m, n \geq N_a$ and $m, n \geq N_b$ (you'll see in a minute why we chose $\varepsilon/2$). Let $N = \max\{N_a, N_b\}$ and suppose that $m, n \geq N$. Then

$$|(a_m + b_m) - (a_n + b_n)| = |(a_m - a_n) + (b_m - b_n)| \leq |a_m - a_n| + |b_m - b_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus $(a_i + b_i)$ is a Cauchy sequence.

For the product, use Proposition 5.7 to find M_a and M_b such that $|a_i| \leq M_a$ and $|b_i| \leq M_b$, and set $M = \max\{M_a, M_b\}$. Since (a_i) and (b_i) are Cauchy, there are natural numbers N_a and N_b (likely different than before) such that

$$|a_m - a_n| < \frac{\varepsilon}{2M} \quad \text{and} \quad |b_m - b_n| < \frac{\varepsilon}{2M}$$

if $m, n \geq N_a$ and $m, n \geq N_b$. Let $N = \max\{N_a, N_b\}$. Then if $m, n \geq N$ we have (cleverly adding 0 again)

$$\begin{aligned} |a_m b_m - a_n b_n| &= |a_m b_m - a_m b_n + a_m b_n - a_n b_n| \\ &\leq |a_m| |b_m - b_n| + |b_n| |a_m - a_n| < M \frac{\varepsilon}{2M} + M \frac{\varepsilon}{2M} < \varepsilon. \end{aligned}$$

Thus $(a_i b_i)$ is Cauchy. $\square$

It should be clear that since the rational numbers satisfy the commutative ring axioms (associativity and commutativity of addition and multiplication, the distributive law, etc.), so do Cauchy sequences. Is there an additive identity element? Sure, take $(0, 0, 0, \dots)$. What about a multiplicative identity element? Yes, use $(1, 1, 1, \dots)$.

What about division? It is tempting to define division of Cauchy sequences by (a_n/b_n) but this only works if $b_n \neq 0$ for all n . Fine, but we don't want to divide by the zero element $(0, 0, 0, \dots)$ anyway. However, there is a catch: the sequences $(1, 1, 1, 1, 1, \dots)$ and $(1, 0, 1, 1, 1, 1, \dots)$ are both Cauchy sequences and they seem to have the same "limit," namely 1. So we should be able to divide by either of them; but we can't divide by the second one in the usual way because of that pesky 0.

A more interesting example comes from the two Cauchy sequences

$$\begin{aligned} &1, 1.4, 1.41, 1.414, 1.4142, 1.41421, \dots, \\ &2, 1.5, 1.42, 1.415, 1.4143, 1.41422, \dots \end{aligned}$$

Both "look like" they are converging to $\sqrt{2}$. But none of the terms in the first sequence are equal to their counterpart in the second sequence, so the sequences are definitely not the same. Our aim is eventually to define the real numbers as limits of Cauchy sequences of rational numbers, so we should consider the two sequences above to be "the same."

Definition 5.9. Two sequences (a_i) and (b_i) are *equivalent* if for each $\varepsilon > 0$ there exists a natural number N (which can depend on ε) such that

$$|a_n - b_n| < \varepsilon \quad \text{whenever} \quad n \geq N.$$

The following two propositions will allow us to define division.

Proposition 5.10. *Suppose that (a_i) is a Cauchy sequence that is not equivalent to the zero sequence $(0, 0, 0, \dots)$. Then there exists a Cauchy sequence (b_i) and a number $m > 0$, with $|b_i| \geq m$ for all i , such that (a_i) and (b_i) are equivalent.*

Proof. Since (a_i) is not equivalent to the zero sequence, there exists a fixed number ε_0 for which the definition of equivalent sequences fails. This means that for each $N \geq 1$ there exists a number $n \geq N$ for which

$$|a_n - 0| \geq \varepsilon_0.$$

Now, (a_i) is a Cauchy sequence, so there exists some (fixed) N_0 for which

$$|a_m - a_n| < \frac{\varepsilon_0}{2} \quad \text{whenever} \quad m, n \geq N_0.$$

And for this fixed N_0 there is a fixed $n_0 \geq N_0$ for which

$$|a_{n_0}| = |a_{n_0} - 0| \geq \varepsilon_0. \quad (5.1)$$

I claim that $|a_n| \geq \varepsilon_0/2$ for every $n \geq N_0$. Why? Suppose, by way of contradiction, that $|a_{n_1}| < \varepsilon_0/2$ for some $n_1 \geq N_0$. Then, by the triangle inequality

$$|a_{n_0}| = |a_{n_0} - a_{n_1} + a_{n_1}| \leq |a_{n_0} - a_{n_1}| + |a_{n_1}| < \frac{\varepsilon_0}{2} + \frac{\varepsilon_0}{2} = \varepsilon_0.$$

But this contradicts (5.1), so we must have $|a_n| \geq \varepsilon_0/2$ for every $n \geq N_0$.

This almost does it. We have shown that the sequence a_n eventually satisfies $|a_n| \geq m$ for $m = \varepsilon_0/2$. We simply define $b_i = m$ for $1 \leq i < N_0$ and $b_i = a_i$ for $i \geq N_0$. It is clear that (a_i) and (b_i) are equivalent because $|a_i - b_i| = 0$ for all $i \geq N_0$. This finishes the proof. $\square$

Sequences such as (b_i) from the previous proposition which satisfy $|b_i| \geq m$ for some $m > 0$ are said to be *bounded away from zero*.

Proposition 5.11. *Suppose that (a_i) is a Cauchy sequence which is bounded away from zero. Then the sequence (a_i^{-1}) is also a Cauchy sequence.*

Proof. Since (a_i) is bounded away from zero, there exists some $m > 0$ for which $|a_i| \geq m$ for all i . Thus

$$|a_m^{-1} - a_n^{-1}| = \left| \frac{a_m - a_n}{a_m a_n} \right| \leq \frac{|a_m - a_n|}{m^2}.$$

So it suffices to find an N for which $m, n \geq N$ implies that $|a_m - a_n| < m^2 \varepsilon$. But (a_i) is a Cauchy sequence, so we can certainly do that. $\square$

Now we can define division of sequences via $(a_i b_i^{-1})$ as long as the sequence (b_i) is bounded away from zero. Furthermore, Proposition 5.10 shows that any sequence (b_i) which is not equivalent to the zero sequence is equivalent to a sequence (c_i) which is bounded away from zero. So we should be able to divide by any sequence which is not equivalent to zero by just swapping (b_i) out for (c_i) . But how do we know that this is well-defined? That is, does the result depend on which sequence (c_i) we choose to swap out for (b_i) ? There will be many choices in general. This opens up a whole new can of worms.

It's time to finally define what we mean by a real number, and then we will be able to answer some of the questions posed above. To do this we will introduce a formal symbol "lim" a_n which you should think of as merely a symbol (hence the quotation marks). We are using "lim" so that your brain thinks of this as the familiar limit of a sequence, but since we haven't defined what we mean by limit yet, this is still just a symbol.

Definition 5.12. A *real number* is defined to be an object of the form "lim" a_n where (a_n) is a Cauchy sequence of rational numbers. Two real numbers "lim" a_n and "lim" b_n are said to be equal if (a_n) and (b_n) are equivalent Cauchy sequences. The set of real numbers is $\mathbb{R}$.

It is a good idea to check that this definition of equal is consistent.

Lemma 5.13. *Suppose that $x = \text{"lim" } a_n$, $y = \text{"lim" } b_n$, and $z = \text{"lim" } c_n$ are real numbers. Then*

- (1) $x = x$,
- (2) if $x = y$ then $y = x$,
- (3) if $x = y$ and $y = z$ then $x = z$.

Proof. Exercise. $\square$

Are the rational numbers also real numbers? In a formal sense, the answer is “no” because real numbers are technically objects of the form “ $\lim$ ” a_n , and rational numbers are not. But there is a sense in which we can *embed* $\mathbb{Q}$ into $\mathbb{R}$: for each rational number x , the sequence $(x, x, x, \dots)$ is Cauchy (why?) and so the object “ $\lim$ ” x is a real number. In this way we can think of the rational numbers as sitting inside the reals, i.e. $\mathbb{Q} \subset \mathbb{R}$.

Since we know how to add, subtract, multiply, and divide sequences, we can do those operations immediately on real numbers. But there is one subtlety that we should take care of: if x, y , and z are real numbers and $x = y$, then is it the case that $x + z = y + z$? Similarly, is it the case that $x \cdot z = y \cdot z$? Since two different Cauchy sequences can represent the same real number, it's not immediately clear. The following lemma takes care of this.

Lemma 5.14. *Addition, multiplication, and reciprocation are well defined. That is, suppose that $x = \text{“lim” } a_n$, $x' = \text{“lim” } a'_n$, and $y = \text{“lim” } b_n$ with $x = x'$. Then we have*

- (1) $x + y = x' + y$,
- (2) $x \cdot y = x' \cdot y$,
- (3) if $x \neq 0$, then $x^{-1} = (x')^{-1}$.

Proof. (1) We need to show that the sequences $(a_n + b_n)$ and $(a'_n + b_n)$ are equivalent. Let $\varepsilon > 0$ be given. Since $x = x'$, the sequences (a_n) and (a'_n) are equivalent, so there exists a number N for which

$$|a_n - a'_n| < \varepsilon \quad \text{whenever} \quad n \geq N.$$

So if $n \geq N$ we have

$$|(a_n + b_n) - (a'_n + b_n)| = |a_n - a'_n| < \varepsilon.$$

Thus $(a_n + b_n)$ is equivalent to $(a'_n + b_n)$.

- (2) Exercise.
- (3) Suppose that $x \neq 0$. Then $x' \neq 0$ by Lemma 5.13. So $x = \text{“lim” } a_n$ and $x' = \text{“lim” } b_n$ for some Cauchy sequences (a_n) and (b_n) which, by Proposition 5.10 we can take to be bounded away from zero. Consider the product P defined by

$$P := x^{-1}x'(x')^{-1} = \text{“lim” } a_n^{-1} \text{“lim” } a_n \text{“lim” } b_n^{-1}.$$

By the definition of multiplication, we have $P = \text{“lim” } (a_n^{-1}a_nb_n^{-1}) = \text{“lim” } b_n^{-1} = (x')^{-1}$. On the other hand, by (2) we have

$$P = x^{-1}x'(x')^{-1} = \text{“lim” } a_n^{-1} \text{“lim” } b_n \text{“lim” } b_n^{-1} = \text{“lim” } (a_n^{-1}b_nb_n^{-1}) = \text{“lim” } a_n^{-1} = x^{-1}.$$

Comparing the two expressions for P , we find that $x^{-1} = (x')^{-1}$. □

We now define division of real numbers via $x/y = x \cdot y^{-1}$. One can show without too much difficulty (and we have almost completely done it already) that the real numbers satisfy all of the field properties on page 7. It remains to define an ordering on the reals as well as the absolute value.

Definition 5.15. A nonzero real number x is *positive* (and we write $x > 0$) if it can be represented by a Cauchy sequence (a_i) with $a_i > 0$ for all i . Similarly, x is *negative* (and we write $x < 0$) if it can be represented by a Cauchy sequence (b_i) with $b_i < 0$ for all i .

Lemma 5.16. *Every real number is either positive, negative, or zero.*

Proof. Let $x \in \mathbb{R}$ and suppose that $x \neq 0$. Then we need to prove that $x > 0$ or $x < 0$. Since $x \neq 0$, Proposition 5.10 implies that there is a Cauchy sequence (a_n) which is bounded away from zero such that $x = \lim a_n$. This means that there is a number $m > 0$ such that $|a_n| \geq m$ for all n . I claim that there is a natural number N such that a_n has the same sign (i.e. a_n is always positive or always negative) for all $n \geq N$.

Indeed, let $\varepsilon = m$. Then there exists a number N such that

$$|a_m - a_n| < m \quad \text{whenever} \quad m, n \geq N. \quad (5.2)$$

Suppose that a_m and a_n have different sign. Without loss of generality, say $a_m > 0$ and $a_n < 0$ (if not, we can just swap the indices). Then, since (a_n) is bounded,

$$|a_m - a_n| = |a_m + |a_n|| = a_m + |a_n| \geq m + m = 2m > m.$$

But this contradicts (5.2). Thus a_m and a_n have the same sign (both are positive or both are negative).

So all of the terms a_n with $n \geq N$ have the same sign. In order to ensure that every term a_n with $n \geq 1$ has the same sign, we can change the beginning of the sequence (since changing finitely many terms at the beginning of a sequence yields an equivalent sequence) to have this property. For instance, if $a_N > 0$ we can define a new sequence (b_n) with $b_n = |a_n|$. This new sequence is equivalent to a_n and has the property that $b_n > m$ for all n . $\square$

We can now define absolute value in the same way we did before:

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

Note that this definition *extends* the original one in the sense that if we think of $r \in \mathbb{Q}$ and $x = (r, r, r, \dots) \in \mathbb{R}$ we have $|x| = (|r|, |r|, |r|, \dots)$. In fact, this is true for every Cauchy sequence of rationals.

Lemma 5.17. *If $x = \lim a_n \in \mathbb{R}$ then $|x| = \lim |a_n|$.*

Proof. Exercise. $\square$

We also have all of the familiar ordering properties that we had for the rational numbers. Unfortunately we don't have time to prove each one of these, so you should check them on your own.

Proposition 5.18. *The ordering on the real numbers satisfies all of the properties that the ordering on the rationals satisfies.*

The following is another important property which shows how $\mathbb{Q}$ sits inside $\mathbb{R}$.

Proposition 5.19. *$\mathbb{Q}$ is dense in $\mathbb{R}$; i.e. for every pair of real numbers x, y with $x < y$ there exists a rational number q such that*

$$x < q < y.$$

Proof. Exercise. $\square$

6. SEQUENCES OF REAL NUMBERS AND LIMITS

To sum up the previous section, $\mathbb{R}$ is an ordered field which contains $\mathbb{Q}$ as a dense subset. Recall that what we really want out of $\mathbb{R}$ is a number system which has no holes (e.g. we want $\sqrt{2} \in \mathbb{R}$). But so far all we've done is constructed something that may not even be better than $\mathbb{Q}$. The purpose of this section is to define limits of sequences of real numbers and to show that $\mathbb{R}$ is *complete*; i.e.

every Cauchy sequence of real numbers converges to a real number.

The following definition should look familiar.

Definition 6.1. Let (a_i) be a sequence of real numbers. We say that (a_i) is a Cauchy sequence if for every $\varepsilon > 0$ there exists a natural number N for which

$$|a_m - a_n| < \varepsilon \quad \text{whenever} \quad m, n \geq N.$$

The concept of convergence is quite similar.

Definition 6.2. Let (a_i) be a sequence of real numbers. We say that (a_i) converges to $L \in \mathbb{R}$ (and we write $\lim_{n \rightarrow \infty} a_n = L$) if for every $\varepsilon > 0$ there exists a natural number N for which

$$|a_n - L| < \varepsilon \quad \text{for all} \quad n \geq N.$$

It should come as no surprise that these notions are related.

Proposition 6.3. *Let (a_i) be a sequence of real numbers. If (a_i) converges to L then (a_i) is a Cauchy sequence.*

Proof. Let $\varepsilon > 0$. Since (a_i) converges to L there is a number N such that

$$|a_n - L| < \frac{\varepsilon}{2} \quad \text{for all} \quad n \geq N.$$

It follows that if $m, n \geq N$ we have (by the triangle inequality)

$$|a_n - a_m| = |a_n - L - (a_m - L)| \leq |a_n - L| + |a_m - L| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus (a_i) is Cauchy. □

Perhaps more interesting is the converse of the previous proposition, namely that every Cauchy sequence converges. This implies that there are no analogues of $\sqrt{2}$ for the real numbers; i.e. there are no “holes” in $\mathbb{R}$. To prove this, we first show that every Cauchy sequence of rationals converges to a real number (the following statement might look completely trivial, but this is likely an artifact of our choice of notation; now you can see why we are using “lim” a_n to denote a real number).

Proposition 6.4. *If (a_i) is a Cauchy sequence of rational numbers then $\lim_{n \rightarrow \infty} a_n = \text{“lim” } a_n$.*

Proof. Let $L = \text{“lim” } a_k$ (we will use k here so as to avoid confusion with $\lim_{n \rightarrow \infty} a_n$) and let $\varepsilon > 0$ be given. Since (a_k) is Cauchy, there is a number $N \geq 1$ such that

$$|a_k - a_\ell| < \varepsilon$$

for all $k, \ell \geq N$.

To prove that $\lim_{n \rightarrow \infty} a_n = L$ we need to show that

$$|a_n - L| < \varepsilon$$

for sufficiently large n (say $n \geq N$). Let's think about what this means: the real number $a_n - L$ is the difference of the rational number a_n and the real number L . To add/subtract real numbers, we need to represent each one by a Cauchy sequence. This is already done for L , and for a_n we can use the constant sequence $(a_n, a_n, a_n, \dots)$. Thus

$$\begin{aligned} |a_n - L| &= |\text{"lim"}(a_n, a_n, a_n, \dots) - \text{"lim"}(a_1, a_2, a_3, \dots, a_k, \dots)| \\ &= |\text{"lim"}(a_n - a_1, a_n - a_2, \dots, a_n - a_k, \dots)| \\ &= \text{"lim"}(|a_n - a_1|, |a_n - a_2|, \dots, |a_n - a_k|, \dots), \end{aligned}$$

where we have used Lemma 5.17 in the last step. For all $k \geq N$ we have

$$|a_n - a_k| < \varepsilon,$$

so the real number $|a_n - L|$ is less than ε (think carefully about why this is). Therefore $\lim_{n \rightarrow \infty} a_n = L$. $\square$

We can now prove that $\mathbb{R}$ is complete.

Theorem 6.5. *$\mathbb{R}$ is complete. Equivalently, every Cauchy sequence of real numbers converges to a real number.*

Proof. Suppose that (x_n) is a Cauchy sequence of real numbers. Let $\varepsilon > 0$ be given. For each n , choose a rational number q_n such that

$$|x_n - q_n| < \frac{\varepsilon}{3}.$$

This is possible because $\mathbb{Q}$ is dense in $\mathbb{R}$. Then the sequence (q_n) is Cauchy because

$$|q_m - q_n| = |q_m - x_m + x_m - x_n + x_n - q_n| \leq |q_m - x_m| + |x_m - x_n| + |x_n - q_n|$$

and for m, n sufficiently large this quantity is less than ε because we can make $|x_m - x_n| < \varepsilon/3$.

Let $L = \text{"lim"} q_n$. We will show that $\lim_{n \rightarrow \infty} x_n = L$. Let $\varepsilon > 0$ be given. By the previous proposition, (q_n) converges to L , so there is some $N \geq 1$ such that

$$|q_n - L| < \frac{\varepsilon}{2}$$

for all $n \geq N$. It follows that

$$|x_n - L| = |x_n - q_n + q_n - L| \leq |x_n - q_n| + |q_n - L| < \frac{\varepsilon}{3} + \frac{\varepsilon}{2} < \varepsilon$$

for all $n \geq N$, as desired. $\square$

We conclude this section by listing some additional properties of convergent sequences. First, since convergent sequences are Cauchy, we can extend the properties of Cauchy sequences of rational numbers to convergent sequences of real numbers (via essentially the same proofs). In particular, we have the following.

Proposition 6.6. *Every convergent sequence of real numbers is bounded.*

Is the converse true? That is, is every bounded sequence convergent? Here is a quick counterexample: consider the sequence $1, 0, 1, 0, 1, 0, 1, 0, \dots$. This sequence is bounded because $|a_n| \leq 1$ for all n , but it is clearly not convergent. However, it is the case that bounded *monotonic* sequences are convergent. There are two versions of this: increasing sequences which are bounded above are convergent, and decreasing sequences which are bounded below are convergent. We will state and prove the first version in the next section.

We can also do arithmetic with limits. One way to prove the following properties is to replace every convergent sequence of reals with a convergent sequence of rationals that has the same limit (see the proof of the previous theorem). Then arithmetic of limits is the same as arithmetic of real numbers of the form “ $\lim$ ” a_n , which we have already covered.

Proposition 6.7. *Suppose that (a_n) and (b_n) are convergent sequences of real numbers. Then the following are true.*

$$(1) \lim_{n \rightarrow \infty} (a_n + b_n) = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n$$

$$(2) \lim_{n \rightarrow \infty} (a_n b_n) = \lim_{n \rightarrow \infty} a_n \lim_{n \rightarrow \infty} b_n$$

$$(3) \text{ if } c \in \mathbb{R} \text{ then } \lim_{n \rightarrow \infty} c a_n = c \lim_{n \rightarrow \infty} a_n$$

$$(4) \text{ if } b_n \neq 0 \text{ for all } n \text{ and if } \lim_{n \rightarrow \infty} b_n \neq 0 \text{ then } \lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{\lim_{n \rightarrow \infty} a_n}{\lim_{n \rightarrow \infty} b_n}$$

The following proposition gives a list of basic limits which will come in handy later on. See Theorem 9.7 in your book for a proof.

Proposition 6.8. *The following are true.*

$$(1) \text{ If } p > 0 \text{ then } \lim_{n \rightarrow \infty} \frac{1}{n^p} = 0.$$

$$(2) \text{ If } |a| < 1 \text{ then } \lim_{n \rightarrow \infty} a^n = 0. \text{ If } a = 1 \text{ then } \lim_{n \rightarrow \infty} a^n = 1. \text{ If } a = -1 \text{ or if } |a| \geq 1 \text{ then the sequence } (a^n) \text{ does not converge.}$$

$$(3) \text{ If } a > 0 \text{ then } \lim_{n \rightarrow \infty} a^{1/n} = 1.$$

Finally, the ordering on the reals plays nicely with limits, too. We will frequently use this property without stating it.

Lemma 6.9. *Suppose that (a_n) is a convergent sequence. If $a_n \leq M$ for all $n \geq 1$ then*

$$\lim_{n \rightarrow \infty} a_n \leq M.$$

Similarly, if $a_n \geq m$ for all $n \geq 1$ then

$$\lim_{n \rightarrow \infty} a_n \geq m.$$

Proof. Exercise. □

7. THE LEAST UPPER BOUND PROPERTY

We constructed the real numbers by filling in the holes in the rational numbers. But one other advantage to the reals over the rationals is the existence of a *least upper bound* for any subset of the real numbers. (Your textbook [2] treats this property as an axiom; see §4, p. 23. We will show that it follows from the completeness of the real numbers.)

Definition 7.1. Suppose that $E \subseteq \mathbb{R}$ is a set of real numbers and let $M, m \in \mathbb{R}$. We say that M is an *upper bound* for E if $x \leq M$ for every $x \in E$. Similarly, m is a *lower bound* for E if $x \geq m$ for every $x \in E$.

Consider the interval $E = [0, 1] = \{x \in \mathbb{R} : 0 \leq x \leq 1\}$. This set has many upper bounds; for example, 1, 2, and 496 are upper bounds. Also $\frac{172893}{172892}$ is an upper bound. It also has many lower bounds; for example, 0 and -5 . But if I asked you to name an upper bound and a lower bound for E , you would likely say 1 and 0, respectively. This is because these are the least upper bound and greatest lower bound (and so they tell you the most information about the set E).

Definition 7.2. Suppose that $E \subseteq \mathbb{R}$ is a set of real numbers and let $M, m \in \mathbb{R}$. We say that M is a *least upper bound* for E if (a) M is an upper bound for E and (b) if M' is another upper bound for E , then $M' \geq M$. Similarly, we say that m is a *greatest lower bound* for E if (a) m is a lower bound for E and (b) if m' is another lower bound for E then $m' \leq m$.

It should come as no surprise that least upper bounds are unique.

Proposition 7.3. Suppose that M and M' are both least upper bounds for E . Then $M = M'$.

Proof. By the definition, since M is a least upper bound and M' is another upper bound, we must have $M' \geq M$. Reversing the roles of M and M' (i.e. by symmetry) we find that $M \geq M'$. Thus $M = M'$. $\square$

If E has a least upper bound then we write $\sup(E)$ for the *supremum*, or least upper bound, of E . Similarly, if E has a greatest lower bound then we write $\inf(E)$ for the *infimum*, or greatest lower bound.

One of the most important properties of the real numbers is the existence of a least upper bound for every subset of $\mathbb{R}$ that you would expect to have a least upper bound. That is, if a set E doesn't have an upper bound, then we shouldn't expect it to have a least upper bound. The proof of this theorem is somewhat involved, and I won't give all of the details; you should fill these in on your own.

Theorem 7.4. Let E be a non-empty subset of $\mathbb{R}$. If E has an upper bound, then it has a least upper bound.

Proof. Since E is non-empty, there is some element $a_1 \in E$, and since E has an upper bound, we can choose some upper bound b_1 . Clearly $a_1 \leq b_1$. We will define two sequences (a_i) and (b_i) inductively such that $a_n \leq b_n$ for all n .

For each $n \geq 1$, suppose that we have constructed a_n and b_n such that $a_n \leq b_n$. Let $K = \frac{a_n + b_n}{2}$ be the average of a_n and b_n . Note that $a_n \leq K \leq b_n$. If K is an upper bound for E , let $a_{n+1} = a_n$ and let $b_{n+1} = K = \frac{a_n + b_n}{2}$. Then $|b_{n+1} - a_{n+1}| = |\frac{a_n + b_n}{2} - a_n| = \frac{b_n - a_n}{2}$. If K is not an upper bound for E , then there is some element $x \in E$ such that $x > K$ (otherwise K would be an upper bound). In that case, let $a_{n+1} = x$ and let $b_{n+1} = b_n$. Then $|b_{n+1} - a_{n+1}| = b_n - x < b_n - K = \frac{b_n - a_n}{2}$. Note that in either case $|b_{n+1} - a_{n+1}| \leq \frac{b_n - a_n}{2}$. Also in either case a_n is increasing (or staying the same) and b_n is decreasing (or staying the same).

Since $a_n \in E$ for all n and b_n is an upper bound for E for all n , we have

$$a_1 \leq a_2 \leq a_3 \leq \cdots \leq b_3 \leq b_2 \leq b_1.$$

Using the inequality $|b_{n+1} - a_{n+1}| \leq \frac{b_n - a_n}{2}$ one can show by induction (check this!) that

$$|b_{n+1} - a_{n+1}| \leq \frac{b_1 - a_1}{2^n}. \quad (7.1)$$

We will show that the sequences (a_n) and (b_n) are Cauchy.

Let $\varepsilon > 0$ be given. Choose $N \geq 1$ such that $2^N > \frac{b_1 - a_1}{\varepsilon}$ (why is this possible?). Let $m, n \geq N$. Without loss of generality, assume that $m \geq n$ (otherwise just swap them). Then $a_m \geq a_n$, and we have

$$|a_m - a_n| = a_m - a_n \leq b_n - a_n \leq \frac{b_1 - a_1}{2^n} < \varepsilon.$$

Thus (a_n) is Cauchy. A very similar argument shows that (b_n) is Cauchy (check this!).

Using (7.1) one can show that (a_n) and (b_n) are equivalent sequences (check this!). Therefore they have the same limit, say L . I claim that L is the least upper bound of E . On the one hand, if $x \in E$ then $x \leq b_n$ for all n since every b_n is an upper bound for E . Thus $x \leq L$. Since x was an arbitrary element of E , we conclude that L is an upper bound for E . On the other hand, if L' is any other upper bound, then $L' \geq a_n$ for all n since every a_n is an element of E . Thus $L' \geq L$. It follows that L is the least upper bound of E . $\square$

There is a version of this for lower bounds as well.

Corollary 7.5. *Let E be a non-empty subset of $\mathbb{R}$. If E has a lower bound, then it has a greatest lower bound.*

Proof. Exercise. $\square$

We can also talk about the supremum and infimum of a sequence. To each sequence (a_n) we associate the set of numbers $E = \{a_n : n \geq 1\}$, and we define $\sup a_n := \sup E$ and $\inf a_n := \inf E$ if those quantities exist.

For example, let $a_n = (-1)^n$; then (a_n) is the sequence $-1, 1, -1, 1, -1, 1, \dots$. The associated set is the two-element set $E = \{-1, 1\}$. This has least upper bound 1 and greatest lower bound -1 . Thus $\sup a_n = 1$ and $\inf a_n = -1$.

As another example, let $a_n = 1/n$. Then the sequence is $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots$ and the set is $E = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots\}$. The least upper bound of E is 1 and the greatest lower bound of E is 0. Note that 0 is not in the set, nor is it in the sequence, but it is the largest number less than or equal to every element of the set/sequence.

The existence of a least upper bound for bounded sets implies the existence of a least upper bound for bounded sequences. We will use this fact to prove the following proposition.

Proposition 7.6. *Bounded monotonic sequences are convergent. That is, if (a_n) is a sequence which satisfies $a_n \leq M$ for some $M \in \mathbb{R}$, and if a_n is increasing ($a_{n+1} \geq a_n$ for all n) then (a_n) converges.*

Proof. Suppose that (a_n) is an increasing sequence which is bounded above. Let E be the set of values $\{a_n : n \geq 1\}$. Then (a_n) has a least upper bound which we will call $s = \sup a_n$. We will show that $\lim_{n \rightarrow \infty} a_n = s$.

Let $\varepsilon > 0$ be given. Since s is the least upper bound, $s - \varepsilon$ is not an upper bound for E . So there exists some $N \geq 1$ such that $a_N > s - \varepsilon$. Since (a_n) is increasing, it follows that $a_n \geq a_N > s - \varepsilon$ for all $n \geq N$. Thus, for all $n \geq N$ we have

$$|a_n - s| = s - a_n \leq s - a_N < s - (s - \varepsilon) = \varepsilon.$$

Therefore (a_n) converges to s . $\square$

There is a version of this for decreasing sequences which are bounded below.

Proposition 7.7. *If (a_n) is a sequence which satisfies $a_n \geq m$ for some $m \in \mathbb{R}$, and if a_n is decreasing ($a_{n+1} \leq a_n$ for all n) then (a_n) converges.*

Proof. Exercise. □

8. SUBSEQUENCES AND THE BOLZANO-WEIERSTRASS THEOREM

Consider the sequence (a_n) defined by $a_n = (-1)^n$. The terms of this sequence are

$$-1, 1, -1, 1, -1, 1, -1, 1, \dots$$

It should be clear that this is not a convergent sequence. But it's kind of close to being convergent in the sense that it seems like the sequence wants to converge to two different limits: 1 and -1 . In this case we say that there is a *subsequence* (say, take all of the -1 terms) that converges.

Definition 8.1. A *subsequence* of a sequence (a_n) is a sequence of the form (a_{n_k}) , where n_k is a strictly increasing sequence of natural numbers.

In plain English, this says that a subsequence is a selection of some (maybe all) of the terms of (a_n) taken in order. So in our example above, say we take $n_1 = 2$, $n_2 = 3$, $n_3 = 5$, $n_4 = 6$, $n_5 = 8$, and $n_k = 2k$ for all $k \geq 6$. Then the terms of the subsequence (a_{n_k}) is

$$1, -1, -1, 1, 1, 1, 1, \dots$$

Since the terms are all 1 after the fourth term, this subsequence is convergent. See §11 of the textbook for more examples of subsequences. The purpose of this section is to prove the Bolzano-Weierstrass theorem, which states that every bounded sequence has a convergent subsequence. We will prove this in a few steps.

Note: we have not explicitly defined what it means for $\lim a_n = \pm\infty$, so if you are not already familiar with this idea, please see Definition 9.8 on p. 50 of your textbook.

Proposition 8.2. Suppose that (a_n) is a sequence.

(1) Let $t \in \mathbb{R}$. There is a subsequence of (a_n) converging to t if and only if the set

$$\{n \in \mathbb{N} : |a_n - t| < \varepsilon\} \tag{8.1}$$

is infinite for each $\varepsilon > 0$.

(2) If (a_n) is unbounded above then it has a subsequence with limit $+\infty$.

(3) If (a_n) is unbounded below then it has a subsequence with limit $-\infty$.

In each case, we can take the subsequence to be monotonic.

Proof. (1) $(\Rightarrow)$ Suppose that there is a subsequence (a_{n_k}) converging to t . Then for each $\varepsilon > 0$ there exists a number N such that

$$|a_{n_k} - t| < \varepsilon$$

for every $n_k \geq N$. Since N is a finite number, there are infinitely many $n_k \geq N$, so the set (8.1) is infinite.

$(\Leftarrow)$ Suppose that, for each $\varepsilon > 0$, the set (8.1) is infinite. There is one special case that we should consider first. It might be that $a_n = t$ for infinitely many t (then certainly (8.1) would be infinite for any $\varepsilon > 0$). But then we can just take the subsequence consisting of terms with $a_{n_k} = t$ and that's a monotonic convergent subsequence. It's more interesting when that's not the case.

So suppose that $a_n = t$ for only finitely many n . Then, for each $\varepsilon > 0$, the set

$$\{n \in \mathbb{N} : 0 < |a_n - t| < \varepsilon\}$$

is infinite. Let's unpack what this means:

$$0 < |a_n - t| < \varepsilon \quad \text{if and only if} \quad 0 < a_n - t < \varepsilon \quad \text{or} \quad -\varepsilon < a_n - t < 0.$$

The last two inequalities can be written

$$t < a_n < t + \varepsilon \quad \text{or} \quad t - \varepsilon < a_n < t.$$

Let

$$S_-(\varepsilon) := \{n \in \mathbb{N} : t - \varepsilon < a_n < t\} \quad \text{and} \quad S_+(\varepsilon) := \{n \in \mathbb{N} : t < a_n < t + \varepsilon\}.$$

Note that if $\varepsilon_2 \leq \varepsilon_1$ then $S_\pm(\varepsilon_2) \subseteq S_\pm(\varepsilon_1)$ (think about this carefully until you are convinced). It follows that one of the following two statements is true (both could be true):

$$S_+(\varepsilon) \quad \text{is infinite for every } \varepsilon > 0 \tag{8.2}$$

or

$$S_-(\varepsilon) \quad \text{is infinite for every } \varepsilon > 0. \tag{8.3}$$

Let us assume that (8.2) is true. The proof in the case that (8.3) holds is quite similar (and is the case your book covers). We will construct a sequence $n_1, n_2, n_3, n_4, \dots$ in a similar way as in the proof of Theorem 7.4. Since $S_+(1)$ is infinite, there exists some $n_1 \in S_+(1)$. For this n_1 we have

$$t < a_{n_1} < t + 1.$$

Now suppose we have chosen n_1 through n_k such that $n_1 < n_2 < n_3 < \dots < n_k$ and such that

$$t < a_{n_j} \leq \min\{t + \frac{1}{j}, a_{n_{j-1}}\} \quad \text{for } j = 1, 2, 3, \dots, k.$$

(We use the min so that the subsequence a_{n_k} is decreasing.) With $\varepsilon_1 = \frac{1}{k+1}$ we can use that $S_+(\varepsilon_1)$ is infinite to find the next index n_{k+1} . But we also want this minimum property to hold. So write $a_{n_k} = t + (a_{n_k} - t)$. Since $a_{n_k} > t$ the number $\varepsilon_2 = a_{n_k} - t$ is positive. So the set $S_+(\varepsilon_2)$ is infinite. Taking ε to be the smaller of $\varepsilon_1, \varepsilon_2$ we can find some number $n_{k+1} \in S_+(\varepsilon)$ such that $n_{k+1} > n_k$ and

$$t < a_{n_{k+1}} \leq \min\{t + \frac{1}{k+1}, a_{n_k}\}.$$

Thus we have constructed a subsequence (a_{n_k}) which is decreasing and bounded below (and therefore convergent).

The proofs of (2) and (3) are much easier; see the proof of Theorem 11.2 in your book if you are interested. $\square$

As an example of what we can do with this proposition, let's prove that every sequence of positive numbers with $\inf a_n = 0$ has a subsequence converging to 0. (Can you write down an example of a sequence such that $\inf a_n = 0$ but (a_n) is not convergent?) By the proposition, it suffices to show that the sets $\{n \in \mathbb{N} : a_n < \varepsilon\}$ are infinite for each $\varepsilon > 0$. Suppose, by way of contradiction, that there is some $\varepsilon_0 > 0$ such that $a_n < \varepsilon_0$ for only finitely many n . Then $\inf a_n = \min\{a_n : a_n < \varepsilon_0\}$ (think carefully about why this is). But each a_n is positive, so this means that $\inf a_n > 0$, a contradiction. So there must exist a subsequence (a_{n_k}) converging to 0.

Proposition 8.3. *If the sequence (a_n) converges, then every subsequence converges to the same limit.*

Proof. This is pretty straightforward. Let (a_{n_k}) be a subsequence of (a_n) . Note that $n_k \geq k$ for all $k \geq 1$ (convince yourself that this is true or look in Theorem 11.3 in your book for an inductive justification). Let $L = \lim a_n$. Then for each $\varepsilon > 0$ there is some $N \geq 1$ such that $|a_n - L| < \varepsilon$. Thus, for $k \geq N$ we have $n_k \geq k \geq N$, which implies that $|a_{n_k} - L| < \varepsilon$. Thus (a_{n_k}) converges to L . $\square$

We are almost ready to prove Bolzano-Weierstrass, but first we need one more intermediate result.

Proposition 8.4. *Every sequence has a monotonic subsequence.*

Proof. Let (a_n) be a sequence. For the purposes of this proof, let's say that a term a_n is *dominant* if it is larger than every term after it:

$$a_n > a_m \quad \text{for all } m > n.$$

If there are infinitely many dominant terms, let (a_{n_k}) be the subsequence consisting of all the dominant terms. Then (a_{n_k}) is decreasing.

Now suppose that there are only finitely many dominant terms. Then we can choose n_1 such that a_{n_1} is after every dominant term. Then for every $k \geq 2$, choose n_k such that $n_k > n_{k-1}$ and $a_{n_k} \geq a_{n_{k-1}}$. Why can we always do this? Because if we couldn't, then there would be another dominant term; but we started after all the dominant terms. So we have constructed an increasing subsequence (a_{n_k}) . $\square$

Our main theorem now follows quickly.

Theorem 8.5 (Bolzano-Weierstrass). *Every bounded sequence has a convergent subsequence.*

Proof. Let (a_n) be a bounded sequence. By Proposition 8.4 there is a monotonic subsequence (a_{n_k}) . This subsequence is clearly also bounded, so it is convergent. $\square$

As an example, consider the sequence (a_n) defined by $a_n = \cos(\frac{\pi n}{3})$. The terms of this sequence are

$$\frac{1}{2}, -\frac{1}{2}, -1, -\frac{1}{2}, \frac{1}{2}, 1, \dots$$

This is a bounded sequence, so it contains a convergent subsequence. For example, we can take the constant subsequences consisting of all terms of the form $\frac{1}{2}$, or $-\frac{1}{2}$, or -1 , or 1 . Actually, this example illustrates an important idea: the values $\frac{1}{2}, -\frac{1}{2}, 1, -1$ are called the *limit points* of the sequence (a_n) . We will discuss this more in the next section.

9. lim sup AND lim inf

Definition 9.1. Let (a_n) be a sequence. A *limit point* of (a_n) is any real number (or symbol $\pm\infty$) that is the limit of some subsequence (a_{n_k}) . Sometimes limit points are referred to as *cluster points* or *accumulation points* or *condensation points*, and your book calls them *subsequential limits*. I think that the term *limit point* is used most throughout mathematics, so we will use it here.

If (a_n) is a convergent sequence, then since all subsequences converge to the same limit (say, L), the set of all limit points of (a_n) is just $\{L\}$. There are several examples of limit points on p. 73 of your book.

There are two particularly important limit points for any sequence, called the $\limsup$ and the $\liminf$. These are defined as

$$\begin{aligned}\limsup a_n &= \lim_{N \rightarrow \infty} \sup\{a_n : n > N\}, \\ \liminf a_n &= \lim_{N \rightarrow \infty} \inf\{a_n : n > N\}.\end{aligned}$$

We allow $\limsup a_n = +\infty$ and $\liminf a_n = -\infty$. Note that $\limsup a_n$ is not necessarily the supremum of the set $\{a_n : n \in \mathbb{N}\}$, but it is the largest value that *infinitely many* of the terms a_n get close to. The following proposition shows that for convergent sequences, the $\liminf$ and the $\limsup$ are equal. We will not prove it here, but a proof is given in Theorem 10.7 of your book.

Proposition 9.2. *If (a_n) is a convergent sequence then*

$$\liminf a_n = \lim_{n \rightarrow \infty} a_n = \limsup a_n.$$

Now let's see how limit points and $\limsup$ / $\liminf$ are related.

Proposition 9.3. *Let (a_n) be a sequence. There exists a monotonic subsequence whose limit is $\limsup a_n$ and a monotonic sequence whose limit is $\liminf a_n$.*

Proof. If (a_n) is unbounded above or below, then we just apply Proposition 8.2 part (2) or (3). So suppose that (a_n) is bounded either above or below (actually let's just pick one since the other case is similar). Suppose that (a_n) is bounded above, then $t = \limsup a_n$ is a finite number. Let $\varepsilon > 0$. Then there exists $N \geq 1$ such that (by definition of $\limsup$)

$$|\sup\{a_n : n > N\} - t| < \varepsilon.$$

But the $\sup$ is larger than or equal to t , so we can say that

$$\sup\{a_n : n > N\} < t + \varepsilon.$$

In particular, $a_n < t + \varepsilon$ for all $n > N$ (since $a_n \leq \sup\{a_n : n > N\}$). I claim that the set

$$\{n \in \mathbb{N} : |a_n - t| < \varepsilon\} \tag{9.1}$$

is infinite. If not, then for some N' we have $a_n \leq t - \varepsilon$ for all $n \geq N'$. But this means that $\limsup a_n < t$, a contradiction. So the set (9.1) is infinite; by Proposition 8.2 there is a monotonic subsequence of (a_n) converging to $t = \limsup a_n$. $\square$

As an immediate consequence of the previous proposition, we see that the set of limit points of any sequence (a_n) contains at least one element (both the $\limsup$ and $\liminf$ are in the set, but they might be equal). Not only that, but it turns out that $\limsup a_n$ is the supremum of the set of limit points (and similarly with the $\liminf$). See Theorem 11.8 of your book for a proof of this last fact. Also see p. 75 for some examples.

There are more properties of $\limsup$ and $\liminf$ in §12 of your book, but we will not need them here (and if we need them later, we will come back and prove them).

10. SERIES

When you saw sequences and series in your Calculus courses previously, you likely spent very little time talking about sequences and a lot of time talking about series. Maybe, because of this, you think of series as being “more complicated” than sequences. However, since we have taken such a long time carefully building up the theory of sequences and limits,

it turns out that we can prove many results about series very quickly. But first we need some rigorous definitions.

Definition 10.1. An infinite series is an object of the form

$$\sum_{k=1}^{\infty} a_k \quad (10.1)$$

where (a_k) is a sequence of real numbers. To give a precise meaning to (10.1), for each $n \geq 1$ let

$$s_n = \sum_{k=1}^n a_k = a_1 + a_2 + \dots + a_n$$

be the n th partial sum. We say that (10.1) *converges* if the sequence (s_n) is a convergent sequence. In this case, if $S = \lim s_n$, we write

$$\sum_{k=1}^{\infty} a_k = S.$$

If the sequence of partial sums is not convergent, then we say the series *diverges*. Sometimes we can be more precise about how it diverges by saying that it *diverges to $+\infty$* if $\lim s_n = +\infty$ and that it *diverges to $-\infty$* if $\lim s_n = -\infty$. Lastly, if $\sum |a_n|$ is convergent, we say the series $\sum a_n$ is *absolutely convergent*.

Note: above I have written the series as starting with the index $k = 1$. In fact we can start at any index $k = m$ and modify the definition appropriately by defining $s_n = a_m + a_{m+1} + \dots + a_n$. Often it will not matter where the index begins, and so I will just write $\sum a_n$ for a “generic” series.

By transferring what we know about sequences to series, we can prove some very useful results without too much effort. We’ll start by using the fact that a sequence converges if and only if it is Cauchy. We say that a series satisfies the *Cauchy criterion* if its sequence of partial sums is Cauchy, and we immediately get the following proposition.

Proposition 10.2 (Cauchy criterion). *A series $\sum a_n$ is convergent if and only if for each $\varepsilon > 0$ there exists a number $N \geq 1$ such that*

$$\left| \sum_{k=m}^n a_k \right| < \varepsilon$$

whenever $n \geq m \geq N$.

From this we (almost) immediately get the following familiar corollary.

Corollary 10.3. *If $\sum a_n$ is convergent then $\lim_{n \rightarrow \infty} a_n = 0$.*

Proof. Apply the Cauchy criterion with $m = n$ to get: for every $\varepsilon > 0$ there exists a number $N \geq 1$ such that $|a_n| < \varepsilon$ whenever $n \geq N$. $\square$

You may be more familiar with the previous corollary under the name “the divergence test.” Indeed, one (usually easy) way to check if a series diverges is to check whether its terms go to zero or not. If not, then the series diverges. Of course, even if the terms do go to zero, that’s not a guarantee that the series converges.

We can also do arithmetic on series (addition and multiplication by a constant) by applying the limit laws for sequences: if $\sum a_n$ and $\sum b_n$ are both convergent then

$$\sum (a_n + b_n) = \sum a_n + \sum b_n$$

and if $\sum a_n$ is convergent and c is any constant then

$$\sum ca_n = c \sum a_n.$$

In order to determine the convergence of a given series, it is often useful to compare it to another “simpler” series which you already know converges or diverges. Two such “simple” series are the *geometric* series and the *p*-series.

A geometric series is a series of the form $\sum ar^n$ for some constant a and some constant ratio r . Since we can always pull out the constant a , we will only need to think about geometric series of the form $\sum r^n$.

Proposition 10.4. *If $|r| < 1$ then the geometric series $\sum r^n$ is absolutely convergent; furthermore,*

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}. \quad (10.2)$$

If $|r| \geq 1$ then $\sum r^n$ diverges.

Proof. It is not too hard to show by induction (try it!) that for any $r \neq 1$ we have

$$\sum_{n=0}^m r^n = \frac{1-r^{m+1}}{1-r}.$$

(If $r = 1$ then it is clear that the series diverges.) By the limit laws, it follows that the series converges if and only if $\lim_{m \rightarrow \infty} r^{m+1}$ is finite, which is true if and only if $|r| < 1$. In that case, we have $\lim_{m \rightarrow \infty} r^{m+1} = 0$, which proves equation (10.2). Finally, if $|r| < 1$ then $\sum |r^n| = \sum |r|^n$ is convergent, so $\sum r^n$ is absolutely convergent. $\square$

A *p*-series is a series of the form $\sum \frac{1}{n^p}$ for some fixed p . We already showed that for $p = 1$, this series is divergent (this was way back when we first learned about Cauchy sequences; see Proposition 5.4). Since this case is so important and is used as a counterexample to many theorems in analysis, let's give it its own proposition.

Proposition 10.5. *The harmonic series $\sum \frac{1}{n}$ is divergent.*

What happens for other values of p ? It turns out that the harmonic series is the dividing line between convergent *p*-series and divergent *p*-series, as the next proposition shows.

Proposition 10.6. *If $p \leq 1$ then the *p*-series $\sum \frac{1}{n^p}$ diverges. If $p > 1$ then it converges.*

This is usually proved using the integral test, but since we haven't covered integrals yet, we're going to use a different method. We will, however, need to prove the comparison test first.

Proposition 10.7 (The comparison test). *Let $\sum a_n$ be a series of nonnegative terms ($a_n \geq 0$ for all n).*

- (1) *If $\sum a_n$ converges and $|b_n| \leq a_n$ for all n then $\sum b_n$ converges.*
- (2) *If $\sum a_n = +\infty$ and $b_n \geq a_n$ for all n then $\sum b_n$ diverges.*

Proof. (1) Suppose that $\sum a_n$ converges. Then it satisfies the Cauchy criterion. Given $\varepsilon > 0$ there exists a number $N \geq 1$ such that

$$\sum_{k=n}^m a_k = \left| \sum_{k=n}^m a_k \right| < \varepsilon$$

for all $n \geq m \geq N$ (note $a_k \geq 0$ for all k). So for $m \geq n \geq N$ we have

$$\left| \sum_{k=n}^m b_k \right| \leq \sum_{k=n}^m |b_k| \leq \sum_{k=n}^m a_k < \varepsilon,$$

where we used the Triangle inequality in the first step. Thus $\sum b_n$ satisfies the Cauchy criterion, and therefore it converges.

(2) Let (s_n) be the sequence of partial sums for $\sum a_n$ and let (t_n) be the sequence of partial sums for $\sum b_n$. Then $t_n \geq s_n$ for all n . Since $\lim s_n = +\infty$ we have $\lim t_n = +\infty$ too. $\square$

Before we prove Proposition 10.6, I should mention a quick corollary of the comparison test which you are likely familiar with.

Corollary 10.8. *Absolutely convergent series are convergent.*

Proof. Suppose that $\sum b_n$ is absolutely convergent. Then $\sum |b_n|$ converges. Since (trivially) $|b_n| \leq |b_n|$, the comparison test says that $\sum b_n$ converges. $\square$

Now we can prove Proposition 10.6.

Proof of Proposition 10.6. Suppose first that $p \leq 1$. Then $\frac{1}{n^p} \geq \frac{1}{n}$ for all $n \geq 1$. By the comparison test, $\sum \frac{1}{n^p}$ diverges.

Now suppose that $p > 1$. Let $(s_n(p))$ denote the sequence of partial sums, i.e.

$$s_n(p) = \sum_{k=1}^n \frac{1}{k^p} = 1 + \frac{1}{2^p} + \frac{1}{3^p} + \cdots + \frac{1}{n^p}.$$

Consider the $2n$ -th partial sum and split into even and odd terms:

$$s_{2n}(p) = 1 + \left(\frac{1}{2^p} + \frac{1}{4^p} + \frac{1}{6^p} + \cdots + \frac{1}{(2n)^p} \right) + \left(\frac{1}{3^p} + \frac{1}{5^p} + \frac{1}{7^p} + \cdots + \frac{1}{(2n-1)^p} \right).$$

Since $3^p > 2^p$ and $5^p > 4^p$ and $7^p > 6^p$, etc. we have

$$\begin{aligned} s_{2n}(p) &< 1 + 2 \left(\frac{1}{2^p} + \frac{1}{4^p} + \frac{1}{6^p} + \cdots + \frac{1}{(2n)^p} \right) \\ &= 1 + \frac{2}{2^p} \left(\frac{1}{1^p} + \frac{1}{2^p} + \frac{1}{3^p} + \cdots + \frac{1}{n^p} \right) \\ &= 1 + \frac{2}{2^p} s_n(p). \end{aligned}$$

Since $(s_n(p))$ is an increasing sequence (we are adding positive terms every time) we have

$$s_n(p) < 1 + \frac{2}{2^p} s_n(p).$$

Solving this inequality for $s_n(p)$ yields

$$s_n(p) < \left(1 - \frac{2}{2^p} \right)^{-1}.$$

Thus $(s_n(p))$ is an increasing sequence which is bounded above. Therefore it converges. $\square$

We will only talk about one more convergence test before we finish our discussion of series. You probably remember several other convergence tests from previous calculus courses; these are certainly interesting, but won't help us much in our ultimate goal for this course. If you are interested in learning more about infinite series, I suggest you take a course in complex analysis or analytic number theory.

Proposition 10.9 (Alternating series test). *If (a_n) is a decreasing sequence $(a_{n+1} \leq a_n)$ of nonnegative real numbers $(a_n \geq 0)$ and $\lim a_n = 0$ then the alternating series $\sum (-1)^{n+1} a_n$ converges.*

Proof. Let (s_n) denote the sequence of partial sums. Let (e_n) be the subsequence of even-indexed terms and (o_n) the subsequence of odd-indexed terms. Then $o_n = s_{2n-1}$ and $e_n = s_{2n}$:

$$(s_n) = o_1, e_1, o_2, e_2, o_3, e_3, \dots$$

The even-indexed terms are increasing since

$$e_{n+1} - e_n = s_{2n+2} - s_{2n} = -a_{2n+2} + a_{2n+1} \geq 0.$$

Similarly, the odd-indexed terms are decreasing since

$$o_{n+1} - o_n = s_{2n+1} - s_{2n-1} = a_{2n+1} - a_{2n} \leq 0.$$

I claim that every even-indexed term is less than or equal to every odd-indexed term:

$$e_m \leq o_n \quad \text{for all } m, n \in \mathbb{N}.$$

First, $e_n \leq o_{n+1} \leq o_n$ because $o_{n+1} - e_n = s_{2n+1} - s_{2n} = a_{2n+1} \geq 0$ and (o_n) is decreasing. Now, if $m \leq n$, then because (e_n) is increasing, $e_m \leq e_n \leq o_n$. On the other hand, if $m \geq n$ then $o_n \geq o_m \geq e_m$. To summarize,

$$e_1 \leq e_2 \leq e_3 \leq \dots \leq o_3 \leq o_2 \leq o_1.$$

Since (e_n) is increasing and bounded above, it converges, say to L_e . Similarly, (o_n) converges, say to L_o . But

$$L_o - L_e = \lim_{n \rightarrow \infty} o_{n+1} - \lim_{n \rightarrow \infty} e_n = \lim_{n \rightarrow \infty} (o_{n+1} - e_n) = \lim_{n \rightarrow \infty} a_{2n+1} = 0.$$

Thus $L_o = L_e$. It follows that (s_n) converges. $\square$

As an example, the alternating harmonic series

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$$

converges since $(\frac{1}{n})$ is a decreasing sequence and $\lim \frac{1}{n} = 0$.

11. CONTINUOUS FUNCTIONS

We are finally ready to begin what most of you probably think of as *calculus*. That is, we are going to discuss continuity, limits, derivatives, etc. To do this we will need to have a precise definition of what a *function* is. (This definition will be precise enough for our purposes, but to be totally rigorous we should probably talk about Cartesian products of sets; we will not do this here.)

Definition 11.1. A function is a pair (X, f) , where $X \subseteq \mathbb{R}$ is the *domain* and f is a rule which specifies the value $f(x)$ for every $x \in X$. We will usually write our functions using the notation $f : X \rightarrow Y$. This means that to each $x \in X$, f assigns a value $f(x)$ which is an element of the *codomain* Y . Note: f must make an assignment for every $x \in X$, but it need not map to every element of Y .

For example, the function $f : \mathbb{R} \rightarrow \mathbb{R}$ given by $f(x) = x^2$ assigns a value to $f(x)$ for every $x \in \mathbb{R}$ (the domain) but does not hit every value of the codomain $\mathbb{R}$ (just the nonnegative reals). The function $g : \mathbb{N} \rightarrow \mathbb{R}$ given by $g(x) = x^2$ is technically a different function (the domain is not $\mathbb{R}$).

Sometimes, we will just write a function as a rule, e.g. $f(x) = 1/x$. In this case, the domain is implicitly understood to be the largest subset of the real numbers for which the assignment is a well-defined real number (in this case the domain is $X = \{x \in \mathbb{R} : x \neq 0\}$). Another example is $g(x) = \sqrt{x}$ in which case the domain is $X = \{x \in \mathbb{R} : x \geq 0\}$.

Definition 11.2. Let $X \subseteq \mathbb{R}$. A function $f : X \rightarrow \mathbb{R}$ is *continuous at the point* $a \in X$ if, for every sequence (a_n) (with $a_n \in X$) converging to a , the sequence $(f(a_n))$ converges to $f(a)$. Put another way, f is continuous at a if

$$\lim_{n \rightarrow \infty} f(a_n) = f\left(\lim_{n \rightarrow \infty} a_n\right)$$

for every sequence (a_n) converges to a .

This definition essentially says that whenever x is close to a , $f(x)$ is close to $f(a)$. There is another way to say this which looks more epsilon-y.

Definition 11.3. Let $X \subseteq \mathbb{R}$. A function $f : X \rightarrow \mathbb{R}$ is *continuous at the point* $a \in X$ if, for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - f(a)| < \varepsilon \quad \text{whenever} \quad x \in X \text{ and } |x - a| < \delta.$$

Is it alright for a mathematical concept to have two definitions? Usually the answer is no; in math, we like to make sure that everything is perfectly well-defined. In this case, since there are two definitions for what it means for a function to be continuous, it's important that we show that they give us the same thing, i.e., that every function which satisfies Definition 11.2 also satisfies Definition 11.3, and vice versa.

Proposition 11.4. *Definitions 11.2 and 11.3 are equivalent.*

Proof. Suppose that $f : X \rightarrow \mathbb{R}$ satisfies Definition 11.3 at the point $a \in X$. Let (a_n) be a sequence in X converging to a . Let $\varepsilon > 0$ be given. Then there exists a $\delta > 0$ such that

$$|f(x) - f(a)| < \varepsilon \quad \text{whenever} \quad x \in X \text{ and } |x - a| < \delta.$$

Since (a_n) converges to a , there is a number $N \geq 1$ such that

$$|a_n - a| < \delta$$

whenever $n \geq N$. Thus, $n \geq N$ we have $|f(a_n) - f(a)| < \varepsilon$, so $(f(a_n))$ converges to $f(a)$.

We argue the other direction by contrapositive. Suppose that Definition 11.3 fails for $f : X \rightarrow \mathbb{R}$ at the point $a \in X$. Then there exists $\varepsilon > 0$ such that for every $\delta > 0$, there exists a point $x \in X$ such that

$$|x - a| < \delta \quad \text{but} \quad |f(x) - f(a)| \geq \varepsilon. \quad (11.1)$$

We will use this to construct a sequence (x_n) which converges to a but for which $(f(x_n))$ does not converge to $f(a)$. For each $n \geq 1$ we apply (11.1) with $\delta = \frac{1}{n}$. Thus, for each $n \geq 1$, there exists a point $x_n \in X$ such that

$$|x_n - a| < \frac{1}{n} \quad \text{but} \quad |f(x_n) - f(a)| \geq \varepsilon.$$

The sequence (x_n) clearly converges to a but the inequality above shows that $(f(x_n))$ does not converge to $f(a)$. $\square$

We say that a function $f : X \rightarrow \mathbb{R}$ is *continuous on* $A \subseteq \mathbb{R}$ if f is continuous at every point $a \in A$. Similarly, we say that $f : X \rightarrow \mathbb{R}$ is *continuous* if f is continuous at a for every $a \in X$.

As an example, let's prove that $f(x) = \sqrt{x}$ is continuous on its domain $X = [0, \infty)$. We can use either definition. Using the sequential definition: given $a \in X$, let (a_n) be a sequence in X converging to a . Then

$$\lim_{n \rightarrow \infty} f(a_n) = \lim_{n \rightarrow \infty} \sqrt{a_n} = \sqrt{\lim_{n \rightarrow \infty} a_n} = \sqrt{a}.$$

Here we're using the fact that $\lim \sqrt{a_n} = \sqrt{\lim a_n}$ which we never actually proved (though you did use it on that one homework problem). Let's prove that $\sqrt{x}$ is continuous using the ε - δ definition.

Let $a \in X$. Given $\varepsilon > 0$ let $\delta = \varepsilon\sqrt{a}$ (I chose this after looking at the inequalities below, then I came back and wrote it in). Then for every $x \in X$ such that $|x - a| < \delta$ we have

$$|f(x) - f(a)| = |\sqrt{x} - \sqrt{a}| = \left| \frac{(\sqrt{x} - \sqrt{a})(\sqrt{x} + \sqrt{a})}{\sqrt{x} + \sqrt{a}} \right| = \frac{|x - a|}{\sqrt{x} + \sqrt{a}} < \frac{\varepsilon\sqrt{a}}{\sqrt{x} + \sqrt{a}}.$$

Since $\sqrt{x} \geq 0$ the latter expression is $< \varepsilon$. Thus $f(x)$ is continuous at $x = a$. Two things to note: first, this choice of δ depends on a (that's fine right now, but later it's going to be important). Second, this doesn't work if $a = 0$ (because then $\delta = 0$ and we need $\delta > 0$). So suppose that $a = 0$. Given $\varepsilon > 0$ let $\delta = \varepsilon^2$. Then for every $x \in X$ such that $|x - 0| < \delta$ we have $|\sqrt{x} - 0| < \sqrt{\delta} = \varepsilon$. Thus $f(x)$ is continuous at $x = 0$.

As another example, consider the function

$$f(x) = \begin{cases} x \sin(1/x) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

We will prove that $f(x)$ is continuous at $x = 0$. Given $\varepsilon > 0$ let $\delta = \varepsilon$. Then for all $x \in \mathbb{R}$ such that $|x - 0| < \delta$ we have

$$|f(x) - f(0)| = |x \sin(1/x)| \leq |x| < \delta = \varepsilon.$$

So $f(x)$ is continuous at $x = 0$.

Now let's look at the function g defined by $g(x) = \sin(1/x)$ if $x \neq 0$ and $g(0) = 0$. By looking at the graph of the function $g(x)$ it seems clear that g cannot possibly be continuous at $x = 0$. Let's prove that g is discontinuous at $x = 0$. We will exhibit two sequences (a_n) and (b_n) both converging to 0 for which $(g(a_n))$ and $(g(b_n))$ converge to different limits. Let

$$a_n = \frac{2}{\pi(4n-3)}, \quad b_n = \frac{1}{\pi n}.$$

Then

$$g(a_n) = \sin\left(\frac{\pi(4n-3)}{2}\right) = \sin\left(\frac{\pi}{2}\right) = 1 \quad \text{for all } n \geq 1$$

and

$$g(b_n) = \sin(\pi n) = 0 \quad \text{for all } n \geq 1.$$

Thus g fails to satisfy Definition 11.2 at the point $x = 0$.

One of the easiest ways to tell if a function is continuous is to build it up from simpler functions which you already know are continuous. Here are some familiar operations on functions (f and g are functions and k is a real constant):

$$\begin{aligned} f + g: & \quad (f + g)(x) = f(x) + g(x) \\ fg: & \quad (fg)(x) = f(x)g(x) \\ f/g: & \quad (f/g)(x) = f(x)/g(x) \\ kf: & \quad (kf)(x) = k \cdot f(x) \\ |f|: & \quad (|f|)(x) = |f(x)| \\ g \circ f: & \quad (g \circ f)(x) = g(f(x)) \end{aligned}$$

Since we have proved a lot of useful properties of limits of sequences, we can get a lot of this essentially for free.

Proposition 11.5. *Suppose that f, g are continuous at $a \in \mathbb{R}$ and let $k \in \mathbb{R}$ be a constant. Then the functions $f + g$, fg , kf , and f/g (provided $g(a) \neq 0$) are continuous at a .*

Proof. These all follow from the corresponding properties for limits. For example, let (a_n) be any sequence converging to a . Since f and g are continuous at a , we have $\lim_{n \rightarrow \infty} f(a_n) = f(a)$ and $\lim_{n \rightarrow \infty} g(a_n) = g(a)$. Therefore

$$\lim_{n \rightarrow \infty} (f + g)(a_n) = \lim_{n \rightarrow \infty} (f(a_n) + g(a_n)) = \lim_{n \rightarrow \infty} f(a_n) + \lim_{n \rightarrow \infty} g(a_n) = f(a) + g(a).$$

The others are similar. □

Corollary 11.6. *Polynomial functions are continuous on $\mathbb{R}$.*

Proof. Polynomials can be built up from the functions x^n , $n \in \mathbb{N}$ by addition and scalar multiples (both continuity-preserving operations). Since multiplication of continuous functions is continuous, by induction we have that x^n is continuous for every $n \in \mathbb{N}$ (I'll leave it as an exercise for you to prove the base case, i.e. that $f(x) = x$ is continuous). □

Proposition 11.7. *The function $f(x) = |x|$ is continuous on $\mathbb{R}$.*

Proof. Let $a \in \mathbb{R}$. The statement is clear for $a \neq 0$ since $f(x) = x$ when $x > 0$ and $f(x) = -x$ when $x < 0$. So suppose that $a = 0$. Given $\varepsilon > 0$ let $\delta = \varepsilon$. Then, for any $x \in \mathbb{R}$ such that $|x - 0| < \delta$ we have

$$|f(x) - f(0)| = ||x| - 0| = |x| < \delta = \varepsilon.$$

So $f(x)$ is continuous at $x = 0$, and thus on all of $\mathbb{R}$. □

Proposition 11.8. *If f is continuous at $x = a$ and if g is continuous at $f(a)$ then the composite function $g \circ f$ is continuous at a .*

Proof. Let (a_n) be any sequence in the domain of f such that (a_n) converges to a and $f(a_n)$ is in the domain of g for all n . Then, since f is continuous at a , the sequence $(f(a_n))$ converges to $f(a)$. Since $(f(a_n))$ is a sequence converging to $f(a)$, and since g is continuous at $f(a)$, we conclude that the sequence $(g(f(a_n)))$ converges to $g(f(a))$. □

The two most important properties of continuous functions are: that continuous functions attain their maximum and minimum values on closed intervals, and the Intermediate Value Theorem. I cannot overstate the importance of these results; they are essential to the proof of the Mean Value Theorem (once we get to derivatives), which is a key ingredient in the proof of the Fundamental Theorem of Calculus.

Theorem 11.9. *Let $f : [a, b] \rightarrow \mathbb{R}$ be a continuous function. Then f is bounded (i.e. there exists a real number M such that $|f(x)| \leq M$ for all $x \in [a, b]$). Furthermore, f attains its maximum and minimum values on $[a, b]$ (i.e. there exist $x_0, y_0 \in [a, b]$ such that $f(x_0) \leq f(x) \leq f(y_0)$ for all $x \in [a, b]$).*

Proof. We argue by contradiction. Suppose that f is not bounded on $[a, b]$. Then for each $n \in \mathbb{N}$ there exists a number $x_n \in [a, b]$ such that $|f(x_n)| > n$. So we get a sequence (x_n) which is bounded (it's always inside $[a, b]$). By the Bolzano-Weierstrass theorem, there exists a convergent subsequence (x_{n_k}) of (x_n) , say that $\lim x_{n_k} = L$. Then $L \in [a, b]$ (why?), so by assumption f is continuous at L . So we have a sequence (x_{n_k}) which converges to L but the sequence $(f(x_{n_k}))$ doesn't converge to $f(L)$ (since $|f(x_{n_k})| \rightarrow +\infty$). This contradicts the continuity of f , so we must have that f is bounded.

Now let $M := \sup\{f(x) : x \in [a, b]\}$. We just proved that M is a finite number. For each $n \in \mathbb{N}$ there exists a number $y_n \in [a, b]$ such that $M - \frac{1}{n} \leq f(y_n) \leq M$ (otherwise M wouldn't be the sup). By the squeeze theorem, $\lim f(y_n) = M$. We aren't guaranteed that (y_n) is a convergent sequence, but it is bounded (always in $[a, b]$) so Bolzano-Weierstrass gives us a convergent subsequence (y_{n_k}) . Say $\lim y_{n_k} = y_0$. Since f is continuous at y_0 we have

$$f(y_0) = f(\lim y_{n_k}) = \lim f(y_{n_k}) = \lim f(y_n) = M.$$

(In the third equality we used that if a sequence is convergent, then every subsequence converges to the same limit.) So f attains its maximum M . The assertion for the minimum is similar. $\square$

Theorem 11.10 (Intermediate Value Theorem). *Let $I \subseteq \mathbb{R}$ be an interval and suppose that $f : I \rightarrow \mathbb{R}$ is continuous. Let $a, b \in I$ such that $a < b$. Then for every y between $f(a)$ and $f(b)$ (i.e. either $f(a) < y < f(b)$ or $f(a) > y > f(b)$) there exists at least one $x \in (a, b)$ such that $f(x) = y$.*

Proof. Assume that $f(a) < y < f(b)$ (the other case is similar). We will construct two sequences (a_n) and (b_n) which converge to x_0 such that $f(a_n)$ is above y and $f(b_n)$ is below y (hence the letters a for above and b for below).

Let $S = \{x \in [a, b] : f(x) < y\}$. Then S is not empty because $a \in S$. Since S is also bounded above (by b) it has a least upper bound, say $x_0 = \sup S$. For each $n \in \mathbb{N}$, the number $x_0 - \frac{1}{n}$ is not an upper bound for S so there exists $b_n \in S$ such that $x_0 - \frac{1}{n} < b_n \leq x_0$. By the squeeze theorem, $\lim b_n = x_0$. Since f is continuous and since $f(b_n) < y$ we have

$$f(x_0) = f(\lim b_n) = \lim f(b_n) \leq y.$$

Now for each $n \in \mathbb{N}$ let $a_n = \min\{b, x_0 + \frac{1}{n}\}$. Then for each n , $a_n \notin S$ so $f(a_n) \geq y$. Again by the squeeze theorem, $\lim a_n = x_0$ so we have

$$f(x_0) = f(\lim a_n) = \lim f(a_n) \geq y.$$

Thus $f(x_0) = y$. $\square$

Here's a neat example of the intermediate value theorem in action. It's an example of a *fixed point theorem*, and it's quite surprising at first glance that it should even be true (the hypotheses seem pretty weak). Fixed point theorems are ubiquitous in mathematics, especially in analysis and topology.

Let $f : [0, 1] \rightarrow [0, 1]$ be a continuous function. Then f has a fixed point; i.e. there exists a point $x_0 \in [0, 1]$ such that $f(x_0) = x_0$.

Proof. This assertion is equivalent to saying that the graph of f crosses the line $y = x$. In order to use the intermediate value theorem, we will “change coordinates” so that the line $y = x$ is “horizontal.” Don't worry, there's nothing fancy going on here, but this is a trick that we will use again several times. Just let $g(x) = f(x) - x$. Then g is also continuous. Since $0 \leq f(x) \leq 1$ we have $-x \leq g(x) \leq 1 - x$. Thus $g(0) \geq 0$ and $g(1) \leq 0$. It follows that there is some point $x_0 \in [0, 1]$ such that $g(x_0) = 0$. Of course, at that point we then have $f(x_0) - x_0 = 0$, i.e. $f(x_0) = x_0$. $\square$

Now let's go back and think about the two examples above,

$$f(x) = \begin{cases} x \sin(1/x) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0, \end{cases}$$

and

$$g(x) = \begin{cases} \sin(1/x) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Both of the functions f and g are continuous on the interval $(0, 1]$ since they are built out of functions which are continuous on that interval (here we are taking for granted that $\sin x$ is a continuous function, but we will not prove that in this class). But above we showed that $f(x)$ is continuous at 0, so we can say that it is continuous on the interval $[0, 1]$. The function $g(x)$ is not continuous at 0 (and there is no way to redefine $g(0)$ so that it is continuous at 0 because we found two sequences (a_n) and (b_n) converging to 0 such that $(g(a_n))$ and $(g(b_n))$ converge to different values). So g is not continuous on the interval $[0, 1]$.

What's the difference between these examples? Why can we define $f(0) = 0$ and make f continuous on $[0, 1]$ but we can't do the same for g ? It turns out that the answer is something called *uniform continuity*. In the remainder of this section we will define this term and prove the following proposition.

Proposition 11.11. *A function $f : (a, b) \rightarrow \mathbb{R}$ can be extended to a continuous function on $[a, b]$ if and only if it is uniformly continuous on (a, b) .*

I should first explain what it means to extend a function from (a, b) to $[a, b]$. This definition should match your intuition about what it should mean for a function to be extended to a larger domain. We say a function $\tilde{f} : Y \rightarrow \mathbb{R}$ is an extension of $f : X \rightarrow \mathbb{R}$ if

$$X \subseteq Y \quad \text{and} \quad \tilde{f}(x) = f(x) \text{ for all } x \in X.$$

The definition of uniform continuity will look a lot like the definition of continuity, with one important difference. Instead of fixing a point $a \in X$ and letting x be a point close to (i.e. within δ of) a , we let x, y both be free to move about in such a way that they remain close to each other. This difference is similar to the difference between the definition of a sequence converging to a number L and the definition of a Cauchy sequence.

Definition 11.12. A function $f : X \rightarrow \mathbb{R}$ is *uniformly continuous* on X if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - f(y)| < \varepsilon \quad \text{whenever} \quad x, y \in X \text{ and } |x - y| < \delta.$$

In this definition, δ is allowed to depend on ε but it cannot depend on x or y . As an example consider the function $f : [-4, 2] \rightarrow \mathbb{R}$ defined by $f(x) = x^2$. Given $\varepsilon > 0$, let $\delta = \square$ (we'll figure out in a minute what δ should be). Note that

$$|f(x) - f(y)| = |x^2 - y^2| = |x - y| \cdot |x + y|.$$

For $x, y \in [-4, 2]$ we have $|x + y| \leq |x| + |y| \leq 4 + 4 = 8$. Thus

$$|f(x) - f(y)| = 8|x - y|.$$

Now let's choose $\delta = \frac{\varepsilon}{8}$ so that when $|x - y| < \delta$ we have

$$|f(x) - f(y)| = 8|x - y| < 8\delta = \varepsilon.$$

Thus f is uniformly continuous on $[-4, 2]$.

The previous example is no accident, as the next proposition shows.

Proposition 11.13. *If f is continuous on a closed interval $[a, b]$ then f is uniformly continuous on $[a, b]$.*

Proof. We argue by contradiction. Suppose that f is continuous on $[a, b]$ but not uniformly continuous on $[a, b]$. Then there exists $\varepsilon > 0$ such that for every $\delta > 0$ there exist $x_\delta, y_\delta \in [a, b]$ such that

$$|x_\delta - y_\delta| < \delta \quad \text{but} \quad |f(x_\delta) - f(y_\delta)| \geq \varepsilon.$$

We will use this to construct sequences (x_n) and (y_n) . For each n , there exist $x_n, y_n \in [a, b]$ such that

$$|x_n - y_n| < \frac{1}{n} \quad \text{but} \quad |f(x_n) - f(y_n)| \geq \varepsilon.$$

By Bolzano-Weierstrass, since (x_n) and (y_n) are bounded, they each have convergent subsequences (x_{n_k}) and (y_{n_k}) . Since we chose x_n, y_n so that $|x_n - y_n| < \frac{1}{n}$ we must have $\lim x_{n_k} = \lim y_{n_k} = c$, say. Since f is continuous at c , we have

$$\lim_{k \rightarrow \infty} f(x_{n_k}) = f(c) = \lim_{k \rightarrow \infty} f(y_{n_k}),$$

so

$$\lim_{k \rightarrow \infty} (f(x_{n_k}) - f(y_{n_k})) = 0.$$

But this contradicts that $|f(x_n) - f(y_n)| \geq \varepsilon$ for all n . Thus f is uniformly continuous on the interval $[a, b]$. $\square$

To make the connection between uniformly continuous functions and Cauchy sequences even more solid, we have the following proposition.

Proposition 11.14. *If $f : X \rightarrow \mathbb{R}$ is uniformly continuous on X and if (a_n) is a Cauchy sequence then $(f(a_n))$ is a Cauchy sequence.*

Proof. Let (a_n) be a Cauchy sequence and let $\varepsilon > 0$. Since f is uniformly continuous on X , there exists a $\delta > 0$ such that

$$|f(x) - f(y)| < \varepsilon$$

whenever $x, y \in X$ and $|x - y| < \delta$. Since (a_n) is Cauchy, there exists an $N \geq 1$ such that (now we're thinking of δ as being the ϵ in the definition of a Cauchy sequence)

$$|a_m - a_n| < \delta$$

for all $m, n \geq N$. Thus, for such m, n we have (thinking of $x = a_m$ and $y = a_n$)

$$|f(a_m) - f(a_n)| < \varepsilon$$

so $(f(a_n))$ is a Cauchy sequence. □

We can use the previous proposition to prove that $g(x) = \sin(1/x)$ is not uniformly continuous on $(0, 1)$. Consider the Cauchy sequence (a_n) given by $a_n = \frac{2}{\pi n}$. Then the first few values of the sequence $(g(a_n))$ are

$$(g(a_n)) = 1, 0, -1, 0, 1, 0, -1, 0, \dots$$

It should be clear that $(g(a_n))$ is not Cauchy, so g is not uniformly continuous on $(0, 1)$.

We can now prove our main proposition on uniform continuity.

Proposition 11.15. *A function $f : (a, b) \rightarrow \mathbb{R}$ can be extended to a continuous function on $[a, b]$ if and only if it is uniformly continuous on (a, b) .*

Proof. ($\Rightarrow$) Suppose that f can be extended to a continuous function $\tilde{f} : [a, b] \rightarrow \mathbb{R}$. Then by Proposition 11.13 $\tilde{f}$ is uniformly continuous on $[a, b]$. It follows immediately that f is uniformly continuous on (a, b) .

($\Leftarrow$) Suppose that f is uniformly continuous on (a, b) . We must decide how to define $\tilde{f}(a)$ and $\tilde{f}(b)$ to make $\tilde{f} : [a, b] \rightarrow \mathbb{R}$ continuous. It suffices to show how to do this for $\tilde{f}(a)$ (since $\tilde{f}(b)$ is similar). There is really only one reasonable way to define $\tilde{f}(a)$, and it's as the limit of a sequence:

$$\tilde{f}(a) = \lim_{n \rightarrow \infty} f(a_n) \text{ for any sequence } (a_n) \text{ in } (a, b) \text{ converging to } a. \quad (11.2)$$

There are two immediate issues here: first, how do we even know the limit exists? and second, how do we know that this definition is unambiguous? (i.e. if you pick another sequence (a'_n) converging to a , does that potentially change the definition of $\tilde{f}(a)$?) We will deal with these one at a time.

Claim 1: if (a_n) is a sequence in (a, b) converging to a then the sequence $(f(a_n))$ converges. To prove this, note that (a_n) is a Cauchy sequence, so by Proposition 11.14, $(f(a_n))$ is Cauchy, so it converges.

Claim 2: if (a_n) and (a'_n) both converge to a then $\lim f(a_n) = \lim f(a'_n)$. To prove this, form the sequence (z_n) (z is for zipper) via

$$(z_n) = (a_1, a'_1, a_2, a'_2, a_3, a'_3, \dots).$$

It should be clear that $\lim z_n = \lim a_n = \lim a'_n = a$ so by Claim 1, $\lim f(z_n)$ exists. But now $(f(a_n))$ and $(f(a'_n))$ are subsequences of the convergent sequence $(f(z_n))$ so they converge to the same limit, i.e. $\lim f(a_n) = \lim f(a'_n)$.

Now that these claims are proven, the definition (11.2) makes sense, and it should also be clear that $\tilde{f}$ is continuous at a (we forced it to be continuous there by defining $\tilde{f}(a)$ to be equal to the limit of $(f(a_n))$ for every sequence (a_n) converging to a). □

12. LIMITS OF FUNCTIONS

In our proof of Proposition 11.15 we extended the function $f : (a, b) \rightarrow \mathbb{R}$ to a continuous function on $[a, b]$ by defining

$$\tilde{f}(a) = \lim_{n \rightarrow \infty} f(a_n) \text{ for any sequence } (a_n) \text{ in } (a, b) \text{ converging to } a.$$

One way to think of this is that $\tilde{f}(a)$ is the value that f “wants” to take at $x = a$, i.e. the behavior of f near $x = a$ is what determines the value of $\tilde{f}(a)$. This is essentially the idea behind limits of functions. We will start by defining left-hand and right-hand limits, then we’ll define the usual (two-sided) limit.

Definition 12.1. Let $f : X \rightarrow \mathbb{R}$ be a function and let $a \in X$. Suppose that X contains the interval (a, b) for some $b > a$. We say that $\lim_{x \rightarrow a^+} f(x) = L$ if

$$\lim_{n \rightarrow \infty} f(a_n) = L \quad \text{for every sequence } (a_n) \text{ in } (a, b) \text{ converging to } a.$$

In this case, the number L is called the *right-hand limit* of f at a .

Suppose that X contains the interval (c, a) for some $c < a$. We say that $\lim_{x \rightarrow a^-} f(x) = L$ if

$$\lim_{n \rightarrow \infty} f(a_n) = L \quad \text{for every sequence } (a_n) \text{ in } (c, a) \text{ converging to } a.$$

In this case, the number L is called the *left-hand limit* of f at a .

Lastly, if

$$\lim_{x \rightarrow a^-} f(x) = L = \lim_{x \rightarrow a^+} f(x),$$

then we write

$$\lim_{x \rightarrow a} f(x) = L$$

and we say that L is *the limit* (or sometimes *the two-sided limit*) of f at a . We can define this last case directly without talking about the left-hand or right-hand limit as follows. Suppose that X contains the interval (c, b) for some $c < a < b$. We say that $\lim_{x \rightarrow a} f(x) = L$ if

$$\lim_{n \rightarrow \infty} f(a_n) = L \quad \text{for every sequence } (a_n) \text{ in } (c, a) \cup (a, b) \text{ converging to } a.$$

(Note that the sequence (a_n) should always be contained in a set which does not contain a .)

A few remarks are in order. First, f need not be defined at a , and even if it is defined at a , the value $f(a)$ need not equal $\lim_{x \rightarrow a} f(x)$. In fact, $f(a) = \lim_{x \rightarrow a} f(x)$ if and only if f is defined on an open interval (c, b) containing a and f is continuous at a . Lastly, limits, when they exist, are unique. See pp. 154-156 of your textbook for examples of limits of functions.

The following proposition shows that limits of functions obey the usual algebraic rules. These follow almost immediately from the corresponding rules for limits of sequences. I will use the nonstandard notation $\lim_{x \rightarrow a^*}$ to mean the right-hand or left-hand limit if $*$ is one of the symbols $+$ or $-$, respectively, or the two-sided limit if $*$ is no symbol.

Proposition 12.2. Let f and g be functions for which

$$\lim_{x \rightarrow a^*} f(x) = L \quad \text{and} \quad \lim_{x \rightarrow a^*} g(x) = M$$

for some $L, M \in \mathbb{R}$. Then

$$(1) \quad \lim_{x \rightarrow a^*} (f + g)(x) = L + M,$$

- (2) $\lim_{x \rightarrow a^*} (fg)(x) = LM$, and
 (3) $\lim_{x \rightarrow a^*} (f/g)(x) = L/M$ provided $M \neq 0$ and $g(x) \neq 0$ for all x in some open interval containing a .

We can also pull limits inside of continuous functions, as we did in the continuity section. I will state this for two-sided limits, but the analogue for one-sided limits also holds.

Proposition 12.3. *Let f be a function defined on the set $(c, a) \cup (a, b)$ for some $c < a < b$, such that $\lim_{x \rightarrow a} f(x) = L \in \mathbb{R}$. Let $Y = \{f(x) : x \in (c, a) \cup (a, b)\} \cup L$. If $g : Y \rightarrow \mathbb{R}$ is continuous at L then*

$$\lim_{x \rightarrow a} (g \circ f)(x) = g(L).$$

There are actually two reasonable definitions for limits, as there were when we talked about continuity. The second definition involves ε and δ and is equivalent to the first (but I won't prove that here; the proof is quite similar to the one we did in the previous section). Again, I'll state this definition for two-sided limits, but the definition is easily modified to deal with one-sided limits (see Corollary 20.8 on p. 160 of your textbook).

Definition 12.4. Let f be a function defined on the set $(c, a) \cup (a, b)$ for some $c < a < b$. Then $\lim_{x \rightarrow a} f(x) = L$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - L| < \varepsilon \quad \text{whenever } 0 < |x - a| < \delta.$$

As an example, let's prove that

$$\lim_{x \rightarrow a} \frac{x^2 - a^2}{x - a} = 2a$$

for any $a \in \mathbb{R}$. This will be useful when we talk about derivatives in the next lecture. Let $\varepsilon > 0$ be given, and let $\delta = \square$ (not sure what this should be yet, so let's do some work first). Then for $0 < |x - a| < \delta$ we have

$$|f(x) - 2a| = \left| \frac{x^2 - a^2}{x - a} - 2a \right| = |x + a - 2a| = |x - a| < \delta.$$

Since we want this to be $< \varepsilon$ we can take $\delta = \varepsilon$.

We can also do this computation using the fact that, for limits, we only consider values of $f(x)$ for x close to a , but never for x equal to a . So we can write

$$\lim_{x \rightarrow a} \frac{x^2 - a^2}{x - a} = \lim_{x \rightarrow a} \frac{(x - a)(x + a)}{x - a} = \lim_{x \rightarrow a} (x + a)$$

since the limit only cares about $x \in (c, a) \cup (a, b)$, but not about $x = a$. Since the function $(x + a)$ is continuous at $x = a$, we can use the previous proposition to "plug in" $x = a$ and conclude that

$$\lim_{x \rightarrow a} \frac{x^2 - a^2}{x - a} = \lim_{x \rightarrow a} (x + a) = 2a.$$

As another example, let's do Exercise 20.16. Suppose that $\lim_{x \rightarrow a^+} f(x) = L$ and $\lim_{x \rightarrow a^+} g(x) = M$. If $f(x) \leq g(x)$ for all $x \in (a, b)$ for some $b > a$ then we have $L \leq M$.

Proof. Given $\varepsilon > 0$ there exists a $\delta_1 > 0$ such that

$$|f(x) - L| < \frac{\varepsilon}{2} \quad \text{whenever } 0 < x - a < \delta_1.$$

(Two remarks: first, you'll see why we want $\frac{\varepsilon}{2}$ shortly; second, the condition above usually looks like $|x - a| < \delta$ but since this is the right-hand limit we are only thinking about $x > a$ so $|x - a| = x - a$.) Similarly, there exists a $\delta_2 > 0$ such that

$$|g(x) - M| < \frac{\varepsilon}{2} \quad \text{whenever } 0 < x - a < \delta_2.$$

So for $0 < x - a < \min\{\delta_1, \delta_2\}$ we have

$$L - \frac{\varepsilon}{2} < f(x) < L + \frac{\varepsilon}{2} \quad \text{and} \quad M - \frac{\varepsilon}{2} < g(x) < M + \frac{\varepsilon}{2}.$$

Since $f(x) \leq g(x)$ for x near a we have

$$L - \frac{\varepsilon}{2} < f(x) \leq g(x) < M + \frac{\varepsilon}{2}$$

from which it follows that $L < M + \varepsilon$. Since ε was arbitrary, this implies that $L \leq M$. $\square$

13. THE DERIVATIVE

Derivatives should be quite familiar to you from your calculus courses. I expect that you can differentiate most, if not all, functions that are built out of elementary functions (polynomials, trig functions, logs, exponents) using algebra and composition. So this section won't focus much, if at all, on these sorts of things. Instead we'll give a rigorous definition of the derivative and use it to prove the properties that will be the most useful to us.

Definition 13.1. Let $f : X \rightarrow \mathbb{R}$ be a function and let $a \in X$ such that there is an interval $(c, b) \subseteq X$ containing a . (Remember that limits and derivatives are *local* ideas, so we need an open interval around a to talk about the derivative.) We say that f is *differentiable* at a if the limit

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$$

exists and is finite, and we call this limit *the derivative of f at a* . We usually write $f'(a)$ for the derivative of f at a .

We will often think of f' as a function in its own right. In the previous section we computed $f'(a)$ for the function $f(x) = x^2$ and found that $f'(a) = 2a$. This should be no surprise to you. Since it doesn't matter what letter we use, we will usually write $f'(x) = 2x$.

Proposition 13.2 (Differentiability implies continuity). *If f is differentiable at a then f is continuous at a .*

Proof. Recall from the last section that a function f is continuous at a if and only if it is defined on an open interval containing a (automatically true here because differentiability also requires f to be defined on an open interval containing a) and if $\lim_{x \rightarrow a} f(x) = f(a)$. Recall that when thinking about the limit of a function at a we consider points x close to a but not equal to a . So we are justified in writing

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} \left[(x - a) \frac{f(x) - f(a)}{x - a} + f(a) \right] = f(a) + \lim_{x \rightarrow a} (x - a) \frac{f(x) - f(a)}{x - a}.$$

Using our assumption that f is differentiable at a , the limit of the quotient on the right-hand side above exists and is finite. So we have

$$\lim_{x \rightarrow a} (x - a) \frac{f(x) - f(a)}{x - a} = \lim_{x \rightarrow a} (x - a) \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = 0 \cdot [\text{something finite}] = 0.$$

This shows that $\lim_{x \rightarrow a} f(x) = f(a)$, as desired. $\square$

As you know, derivatives respect sums and scalar multiples (i.e. differentiation is a linear transformation on the vector space of differentiable functions). I won't prove this here because it's straightforward (see Theorem 28.3 of your textbook).

Proposition 13.3. *Let f and g be functions that are differentiable at a . Then $f + g$ and cf ($c \in \mathbb{R}$) are differentiable at a , and we have the formulas*

- (1) $(cf)'(a) = c \cdot f'(a)$
- (2) $(f + g)'(a) = f'(a) + g'(a)$.

For products and quotients, things are a bit more interesting (as you already know). I will state both the product and quotient rules, but only prove the product rule (since that's the only one we'll need in this class).

Proposition 13.4 (Product rule). *Let f and g be functions that are differentiable at a . Then fg is differentiable at a and*

$$(fg)'(a) = f(a)g'(a) + f'(a)g(a).$$

Proof. Adding zero in a clever way, we compute

$$\begin{aligned} \lim_{x \rightarrow a} \frac{f(x)g(x) - f(a)g(a)}{x - a} &= \lim_{x \rightarrow a} \frac{f(x)g(x) - f(x)g(a) + f(x)g(a) - f(a)g(a)}{x - a} \\ &= \lim_{x \rightarrow a} f(x) \frac{g(x) - g(a)}{x - a} + g(a) \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} \\ &= f(a)g'(a) + f'(a)g(a), \end{aligned}$$

as desired. $\square$

Proposition 13.5 (Quotient rule). *Let f and g be functions that are differentiable at a and suppose that $g(a) \neq 0$. Then f/g is differentiable at a and*

$$(f/g)'(a) = \frac{g(a)f'(a) - f(a)g'(a)}{[g(a)]^2}.$$

Using the product rule and a bit of induction, we can quickly prove that $(x^n)' = nx^{n-1}$ for all integers $n \geq 1$. Indeed, this is true for $n = 1$ since

$$\lim_{x \rightarrow a} \frac{x - a}{x - a} = 1.$$

Now suppose that it is true for n , we will prove it for $n + 1$. We have

$$(x^{n+1})' = (x \cdot x^n)' = x \cdot nx^{n-1} + 1 \cdot x^n = (n + 1)x^n.$$

Thus it is true that $(x^n)' = nx^{n-1}$ for all $n \geq 1$.

Lastly, we want to be able to differentiate the compositum $g \circ f$ of two functions f and g . As you know, this procedure is called the *chain rule*, and its proof is somewhat technical.

Proposition 13.6 (Chain Rule). *If f is differentiable at a and g is differentiable at $f(a)$ then the composite function $g \circ f$ is differentiable at a and we have*

$$(g \circ f)'(a) = g'(f(a))f'(a).$$

Proof. We first need to check that $g \circ f$ is defined on an open interval containing a . I'm not going to make a big deal about this; you can see Exercise 28.13 of your textbook (along with the solution on p. 390) if you're interested.

Suppose first that there exists an open interval I containing a such that $f(x) \neq f(a)$ for all $x \in I \setminus \{a\}$. Then we can write

$$\lim_{x \rightarrow a} \frac{g(f(x)) - g(f(a))}{x - a} = \lim_{x \rightarrow a} \frac{g(f(x)) - g(f(a))}{f(x) - f(a)} \cdot \frac{f(x) - f(a)}{x - a}. \quad (13.1)$$

Both limits exist since g is differentiable at $f(a)$ and since f is differentiable at a . Thus

$$(g \circ f)'(a) = g'(f(a))f'(a).$$

Now suppose that $f(x) = f(a)$ for x arbitrarily close to a (that is, suppose we can't find any open intervals containing a such that $f(x) \neq f(a)$ for all $x \neq a$). Then there is a sequence (x_n) converging to a such that $f(x_n) \neq f(a)$ for all n . Using the sequential definition of the limit, we find that

$$f'(a) = \lim_{n \rightarrow \infty} \frac{f(x_n) - f(a)}{x_n - a} = 0.$$

Thus, we want to prove that $(g \circ f)'(a) = 0$. We will use the squeeze theorem to show this. By assumption $g'(f(a))$ exists, so the limit

$$\lim_{x \rightarrow a} \frac{g(f(x)) - g(f(a))}{f(x) - f(a)}$$

exists, and therefore the quotients $\frac{g(f(x)) - g(f(a))}{f(x) - f(a)}$ are bounded, say by M , for $f(x)$ sufficiently close to $f(a)$, i.e.

$$\left| \frac{g(f(x)) - g(f(a))}{f(x) - f(a)} \right| \leq M.$$

Thus, for x sufficiently close to a ,

$$\left| \frac{g(f(x)) - g(f(a))}{x - a} \right| \leq M \left| \frac{f(x) - f(a)}{x - a} \right|. \quad (13.2)$$

(To see this, note that if $f(x) = f(a)$ then both sides are zero; otherwise we can use the same sort of trick as in (13.1)). By the squeeze theorem, since the limit of the right-hand side of (13.2) as $x \rightarrow a$ is zero, so is the limit of the left-hand side. $\square$

Perhaps the most important result (for us, at least) concerning derivatives is the Mean Value Theorem. Indeed, we will see later that it is an essential ingredient in the proof of the Fundamental Theorem of Calculus. To prove the MVT, we will first prove an intermediate result called Rolle's theorem.

Theorem 13.7 (Rolle's Theorem). *Suppose that f is continuous on $[a, b]$ and differentiable on (a, b) and that $f(a) = f(b)$. Then there exists a point $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. By the Extreme Value Theorem, since f is continuous on the closed interval $[a, b]$ it attains its maximum and minimum on $[a, b]$. So there exist $y, z \in [a, b]$ such that $f(y) \leq f(x) \leq f(z)$ for all $x \in [a, b]$. Suppose that z is an interior point, i.e. $z \in (a, b)$. Let (z_n)

be a sequence in $(a, z) \cup (z, b)$ converging to z and let (z_n^+) and (z_n^-) be the subsequences of (z_n) such that $z_n^+ > z$ and $z_n^- < z$. Then (since $f(z)$ is a maximum)

$$\lim_{n \rightarrow \infty} \frac{f(z_n^+) - f(z)}{z_n^+ - z} \leq 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{f(z_n^-) - f(z)}{z_n^- - z} \geq 0.$$

(Both limits exist because f is differentiable at z .) Since subsequences of a convergent sequence converge to the same limit, we see that

$$\lim_{n \rightarrow \infty} \frac{f(z_n) - f(z)}{z_n - z} \leq 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{f(z_n) - f(z)}{z_n - z} \geq 0.$$

It follows that $f'(z) = 0$. By the same reasoning, we have $f'(y) = 0$ if $y \in (a, b)$. The only case left is if y and z are both endpoints of $[a, b]$. But since $f(a) = f(b)$ this forces $f(y) = f(x) = f(z)$ for all $x \in [a, b]$, so $f'(x) = 0$ for all $x \in (a, b)$. $\square$

Before we prove the Mean Value Theorem, here is a neat example of how you might use Rolle's theorem to get information about roots of differentiable functions. Consider the polynomial $f(x) = x^3 + 3x + 1$. Since $f(0) = 1$ and $f(-1) = -3$ there must be a root of f in the interval $(-1, 0)$ (this is the Intermediate Value Theorem). Now suppose that f has two distinct roots a and b . Then $f(a) = f(b) = 0$ and f is clearly continuous on $[a, b]$ and differentiable on (a, b) . It follows from Rolle's theorem that $f'(c) = 0$ for some $c \in (a, b)$. But $f'(x) = 3x^2 + 3 > 0$ for all $x \in \mathbb{R}$, which is a contradiction. Thus f has exactly one root.

Theorem 13.8 (The Mean Value Theorem). *Suppose that f is continuous on $[a, b]$ and differentiable on (a, b) . Then there exists a $c \in (a, b)$ such that*

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

Proof. You should think of the MVT as Rolle's theorem where you've tilted your head to the side (or maybe you rotated the plane using some linear transformation). We will use this idea to prove the MVT using Rolle's Theorem. Let

$$g(x) = f(x) - \left(\frac{f(b) - f(a)}{b - a} (x - a) + f(a) \right).$$

The function in parentheses is just the equation of the secant line connecting $(a, f(a))$ and $(b, f(b))$. We compute

$$g(a) = 0 \quad \text{and} \quad g(b) = 0.$$

Clearly g is continuous on $[a, b]$ and differentiable on (a, b) , so Rolle's theorem applies. So there is some point $c \in (a, b)$ such that $g'(c) = 0$. On the other hand,

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

This proves the theorem. $\square$

There are three simple but important corollaries to the Mean Value Theorem. You used these implicitly in your Calculus courses, likely without realizing it.

Corollary 13.9. *Suppose that f is differentiable on (a, b) and that $f'(x) = 0$ for all $x \in (a, b)$. Then f is a constant function on (a, b) .*

Proof. Let $x, y \in (a, b)$ such that $x < y$. Applying the Mean Value Theorem to the closed interval $[x, y]$, we find that, for some $c \in (x, y)$,

$$\frac{f(y) - f(x)}{y - x} = f'(c) = 0$$

since $f'(x) = 0$ for all $x \in (a, b)$. Therefore $f(y) = f(x)$. Let $c = f(y) = f(x)$. Since x, y were arbitrary, it follows that $f(x) = c$ for all $x \in (a, b)$. $\square$

Corollary 13.10. *Suppose that f and g are differentiable on (a, b) and that $f' = g'$ on (a, b) . Then there exists a constant c such that $f(x) = g(x) + c$ for all $x \in (a, b)$.*

Proof. Apply the previous corollary to the function $f - g$. $\square$

This result is what allows us to find a family of antiderivatives of a function by finding a single antiderivative and adding an arbitrary constant ($+C$).

For the next corollary, we need a few quick definitions. We say that f is *increasing* on an interval (a, b) if $f(y) \geq f(x)$ whenever $y > x$ in (a, b) . Similarly, f is *decreasing* on an interval (a, b) if $f(y) \leq f(x)$ whenever $y > x$ in (a, b) . We add the word *strictly* when we can replace the inequalities $\geq$ or $\leq$ by $>$ or $<$, respectively.

Corollary 13.11. *Suppose that f is differentiable on (a, b) . Then*

- (1) *if $f'(x) > 0$ for all $x \in (a, b)$ then f is strictly increasing;*
- (2) *if $f'(x) < 0$ for all $x \in (a, b)$ then f is strictly decreasing;*
- (3) *if $f'(x) \geq 0$ for all $x \in (a, b)$ then f is increasing;*
- (4) *if $f'(x) \leq 0$ for all $x \in (a, b)$ then f is decreasing.*

Proof. We will prove (1); the others are similar. Suppose that $a < x < y < b$. We will apply the Mean Value Theorem to the closed interval $[x, y]$. There exists some $c \in (x, y)$ such that

$$\frac{f(y) - f(x)}{y - x} = f'(c) > 0.$$

Since $y - x > 0$, this implies that $f(y) - f(x) > 0$, or $f(x) < f(y)$. So f is strictly increasing. $\square$

14. INTEGRATION

We are finally ready to talk about integration and the Fundamental Theorem of Calculus. This section can get somewhat technical so it's important to keep a big-picture in mind as we're going through this. The main idea is that the integral of f is essentially defined as the (signed) area under the graph of f (where we give the area a negative sign if the graph of f is below the x -axis). It's important to point out here that *this definition of integration says nothing whatsoever about derivatives*. The connection between integrals and antiderivatives is the beautiful Fundamental Theorem of Calculus, but you can talk about integrals without ever having defined the derivative. It turns out that the Mean Value Theorem will supply us with the glue that connects the two ideas.

You might remember Riemann sums from your introductory calculus course. While these are a perfectly good way of talking about integrals, it turns out to be somewhat more convenient to use a slightly different (but ultimately equivalent) approach using *Darboux sums*.

Definition 14.1. Suppose that f is bounded on $[a, b]$. A *partition* P of $[a, b]$ is a finite ordered set

$$P = \{a = x_0 < x_1 < x_2 < \dots < x_n = b\}.$$

The intervals $[x_{k-1}, x_k]$ are called *subintervals* of P . For each subinterval, let

$$m_k = \inf\{f(x) : x \in [x_{k-1}, x_k]\},$$

$$M_k = \sup\{f(x) : x \in [x_{k-1}, x_k]\}.$$

The *lower Darboux sum* $L(f, P)$ of f with respect to P is

$$L(f, P) = \sum_{k=1}^n m_k(x_k - x_{k-1})$$

and the *upper Darboux sum* $U(f, P)$ of f with respect to P is

$$U(f, P) = \sum_{k=1}^n M_k(x_k - x_{k-1}).$$

The next two lemmas show the relationship between upper and lower sums for different partitions.

Lemma 14.2. Let f be a bounded function on $[a, b]$. If P and Q are partitions of $[a, b]$ and $P \subseteq Q$, then

$$L(f, P) \leq L(f, Q) \leq U(f, Q) \leq U(f, P).$$

Proof. The middle inequality $L(f, Q) \leq U(f, Q)$ should be clear from the definition. We will prove that $L(f, P) \leq L(f, Q)$; the remaining inequality is similar. Suppose that Q has exactly one more point than P , say that u was inserted into the subinterval $[x_{k-1}, x_k]$ to create two subintervals $[x_{k-1}, u]$ and $[u, x_k]$. Then the summands of the lower Darboux sums for P and Q are exactly the same except that $L(f, P)$ includes the term

$$m_k(x_k - x_{k-1}),$$

while $L(f, Q)$ includes both of the terms

$$m'_k(x_k - u) + m''_k(u - x_{k-1}),$$

where $m'_k = \inf\{f(x) : x \in [x_{k-1}, u]\}$ and $m''_k = \inf\{f(x) : x \in [u, x_k]\}$. Since we must have $m'_k \geq m_k$ and $m''_k \geq m_k$ (the inf can't go down if you are talking about a smaller set) it follows that

$$m_k(x_k - x_{k-1}) = m_k(x_k - u) + m_k(u - x_{k-1}) \leq m'_k(x_k - u) + m''_k(u - x_{k-1}).$$

Therefore $L(f, P) \leq L(f, Q)$.

In more generality, we can get Q from P by adding one point at a time, so an induction argument (which I'll omit here) shows that $L(f, P) \leq L(f, Q)$ in general. $\square$

Lemma 14.3. If f is bounded on $[a, b]$ and if P and Q are any partitions of $[a, b]$ then $L(f, P) \leq U(f, Q)$.

Proof. The set $P \cup Q$ is also a partition of $[a, b]$. Since $P \subseteq P \cup Q$ and $Q \subseteq P \cup Q$, we can apply the previous proposition to get

$$L(f, P) \leq L(f, P \cup Q) \leq U(f, P \cup Q) \leq U(f, Q).$$

$\square$

Intuitively, it helps to think of $U(f, P)$ as an overestimate for the area under f and $L(f, P)$ as an underestimate. Then as we *refine* the partition by adding more points, the approximation gets better (i.e. the lower sums get bigger and the upper sums get smaller). Then we say that the integral exists if these two sums meet somewhere in the middle. To make this precise, we define

$$L(f) = \sup\{L(f, P) : P \text{ is a partition of } [a, b]\}$$

and

$$U(f) = \inf\{U(f, P) : P \text{ is a partition of } [a, b]\},$$

and we say that f is *integrable* on $[a, b]$ if $L(f) = U(f)$. In that case we write

$$\int_a^b f = L(f) = U(f).$$

Proposition 14.4. *If f is bounded on $[a, b]$ then $L(f) \leq U(f)$.*

Proof. Let P be a partition of $[a, b]$. The previous lemma shows that $L(f, P)$ is a lower bound for $U(f, Q)$ for every Q . Thus $L(f, P) \leq U(f)$. But this implies that $U(f)$ is an upper bound for $L(f, P)$ for every P , so we have $U(f) \geq L(f)$. $\square$

As with almost every topic in this course, we have a Cauchy-like criterion for integrability.

Proposition 14.5. *Suppose that f is bounded on $[a, b]$. Then f is integrable if and only if for every $\varepsilon > 0$ there exists a partition P of $[a, b]$ such that*

$$U(f, P) - L(f, P) < \varepsilon.$$

Proof. ($\Leftarrow$) Suppose that such a partition P exists for each $\varepsilon > 0$. Then

$$U(f) - L(f) \leq U(f, P) - L(f, P) < \varepsilon.$$

Since ε is arbitrary, this implies that $U(f) = L(f)$.

($\Rightarrow$) Now suppose that f is integrable. Let $\varepsilon > 0$ be given. Since $U(f)$ is the greatest lower bound of the $U(f, P)$ over all partitions P , the number $U(f) + \frac{\varepsilon}{2}$ is no longer a lower bound, so there exists a partition P_1 such that

$$U(f, P_1) < U(f) + \frac{\varepsilon}{2}.$$

Similarly, there exists a partition P_2 such that

$$L(f, P_2) > L(f) - \frac{\varepsilon}{2}.$$

Let $P = P_1 \cup P_2$. Since integrability means that $U(f) = L(f)$, we have

$$\begin{aligned} U(f, P) - L(f, P) &\leq U(f, P_1) - L(f, P_2) \\ &= U(f, P_1) - U(f) + L(f) - L(f, P_2) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

This completes the proof. $\square$

Now that we have defined integration, a fundamental question arises: which functions are integrable? We won't be able to answer this question in full generality in this course (the answer is quite interesting, though) but we will be able to determine that many familiar functions are integrable.

Proposition 14.6. *Every continuous function on $[a, b]$ is integrable.*

Proof. First, we observe that because f is continuous on the closed interval $[a, b]$ it is uniformly continuous on $[a, b]$. So, given $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - f(y)| < \frac{\varepsilon}{b-a} \quad \text{whenever } |x - y| < \delta.$$

Let P be a partition of $[a, b]$ such that $x_k - x_{k-1} < \delta$ for every subinterval. By the extreme value theorem applied to $[x_{k-1}, x_k]$, the function f attains its maximum M_k and its minimum m_k for some points z_k, y_k in the interval $[x_{k-1}, x_k]$. And whatever those points z_k, y_k are, they are separated by a distance of at most δ . So

$$M_k - m_k = f(z_k) - f(y_k) < \frac{\varepsilon}{b-a}.$$

It follows that

$$U(f, P) - L(f, P) = \sum_{k=1}^n (M_k - m_k)(x_k - x_{k-1}) < \frac{\varepsilon}{b-a} \sum_{k=1}^n (x_k - x_{k-1}) = \varepsilon$$

since the last sum is telescoping and equals $b - a$. By the Cauchy criterion, f is integrable on $[a, b]$. $\square$

Proposition 14.7. *Every monotone function on $[a, b]$ is integrable.*

Proof. Suppose that f is increasing (decreasing is similar). The proof is very similarly to the previous proof if you make the following observation: on $[x_{k-1}, x_k]$ we have $m_k = f(x_{k-1})$ and $M_k = f(x_k)$. Given $\varepsilon > 0$ we pick P such that $x_k - x_{k-1} < \frac{\varepsilon}{f(b) - f(a)}$ for all k ; then

$$\begin{aligned} U(f, P) - L(f, P) &= \sum_{k=1}^n (M_k - m_k)(x_k - x_{k-1}) \\ &= \sum_{k=1}^n (f(x_k) - f(x_{k-1}))(x_k - x_{k-1}) \\ &< \frac{\varepsilon}{f(b) - f(a)} \sum_{k=1}^n (f(x_k) - f(x_{k-1})) = \varepsilon. \end{aligned}$$

Again by the Cauchy criterion, we are done. $\square$

Since many of the functions that you know are either continuous or (piecewise) monotonic, it begs the question: can we find a non-integrable function? One good candidate is this crazy function we've seen before:

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

Let P be any partition of $[0, 1]$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$ (and $\mathbb{R} \setminus \mathbb{Q}$ is also dense in $\mathbb{R}$), every subinterval contains a point x where $f(x) = 1$ and a point y where $f(y) = 0$. Therefore $U(f, P) = 1$ and $L(f, P) = 0$. Since this is the case for any partition P , it follows that $L(f) \neq U(f)$. So f is not integrable on $[0, 1]$.

The following proposition shows that the set of integrable functions on $[a, b]$ forms a vector space and the integral is a linear operator on that space.

Proposition 14.8. *Suppose that f, g are integrable functions on $[a, b]$ and let $c \in \mathbb{R}$. Then*

- (1) cf is integrable and $\int_a^b cf = c \int_a^b f$.
 (2) $f + g$ is integrable and $\int_a^b (f + g) = \int_a^b f + \int_a^b g$.

Proof. (1) To prove this, we first consider the case when $c \geq 0$ (note that it is trivial when $c = 0$). This case is pretty straightforward; you can read it in your book if you are interested, but I'll skip it for the sake of time. Then we consider the case $c = -1$. For this we use that, for $X \subset \mathbb{R}$, we have $\sup(-X) = -\inf X$. Thus, if P is any partition of $[a, b]$, we have $U(-f, P) = -L(f, P)$. It follows that

$$U(-f) = \inf_P U(-f, P) = \inf_P (-L(f, P)) = -\sup_P L(f, P) = -L(f).$$

Replacing f by $-f$, we also have $L(-f) = -U(f)$. Since f is integrable, we have

$$U(-f) = -L(f) = -U(f) = L(-f).$$

Thus f is integrable, and $\int_a^b -f = -\int_a^b f$. Now, if $c < 0$, we just write $c = -(-c)$ and apply the cases $c > 0$ and $c = -1$.

(2) We will use the Cauchy criterion. Given $\varepsilon > 0$ there exist partitions P_1 and P_2 such that

$$U(f, P_1) - L(f, P_1) < \frac{\varepsilon}{2} \quad \text{and} \quad U(g, P_2) - L(g, P_2) < \frac{\varepsilon}{2}$$

since f and g are integrable. Let $P = P_1 \cup P_2$. There was a homework problem several weeks ago where we proved that $\sup(A + B) \leq \sup A + \sup B$. From this fact it follows that $U(f + g, P) \leq U(f, P) + U(g, P)$. By a similar argument, we have $L(f + g, P) \geq L(f, P) + L(g, P)$. Thus we have

$$U(f + g, P) - L(f + g, P) \leq U(f, P) + U(g, P) - L(f, P) - L(g, P) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Therefore $f + g$ is integrable. To compute $\int_a^b f$ we write

$$\begin{aligned} \int_a^b (f + g) &= U(f + g) \leq U(f + g, P) \leq U(f, P) + U(g, P) \\ &< L(f, P) + L(g, P) + \varepsilon \leq L(f) + L(g) + \varepsilon = \int_a^b f + \int_a^b g + \varepsilon. \end{aligned}$$

A similar argument shows that $\int_a^b (f + g) > \int_a^b f + \int_a^b g - \varepsilon$. Since ε is arbitrary, this proves that $\int_a^b (f + g) = \int_a^b f + \int_a^b g$. $\square$

We also have the following useful properties which we probably won't have time to prove in class. However, we will need them for the proof of the Fundamental Theorem of Calculus, so I'll sketch the proofs here.

- Proposition 14.9.** (1) If f and g are integrable on $[a, b]$ and $f(x) \leq g(x)$ for $x \in [a, b]$ then $\int_a^b f \leq \int_a^b g$.
 (2) If g is a continuous, nonnegative function on $[a, b]$ and if $\int_a^b g = 0$ then $g(x) = 0$ for all $x \in [a, b]$.
 (3) (Triangle inequality for integrals) If f is integrable on $[a, b]$, then so is $|f|$, and $|\int_a^b f| \leq \int_a^b |f|$.

(4) Let $a < c < b$. If f is integrable on $[a, c]$ and on $[c, b]$ then f is integrable on $[a, b]$ and

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

(This also holds even if $c \notin (a, b)$).

Proof. (1) Let $h = g - f$. Then h is integrable and $h(x) \geq 0$ for all $x \in [a, b]$. It follows that $L(h, P) \geq 0$ for any partition P of $[a, b]$, thus $L(h) \geq 0$. So $\int_a^b h \geq 0$, from which we have $\int_a^b f \leq \int_a^b g$.

(2) If $g(c) > 0$ for some $c \in [a, b]$ then by continuity (I'm skipping a few steps here) we have $g(x) \geq \varepsilon$ for some $\varepsilon > 0$ for all x in some small interval (α, β) around c . It follows that $\int_a^b g \geq \int_\alpha^\beta g \geq \varepsilon(\beta - \alpha) > 0$, which contradicts our assumption. So $g(x) = 0$ for all $x \in [a, b]$.

(3) Again, this is a sketch. Let P be a partition of $[a, b]$ and consider the subinterval $[x_{k-1}, x_k]$. Let $M_k(f)$ and $M_k(|f|)$ denote the suprema of f and $|f|$ on this subinterval, with $m_k(f)$, $m_k(|f|)$ defined similarly. It's not obvious, but you can check that

$$M_k(|f|) - m_k(|f|) \leq M_k(f) - m_k(f).$$

It follows that

$$U(|f|, P) - L(|f|, P) \leq U(f, P) - L(f, P)$$

and therefore by the Cauchy criterion we have that $|f|$ is integrable. Now we apply (1) to the inequality $-|f| \leq f \leq |f|$ to get

$$-\int_a^b |f| \leq \int_a^b f \leq \int_a^b |f|,$$

from which the triangle inequality for integrals follows.

(4) The idea is to apply the Cauchy criterion (with $\varepsilon/2$ each time) to a partition P_1 of $[a, c]$ and a partition P_2 of $[c, b]$. Then $P = P_1 \cup P_2$ is a partition of $[a, b]$ and we have $U(f, P) = U(f, P_1) + U(f, P_2)$ (and similarly for lower sums). It follows that f is integrable on $[a, b]$. The formula for $\int_a^b f$ follows similarly. $\square$

We are now ready to prove the Fundamental Theorem of Calculus. This theorem is traditionally broken into two parts, and some books swap part I and part II from what I have written below. It doesn't really matter as long as you know both of them. I have followed the order given in your textbook in the name of consistency.

Theorem 14.10 (The Fundamental Theorem of Calculus I). *If g is continuous on $[a, b]$ and differentiable on (a, b) , and if g' is integrable on (a, b) then*

$$\int_a^b g' = g(b) - g(a).$$

Proof. Let P be a partition of $[a, b]$. For each k , apply the Mean Value Theorem to the interval $[x_{k-1}, x_k]$ to find a point $c_k \in (x_{k-1}, x_k)$ such that $(x_k - x_{k-1})g'(c_k) = g(x_k) - g(x_{k-1})$. Then

$$g(b) - g(a) = \sum_{k=1}^n (g(x_k) - g(x_{k-1})) = \sum_{k=1}^n g'(c_k)(x_k - x_{k-1}).$$

For each k we clearly have

$$\inf\{g'(x) : x \in [x_{k-1}, x_k]\} \leq g'(c_k) \leq \sup\{g'(x) : x \in [x_{k-1}, x_k]\}.$$

It follows that

$$L(g', P) \leq g(b) - g(a) \leq U(g', P).$$

Since this inequality holds for every partition P (and the middle quantity does not depend on P), we have

$$L(g') \leq g(b) - g(a) \leq U(g').$$

But $U(g') = L(g')$ since g' is integrable, so $\int_a^b g' = U(g') = g(b) - g(a)$. $\square$

Theorem 14.11 (The Fundamental Theorem of Calculus II). *Suppose that f is integrable on $[a, b]$ and define*

$$F(x) = \int_a^x f(t) dt.$$

Then F is continuous on $[a, b]$. If f is continuous at $c \in (a, b)$ then F is differentiable at c and

$$F'(c) = f(c).$$

Proof. Let M be such that $|g(x)| \leq M$ for all $x \in [a, b]$. For $x, y \in [a, b]$ we have

$$|F(y) - F(x)| = \left| \int_a^y f(t) dt - \int_a^x f(t) dt \right| = \left| \int_x^y f(t) dt \right| \leq \int_x^y |f(t)| dt \leq M|x - y|.$$

Therefore F is Lipschitz, so it is uniformly continuous on $[a, b]$.

Suppose that f is continuous at $c \in (a, b)$. Observe that

$$f(c) = \frac{1}{x - c} \int_c^x f(c) dt$$

and

$$\frac{F(x) - F(c)}{x - c} = \frac{1}{x - c} \int_c^x f(t) dt,$$

so

$$\frac{F(x) - F(c)}{x - c} - f(c) = \frac{1}{x - c} \int_c^x (f(t) - f(c)) dt.$$

Let $\varepsilon > 0$ be given. Since f is continuous at c , there exists a $\delta > 0$ such that $|f(t) - f(c)| < \varepsilon$ whenever $|t - c| < \delta$. For the integral above we only care about values of t between c and x , and if $|x - c| < \delta$ then certainly $|t - c| < \delta$. For such x it follows that

$$\left| \frac{F(x) - F(c)}{x - c} - f(c) \right| \leq \frac{1}{|x - c|} \int_c^x |f(t) - f(c)| dt < \frac{1}{|x - c|} \int_c^x \varepsilon dt = \varepsilon.$$

Therefore $F'(c) = f(c)$. $\square$

One neat consequence of FTC II is that every integrable function has an antiderivative, specifically $F(x) = \int_a^x f(t) dt$ is the antiderivative of f . This holds even if you can't find a nice expression for an antiderivative of f .

REFERENCES

- [1] Stephen Abbott. *Understanding analysis*. Undergraduate Texts in Mathematics. Springer, New York, second edition, 2015.
- [2] Kenneth A. Ross. *Elementary analysis*. Undergraduate Texts in Mathematics. Springer, New York, second edition, 2013. The theory of calculus, In collaboration with Jorge M. López.
- [3] Terence Tao. *Analysis. I*, volume 37 of *Texts and Readings in Mathematics*. Hindustan Book Agency, New Delhi, third edition, 2014.