Contents

1 Sets **1**

1.1 Functions and Relations 3

1.2 Mathematical Induction 5

1.3 Equivalence of Sets and Countability 6

1.4 The Real Numbers 8

2 Sequences and Convergence **13**

2.1 The Arithmetic of Sequences 14

2.2 Subsequences 18

2.3 Convergence of Sequences and Series 21

2.4 Open sets, Closed sets 24

3 Limits of Functions **29**

4 Continuitiy **35**

4.1 Uniform Continuity 37

4.2 Properties of Continuous Functions 39

5 Differentiation **43**

5.1 The Mean Value Theorem 46

5.2 L'Hôpital's Rule 48

6 The Riemann Integral **51**

6.1 Integrability 52

6.2 Arithmetic of Integrals 56
6.3 The Fundamental Theorem 59

Chapter 1

Sets

A set is a colletion of mathematical objects. Sets, themselves, are mathematical objects. One way to think of sets is as 'jars' or 'containers' that have mathematical objects, possibly including other sets, inside of them. If X is a set and x is some mathematical object, we write $x \in X$ to denote the fact that x is in the collection X. We read this as * x is an element of X " or as " x is in X". If this is not the case, we write, instead, $x \notin X$ and read it as "x is not in X". We say two sets are equal if they have exactly the same elements. That is, $X = Y$ if and only if for every mathematical object x, we have $x \in X$ if and only if $x \in Y$.

There are two standard ways of describing sets. In the first, we simply list the elements of the set between braces ($\{\}$). So, for example, the empty set $\emptyset = \{\}$. This is the set that has no elements in it. We could also speak of the set $A = \{1, 2, 3, 4\}$ where $1 \in A$, but $5 \notin A$. Some care is needed since, for example, $\{\emptyset\}$ is a set with exactly one element in it (namely, the empty set), so $\emptyset \neq \{\emptyset\}$.

In this way, we have the set $\mathbb{N} = \{1, 2, 3, \cdots\}$ of all natural numbers and the set $\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$ of all integers.

The second way to describe a set is by considering all x that have some property $P(x)$. For this we write the set $A = \{x | P(x)\}$. Then $x \in A$ if and only if $P(x)$ is true. Now, in this $P(x)$ should be a statement that depends on x and

is either true or false for each mathematical object x. So, another description of the emptyset is $\emptyset = \{x | x \neq x\}$. In this, "$x \neq x$" is our property $P(x)$ and this is false for *all* objects x.

In this way, we can define $\mathbb{Q} = \{x : \text{ there are } m, n \in \mathbb{Z} \text{ with } x = \frac{m}{n} \text{ and } n \neq 0\}$, the set of all rational numbers. Then $\frac{1}{3} \in \mathbb{Q}$, for example.

Definition 1.1 *Let A and B be sets. We say A is a subset of B and write $A \subseteq B$ if for all objects x, $x \in A$ implies that $x \in B$. In other words, everything in A is in B.*

Directly from the definitions of equality of sets and of subsets, we have that $A = B$ if and only if both $A \subseteq B$ and $B \subseteq A$. This is a very common way to how that two sets are equal.

Some caution with language is a good thing at this point. If A and B are sets, the phrase "A is contained in B" could mean either $A \in B$ or $A \subseteq B$. These are very different concepts and it should always be made clear which of these two possibilities is meant. For example $\{1\} \subseteq \{1, 2\}$, but $\{1\} \notin \{1, 2\}$.

If X is a set, we define the power set of X by $\mathcal{P}(X) = \{A | A \subseteq X\}$. Hence, $\mathcal{P}(X)$ is the collection of all subsets of X. Notice that $\emptyset \in \mathcal{P}(X)$ and $X \in \mathcal{P}(X)$.

As an example, if $X = \{1, 2\}$, then $\mathcal{P}(X) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$. Notice that, here, we have $\{1\} \in \mathcal{P}(X)$, but $1 \notin \mathcal{P}(X)$.

Next if A and B are sets, we define $A \cap B = \{x | x \in A \text{ and } x \in B\}$ and $A \cup B = \{x | x \in A \text{ or } x \in B\}$. We call these sets the intersection and the union of A and B, respectively. We have the following:

Proposition 1.2 *Let A, B, ad C be sets. then*

1. $A \cap B = B \cap A$
2. $A \cup B = B \cup A$
3. $(A \cap B) \cap C = A \cap (B \cap C)$
4. $(A \cup B) \cup C = A \cup (B \cup C)$
5. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

6. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Next, if A and B are sets, we define the complement of B in A by $A \setminus B = \{x | x \in A \text{ and } x \notin B\}$. We then have the following:

Proposition 1.3 *Let A, B, and C be sets. Then,*

1. $A \setminus (A \setminus B) = A \cap B$
2. $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$
3. $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$
4. $(A \cup B) \setminus C = (A \setminus C) \cup (B \setminus C)$
5. $(A \cap B) \setminus C = (A \setminus C) \cap (B \setminus C)$

1.1 Functions and Relations

If x and y are mathematical objects, we define the ordered pair $(x, y) = \{\{x\}, \{x, y\}\}$. The only property of (x, y) that is required is the following:

Proposition 1.4 *We have that $(x, y) = (a, b)$ if and only if $x = a$ and $y = b$.*

A funtion is defined to be a set f of ordered pairs such that whenever $(x, y) \in f$ and $(x, z) \in f$, we may conclude that $y = z$. In other words, there is at most one second component for each first component for elements of f. In this case, if we have $(x, y) \in f$, we write $y = f(x)$.

If f is a function, we define the domain of f by

$$\text{dom } f = \{x | \text{ there is a } y \text{ such that } (x, y) \in f\}$$

and the image of f by

$$\text{im } f = \{y | \text{there is an } x \text{ such that } (x, y) \in f\}.$$

If, now, $A = \text{dom } f$ and $\text{im } f \subseteq B$, we say that f is a function from the set A to the set B and write $f : A \to B$. Notice that the domain of f must be the

set A, but the image does not have to be the set B. The image need only be a *subset* of B.

Now, let A and B be sets, and define the cross product $A \times B = \{(x, y) : x \in A, y \in B\} = \{z : \text{ there are } x \in A \text{ and } y \in B \text{ such that } z = (x, y)\}$. In particular, if $f : A \to B$, then $f \subseteq A \times B$.

Now suppose that Λ is a set and f is a function with $\Lambda = \text{dom } f$ such that $f(\lambda)$ is a set for each $\lambda \in \Lambda$. Let's write A_λ for $f(\lambda)$. Then, we say that $\{A_\lambda\}_{\lambda\in\Lambda}$ is an indexed family of sets with index set Λ.

For example, we might take $\Lambda = \mathbb{N}$ and define $A_n = \{\frac{m}{n} : m \in \mathbb{Z}\}$. Then A_n consists of those fractions whose demoninators are n. It is very common for the index set of a family of sets to be either $\mathbb{N}$ or some subset of $\mathbb{N}$.

As another example, we might have $\Lambda = \mathbb{Z}$ and define $A_n = \{x \in \mathbb{Z} : x > n\}$. Then $A_0 = \mathbb{N}$, for example.

Now, suppose that $\{A_\lambda\}_{\lambda\in\Lambda}$ is an indexed family of sets. We define the union and intersection of this family by

$$\bigcup_{\lambda\in\Lambda} A_\lambda = \{x| \text{ there exists a } \lambda \in \Lambda \text{ with } x \in A_\lambda\}$$

and

$$\bigcap_{\lambda\in\Lambda} A_\lambda = \{x|x \in A_\lambda \text{ for all } \lambda \in \Lambda\}.$$

As an example, suppose that $\Lambda = \mathbb{N}$ and $A_n = \{\frac{m}{n}|m \in \mathbb{Z}\}$. We claim that $\bigcup_{n\in\mathbb{N}} A_n = \mathbb{Q}$ and $\bigcap_{n\in\mathbb{N}} A_n = \mathbb{Z}$.

As anothr example, if $\Lambda = \mathbb{Z}$ and $A_n = \{x \in \mathbb{Z}|x > n\}$, then $\bigcup_{n\in\mathbb{Z}} A_n = \mathbb{Z}$ and $\bigcap_{n\in\mathbb{Z}} A_n = \emptyset$.

In general, we write $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n\in\mathbb{N}} A_n$ whenever $\{A_n\}_{n\in\mathbb{N}}$ is a family of sets indexed by $\mathbb{N}$ with a similar notation for intersections.

We now have:

Proposition 1.5 *Suppose that X is a set and that $\{A_\lambda\}_{\lambda\in\Lambda}$ is an indexed family of sets. Then*

a) $X \setminus (\bigcup_{\lambda\in\Lambda} A_\lambda) = \bigcap_{\lambda\in\Lambda}(X \setminus A_\lambda)$

b) $X \setminus (\bigcap_{\lambda\in\Lambda} A_\lambda) = \bigcup_{\lambda\in\Lambda}(X \setminus A_\lambda)$

Definition 1.6 *Suppose that $f : X \to Y$ is a function. We say that f is one-to-one if the only way for $f(x_1) = f(x_2)$ is for $x_1 = x_2$. Alternatively, a function is ne-to-one when $x_1 \neq x_2$ implies that $f(x_1) \neq f(x_2)$.*

Next, we say that f is onto if whenever $y \in Y$, there is an $x \in X$ with $y = f(x)$. In other words, $f : X \to Y$ is onto when im $f = Y$.

Finally, we say that f is a one-to-one correspondence (or an equivalence, or a bijection) if f is both one-to-one and onto.

Definition 1.7 *Suppose that $f : X \to Y$ and $g : Y \to Z$ are functions. We define the composition, $g \circ f : X \to Z$ by setting $g \circ f(x) = g(f(x))$ for all $x \in X$.*

Definition 1.8 *A relation is any set of ordered pairs. If R is a relation, say $R \subseteq X \times Y$, then we can define the opposite relation $\hat{R} = \{(x, y)|(y, x) \in R\}$. Notice that $\hat{\hat{R}} = R$.*

Proposition 1.9 *A function $f : X \to Y$ is one-to-one exactly when $\hat{f}$ is a function. In this case, dom $\hat{f}$ = im f and im $\hat{f}$ = dom f.*

In particular, if $f : X \to Y$ is a bijection, then $\hat{f} : Y \to X$ is also a bijection. In this case, we write f^{-1} for $\hat{f}$.

Proposition 1.10 *Suppose that $f : X \to Y$ and $g : Y \to Z$. then we have*

a) If both f and g are one-to-one, so is $g \circ f$.

b) If both f and g are onto, so is $g \circ f$.

c) If both f and g are bijections, so is $g \circ f$.

d) If $g \circ f$ is onto, so is g.

e) If $g \circ f$ is one-to-one, so is f.

1.2 Mathematical Induction

We assume a property of the natural numbers called well-ordering. It says that if $A \subseteq \mathbb{N}$ is non-empty, then there is a smallest element of A. In other words, there is an $a \in A$ such that $a \leq x$ for every $x \in A$.

Theorem 1.11 *Suppose that $P(n)$ is a property that a natural number can either have or not have. Suppose*

1. *$P(0)$ is true.*
2. *For every $n \in \mathbb{N}$, if $P(n)$ is true, then $P(n+1)$ follows.*

Then $P(n)$ is true for all $n \in \mathbb{N}$.

Proof: Let $A = \{n \in \mathbb{N} : P(n) \text{ is true}\}$. Then $0 \in A$, so A is not empty. We have to show that $A = \mathbb{N}$. Suppose this is not the case, so $\mathbb{N} \setminus A \neq \emptyset$. Let a be the smallest element of $\mathbb{N} \setminus A$. Then $a \neq 0$, so $a - 1 \in \mathbb{N}$. But then $a - 1 \in A$, so $P(a-1)$. By the second hypothesis, we can conclude that $P((a-1)+1)$ is true. This show that $a \in A$, contrary to the definition of a.

1.3 Equivalence of Sets and Countability

Definition 1.12 *We say two sets X and Y are equivalent if there exists a bijection $f : X \to Y$. In this case, we write $X \sim Y$.*

Theorem 1.13 *Let X, Y, and Z be sets. Then*

a) $X \sim X$.

b) If $X \sim Y$, then $Y \sim X$.

c) If $X \sim Y$ and $Y \sim Z$, then $X \sim Z$.

Definition 1.14 *We say a set A is finite if it is either empty or there is an $n \in \mathbb{N}$ such that $A \sim \{1, 2, \cdots, n\}$. We say that A is countably infinite if $A \sim \mathbb{N}$ and that A is countable if it is either finite or countably infinite.*

Proposition 1.15 *A set A is countable if and only if there is a subset $B \subseteq \mathbb{N}$ with $A \sim B$.*

Proof:

The forward direction is clear from the defintion of countability. So now suppose that $B \subseteq \mathbb{N}$ and that $A \sim B$. To prove the result, it is enough to show that either$B = \emptyset$, $B \sim \{1, \cdots, n\}$ for some $n \in \mathbb{N}$ or that $B \sim \mathbb{N}$.

If B is not empty, then there is a smallest element of B. We call this element b_1. Now, if $B = \{b_1\}$, then $B \sim \{1\}$, so we are done. Otherwise, there is a smallest element b_2 of $B \setminus \{b_1\}$.

If $B = \{b_1, b_2\}$, then $B \sim \{1, 2\}$, so we are done. We proceed by induction. Either this process stops at some point, in which case $B \sim \{1, \cdots, n\}$ for some $n \in \mathbb{N}$, or $B = \{b_1, b_2, \cdots\}$, so $B \sim \mathbb{N}$.

Proposition 1.16 *Subsets of countable sets are countable.*

Proposition 1.17 $\mathbb{N} \times \mathbb{N}$ *is countable.*

Proof:

Define $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by $f(m, n) = 2^m 3^n$. By unique factorization, f is one-to-one. Thus $\mathbb{N} \times \mathbb{N} \sim \text{im } f \subseteq \mathbb{N}$, so $\mathbb{N} \times \mathbb{N}$ is countable.

Theorem 1.18 *A set A is countable if and only if there is a one-to-one map $f : A \to \mathbb{N}$. Similarly, A is countable if and only if there is an onto map $f : \mathbb{N} \to A$.*

Proof: If $f : A \to \mathbb{N}$ is one-to-one, then $A \sim \text{ im } f \subseteq \mathbb{N}$, so A is countable. On the other hand, if $f : \mathbb{N} \to A$ is onto, then for each $a \in A$, there is an $n_a \in \mathbb{N}$ with $f(n_a) = a$. Define $g : A \to \mathbb{N}$ by $g(a) = n_a$. Then g is one-to-one, so A is countable. Conversely, if A is countable, the definition provides $f : A \to \mathbb{N}$ which is one-to-one. If A is countably infinite, then the definition also gives an f that is onto. Now, if A is finite, choose $f : \{1, \cdots, n\} \to A$ which is one-to-one and onto. Then define $F : \mathbb{N} \to A$ by $F(k) = f(k)$ if $k \leq n$ and $F(k) = f(1)$ if $k > n$. Then F is onto.

Theorem 1.19 *Supose that $\{A_n\}_{n=1}^{\infty}$ is a family of countable sets indexed by $\mathbb{N}$. Then $A = \bigcup_{n=1}^{\infty} A_n$ is countable.*

Proof: Choose, for each n, an onto function $f_n : \mathbb{N} \to A_n$. Now define $f : \mathbb{N} \times \mathbb{N} \to A$ by $f(n, m) = f_n(m)$. Then f is onto, so A is countable.

It is a good question whether every set is countable. That this is not the case follows from the following.

Proposition 1.20 *Let X be a set. Then $X \not\sim \mathcal{P}(X)$.*

Proof: Suppose that $f : X \to \mathcal{P}(X)$ is any map. We will show that f cannot be onto. Hence there is no bijection from X to $\mathcal{P}(X)$.

Let $A = \{x \in X : x \notin f(x)\}$. Then $A \subseteq X$, so $A \in \mathcal{P}(X)$. Suppose, though that $f(y) = A$. There are two choices for y. Either $y \in A$, in which case the definition of A shows that $y \notin f(y) = A$, which is a contradiction, or $y \notin A = f(y)$, n which case $y \in A$. Since either case leads to a contradiction, no such y can exist, so f is not onto.

We will later show that the set of real numbers, $\mathbb{R}$ is uncountable.

1.4 The Real Numbers

We assume the existence of a set $\mathbb{R}$ and two operations $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. We usually write $x + y$ and $x \cdot y$ for $+(x, y)$ and $\cdot(x, y)$, resp. We also assume there is a relation $< \subseteq \mathbb{R} \times \mathbb{R}$ where we write $x < y$ for $(x, y) \in <$. We further assume the following properties:

1. For all $x, y, z \in \mathbb{R}$, we have $(x + y) + z = x + (y + z)$.
2. For all $x, y \in \mathbb{R}$, we have $x + y = y + x$.
3. There is an element $0 \in \mathbb{R}$ such that $x + 0 = x$ for all $x \in \mathbb{R}$.
4. For each $x \in \mathbb{R}$, there is a $y \in \mathbb{R}$ such that $x + y = 0$.
5. For all $x, y, z \in \mathbb{R}$, we have $(x \cdot y) \cdot z = x \cdot (y \cdot z)$.
6. For all $x, y \in \mathbb{R}$, $x \cdot y = y \cdot x$.
7. There is an element $1 \in \mathbb{R}$ with $0 \neq 1$ and $x \cdot 1 = x$ for all $x \in \mathbb{R}$.

8. For each $x \neq 0$ in $\mathbb{R}$, there is a $y \in \mathbb{R}$ with $x \cdot y = 1$.

9. For all $x, y, z \in \mathbb{R}$, we have $x \cdot (y + z) = x \cdot y + x \cdot z$.

10. If $x < y$ and $z \in \mathbb{R}$, then $x + z < y + z$.

11. If $0 < x$ and $y < z$ in $\mathbb{R}$, then $x \cdot y < x \cdot z$.

12. If $x, y, z \in \mathbb{R}$ and $x < y$ and $y < z$, then $x < z$.

13. If $x, y \in \mathbb{R}$, exactly one of the following is true: $x < y$,$y < x$, or $x = y$.

We write $x \leq y$ if $x < y$ or $x = y$. Similarly, we write $x > y$ is $y < x$ and $x \geq y$ if $x > y$ or $x = y$. Notice that all the axioms so far are satisied by $\mathbb{Q}$ with the usual operations. This means that the final axiom is our only way to distinguish $\mathbb{Q}$ from $\mathbb{R}$.

Proposition 1.21 a) The element 0 above is unique.

b) For each x, the element y such that $x + y = 0$ is unique. We call it $-x$.

c) The element 1 aboce is unique.

d) For each $x \neq 0$, the element y with $x \cdot y = 1$ is unique. We call it $\frac{1}{x}$.

e) If $x < y$, then $-y < -x$.

f) If $0 < x < y$, then $0 < \frac{1}{y} < \frac{1}{x}$.

g) $0 < 1$.

h) If $x \in \mathbb{R}$, then $0 \leq x^2$.

Definition 1.22 *For $a < b$ in $\mathbb{R}$, define $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$. Similarly, define $(a, b), [a, b)$, and $(a, b]$. Also, define $[a, \infty) = \{x \in \mathbb{R} : a \leq x\}$. Similarly define $(a, \infty), (-\infty, a)$, and $(-\infty, a]$.*

To understand our final axiom, we need some terminology. Suppose that $\emptyset \neq A \subseteq \mathbb{R}$. We say that A is bounded above if there is an $x \in \mathbb{R}$ such that $a \leq x$ for all $a \in A$. We also say that x is an upper bound for A in this case.

We say that $x_0 \in \mathbb{R}$ is a least upper bound for A if it is an upper bound for A and if x is *any* upper bound for A, we have $x_0 \leq x$. In other words, x_0 is the smallest possible upper bound for A. Our axiom is then

Least Upper Bound Property: Every non-empty subset of $\mathbb{R}$ with an upper bound in $\mathbb{R}$ has a least upper bound in $\mathbb{R}$.

Proposition 1.23 *For each $x \in \mathbb{R}$, there is an $n \in \mathbb{N}$ with $x \leq n$.*

Proof: Assume not. Then there is an $x \in \mathbb{R}$ with $n < x$ for all $n \in \mathbb{N}$. In other words, x is an upper bound for $\mathbb{N}$. By our axiom, there is a least upper bound for $\mathbb{N}$. Call it x_0. Then $n \leq x_0$ for all $n \in \mathbb{N}$. But, since $x_0 - 1 < x_0$, x_0 is NOT an upper bound for $\mathbb{N}$. This means that there is an $n_0 \in \mathbb{N}$ with $x_0 - 1 < n_0$. But then $x_0 < n_0 + 1$, which contradicts that x_0 is an upper bound for $\mathbb{N}$.

Proposition 1.24 *If $x \in \mathbb{R}$, there is an $n \in \mathbb{Z}$ such that $n \leq x < n + 1$.*

Proof: First assume that $x \geq 0$. By the previous result, there is an $n \in \mathbb{N}$ with $x < n$. By the fact that $\mathbb{N}$ is well-ordered, there is a smallest $n_0 \in \mathbb{N}$ such that $x < n_0$. But then $n_0 - 1 \leq x$, so $n = n_0 - 1$ works in the result.

Now suppose that $x < 0$. Then $-x > 0$, so there is an $n_0 \in \mathbb{Z}$ with $n_0 \leq -x < n_0 + 1$. But then $-n_0 - 1 < x \leq -n_0$. If $x \neq -n_0$, let $n = -n_0 - 1$. Otherwise, let $n = -n_0$.

Proposition 1.25 *Let $a < b$ in $\mathbb{R}$. Then, there is an $r \in \mathbb{Q}$ with $a < r < b$.*

Proof: First, $b - a > 0$, so $1/(b-a) > 0$, so there is an $n \in \mathbb{N}$ with $1/(b-a) < n$. now find $m \in \mathbb{Z}$ such that $m < n \cdot b \leq m + 1$. Then we have $1 < n \cdot (b - a) = nb - na$, so $na < nb - 1 \leq m < nb$, so $a < m/n < b$. let $r = m/n$.

Proposition 1.26 *There is no $x \in \mathbb{Q}$ with $x^2 = 2$ (here, $2 = 1 + 1$ by definition). There is an $x \in \mathbb{R}$ with $x^2 = 2$. In fact, for each $a \geq 0$, there is an $x \geq 0$ with $x^2 = a$.*

Proof: First assume that $1 \leq a$. Consider the set $A = \{x \in \mathbb{R} : x > 0, x^2 \leq a\}$. Notice that $1 \in A$, so A is non-empty. Also, suppose that $x \in \mathbb{R}$. Then either

$x < a$, $x = a$, or $x > a$. In the last case, though $x^2 = x \cdot x > x \cdot a > a \cdot a \geq 1 \cdot a = a$, so $x \notin A$. In other words, $x \in A$ implies that $x \leq a$, so a is an upper bound for A. Using our axiom, let x_0 be the least upper bound for A. The claim is that $x_0^2 = a$. To show this, we note that the only other possibilities are that $x_0^2 < a$ or $x_0^2 > a$. We eliminate these two possibilities.

Case I: $x_0^2 < a$.

Let $\varepsilon = \min\{1, (a - x_0^2)/(2x_0 + 1)\}$. Then $\varepsilon > 0$,so $x_0 + \varepsilon > x_0$. But $(x_0+\varepsilon)^2 = x_0^2+2x_0\varepsilon+\varepsilon^2 \leq x_0^2+2x_0\varepsilon+\varepsilon \leq x_0^2+(2x_0+1)\varepsilon \leq x_0^2+(a-x_0^2) = a$. This means that $x_0 + \varepsilon \in A$, which contradicts the fact that x_0 is an upper bound for A.

Case II: $x_0^2 > a$.

In this case, let $\varepsilon = (x_0^2 - a)/(2x_0$. Then $(x_0 - \varepsilon)^2 = x_0^2 - 2x_0\varepsilon + \varepsilon^2 > x_0^2-2x_0\varepsilon = x_0^2-(x_0^2-a) = a$. Now, if $x > x_0-\varepsilon$, we have that $x^2 > (x_0-\varepsilon)^2 > a$, so $x \notin A$. in other words, if $x \in A$, $x \leq x_0 - \varepsilon$. This shows that $x_0 - \varepsilon$ is an upper bound for A. But, $x_0 - \varepsilon < x_0$, which contradicts that x_0 is the least upper bound for A.

The only other possibility is that $x_0^2 = a$. This is our result for $a \geq 1$.

Now, if $a = 0$, $x = 0$ works for $x^2 = a$. If $0 < a < 1$, then $1/a > 1$, so by the previous case there is an x_0 with $x_0^2 = 1/a$. But now $(1/x_0)^2 = a$, so we have the result in this case also.

Proposition 1.27 *Let $x < y$ in $\mathbb{R}$. Then there is a $z \in \mathbb{R} \setminus \mathbb{Q}$ such that $x < z < y$.*

Proof: Since $x < y$, $x - \sqrt{2} < y - \sqrt{2}$. By a previous result, there is a $r \in \mathbb{Q}$ with $x - \sqrt{2} < r < y - \sqrt{2}$. Then $x < r + \sqrt{2} < y$. Then, with $z = r + \sqrt{2}$, we have the result.

Finaly, we make a definition.

Definition 1.28 *For $x \in \mathbb{R}$, we define $|x|$ to be x if $x \geq 0$ and to be $-x$ if $x < 0$. Notice that $|x| \geq 0$ in every case and that $-|x| \leq x \leq |x|$.*

Proposition 1.29 *We have the folowing:*

a) $|xy| = |x| \cdot |y|$

b) $|x + y| \leq |x| + |y|$

c) $|x| < a$ is equivalent to $-a < x < a$.

d) If $b \in \mathbb{R}$, then $|x - b| < a$ is equivalent to $b - a < x < b + a$.

Chapter 2

Sequences and Convergence

Definition 2.1 *A sequence is a function with domain* $\mathbb{N}$*. We will be concerned with real-valued sequences, which are functions* $a : \mathbb{N} \to \mathbb{R}$*. Instead of* $a(n)$*, it is traditional to write* a_n *and to write* $\{a_n\}_{n=1}^{\infty}$ *for the sequence. Typically, a formula is given for* a_n *in terms of* $n \in \mathbb{N}$*.*

For example, if we let $a = \{\frac{1}{n}\}_{n=1}^{\infty}$, we have $a_n = \frac{1}{n}$, so for example $a_3 = \frac{1}{3}$. As another example, let $b = \{1 + (-1)^n\}_{n=1}^{\infty}$. Then $b_1 = 0, b_2 = 2, b_3 = 0, \cdots$.

Definition 2.2 *Let* $\{a_n\}_{n=1}^{\infty}$ *be a sequence and* $A \in \mathbb{R}$*. We say that* $\{a_n\}_{n=1}^{\infty}$ *converges to* A *if for every* $\varepsilon > 0$*, there is an* $N \in \mathbb{N}$ *such that whenever* $n \geq N$*, we have* $|a_n - A| < \varepsilon$*. This will be written as* $a_n \to A$*.*

As an example, we show that $\{\frac{1}{n}\}_{n=1}^{\infty}$ converges to 0. Suppose that $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that $\frac{1}{\varepsilon} < N$. Then, whenever $n \geq N$, we have $|\frac{1}{n} - 0| = \frac{1}{n} \leq \frac{1}{N} < \varepsilon$.

On the other hand, we show that $\{\frac{1}{n}\}_{n=1}^{\infty}$ does NOT converge to 1. To do this, we need to negate the definition of convergence. Hence, we need to show that there is an $\varepsilon > 0$ such that for all $N \in \mathbb{N}$, there is an $n \geq N$ with $\frac{1}{n} - 1| \geq \varepsilon$. I take $\varepsilon = \frac{1}{2}$. For any $N \in \mathbb{N}$, pick $n \geq \max\{2, N\}$. Then $|\frac{1}{n} - 1| = 1 - \frac{1}{n} \geq 1 - \frac{1}{2} = \varepsilon$.

In this definition, ε represents how close we want the terms ot he sequence to be to A and N is how far out in the sequence we have to go to guarantee

that the sequence is within this degree of closeness to A. It is clear that if N works for a given ε, any larger natural number will also.

The first thing we want to do is show that a sequence cannot converge to two different real numbers simultaneously.

Proposition 2.3 *Suppose that the sequence $\{a_n\}_{n=1}^{\infty}$ converges to both A and B. Then $A = B$.*

Proof: Suppose not, in other words, suppose that $A \neq B$. Then $\varepsilon = |A-B|/3 > 0$. Since $\{a_n\}_{n=1}^{\infty}$ converges to A, there is an $N_1 \in \mathbb{N}$ such that whenever $n|geqN_1$ we have $|a_n - A| < \varepsilon$. SImilarly, there is an $N_2 \in \mathbb{N}$ such that whenever $n \geq N_2$, we have $|a_n - B| < \varepsilon$. Now, to obtain our contradiction, let $n \in \mathbb{N}$ be chosen so that $n \geq N_1, N_2$. We then have that

$$|A - B| \leq |A - a_n| + |a_n - B| \leq \varepsilon + \varepsilon = 2\varepsilon = \frac{2}{3}|A - B|.$$

This is a contradiction since $|A - B| > 0$.

Definition 2.4 *We say a sequence $\{a_n\}_{n=1}^{\infty}$ converges if there exists an $A \in \mathbb{R}$ such that $\{a_n\}_{n=1}^{\infty}$ converges to A. By the previous result, if such an A exists, it is unique.*

As another example, we show that the sequence $\{a_n = 1 + (-1)^n\}_{n=1}^{\infty}$ does not converge. In fact, suppose that $a_n \to A$ and let $\varepsilon = 1$. Then, there is an $N \in \mathbb{N}$ such that whenever $n \geq \mathbb{N}$ we have $|a_n - A| < \varepsilon$. Let n be any odd integer with $n \geq N$. Then $n + 1 \geq N$ also, so we have $|a_n - A|, |a_{n+1} - A| < \varepsilon$. But $a_n = 0$ and $a_{n+1} = 2$, so $|A| < \varepsilon$ and $|2 - A| < \varepsilon$. But this shows that $2 = |2| = |2 - A| + |A| < \varepsilon + \varepsilon = 1 + 1 = 2$, a contradiction.

We will find many more cases of sequences that fail to converge in this chapter.

2.1 The Arithmetic of Sequences

Proposition 2.5 *Suppose that $a_n \to A$ and $b_n \to B$. Then $a_n + b_n \to A + B$.*

Proof: Let $\varepsilon > 0$. Choose $N_1 \in \mathbb{N}$ such that whenever $n \geq N_1$, we have $|a_n - A| < \varepsilon/2$. Similarly, choose $N_2 \in \mathbb{N}$ such that whenever $n \geq N_2$, we have $|b_n - B| < \varepsilon/2$. Let $N - \max\{N_1, N_2\}$. Then, if $n|geqN$, we have $|(a_n + b_n) - (A + B)| = |a_n - A + b_n - B| \leq |a_n - A| + |b_n - B| < \varepsilon/2 + \varepsilon/2 = \varepsilon$.

It should be pointed out that essentially the same proof shows that $a_n - b_n \to A - B$.

Now let's consider what it will take to show that $a_n b_n \to AB$. Clearly, we have to consider the expression $|a_n b_n - AB|$. The quetion is how to make this small (i.e., less than ε) given the knowledge that $|a_n - A|$ and $|b_n - B|$ are both small. After some playing around, and remembering a trick used in the proof of the product rule for derivatives, we add and subtract $a_n B$ and argue that $|a_n b_n - AB| = |a_n b_n - a_n B + a_n B - AB| \leq |a_n| \cdot |b_n - B| + |B| \cdot |a_n - A|$. Now, in both terms we have some expressions that are small. In the second, the expression is multiplied by $|B|$, so we can deal with it by using $\varepsilon/|B|$ instead of ε in the defintion of convergence of a_n. But the first term has $|a_n|$, which depends on n and so cannot simply be divided off the ε. But, if we can show that $|a_n|$ never gets too large, we can still make things work. We give a definition first.

Definition 2.6 *We say a sequence $\{a_n\}_{n=1}^{\infty}$ is bounded if there exists $M \in \mathbb{R}$ such that $|a_n| \leq M$ for all $n \in \mathbb{N}$.*

As an example, the sequence $\{(-1)^n\}_{n=1}^{\infty}$ is bounded (we can take $M = 1$), but it not convergent. However, we do have the following positive result:

Proposition 2.7 *Suppose that $\{a_n\}_{n=1}^{\infty}$ converges. Then it is bounded.*

Proof: Suppose that $a_n \to A$. In the definition of convergence, take $\varepsilon = 1$ and find $N \in \mathbb{N}$ such that whenever $n \geq N$, we have $|a_n - A| < \varepsilon = 1$. Then, whenever $n \geq N$, we also have $|a_n| = |a_n - A + A| \leq |a_n - A| + |A| < 1 + |A|$. Now, if we let $M = \max\{|a_1|, |a_2|, \cdots, |a_{N-1}|, 1 + |A|\}$, we have that $|a_n| \leq M$ for *all* $n \in \mathbb{N}$.

Now we can prove

Proposition 2.8 *Suppose that $a_n \to A$ and $b_n \to B$. Then $a_n b_n \to AB$.*

Proof: Let $\varepsilon > 0$. Since $\{a_n\}_{n=1}^{\infty}$ converges (to A), it is bounded. Let $M \in \mathbb{R}$ with $|a_n| \leq M$ for all $n \in \mathbb{N}$. We may assume $M \neq 0$ (if it was, $M = 1$ works just as well).

Now, find $N_1 \in \mathbb{N}$ such that whenever $n \geq N_1$ we have $|a_n - A| < \varepsilon/2(|B| + 1)$. We use $|B| + 1$ instead of $|B|$ just in case $B = 0$. Also, pick $N_2 \in \mathbb{N}$ such that whenever $n \geq N_2$, we have $|b_n - B| < \varepsilon/2M$. Let $N = \max\{N_1, N_2\}$.

Then, for any $n \geq N$, we have $|a_n b_n - AB| = |a_n b_n - a_n B + a_n B - AB| \leq |a_n| \cdot |b_n - B| + |B| \cdot |a_n - A| < M \cdot \varepsilon/2M + |B| \cdot \varepsilon/2(|B| + 1) < \varepsilon/2 + \varepsilon/2 = \varepsilon$ and we are done.

If we try to apply this result to the sequence $\{\frac{(-1)^n}{n}\}_{n=1}^{\infty}$, by writing $a_n = (-1)^n$ and $b_n = \frac{1}{n}$, we find that it doesn't work because $\{(-1)^n\}_{n=1}^{\infty}$ doesn't converge. Nonetheless, we can show that $\{\frac{(-1)^n}{n}\}_{n=1}^{\infty}$ converges to 0 as follows:

Proposition 2.9 *Suppose that $\{a_n\}_{n=1}^{\infty}$ is bounded and $b_n \to 0$. Then $a_n b_n \to 0$.*

Proof: Let $\varepsilon > 0$ and $M \in \mathbb{R}$ such that $|a_n \leq M$ for all $n \in \mathbb{N}$. We may assume $M \neq 0$. Choose $N \in \mathbb{N}$ such that whenever $n \geq N$, we have $|b_n - 0| < eps/M$. Then, for these same n, we have $|a_n b_n - 0| = |a_n| \cdot |b_n| \leq M \cdot |b_n - 0| < M \cdot \varepsilon/M = \varepsilon$.

If $\{a_n\}_{n=1}^{\infty}$ is not bounded, this result fails. For example, take $a_n = n$ and $b_n = 1/n$. Then $b_n \to 0$, yet $a_n b_n \to 1$.

Of course, our next step is to show that division works for limits of sequences. We have to be sure both that the sequence we divide by is never 0 and that the limit is not 0 just to be sure everything we want is defined. Our first step is to show that reciprocals work well. So, suppose that $a_n \neq 0$ for all n and that $a_n \to A$ with $A \neq 0$. How do we show that $\frac{1}{a_n} \to \frac{1}{A}$? We have to consider $|\frac{1}{a_n} - \frac{1}{A}| = \frac{|A - a_n|}{|a_n| \cdot |A|}$. Now, the numerator converges to 0 by assumption, so if we can show that $\frac{1}{|a_n| \cdot |A|}$ is bounded, we will be done by the following result.

Proposition 2.10 *Let $\{a_n\}_{n=1}^{\infty}$ be a sequence. Then $a_n \to A$ if and only if $a_n - A \to 0$.*

This is an assigned homework problem.

Another useful, if easy result:

Proposition 2.11 *Suppose that* $|a_n| \leq b_n$ *for all* $n \in \mathbb{N}$ *and suppose that* $b_n \to 0$. *Then* $a_n \to 0$.

Proof: Suppose $\varepsilon > 0$ and choose $N \in \mathbb{N}$ such that $n \geq N$ gives $b_n - 0| < \varepsilon$. Then $|a_n - 0| = |a_n| \leq b_n < \varepsilon$.

Lemma 2.12 *Suppose that* $a_n \neq 0$ *for all* $n \in \mathbb{N}$ *and* $a_n \to A$ *with* $A \neq 0$. *Then* $\{\frac{1}{a_n}\}_{n=1}^{\infty}$ *is bounded.*

Proof: In the definition of convergence, choose $\varepsilon = |A|/2 > 0$ and find $N \in \mathbb{N}$ so that $n \geq N$ implies that $|a_n - A| < \varepsilon = |A|/2$. Then, for these n, we have $|A| = |A - a_n + a_n| \leq |a_n - A| + |a_n| < |A|/2 + |a_n|$, so $|a_n| > |A| - |A|/2 = |A|/2$. But then $|\frac{1}{a_n}| < 2/|A|$. Now, let $M = \max\{|\frac{1}{a_1}|, \cdots, |\frac{1}{a_{N-1}}|, 2/|A|\}$ to se that $|\frac{1}{a_n}| \leq M$ for all $n \in \mathbb{N}$.

Proposition 2.13 *Suppose that* $a_n \neq 0$ *for all* $n \in \mathbb{N}$ *and that* $a_n \to A$ *with* $A \neq 0$. *Then* $\frac{1}{a_n} \to \frac{1}{A}$.

Proof: We need to show that $\frac{1}{a_n} - \frac{1}{A} \to 0$. So, $|\frac{1}{a_n} - \frac{1}{A}| = \frac{|A - a_n|}{|a_n| \cdot |A|} = \frac{1}{A} \cdot \frac{1}{a_n} \cdot |a_n - A|$. Since $\{\frac{1}{a_n}\}_{n=1}^{\infty}$ is bounded and $\frac{1}{A}$ is constant, and $|a_n - A| \to 0$, we obtain our result.

Proposition 2.14 *Suppose that* $a_n \neq 0$ *for all* $n \in \mathbb{N}$ *and* $a_n \to A \neq 0$. *Suppose also that* $b_n \to B$. *Then* $\frac{b_n}{a_n} \to \frac{B}{A}$.

Proof: $\frac{b_n}{a_n} = b_n \cdot \frac{1}{a_n} \to B \cdot \frac{1}{A} = \frac{B}{A}$.

It is often the case that none of our results so far applies to a given sequence, but the sequence can be manipulated so that these results apply.

Example: Let $a_n = \frac{2n+1}{3n+2}$. Since neither $\{2n+1\}_{n=1}^{\infty}$ nor $\{3n+2\}_{n=1}^{\infty}$ converge, the result on division of convergent sequences does not apply. However, if we multiply and divide by $\frac{1}{n}$, we see that $\frac{2n+1}{3n+2} = \frac{2+\frac{1}{n}}{3+\frac{2}{n}}$. SInce $\frac{1}{n} \to 0$, we see

that $2 + \frac{1}{n} \to 2 + 0 = 2$ and $3 + \frac{2}{n} \to 3 + 2 \cdot 0 = 3$, so we find that $\frac{2n+1}{3n+2} \to \frac{2}{3}$. Much more complicated results can be obtained in this way.

Finally, we give a result showing how convergence of sequences relates to the order properties of $\mathbb{R}$.

Proposition 2.15 *Suppose that $a_n \to A$, $b_n \to B$ and $a_n \leq b_n$ for all $n \in \mathbb{N}$. Then $A \leq B$.*

Proof: Suppose not, that is $B < A$. Let $\varepsilon = (A - B)/2 > 0$. Find $N \in \mathbb{N}$ such that whenever $n \geq N$, we have both $|a_n - A| < \varepsilon$ and $|b_n - B| < \varepsilon$. Then, for every $n \geq N$, we have $b_n < B + \varepsilon = B + (A - B)/2 = (A + B)/2 = A - (A - B)/2 = A - \varepsilon < a_n$, in contradiction to our assumption.

It should be pointed out that the previous result is false if $\leq$ is replaced by $<$ throughout. In fact, let $a_n = 1 - \frac{1}{n}$ and $b_n = 1 + \frac{1}{n}$. Then $a_n < b_n$ for every n. But, $a_n \to 1$ and $b_n \to 1$, so we have equality of the limits.

2.2 Subsequences

Definition 2.16 *Suppose that $\{a_n\}_{n=1}^{\infty}$ is a sequence and $n_1 < n_2 < \cdots$ is a sequence of natural numbers. We say that $\{a_{n_k}\}_{k=1}^{\infty}$ is a subsequence of the sequence $\{a_n\}_{n=1}^{\infty}$.*

For example, suppose that $a_n = (-1)^n$. If we choose $n_k = 2k$, we get that $a_{n_k} = (-1)^{2k} = 1$ for all $k \in \mathbb{N}$. Alternatively, if we chose $m_k = 2k - 1$, then $a_{m_k} = -1$ for all $k \in \mathbb{N}$. So, even though the original sequence fails to converge, these particular subsequences converge. Since the original sequence is a subsequence of itself, it is clear that some subsequences fail to converge. A positive result in this direction is the following.

Proposition 2.17 *Suppose that $a_n \to A$. Then for any subsequence, $a_{n_k} \to A$.*

Proof: First notice that since $n_1 \in \mathbb{N}$, we have $1 \leq n_1$. Also, if $k \leq n_k$ for some $k \in \mathbb{N}$, we have that $k + 1 \leq n_k + 1 \leq n_{k+1}$. Thus, indiction shows that $k \leq n_k$ for every $k \in \mathbb{N}$.

Now, let $\varepsilon > 0$ and find $N \in \mathbb{N}$ such that whenever $n \geq N$, we have $|a_n - A| < \varepsilon$. Then, if $k \geq N$, we have $n_k \geq k \geq N$, so $|a_{n_k} - A| < \varepsilon$. Thus $a_{n_k} \to A$.

This is actually a very handy way to show that some sequences fail to converge. If we can find two subsequences that converge to different limits, the whole sequence cannot converge at all. For example, suppose that $a_n = (-1)^n(1 + \frac{1}{n})$. Then, it is easy to see that $a_{2k} = 1 + \frac{1}{2k} \to 1$ while $a_{2k+1} = -(1 + \frac{1}{2k+1}) \to -1$. Hence $\{a_n\}_{n=1}^{\infty}$ cannot converge.

Definition 2.18 *Let $\{a_n\}_{n=1}^{\infty}$ be a sequence and $A \in \mathbb{R}$. We say that A is a cluster point of $\{a_n\}_{n=1}^{\infty}$ if for all $\varepsilon > 0$, the set $\{n \in \mathbb{N} : |a_n - A| < \varepsilon\}$ is infinite.*

For example, -1 is a cluster point of the sequence $\{(-1)^n\}_{n=1}^{\infty}$. On the other hand, the sequence $\{n\}_{n=1}^{\infty}$ has no cluster points at all. More interestingly, we know that $\mathbb{Q}$ is a countable set, so there is an onto function $a : \mathbb{N} \to \mathbb{Q}$. This function is, by definition a sequence. Furthermore, if $A \in \mathbb{R}$ is *any* real number and $\varepsilon > 0$, we know that there are infinitely many rational numbers in the open interval $(A - \varepsilon, A + \varepsilon)$, so the set $\{n \in \mathbb{N} : |a_n - A| < \varepsilon\}$ is infinite. Hence *every* real number is a cluster point of this sequence!

Theorem 2.19 *A point A is a cluster point of the sequence $\{a_n\}_{n=1}^{\infty}$ if and only if there is a subsequence $\{a_{n_k}\}_{k=1}^{\infty}$ such that $a_{n_k} \to A$.*

Proof: First assume that there is a subsequence with $a_{n_k} \to A$. Let $\varepsilon > 0$ and choose N so that $k \geq N$ implies that $|a_{n_k} - A| < \varepsilon$. Then every n_k with $k \geq N$ is in the set $\{n \in \mathbb{N} : |a_n - A| < \varepsilon\}$, so this set is infinite. This A is a cluster point of the sequence $\{a_n\}_{n=1}^{\infty}$.

Conversely, suppose that A is a cluster point of $\{a_n\}_{n=1}^{\infty}$. We need to find a subsequence which converges to A. First, take $\varepsilon = 1$ in the definition of cluster point to see that there is an n_1 such that $|a_{n_1} - A| < 1$ (in fact, there are infinitely many!).

Now, suppose that $n_1 < n_2 < \cdots n_k$ have been chosen. Choose $\varepsilon = \frac{1}{k+1}$ in the definition of cluster point to see that the set $\{n \in \mathbb{N} : |a_n - A| < \frac{1}{k+1}\}$ is

infinite. In particular, there is an n_{k+1} in this set and with $n_k < n_{k+1}$. In this way, we inductively find $n_1 < n_2 < \cdots$ such that $|a_{n_k} - A| < \frac{1}{k}$ for all $k \in \mathbb{N}$. Since $\frac{1}{k} \to 0$, we see that $|a_{n_k} - A| \to 0$, so $a_{n_k} \to A$.

Theorem 2.20 (Bolzano-Weierstrass Theorem) *Every bounded sequence has a cluster point.*

Proof: Suppose that $\{a_n\}_{n=1}^{\infty}$ is a bounded sequence with $|a_n| \leq M$ for all $n \in \mathbb{N}$. Let $x_0 = -M$ and $y_0 = M$. Then $x_0 \leq a_n \leq y_0$ for all $n \in \mathbb{N}$. In particular, the set $\{n \in \mathbb{N} : x_0 \leq a_n \leq y_0\}$ is infinite.

Let $z_0 = (x_0 + y_0)/2$ and notice that $\{n \in \mathbb{N} : x_0 \leq a_n \leq y_0\} = \{n \in \mathbb{N} : x_0 \leq a_n \leq z_0\} \cup \{n \in \mathbb{N} : z_0 \leq a_n \leq y_0\}$. Hence at least one of the two sets on the right is infinite. If the first one is, let $x_1 = x_0$ and $y_1 = z_0$. If not, let $x_1 = z_0$ and $y_1 = y_0$. Then, we have that $\{n \in \mathbb{N} : x_1 \leq a_n \leq y_1\}$ is infinite and that $y_1 - x_1 = (y_0 - x_0)/2$.

If we assume that we have chosen $x_0 \leq \cdots x_k < y_k \leq \cdots y_0$ such that $\{n \in \mathbb{N} : x_k \leq a_n \leq y_k\}$ is infintie and such that $y_k - x_k = (y_0 - x_0)/2^k$, then let $z_k = (x_k + y_k)/2$ so that $z_k - x_k = y_k - z_k = (y_k - x_k)/2 = (y_0 - x_0)/2^{k+1}$. Now $\{n \in \mathbb{N} : x_k \leq a_n \leq y_k\} = \{n \in \mathbb{N} : x_k \leq a_n \leq z_k\} \cup \{n \in \mathbb{N} : z_k \leq a_n \leq y_k\}$, so at least one of the sets on the right is infinite. if it is the first, let $x_{k+1} = x_k$ and $y_{k+1} = z_k$, and otherwise let $x_{k+1} = z_k$ and $y_{k+1} = y_k$. In this way, the induction continues.

In this way, we obtain sequences $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ such that

$$x_0 \leq x_1 \leq \cdots x_k \cdots \leq \cdots \leq y_k \cdots \leq y_1 \leq y_0$$

such that the set $\{n \in \mathbb{N} : x_k \leq a_n \leq y_k\}$ is infinite and such that $y_k - x_k = (y_0 - x_0)/2^k$ for each $k \in \mathbb{N}$.

There are now two things to notice: the first is that the set $\{x_k : k \in \mathbb{N}\}$ is bounded above (by y_0, for example). The second is that $y_k - x_k = (y_0 - x_0)/2^k \to 0$.

Let $A = \sup\{x_k : k \in \mathbb{N}\}$. The claim is that A is a cluster point of the original sequence $\{a_n\}_{n=1}^{\infty}$. To see this, let $\varepsilon > 0$. Then $A - \varepsilon$ is not an upper

bound for $\{x_k : k \in \mathbb{N}\}$, so there is a $K_1 \in \mathbb{N}$ with $A - \varepsilon < x_{K_1} \leq A$. Notice that for all $k \geq K_1$, we have $A - \varepsilon < x_{K_1} leq x_k \leq A$.

Next, there is a K_2 such that $k \geq K_2$ implies that $y_k - x_k < \varepsilon$. Choose any $k \geq K_1, K_2$. Then, we have $A - \varepsilon < x_k < y_k = x_k + (y_k - x_k) < x_k + \varepsilon \leq A + \varepsilon$. In other words, $[x_k, y_k] \subseteq (A - \varepsilon, A + \varepsilon)$. But then, $\{n \in \mathbb{N} : x_k \leq a_n \leq y_k\} \subseteq \{n \in \mathbb{N} : |a_n - A| < \varepsilon\}$. Since the first set is infinite, so is the second. This proves that A is a cluster point, proving the result.

Corollary 2.21 (Sequential comleteness of bounded sets) *Every bounded sequence has a convergent subsequence.*

Theorem 2.22 *A bounded sequence converges if and only if every convergent subsequence converges to the same value.*

Proof: One direction is easy since if $\{a_n\}_{n=1}^{\infty}$ converges to A, every subsequence converges to A also. Hence every convergent subsequence converges to A.

Now, suppose that $\{a_n\}_{n=1}^{\infty}$ is a bounded sequence and that every convergent subsequence converges to A. We need to show that $a_n \to A$.

Assume not. Then, there is an $\varepsilon > 0$ such that for every $N \in \mathbb{N}$, there is a $n \geq N$ with $|a_n - A| \geq \varepsilon$. Pick n_1 such that $|a_{n_1} - A| \geq \varepsilon$. Next, pick $n_2 \geq n_1 + 1$ with $|a_{n_2} - A| \geq \varepsilon$. Continue in the way to find $n_1 < n_2 < \cdots$ with $|a_{n_k} - A| \geq \varepsilon$ for all $k \in \mathbb{N}$.

The problem is that the sequence $\{a_{n_k}\}_{k=1}^{\infty}$ probably does not converge. However, it is a bounded sequence, so it has a subsequence $\{a_{n_{k_j}}\}$ which does converge. But this is a subsequence of $\{a_n\}_{n=1}^{\infty}$ which does not converge to A, which violates our assumption.

2.3 Convergence of Sequences and Series

Definition 2.23 *We say a sequence $\{a_n\}_{n=1}^{\infty}$ is non-decreasing if $a_n \leq a_{n+1}$ for all $n \in \mathbb{N}$. Similarly, we say the sequence is non-increasing if $a_{n+1} \leq a_n$ for all $n \in \mathbb{N}$. Finally, we say $\{a_n\}_{n=1}^{\infty}$ is monotone if it is either non-decreasing or non-increasing.*

Theorem 2.24 *A non-decreasing sequence is convergent if and only if it is bounded.*

Proof: We know that every convergent sequence is bounded, so one direction is already known. For the converse, assume that $\{a_n\}_{n=1}^{\infty}$ is non-decreasing and bounded. Consider the set $\{a_n : n \in \mathbb{N}\}$. It is non-empty and bounded above, so it has a least upper bound. Let $A = \sup\{a_n : n \in \mathbb{N}\}$. We claim that $a_n \to A$.

In fact, suppose that $\varepsilon > 0$. Since $A - \varepsilon$ is not an upper bound, there is a $N \in \mathbb{N}$ such that $A - \varepsilon < a_N \leq A$. But then, for each $n \geq N$, we have $A - \varepsilon < a_N \leq a_n \leq A < A + \varepsilon$, so $|a_n - A| < \varepsilon$.

As an example, let $a_n = 1 + \frac{1}{2} + \cdots \frac{1}{2^n}$. Using induction, it is easy to see that $a_n = 2 - \frac{1}{2^n}$. Hence, $a_n \to 2$. But now, consider $b_n = 1 + \frac{1}{2} + \cdots \frac{1}{n!}$. Then $\{b_n\}_{n=1}^{\infty}$ is clearly non-decreasing and $b_n \leq a_n < 2$, so it is bounded also. Hence $\{b_n\}_{n=1}^{\infty}$ converges. This is a very useful way to show sequences correspoonding to infinite sums are convergent.

Definition 2.25 *We say that a sequence $\{a_n\}_{n=1}^{\infty}$ is Cauchy if whenever $\varepsilon > 0$, there is a $N \in \mathbb{N}$ such that whenever $n, m \geq N$, we have $|a_n - a_m| < \varepsilon$.*

The difference between convergence to a point and being Cauchy is that for $a_n \to A$, we know the limit point (A), and the terms of the sequence are getting close to A. For Cauchy sequences, we only know that the terms are getting closer to each other.

Proposition 2.26 *Every convergent sequence is Cauchy.*

Proof: Suppose that $a_n \to A$ and that $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that for all $n \geq N$ we have $|a_n - A| < \varepsilon/2$. Then, whenever $m, n \geq N$, we have $|a_n - a_m| = |(a_n - A) - (a_m - A)| \leq |a_n - A| + |a_m - A| < \varepsilon/2 + \varepsilon/2 = \varepsilon$.

Proposition 2.27 *A Cauchy sequence with a convergent subsequence is, itself, convergent.*

Proof: Suppose that $\{a_n\}_{n=1}^{\infty}$ is Cauchy and that $a_{n_k} \to A$. We show that $a_n \to A$.

Let $\varepsilon > 0$. Choose $N_1 \in \mathbb{N}$ such that $m, n \geq N$ implies that $|a_n - a_m| < \varepsilon/2$. Also, choose $N_2 \in \mathbb{N}$ such that $k \geq N_2$ implies that $|a_{n_k} - A| < \varepsilon/2$.

Now, let $n \neq N$ be any natural number. Choose any $k \geq N, N_2$. Then $n_k \geq N$ so $|a_n - A| \leq |a_n - a_{n_k}| + |a_{n_k} - A| < \varepsilon/2 + \varepsilon/2 = \varepsilon$ and we are done.

Proposition 2.28 *Every Cauchy sequence is bounded.*

Proof: Suppose that $\{a_n\}_{n=1}^{\infty}$ is a Cauchy equence. Choose $\varepsilon = 1$ in the definition of Cauchy to find $N \in \mathbb{N}$ such that whenever $m, n \geq N$ we have $|a_n - a_m| < 1$. Then, for each $n \geq N$ we have $|a_n| \leq |a_N| + |a_N - a_n| < |a_N| + 1$. Now let $M = \max\{|a_1|, \cdots, |a_{N-1}|, |a_N| + 1\}$ to see that $|a_n| \leq M$ for all $n \in \mathbb{N}$.

Theorem 2.29 *A sequence is convergent if and only if it is Cauchy.*

Proof: We have shown that convergent sequences are Cauchy. Conversely, A Cauchy sequence is bounded, and so has a convergent subsequence. But then, the original sequence must converge.

While the previus proof was easy, it required the very difficult Bolzano-Weiestrass theorem for its proof. This makes it a major result.

Definition 2.30 *Let $\{a_n\}_{n=1}^{\infty}$ be a sequence. We can construct another sequence $\{s_n\}_{n=1}^{\infty}$ of 'partial sums' by setting $s_n = a_1 + a_2 + \cdots + a_n$. We say the series $\sum_{n=1}^{\infty} a_n$ converges if and only if $\{s_n\}_{n=1}^{\infty}$ converges. Furthermore, if $s_n \to A$, we write $\sum_{n=1}^{\infty} a_n = A$.*

Proposition 2.31 *If $\sum_{n=1}^{\infty} a_n$ converges, then $a_n \to 0$.*

Proof: Let $\{s_n\}_{n=1}^{\infty}$ be the sequence of partial sums and assume that $s_n \to A$. Then $a_n = s_n - s_{n-1} \to A - A = 0$.

Proposition 2.32 *Suppose that $a_n \geq 0$ for all $n \in \mathbb{N}$. Then $\sum_{n=1}^{\infty} a_n$ converges if and only if the sequence of partial sums is bounded.*

Proof: Notice that the sequence of partial sums $\{s_n\}_{n=1}^{\infty}$ is non-decreasing since $a_n \geq 0$. Hence, it converges if and only if it is bounded.

Proposition 2.33 *Supose that* $|a_n| \leq b_n$ *and that* $\sum_{n=1}^{\infty} b_n$ *converges. Then* $\sum_{n=1}^{\infty} a_n$ *converges.*

Proof: Let $\{s_n\}_{n=1}^{\infty}$ be the sequence of partial sums for $\sum_{n=1}^{\infty} a_n$ and $\{t_n\}_{n=1}^{\infty}$ the sequence of partial sums for $\sum_{n=1}^{\infty} b_n$. Then, since $\{t_n\}_{n=1}^{\infty}$ converges, it is Cauchy. We show that $\{s_n\}_{n=1}^{\infty}$ is also Cauchy, so it, in turn, converges.

Let $\varepsilon > 0$. Choose N so that $m, n \geq N$ implies that $|t_n - t_m| < \varepsilon$. We may assume $m < n$. But then $|s_n - s_m| = |a_{m+1} + \cdots a_n| \leq |a_{m+1}| + \cdots |a_n| \leq b_{m+1} + \cdots b_n = |t_n - t_m| < \varepsilon$ and we are done.

Theorem 2.34 (Alternating Series Test) *Suppose that* $\{a_n\}_{n=1}^{\infty}$ *is non-increasing with* $a_n \geq 0$ *for all* n. *Then* $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ *converges if and only if* $a_n \to 0$.

Proof: Let $\{s_n\}_{n=1}^{\infty}$ be the sequence of partial sums and let $x_n = s_{2n-1}$ and $y_n = s_{2n}$. Then, $x_{n+1} = s_{2n+1} = S_{2n-1} - a_{2n} + a_{2n+1} \leq x_n$ and $y_{n+1} = s_{2n+2} = s_{2n} + a_{2n+1} - a_{2n+2} \geq s_{2n} = y_n$ since a_n is non-increasing. Thus, $\{y_n\}_{n=1}^{\infty}$ is non-decreasing and $\{x_n\}_{n=1}^{\infty}$ is non-increasing. Next, $y_n = s_{2n} = a_1 - a_2 + a_3 - \cdots - a_{2n} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \cdots - a_{2n} \leq a_1$, so $\{y_n\}_{n=1}^{\infty}$ is bounded above. Similarly, $\{x_n\}_{n=1}^{\infty}$ is bounded below. Hence, both $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ converge. Let $x_n \to A$ and $y_n \to B$. Then $x_n - y_n = a_{2n} \to 0$, so $A - B = 0$, so $A = B$. Thus, s_{2n} and s_{2n-1} both convege to A, so $s_n \to A$. In other words, our series converges.

As an example, $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ converrges even though $\sum_{n=1}^{\infty} \frac{1}{n}$ does not.

2.4 Open sets, Closed sets

Definition 2.35 *We say that a subset* $U \subseteq \mathbb{R}$ *is open if, for every* $x \in U$, *there is an* $\varepsilon > 0$ *such that* $(x - \varepsilon, x + \varepsilon) \subseteq U$. *We say a subset* $C \subseteq \mathbb{R}$ *is closed if* $\mathbb{R} \setminus C$ *is open.*

For example, let $a \in \mathbb{R}$ and set $U = (a, +\infty)$. We show that U is open. In fact, if $x \in U$, then $a < x$. Let $\varepsilon = x - a > 0$. Then $(x-\varepsilon, x+\varepsilon) = (a, 2x-a) \subseteq (a, +\infty) = U$. Similarly, $(-\infty, a)$ is also open.

Proposition 2.36 *Suppose that U and V are open sets. Then $U \cap V$ is open.*

Proof: Let $x \in U \cap V$. Then, there is an $\varepsilon_1 > 0$ such that $(x - \varepsilon_1, x + \varepsilon_2) \subseteq U$ and an $\varepsilon_2 > 0$ such that $(x - \varepsilon_2, x + \varepsilon_2) \subseteq V$. Let $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$. Then $(x - \varepsilon, x + \varepsilon) \subseteq U \cap V$.

In particular, if $a < b$, then $(a, b) = (a, +\infty) \cap (-\infty, b)$ is open. Hence open intervals are open.

Proposition 2.37 *Suppose that $\{U_\alpha\}_{\alpha \in I}$ is an indexed family of open sets. Then $\bigcup_{\alpha \in I} U_\alpha$ is open.*

Proof: Suppose that $x \in U = \bigcup U_\alpha$. Then there is an $\alpha \in I$ with $x \in U_\alpha$. Since U_α is open, ther eis an $\varepsilon > 0$ such that $(x - \varepsilon, x + \varepsilon) \subseteq U_\alpha \subseteq U$.

Thus, every union of open intervals is an open set.

Next, if $a \in \mathbb{R}$, then $[a, +\infty)$ is closed because $\mathbb{R} \setminus [a, +\infty) = (-\infty, a)$ is open. Also, $(-\infty, a]$ is closed. We have the following using DeMorgan's theorem.

Proposition 2.38 *Suppose that C and D are closed sets. Then $C \cup D$ is closed.*

Proposition 2.39 *Supose that $\{C_\alpha\}$ is an indexed family of closed sets. Then $\bigcap C_\alpha$ is closed.*

The investigation of open and closed sets is part of the subject matter of topology. It turns out to be very useful to use convergence of sequences to determine when a set is open or closed. The following provides a route to do this.

Theorem 2.40 *Let $C \subseteq \mathbb{R}$. The following are equivalent:*

a) C is closed.

b) Whenever $\{a_n\}_{n=1}^{\infty}$ is a sequence with $a_n \in C$ for all $n \in \mathbb{N}$ and $a_n \to x$, then we can conclude that $x \in C$.

Proof: Suppose that C is closed and that $\{a_n\}_{n=1}^{\infty}$ is a sequence with $a_n \in C$ and $a_n \to x$. Suppose also that $x \notin C$. Then $x \in \mathbb{R} \setminus C$, which is an open set. Hence, there is an $\varepsilon > 0$ with $(x-\varepsilon, x+\varepsilon) \subseteq \mathbb{R} \setminus C$. Since $a_n \to x$, there is an $N \in \mathbb{N}$ such that whenever $n \geq N$, we have $|a_n - x| < \varepsilon$, which implies that $x - \varepsilon < a_n < x + \varepsilon$. But this implies that $a_n \in \mathbb{R} \setminus C$ for all $n \geq N$, in contradiction to our assumption.

For the converse, suppose that C is not closed. Then $\mathbb{R} \setminus C$ is not open, which means there is a $x \in \mathbb{R} \setminus C$ such that whenever $\varepsilon > 0$, the set $(x-\varepsilon, x+\varepsilon)$ is not contained in $\mathbb{R} \setminus C$. But this means that for each $\varepsilon > 0$, there is an $a \in C$ with $|a - x| < \varepsilon$. Now, for each $n \in \mathbb{N}$, choose $a_n \in C$ with $|a_n - x| < \frac{1}{n}$. Then, $a_n \to x$ and $a_n \in C$ for all $n \in \mathbb{N}$ but $x \notin C$. In other words, if C is not closed, the second condition fails also.

Definition 2.41 *Let $A \subseteq \mathbb{R}$ and let $x \in \mathbb{R}$. We say that x is an accumulation point of A if for each $\varepsilon > 0$, the set $A \cap (x-\varepsilon, x+\varepsilon)$ is infinite.*

As an example 1 is an accumulation point of the set $(0,1)$ since $(0,1) \cap (1-\varepsilon, 1+\varepsilon) = (1-\varepsilon, 1)$ is infinite for every $\varepsilon > 0$.

Similarly, *every* real number is an accumulation point of the set $\mathbb{Q}$ and of the set $\mathbb{R} \setminus \mathbb{Q}$.

Theorem 2.42 *Let $A \subseteq \mathbb{R}$ and $x \in \mathbb{R}$. Then the following are equivalent:*

a) x is an accumulation point of A.

b) For each $\varepsilon > 0$, there is an element of $A \cap (x-\varepsilon, x+\varepsilon)$ different from x.

c) There is a sequence $\{a_n\}_{n=1}^{\infty}$ of distinct points in A with $a_n \to x$.

Proof: a) implies b):Suppose that x is an accumulation point of A. Then for each $\varepsilon > 0$, the set $A \cap (x-\varepsilon, x+\varepsilon)$ is infinite, so has at least one point different from x.

c) implies a): If $\{a_n\}_{n=1}^{\infty}$ is a sequence of distinct points in A with $a_n \to x$, let $\varepsilon > 0$. Then, there is an $N \in \mathbb{N}$ such that $|a_n - x| < \varepsilon$ for all $n \geq N$. In other

words, the infinite set $\{a_n : n \geq N\} \subseteq A \cap (x-\varepsilon, x+\varepsilon)$. Thus $A \cap (x-\varepsilon, x+\varepsilon)$ is infinite for every $\varepsilon > 0$, so x is an accumulation point of A.

b) implies c): Let $\varepsilon_1 = 1$. There is an $a_1 \in A$ with $a_1 \neq x$ and $|a_1 - x| < 1$.

Now, let $\varepsilon_2 = \min\{\frac{1}{2}, |a_1 - x|\}$. Then $\varepsilon_2 > 0$ and so there is an $a_2 \neq x$ with $|a_2 - x| < \varepsilon_2$. Notice that $a_1 \neq a_2$ by construction.

Inductively, suppose we have $a_1, \cdots a_k \in A$ all distinct and different from x with $|a_j - x| < \frac{1}{j}$ for $1 \leq j \leq k$. Set $\varepsilon_{k+1} = \min\{\frac{1}{k+1}, |a_1 - x|, \cdots, |a_k - x|\}$. Then $\varepsilon_{k+1} > 0$ so there is an $a_{k+1} \in A$ with $a_{k+1} \neq x$ and $|a_{k+1} - x| < \varepsilon_{k+1}$. Then a_{k+1} is different than all the previous a_j and $|a_{k+1} - x| < \frac{1}{k+1}$, so the induction continues.

This gives us a sequence $\{a_n\}_{n=1}^{\infty}$ of distinct terms with $a_n \in A$ for all n and $a_n \to x$.

An important consequence of this is the following:

Proposition 2.43 *Assume $A \subseteq \mathbb{R}$ is bounded above. Then* $\sup A$ *is an element of A or an accumulation point of A.*

Proof: Suppose that $x = \sup A \notin A$. Then, for each $\varepsilon > 0$, there is an $a \in A$ with $x - \varepsilon < a < x < x + \varepsilon$. Hence, x is an accumulation point of A.

Theorem 2.44 (Bolzano-Weierstrass Theorem) *Every bounded infinie set has an accumulation point.*

Proof: Let $A \subseteq [-M, M]$ be infinite. Inductuvely, pick a sequence $\{a_n\}_{n=1}^{\infty}$ of distinct terms with $a_n \in A$ for all n. This sequence is then bounded, so it has a convergent subsequence. This subsequence also has distinct terms, so its limit is an accumulation point of A.

Theorem 2.45 *Let $A \subseteq \mathbb{R}$ and let A' be the set of accumulation points of A. Then $A \cup A'$ is closed.*

Proof: So, let $A \subseteq \mathbb{R}$ and set $B = A \cup A'$. We show that $\mathbb{R} \setminus B$ is open. So, suppose that $x \in \mathbb{R} \setminus B$. Then $x \notin A'$, so there is an $\varepsilon > 0$ so that $A \cap (x - \varepsilon, x + \varepsilon)$ contains no point other than x. But, since $x \notin A$, we must

have $A \cap (x-\varepsilon, x+\varepsilon) = \emptyset$. But then, no point of $(x-\varepsilon, x+\varepsilon)$ can be an accumulation point of A, so $A' \cap (x-\varepsilon, x+\varepsilon) = \emptyset$ also. In other words, $(x-\varepsilon, x+\varepsilon) \cap B = \emptyset$, so $(x-\varepsilon, x+\varepsilon) \subseteq \mathbb{R} \setminus B$. This shows that $\mathbb{R} \setminus B$ is open, so B is closed.

Corollary 2.46 *Let $A \subseteq \mathbb{R}$. Then A is closed if and only if $A' \subseteq A$.*

Proof: If A is closed and $x \in A'$, then there is a sequence $\{a_n\}_{n=1}^{\infty}$ in A of distinct terms with $a_n \to x$. By our characterization of closed sets via sequences, $x \in A$. Hence $A' \subseteq A$.

Conversely, if $A' \subseteq A$, then $A = A \cup A'$ is closed.

Corollary 2.47 *Suppose A is closed and bounded above. Then* $\sup A \in A$.

Exercises:

1. Suppose that $0 < r < 1$. Show that the sequence $\{r^n\}_{n=1}^{\infty}$ converges. Then find its limit. What happens if $-1 < r < 0$?

2. Suppose that $1 < r$. Show that $\{r^n\}_{n=1}^{\infty}$ is unbounded. Hint: if, not, show it converges. What would it have to converge to?

3. Suppose that $r \neq 1$. Show that $1 + r + r^2 + r^3 + \cdots + r^{n-1} = \frac{1-r^n}{1-r}$.

4. Suppose that $|r| < 1$. Show that $\sum_{n=1}^{\infty} r^{n-1} = \frac{1}{1-r}$.

5. Suppose that $0 < a, b$. Show that $\lim_{n\to\infty}(a^n + b^n)^{1/n} = \max\{a, b\}$.

6. Let $x_1 = 1$ and $x_{n+1} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right)$. Show that $|x_{n+1}^2 - 2| \leq \frac{1}{4}|x_n^2 - 2|^2$. From this, show that $x_n \to \sqrt{2}$.

Chapter 3

Limits of Functions

In this chapter, we want to consder limits of functions. In particular, we want to justify calculations like those done in elementary classes such as

$$\begin{aligned}\lim_{x\to 2}\frac{x^2-4}{x-2} &= \lim_{x\to 2}\frac{(x-2)(x+2)}{x-2}\\ &= \lim_{x\to 2} x+2\\ &= 4.\end{aligned}$$

Sevral things should be pointed out in this calculation. First, and most significantly, the function involved, $f(x) = (x^2-4)/(x-2)$ is not defined at the place we are approaching in the limit (i.e. $x = 2$). However, the point $x = 2$ is an accumulation point of the domain of f, i.e. $\{x \in \mathbb{R} : x \neq 2\}$. Next, the second inequality depends on the fact that the x-values we take are not equal to 2, while the last limit seems to 'plug in' $x = 2$. These characteristics of limits made the formal definition a relatively late occurence. We give it here.

Definition 3.1 *Let $f : D \to \mathbb{R}$ be a function with $D \subseteq \mathbb{R}$ and suppose that a is an accumulation point of D. We say that f has a limit $L \in \mathbb{R}$ as x approaches a and write $\lim_{x\to a} f(x) = L$ if for all $\varepsilon > 0$, there is a $\delta > 0$ such that whenever $x \in D$ with $0 < |x-a| < \delta$, we have $|f(x) - L| < \varepsilon$.*

We say that the limit of f as x approaches a exists if there is an $L \in \mathbb{R}$ such that the limit of f as x approaches a is L.

Another thing to point out is that since we use $0 < |x - a| < \varepsilon$, the value of $f(a)$ is irrelevant for the existence of the limit even when $a \in D$.

As an example, we show that $\lim_{x\to 2} x^2 = 4$. Let $\varepsilon > 0$ and choose $\delta = \min\{1, \varepsilon/5\}$. Then $\delta > 0$. Implicitly in this example, we have $D = \mathbb{R}$, so 2 is an accumulation point of D. So, suppose that $x \in \mathbb{R}$ with $0 < |x - 2| < \delta$. Then, first, $|x| \leq |x - 2| + 2 < 1 + 2 = 3$, so $|x + 2| \leq 3 + 2 = 5$. Then, we have $|f(x) - L| = |x^2 - 4| = |(x - 2)(x + 2)| \leq |x - 2| \cdot 5 < \varepsilon$ as required.

Proposition 3.2 *Suppose that the limit of f as x approaches a is both L_1 and L_2. Then $L_1 = L_2$.*

In other words, limits are unique. We have implicitly used this when we write $\lim_{x\to a} f(x) = L$, so some care is required here.

Proof: Suppose $L_1 \neq L_2$ and set $\varepsilon = |L_1 - L_2|/2$. Then pick $\delta_1 > 0$ such that whenever $x \in D$ with $0 < |x - a| < \delta_1$, we have $|f(x) - L_1| < \varepsilon$. Smilarly, pick $\delta_2 > 0$ such that whenever $x \in D$ with $0 < |x - a| < \delta_2$, we have $|f(x) - L_2| < \varepsilon$. Let $\delta = \min\{\delta_1, \delta_2\}$. Since a is an accumulation point of D, there is an $x \in D$ with $x \neq a$ and $|x - a| < \delta$. For this x, we have $|f(x) - L_1| < \varepsilon$ and $|f(x) - L_2| < \varepsilon$. But this implies that $|L_1 - L_2| = |(f(x) - L_2) - (f(x) - L_1)| < \varepsilon + \varepsilon = 2\varepsilon = |L_1 - L_2|$, a contradction.

Examples:

a) We always have that $\lim_{x\to a} x = a$. In fact, if $\varepsilon > 0$, simply choose $\delta = \varepsilon$. Then, if $x \in D$ and $0 < |x - a| < \delta$, we have $|x - a| < \varepsilon$.

b) let $D = \mathbb{R}$. We show that $\lim_{x\to 3} x^2 + x = 12$. In fact, let $\varepsilon > 0$. Choose $\delta = \min\{1, \varepsilon/8\}$. Then, if $x \in D$ with $0 < |x - 3| < \delta$, we have first of all that $|x + 4| = |(x - 3) + 7| \leq |x - 3| + 7 < 1 + 7 = 8$. Now, this shows that $|(x^2 + x) - 12| = |(x - 3)(x + 4)| = |x - 3| \cdot |x + 4| < 8|x - 3| < \varepsilon$.

c) Let $D = \mathbb{R} \setminus \{0\}$ and let $f(x) = \cos(\frac{1}{x})$. Notice that 0 is an accumulation point of D. We claim that f does not have a limit as x approaches 0. In fact, suppose to the contrary that $\lim_{x\to 0} f(x) = L$ and choose $\varepsilon = 1$. Then, there would be a $\delta > 0$ such that whenever $0 < |x - 0| < \delta$, we

have $|f(x) - L| < \varepsilon$. But, we have that the sequence $\frac{1}{\pi n} \to 0$, so there is an $N \in \mathbb{N}$ such that whenever $n \geq N$, we have $|\frac{1}{\pi n} - 0| < \delta$. But then one of N and $N + 1$ is odd and the other is even. Thus one of $\cos(\pi N)$ and $\cos(\pi(N + 1))$ is 1 and the other is -1. letting $x = \frac{1}{\pi N}$ and $x = \frac{1}{\pi(N+1)}$ shows that $|L - 1| < \varepsilon$ and $|L - (-1)| < \varepsilon$. But then, $2 = |1 - (-1)| \leq |1 - L| + |L - (-1)| < 2\varepsilon = 2$. This is a contradiction.

d) Let $D = (0, 1)$ and define $f(x) = 0$ if x is irrational and $f(x) = \frac{1}{q}$ if $x = \frac{p}{q}$ is rational and in lowest terms. We claim that $\lim_{x\to a} f(x) = 0$ for all $a \in [0, 1]$.

In fact, let $a \in [0, 1]$ and let $\varepsilon > 0$. There is an $N \in \mathbb{N}$ such that $\frac{1}{N} < \varepsilon$. Now consider the set of fractions of the form $\frac{p}{q}$ where $q \leq N$. There are only finitely many such fractions. Let $\delta = \min\{|\frac{p}{q} - a| : q \leq N, a \neq \frac{p}{q}\}$ be the smallest distance from a to one of these fractions. Then $\delta > 0$. Now suppose that $x \in D$ and $0 < |x - a| < \delta$. Either x is irrational, in which case $|f(x) - 0| = |0 - 0| = 0 < \varepsilon$, or x is rational, but $x = \frac{p}{q}$ where $q > N$. In the latter case, $|f(x) - 0| = |\frac{1}{q} - 0| = \frac{1}{q} < \frac{1}{N} < \varepsilon$.

At this point, we can prove all of the algebraic properties of limits for functions in a way that is similar to how we showed them for sequences in the last chapter. The following result is an example of this technique.

Proposition 3.3 *Let $f, g : D \to \mathbb{R}$ and a an accumulation point of D. Suppose that $\lim_{x\to a} f(x) = L_1$ and $\lim_{x\to a} g(x) = L_2$. Then $\lim_{x\to a} f(x) + g(x) = L_1 + L_2$.*

Proof: Let $\varepsilon > 0$. Choose $\delta_1 > 0$ such that whenever $x \in D$ with $0 < |x-a| < \delta_1$ we have $|f(x) - L_1| < \varepsilon/2$. Similarly, find $\delta_2 > 0$ such that whenever $x \in D$ with $0 < |x - a| < \delta_2$ we have $|g(x) - L_2| < \varepsilon/2$. let $\delta = \min\{\delta_1, \delta_2\}$, so that $\delta > 0$. Now, if $x \in D$ and $0 < |x - a| < \delta$, we have $|f(x) + g(x) - (L_1 + L_2)| \leq |f(x) - L_1| + |g(x) - L_2| < \varepsilon/2 + \varepsilon/2 = \varepsilon$, as required.

Just as for sequences, we can define boundedness of functions:

Definition 3.4 *Let $f : D \to \mathbb{R}$. We say that f is bounded if there is an $M \in \mathbb{R}$ such that $|f(x)| \leq M$ for all $x \in D$.*

Proposition 3.5 *Suppose that $f : D \to \mathbb{R}$ and a is an accumulation point of D where $lim_{x\to a} f(x)$ exists. Then, there is an $\delta > 0$ such that f is bounded on the set $D \cap (a - \delta, a + \delta)$.*

It is possible to prove all the results for the algebraic properties of limits of functions via techniques similar to those used for the properties of limits of sequences. However, it turns out to be more helpful to *use* the results from sequences to do this. For this we need the following, very important result.

Theorem 3.6 *Let $f : D \to \mathbb{R}$ and a and accumulation point of D. The following are equivalent.*

a) The limit $\lim_{x\to a} f(x) = L$.

b) Whenever $\{a_n\}_{n=1}^{\infty}$ is a sequence in D with $a_n \neq a$ for all $n \in \mathbb{N}$ and $a_n \to a$, we have that $f(a_n) \to L$.

Proof: a) implies b): Suppose that $\lim_{x\to a} f(x) = L$. Also, suppose that $a_n \in D$, $a_n \neq a$, and $a_n \to a$. We need to show that $f(a_n) \to L$.

So let $\varepsilon > 0$. The, there is a $\delta > 0$ such that whenever $x \in D$ with $0 < |x - a| < \delta$, we have $|f(x) - L| < \varepsilon$. But then, there is an $N \in \mathbb{N}$ such that whenever $n \geq N$, we have $|a_n - a| < \delta$. But now, for these n, we have that $a_n \neq a$, so $0 < |a_n - a| < \delta$, which implies that $|f(a_n) - L| < \varepsilon$. In other words, $f(a_n) \to L$.

b) implies a): Suppose that f does not have a limit of L as $x \in D$ approaches a. Then, there is an $\varepsilon > 0$ such that for every $\delta > 0$, there is an $x \in D$ with $0 < |x - a| < \delta$ yet $|f(x) - L| \geq \varepsilon$. Let $\delta = \frac{1}{n}$ and find $x_n \in D$ with $0 < |x - a| < \frac{1}{n}$ but such that $|f(x_n) - L| \geq \varepsilon$. Then $x_n \to a$ with $x_n \neq a$ for all n, but $\{f(x_n)\}_{n=1}^{\infty}$ does not converge to L. This contradicts our assumption and shows that $\lim_{x\to a} f(x) = L$.

As an example of the use of this result, we show that limits of products work the 'right way'.

Proposition 3.7 *Let $f, g : D \to \mathbb{R}$ and a an accumulation point of D. Assume that $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$. Then $\lim_{x\to a} f(x)g(x) = LM$. Similarly, if $g(x) \neq 0$ for all $x \in D$ and $M \neq 0$, then $\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{L}{M}$.*

Proof: Suppose that $\{a_n\}_{n=1}^{\infty}$ in D with $a_n \neq a$ for all n. Then, $f(a_n) \to L$ and $g(a_n) \to M$, so by the result on products of sequences, $f(a_n)g(a_n) \to LM$. Since this happens for all sequences $\{a_n\}_{n=1}^{\infty}$, we have that $\lim_{x\to a} f(x)g(x)$ exists and equals LM. The result for quotients follows similarly.

The next result is a bit more subtle. Instead of assuming convergence to a specific point, it merely assumes convergence.

Theorem 3.8 *Assume $f : D \to \mathbb{R}$ and a is an accumulation point of D. The following are equivalent:*

a) The limit of f as x approaches a exists.

b) Whenever $\{a_n\}_{n=1}^{\infty}$ is a sequence in D with $a_n \neq a$ and $a_n \to a$, then the sequence $\{f(a_n)\}_{n=1}^{\infty}$ converges.

Proof: a) implies b): This follows directly from Theorem 3.6.

b) implies a): The difficulty here is that we only assume that the sequences $\{f(a_n)\}_{n=1}^{\infty}$ converge, not that they all converge to the same thing. Since a is an accumulation point of D, there is at least one sequence $\{a_n\}_{n=1}^{\infty}$ in D with $a_n \neq a$ for all n and $a_n \to a$. By assumption, $\{f(a_n)\}_{n=1}^{\infty}$ converges, say $f(a_n) \to L$.

Now, assume that $\{b_n\}_{n=1}^{\infty}$ is any other sequence in D with $b_n \neq a$ for all n and $b_n \to a$. We need to show that $f(b_n) \to L$ also (we only know this sequence converges, NOT that it converges to L). To do this, let $\{c_n\}_{n=1}^{\infty}$ be defined by $c_{2n-1} = a_n$ and $c_{2n} = b_n$, so c_n alternates between a_n and b_n. Then $c_n \in D$ and $c_n neq x$ for all n. Also, $c_n \to a$. By assumption $\{f(c_n)\}_{n=1}^{\infty}$ converges, say $f(c_n) \to L'$. But now, $\{f(a_n)\}_{n=1}^{\infty}$ is a subsequence of $\{f(c_n)\}_{n=1}^{\infty}$, so $f(a_n) \to L'$ also. Hence, $L = L'$. But now, $\{f(b_n)\}_{n=1}^{\infty}$ is also a subsequence of $\{f(c_n)\}_{n=1}^{\infty}$, so $f(b_n) \to L' = L$.

Now we can apply Theorem 3.6 again to see that $\lim_{x\to a} f(x) = L$.

Exercises:

1. Show that $\lim_{x\to 3} \frac{x^2-9}{x-3} = 6$ using the ε, δ definition.

2. Show that $\lim_{x\to 2} x^2 + 4x + 5 = 17$ using the ε, δ definition.

3. Suppose that $f, g : D \to \mathbb{R}$ and a is an accumulation point of D. Suppose that g is bounded and that $\lim_{x\to a} f(x) = 0$. Show that $\lim_{x\to a} f(x)g(x) = 0$ using the ε, δ definition.

4. Suppose that $f, g, h : D \to \mathbb{R}$ and a is an accumulation point of D. Suppose that $f(x) \leq g(x) \leq h(x)$ for all $x \in D$ and that $\lim_{x\to a} f(x) = \lim_{x\to a} h(x) = L$. Show that $\lim_{x\to a} g(x) = L$.

5. Suppose that $f : D \to \mathbb{R}$ and a is an accumulation point of D. Suppose also that for each $\varepsilon > 0$ there is a $\delta > 0$ such that whenever $x, y \in D$ with $x \neq a \neq y$ and $|x - y| < \delta$, then $|f(x) - f(y)| < \varepsilon$. Show that $\lim_{x\to a} f(x)$ exists. Hint: Suppose that $a_n \in D$ with $a_n \to a$. Show that $\{f(a_n)\}$ is Cauchy.

Chapter 4

Continuitiy

Definition 4.1 *Let $f : D \to \mathbb{R}$ and $a \in D$. We say that f is continuous at a if for every $\varepsilon > 0$, there is a $\delta > 0$ such that whenever $x \in D$ with $|x - a| < \delta$, we have $|f(x) - f(a)| < \varepsilon$.*

We say that f is continuous (on D) if f is continuous at a for every $a \in D$.

The first comment is that when a is not an accumulation point of D, *every* function on D is continuous at a. In fact, there is a $\delta.0$ such that $D \cap (a - \delta, a + \delta) = \{a\}$. Then, no matter what $\varepsilon > 0$ is given, simply choose this one $\delta > 0$. Then $x \in D$ and $|x - a| < \delta$ implies that $x = a$, so $|f(x) - f(a)| = 0 < \varepsilon$.

On the other hand, if a is an accumulation point of D, continuity is equivalent to the statement that $\lim_{x \to a} f(x) = f(a)$. This immediately leads to the following result.

Theorem 4.2 *let $f : D \to \mathbb{R}$ and $a \in D$. The following are equivalent.*

a) f is continuous at a.

b) If a is an accumulation point of D, then $\lim_{x \to a} f(x) = f(a)$.

c) If $\{a_n\}_{n=1}^{\infty}$ is a sequence in D with $a_n \to a$, then $f(a_n) \to f(a)$.

This, in turn, leads immediately to the following.

Theorem 4.3 *Suppose that $f, g : D \to \mathbb{R}$ are continuous at $a \in D$. Then $f+g, f-g, f \cdot g$ are continuous at a. Furthermore, if $g(x) \neq 0$ for all $x \in D$, then $\frac{f}{g}$ is continuous at a.*

By use of this result, polynomials and rational functions are continuous on their domains. We can also see that when $f(x) \geq 0$ for all $x \in D$, then $\sqrt{f}$ is continuous at a when f is.

As another example, let

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

Then, f is not continuous at $a = 0$, but $g(x) = xf(x)$ is continuous at $a = 0$, and so is continuous on $\mathbb{R}$.

We also need to know what happens with compositions of functions. Fortunately, compositions respect continuity.

Theorem 4.4 *Suppose that $f : D \to E$ and $g : E \to \mathbb{R}$. Assume that f is continuous at $a \in D$ and that g is continuous at $f(a) \in E$. Then the composition $g \circ f : D \to \mathbb{R}$ is continuous at a.*

Proof: Let $\varepsilon > 0$. Then, there is a $\delta_1 > 0$ so that whenever $y \in E$ with $|y - f(a)| < \delta_1$, we have $|g(y) - g(f(a)| < \varepsilon$. But now, there is a $\delta > 0$ such that whenever $x \in D$ with $|x - a| < \delta$, we have that $f(x) - f(a)| < \delta_1$. But this then implies that $|g \circ f(x) - g \circ f(a)| = |g(f(x)) - g(f(a))| < \varepsilon$.

We will also need the following, more technical, result later.

Proposition 4.5 *Suppose that $f : D \to \mathbb{R}$ and $a \in D$ with $f(a) > 0$. Then there is a $\delta > 0$ such that $f(x) > 0$ for all $x \in D \cap (a - \delta, a + \delta)$.*

Proof: Let $\varepsilon = f(a) > 0$ in the definition of continuity to find $\delta > 0$ such that $x \in D$ with $||x - a| < \delta$ implies that $|f(x) - f(a)| < \varepsilon = f(a)$. Then, for $x \in D \cap (a - \delta, a + \delta)$, we have $f(x) > 0$ since otherwise $|f(x) - f(a)| = f(a) - f(x) > f(a) = \varepsilon$.

4.1 Uniform Continuity

Let's consider what it means for a function $f : D \to \mathbb{R}$ to be continuous. Following the definition, it says that for every $a \in D$ and every $\varepsilon > 0$, there is a $\delta > 0$ such that for all $x \in D$, we have $|x-a| < \delta$ implies that $|f(x)-f(a)| < \varepsilon$. In this, the δ can depend on both $a \in D$ and $\varepsilon > 0$. It turns out to be useful to consider those cases where δ can be chosen in a way that does not depend on a.

Definition 4.6 *Let $f : D \to \mathbb{R}$. We say that f is uniformly continuous (on D) if for each $\varepsilon > 0$ there is a $\delta > 0$ such that whenever $x, a \in D$ with $|x-a| < \delta$, we have $|f(x) - f(a)| < \varepsilon$.*

Clearly, if f is uniformly continuous, it is continuous. The converse is not always true.

Example: Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x^2$. Then, since f is a polynomial, f is continuous on $\mathbb{R}$. We claim that f is *not* uniformly continuous on $\mathbb{R}$. In fact, let $\varepsilon = 1$ and suppose that $\delta > 0$. Choose $n \in \mathbb{N}$ with $\frac{1}{n} < \delta$ and consider $a = n$ and $x = n + \frac{1}{n}$. Then $|x - a| = \frac{1}{n} < \delta$, but $|f(x) - f(a)| = |(n + \frac{1}{n})^2 - n^2| = 2 + \frac{1}{n} \geq \varepsilon$.

Example: On the other hand, let $f : [-20, 20] \to \mathbb{R}$ be defined by $f(x) = x^2$. The claim is that f is uniformly continuous on $[-20, 20]$. In fact, let $\varepsilon > 0$ and set $\delta = \frac{\varepsilon}{40}$. Then, if $x, a \in [-20, 20]$ with $|x - a| < \delta$, we have $|f(x)-f(a)| = |x^2-a^2| = |(x-a)(x+a)| \leq |x-a|(|x|+|a|) < \delta \cdot (20+20) = \varepsilon$.

Example: Let $f : (0, 1] \to \mathbb{R}$ be defined by $f(x) = \frac{1}{x}$. Then f is continuous on $(0, 1]$. We claim that f is not uniformly continuous on $(0, 1]$. In fact, let $\varepsilon = 1$ and $\delta > 0$. Choose $n \in \mathbb{N}$ so that $\frac{1}{n} < \delta$. let $a = \frac{1}{n}$ and $x = \frac{1}{2n}$. Then $|x - a| = \frac{1}{2n} < \delta$, but $|f(x) - f(a)| = |2n - n| = n \geq \varepsilon$.

Example: On the other hand, define $f : [\frac{1}{3}, 1] \to \mathbb{R}$ be defined by $f(x) = \frac{1}{x}$. We claim f is uniformly continuous. In fact, let $\varepsilon > 0$ and set $\delta = \frac{\varepsilon}{9}$. Then, if $x, a \in [\frac{1}{3}, 1]$ with $|x-a| < \delta$, we have $|f(x)-f(a)| = |\frac{1}{x} - \frac{1}{a}| = \frac{|a-x|}{|xa|} \leq \delta \cdot 3 \cdot 3 = \varepsilon$.

One quick test is available for uniform continuity.

Theorem 4.7 *Suppose that $f : D \to \mathbb{R}$ is uniformly continuous and a is an accumulation point of D. Then $\lim_{x \to a} f(x)$ exists.*

Proof: Suppose that $\{a_n\}_{n=1}^{\infty}$ is a sequence in D with $a_n \neq a$ and $a_n \to a$. We show that $\{f(a_n)\}_{n=1}^{\infty}$ converges by showing it is Cauchy. In fact, let $\varepsilon > 0$. Use uniform continuity of f to find $\delta > 0$ such that whenever $x, y \in D$ with $|x-y| < \delta$, we have $|f(x) - f(y)| < \varepsilon$. Since $a_n \to a$, the sequence $\{a_n\}_{n=1}^{\infty}$ is Cauchy, so there is a $N \in \mathbb{N}$ such that whenever $m, n \geq N$, we have $|a_n - a_m| < \delta$. But then, this implies that $|f(a_n) - f(a_m)| < \varepsilon$.

For example, with $f(x) = \frac{1}{x}$ and $f : (0, 1]$, we may let $a_n = \frac{1}{n}$. Then $a_n \to 0$, but $\{f(a_n)\}_{n=1}^{\infty} = \{n\}_{n=1}^{\infty}$ is not even bounded, so it doesn't converge. Hence (as we already know) f is not uniformly continuous.

In a similar way, leting $a_n = \frac{1}{\pi n}$ shows that $f(x) = \sin(\frac{1}{x})$ is not uniformly continuous on $(0, 1]$.

There is, however, a case where we can conclude uniform continuity given only continuity.

Theorem 4.8 *Suppose that D is a closed and bounded subset of $\mathbb{R}$ and that $f : D \to \mathbb{R}$ is continuous. Then f is uniformly continuous.*

Proof: Suppose not. then, there is an $\varepsilon > 0$ such that for every $\delta > 0$, there are $x, a \in D$ with $|x - a| < \delta$ but $|f(x) - f(a)| \geq \varepsilon$. Choose $\delta = \frac{1}{n}$ and obtain sequences $\{x_n\}_{n=1}^{\infty}$ and $\{a_n\}_{n=1}^{\infty}$ in D wih $|x_n - a_n| < \frac{1}{n}$ for each n and $|f(x_n) - f(a_n)| \geq \varepsilon$.

Now, since D is bounded, both sequences $\{x_n\}_{n=1}^{\infty}$ and $\{a_n\}_{n=1}^{\infty}$ are bounded. Thus, the first has a convergent subsequence. From that subsequence, there is a convergent subsequence of the second. Hence we find $n_k \in \mathbb{N}$ with $x_{n_k} \to x$ and $a_{n_k} \to a$ for some x, a.

Since $x_n, a_n \in D$ and since D is closed, $x, a \in D$. Next, we have that $f(x_{n_k}) \to f(x)$ and $f(a_{n_k}) \to f(a)$ since f is continuous. Since $|f(x_{n_k}) - f(a_{n_k})| \geq \varepsilon$ for all k, this implies that $|f(x) - f(a)| \geq \varepsilon$.

But, $|x_{n_k} - a_{n_k}| \to |x - a|$ and $|x_{n_k} - a_{n_k}| < \frac{1}{n_k} \to 0$ shows that $|x - a| = 0$, so $x = a$. But this implies that $|f(x) - f(a)| = 0 < \varepsilon$, a contradiction.

4.2 Properties of Continuous Functions

Theorem 4.9 *Let D be a closed and bunded subset of $\mathbb{R}$ and suppose that $f : D \to \mathbb{R}$ is continuous. Then, the range of f, $\{f(x) : x \in D\}$ is closed and bounded.*

Proof: Closed: Suppose that $\{y_n\}_{n=1}^{\infty}$ is a sequence in the range of f and $y_n \to y \in \mathbb{R}$. We need to show that y is in the range of f. We can write $y_n = f(x_n)$ for some $x_n \in D$. Now, since D is bounded, the sequence $\{x_n\}_{n=1}^{\infty}$ has a convergent subsequence, say $x_{n_k} \to x$. Since D is closed, we see that $x \in D$. But now, $y = \lim y_{n_k} = \lim f(x_{n_k}) = f(x)$ by continuity of f, so y is in the range of f. This shows that the range of f is closed.

Now, suppose that the range of f is not bounded. Then, for each $n \in \mathbb{N}$ there is an $x_n \in D$ such that $|f(x_n)| > n$. Again, there is a convergent subsequence $x_{n_k} \to x$ and we have that $x \in D$. But now, $|f(x)| = \lim |f(x_{n_k})$. Since $\{f(x_{n_k})$ is unbounded, it does not converge. This is a contradiction.

Corollary 4.10 *Let D be closed and bounded in $\mathbb{R}$ and $f : D \to \mathbb{R}$ continuous. Then, there is an $a \in D$ such that $f(x) \leq f(a)$ for all $x \in D$. In other words, f attains it's maximum value on D. Similarly, f attains its minmum value on D.*

A common intuition for a function to be continuous is that it can be graphed 'without lifting the pencil from the paper'. While this is far from a formal definition, it does give an intuition concerning the next result.

Theorem 4.11 (Intermediate Value Theorem) *Let $a < b$ and $f : [a, b] \to \mathbb{R}$ is continous. Assume also that $f(a) < 0$ and $f(b) > 0$. Then, there is a $c \in (a, b)$ such that $f(c) = 0$.*

Proof: Let $A = \{x \in [a, b] : f(x) \leq 0\}$. Then $a \in A$, so A is non-empty. Also, A is bounded above by b, so $c = \sup A$ exists. We claim that $f(c) = 0$.

Suppose not, that is $f(c) \neq 0$. If $f(c) > 0$, use proposition 4.5 to find $\delta > 0$ so that $f(x) > 0$ for all $x \in [a, b] \cap (c - \delta, c + \delta)$. But this means that $c - \delta$

is an upper bound for the set A, which contradicts that c is the least upper bound. On the other hand, if $f(c) < 0$, there is a $\delta > 0$ so that $f(x) < 0$ for all $x \in [a, b] \cap (c - \delta, c + \delta)$. But this shows that $(c - \delta, c + \delta) \subseteq A$, so c is not an upper bound for A, another contradiction.

Corollary 4.12 *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous and $f(a) < y < f(b)$. Then, there is a $c \in [a, b]$ with $f(c) = y$. Similarly if $f(b) < y < f(a)$.*

Proof: In the first case, consider $g(x) = f(x) - y$ and apply the theorem to g. In the second, consider $g(x) = y - f(x)$.

As an easy consequence, we show the existence of roots other than just square roots.

Proposition 4.13 *Let $a \geq 0$ and $n \in \mathbb{N}$. Then, there is a unique $x \geq 0$ with $x^n = a$.*

Proof: We first show existence. If $a = 0$, choose $x = 0$. If $a = 1$, choose $x = 1$. If $0 < a < 1$, consider the function $f(x) = x^n$ on the closed interval $[0, 1]$. We have $f(0) = 0 < a$ and $a < 1 = f(1)$. The Intermediate Value Theorem then gives an $x \in [0, 1]$ such that $f(x) = a$. If $a > 1$, consider $f(x) = x^n$ on the closed interval $[1, a]$. Then $f(1) = 1 < a$ and $a < a^n = f(a)$, so again we obtain $x \in [1, a]$ with $f(x) = a$.

Uniqueness follows since $x_1 < x_2$ implies that $0 \leq x_1^n < x_2^n$, so f is one-to-one on $[0, +\infty)$.

As another application of the IVT, we have the following.

Proposition 4.14 *Suppose that $f : [a, b] \to [a, b]$ is continuous. Then there is an $x \in [a, b]$ with $f(x) = x$. We say x is a fixed point of f.*

Proof: Let $g(x) = f(x) - x$. Then $a \leq f(a)$ and $f(b) \leq b$, so if $f(a) \neq a$ and $f(b) \neq b$, we have $g(a) < 0$ and $g(b) > 0$. By IVT, there is an $x \in [a, b]$ such that $g(x) = 0$, i.e. so that $f(x) = x$.

Corollary 4.15 *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous. Then, there are $m, M \in \mathbb{R}$ such that the range of f is equal to $[m, M]$.*

It is often the case that we have a one-to-one function and we want to know about the continuity of the inverse function. The following example shows that some care is needed.

Example: Let $D = [0,1) \cap [2,3]$ and define

$$f(x) = \begin{cases} x & 0 \leq x < 1 \\ x-1 & 2 \leq x \leq 3. \end{cases}$$

Then f is one-to-one, continuous, and onto $[0,2]$. The inverse function, $f^{-1} : [0,2] \to D$ is given by

$$f^{-1}(x) = \begin{cases} x & 0 \leq x < 1 \\ x+1 & 1 \leq x \leq 2. \end{cases}$$

Hence, f^{-1} is NOT continuous at $x = 1$.

It turns out that we can recover continuity of the inverse if we put some restrictions on the domain of f.

Theorem 4.16 *Suppose that D is a closed and bounded subset of $\mathbb{R}$ and $f : D \to E$ is continuous, one-to-one, and onto. Then, $f^{-1} : E \to D$ is continuous.*

Proof: First, we know that E is also closed and bounded. We show that f^{-1} is continuous by applying sequences. So, suppose that $y_n \in E$ and $y_n \to y \in E$. We need to show that $f^{-1}(y_n) \to f^{-1}(y)$. The main issue is that we do not know that the sequence $\{f^{-1}(y_n)\}_{n=1}^{\infty}$ even converges. First, write $x_n = f^{-1}(y_n) \in D$ and $x = f^{-1}(y) \in D$. We use Theorem 2.22.

Notice that $\{x_n\}_{n=1}^{\infty}$ is a bounded sequence since $x_n \in D$ for all n. Now, suppose that $\{x_{n_k}\}$ is a *convergent* subsequence, say $x_{n_k} \to z$. Then $z \in D$ since D is closed. Also, $f(z) = \lim f(x_{n_k}) = \lim y_{n_k} = y$. Thus, $z = f^{-1}(y) = x$. In other words, every convergent subsequence of $\{x_n\}_{n=1}^{\infty}$ converges to x. By the mentioned theorem, $x_n \to x$, as required.

Exercises:

1. Show there is exactly one solution of $\cos x = x$ in the interval $[0, \pi/2]$.

Chapter 5

Differentiation

Definition 5.1 *Let $f : D \to \mathbb{R}$ and suppose that $a \in D$ is an accumulation point of D. Consider the* difference quotient *$T : D \setminus \{a\} \to \mathbb{R}$ defined by $T(x) = \frac{f(x)-f(a)}{x-a}$. If $lim_{x\to a}T(x)$ exists, we say that f is differentiable at a and write $f'(a) = \lim_{x\to a} T(x)$.*

We say that f is differentiable on D if it is differentiable at every $a \in D$. Notice that this requires every point of D to be an accumulation point of D.

Example: Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x^2$. Then, for $a \in \mathbb{R}$, we have $f'(a) = \lim_{x\to a} \frac{x^2-a^2}{x-a} = \lim_{x\to a} x + a = 2a$, so f is differentiable everywhere.

Example: Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = |x|$. When $a = 0$, we have $f'(0) = \lim_{x\to 0} \frac{|x|}{x}$, which does not exist. Thus, f is continuous at $a = 0$, but not differentiable there.

Theorem 5.2 *Suppose that $f : D \to \mathbb{R}$ is differnetiable at $a \in D$. Then f is continuous at a.*

Proof: In fact, $\lim_{x\to a} f(x) = \lim_{x\to a} f(a) + \frac{f(x)-f(a)}{x-a}(x-a) = f(a) + f'(a)\cdot 0 = f(a)$. Notice that all required limits exist in this calculation.

Theorem 5.3 *Suppose that $f, g : D \to \mathbb{R}$ are differentiable at $a \in D$. Then the sum and product, $f+g, fg$ are also differentiable at a. Furthermore, $(f+g)'(a) =$*

$f'(a)+g'(a)$ and $(fg)'(a)=f(a)g'(a)+f'(a)g(a)$. Also, if $g(x)\neq 0$ for $x\in D$, then $\frac{f}{g}$ is differentiable at a and $(\frac{f}{g})'(a)=\frac{f'(a)g(a)-f(a)g'(a)}{(g(a))^2}$.

Proof:

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

Similarly,

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

where we have used continuity of f to evaluate the first limit $\lim_{x\to a}f(x)=f(a)$.

Next,

$$\begin{aligned}\lim_{x\to a}\frac{(\frac{1}{g})(x)-(\frac{1}{g})(a)}{x-a} &= \lim_{x\to a}\frac{g(a)-g(x)}{(x-a)g(x)g(a)}\\ &= -\frac{g(x)-g(a)}{x-a}\frac{1}{g(x)g(a)}\\ &= -\frac{g'(a)}{(g(a))^2}\end{aligned}$$

This, together with the product rule gives us the quotient rule for derivatives.

Proposition 5.4 *Let $n\in\mathbb{Z}$ with $n\neq 0$ and set $f_n(x)=x^n$. Then f_n is differentiable wherever it is defined and $f_n'(a)=na^{n-1}$.*

Proof: For $n=1$, we have $f_1(x)=x$ and a simple calculation shows that ${f_1}'(a)=1=1\cdot a^0$ for all $a\in\mathbb{R}$.

Now, suppose the result is true for $n\in\mathbb{N}$. Then we have that $f_{n+1}=f_nf_1$, so $f_{n+1}'(a)=f_n'(a)f_1(a)+f_n(a)f_1'(a)=na^{n-1}\cdot a+a^n\cdot 1=(n+1)a^n$, which is the result for $n+1$.

By induction, we have the result for all $n\in\mathbb{N}$.

Now, suppose that $n \in \mathbb{Z}$ with $n \neq 0$ but $n \notin \mathbb{N}$. Then $-n \in \mathbb{N}$ and $f_n = \frac{1}{f_{-n}}$. The quotient rule then gives that $f_n'(a) = -\frac{f_{-n}'(a)}{(f_{-n}(a))^2} = -\frac{-na^{-n-1}}{(a^{-n})^2} = na^{n-1}$ as desired.

Using this we obtain that all polynomials are differentiable everywhere and all rational functions are differentiable whever they are defined. We now consider some more unusual examples.

Example: Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = \begin{cases} x^2 & x \notin \mathbb{Q} \\ 0 & x \in \mathbb{Q} \end{cases}$ Then, f is discontinous at every $a \neq 0$, but at $a = 0$, we have that $f'(0) = \lim_{x\to 0} \frac{f(x)}{x} = 0$, so f is differentiable at exactly one point.

Example: Define $f(x) = |x|$. Then f is continous everywhere, but $f'(0) = \lim_{x\to 0} \frac{|x|}{x}$ does not exist.

A very important result for derivatives is the chain rule. It involves the derivaive of the composition of two differentiable functions.

Theorem 5.5 *Suppose that $g : D \to E$ is differentiable at $a \in D$ and $f : E \to \mathbb{R}$ is differentiable at $g(a) \in E$. Then the composition $f \circ g$ is differentiable at $a \in D$ and $(f \circ g)'(a) = f'(g(a))g'(a)$.*

A simple, failed proof can be obtained by writing $\frac{(f\circ g)(x)-(f\circ g)(a)}{x-a} = \frac{f(g(x))-f(g(a))}{g(x)-g(a)} \frac{g(x)-g(a)}{x-a}$ and arguing that the first term converges to $f'(g(a))$ and the second to $g'(a)$. This would, in fact, work if we knew that $g(x) \neq g(a)$ for x close to a. Otherwise, the division by 0 beccomes problematic. Our proof is set up to avoid this difficulty.

Proof: First, define $T : E \to \mathbb{R}$ by $T(x) = \frac{f(x)-f(g(a))}{x-g(a)}$ for $x \neq g(a)$ and also letting $T(g(a)) = f'(g(a))$. Since, by definition, $\lim_{x\to g(a)} \frac{f(x)-f(g(a))}{x-g(a)} = f'(g(a))$, we have that T is continuous at $g(a) \in E$.

Next, we note that $\frac{(f\circ)(x)-(f\circ g)(a)}{x-a} = T(g(x))\frac{g(x)-g(a)}{x-a}$ for all $x \in D$. This is seen to be true as above if $g(x) \neq g(a)$, while if $g(x) = g9a)$, both sides are 0. But now, $(f \circ g)'(a) = \lim_{x\to a} T(g(x))\frac{g(x)-g(a)}{x-a}$. By continouity of g at a and of T at $g(a)$, the first term converges to $T(g(a)) = f'(g(a)$. The second term converges to $g'(a)$ by definition of the derivative.

5.1 The Mean Value Theorem

Theorem 5.6 (Rolle's Theorem) *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Suppose also that $f(a) = f(b) = 0$. Then there is a $c \in (a, b)$ such that $f'(c) = 0$.*

Proof: If f is the constant function 0, then $f'(x) = 0$ for *all* $x \in (a, b)$, so we are done. Otherwise, there is an $x \in (a, b)$ with $f(x) \neq 0$. By considering $-f$ in place of f, we may assume that $f(x) > 0$. Since f is continous on $[a, b]$, f attains its maximum value at some point $c \in [a, b]$. Then, $f(c) > 0$ and $f(x) \leq f(c)$ for all $x \in [a, b]$. Since $f(a) = f(b) = 0$, we have that $c \neq a, b$.

Now, for $x < c$, we have that $\frac{f(x)-f(c)}{x-c} \geq 0$, so $f'(c) = \lim_{x\to c} \frac{f(x)-f(c)}{x-c} \geq 0$. Similarly, for $x > c$, we have that $\frac{f(x)-f(x)}{x-c} \leq 0$, showing that $f'(c) = 0$.

This is quickly generalized to the

Theorem 5.7 (Mean Value Theorem) *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b). Then, there is a $c \in (a, b)$ such that $f'(c) = \frac{f(b)-f(a)}{b-a}$.*

Proof: Let $g(x) = f(a) + \frac{f(b)-f(a)}{b-a}(x - a)$. Then, $g(a) = f(a), g(b) = f(b)$ and $g'(x) = \frac{f(b)-f(a)}{b-a}$ for all $x \in [a, b]$. The function $f - g$ satisfies the conditions of Rolle's Theorem, so there is a $c \in (a, b)$ with $(f - g)'(c) = 0$. This is the desired c.

Corollary 5.8 *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous and that $f' = 0$ on (a, b). Then f is a constant function.*

Proof: Let $x \in (a, b]$. Then the hypotheses of the Mean Value Theorem hold for f on the interval $[a, x]$, so there is a $c \in (a, x)$ such that $\frac{f(x)-f(a)}{x-a} = f'(c) = 0$. This shows that $f(x) = f(a)$. Since this holds for every $x \in (a, b]$, we have that f is the constant function $f(a)$.

Corollary 5.9 *Suppose that $f, g[a, b] \to \mathbb{R}$ are continuous on $[a, b]$ and differentiable on (a, b) and that $f'(x) = g'(x)$ for all $x \in (a, b)$. Then, there is a $C \in \mathbb{R}$ such that $f(x) = g(x) + C$ for all $x \in [a, b]$.*

Proof: Let $h9x) = f(x) - g(x)$. Then h is continuous on $[a, b]$ and differentiable on (a, b) with $h'(x) = f'(x) - g'(x) = 0$ for $x \in (a, b)$. By the previous corollary, h is constant.

Proposition 5.10 *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is differentiable and f' is bounded. Then f is uniformly continuous.*

Definition 5.11 *We say $f : D \to \mathbb{R}$ is*

a) non-decreasing if $x < y$ in D implies that $f(x) \leq f(y)$,

b) increasing if $x < y$ in D implies that $f(x) < f(y)$,

c) non-increasing if $x < y$ in D implies that $f(x) \geq f(y)$,

d) decreasing if $x < y$ in D implies that $f(x) > f(y)$.

Proposition 5.12 *Suppose that $f : D \to \mathbb{R}$ is differentiable on D. If f is non-decreasing, then $f'(x) \geq 0$ for all $x \in D$.*

Proposition 5.13 *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b). If $f'(x) > 0$ for all $x \in (a, b)$, then f is increasing on $[a, b]$. Similarly, if $f'(x) \geq 0$ for all $x \in (a, b)$, then f is non-decreasing on $[a, b]$.*

Proof: Suppose that $x, y \in [a, b]$ with $x < y$. Then f is continuous on $[x, y]$ and differentiable on (x, y), so there is a $c \in (x, y) \subseteq (a, b)$ such that $f'(c) = \frac{f(y)-f(x)}{y-x}$. Since $y - x > 0$ and $f'(c) > 0$ (or $f'(c) \geq 0$), this implies that $f(y) - f(x) > 0$ (or $f(y) - f(x) \geq 0$). This gives the result.

Proposition 5.14 *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b). Suppose also that $f'(x) \neq 0$ for all $x \in (a, b)$. Then f is one-to-one on $[a, b]$.*

Proposition 5.15 *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous and one-to-one. Then f is either increasing or it is decreasing.*

Proof: Suppose f is neither increasing nor decreasing. Then there are $x < y$ and $w < z$ such that $f(x) \leq f(y)$ and $f(w) \geq f(z)$. If $f(x) = f(y)$ or $f(w) = f(z)$, then f is clearly not one-to-one. Otherwise, among x, y, w, z there are $s < t < u$ with $f(s), f(u) < f(t)$. Then, by the intermediate value theorem, f cannot be one-to-one on $[s, u]$.

Theorem 5.16 (Intermediate Value Theorem for Derivatives) *Suppose that $f : [a, b] \to \mathbb{R}$ is differentiable on $[a, b]$. Suppose also that $f'(a) < 0 < f'(b)$. Then, there is a $c \in (a, b)$ such that $f'(c) = 0$.*

Proof: If not, then f is one-to-one on $[a, b]$, so f is either increasing or decreasing. But this implies that $f'(x) \geq 0$ for all $x \in [a, b]$ or that $f'(x) \leq 0$ for all $x \in [a, b]$. Both of these are against the assumption on f'.

It is important to point out that f' need not be continuous on $[a, b]$ under the given hypotheses. Consider, for example, $f(x) = x^2 sin(\frac{1}{x}) - x$ for $x \neq 0$ and $f(0) = 0$. Note that $f'(0) = -1$ and that f' is not continuous at 0.

5.2 L'Hôpital's Rule

We begin this section with a generalization of the Mean Value Theorem for derivatives.

Theorem 5.17 (Generalized Mean Value Theorem) *Suppose that $f, g : [a, b] \to \mathbb{R}$ are continous on $[a, b]$ and differentiable on (a, b). Then, there is a $c \in (a, b)$ such that $f'(c)(g(b) - g(a)) = g'(c)(f(b) - f(a))$.*

Note tht $g(x) = x$ gives the standard Mean velue Theorem. *Proof*: Define $h(x) = f(x)(g(b) - g(a)) - g(x)(f(b) - f(a))$. Then, h is continuous on $[a, b]$ and differentiable on (a, b) with $h'(x) = f'(x)(g(b) - g(a)) - g'(x)(f(b) - f(a))$. Now, compute that $h(a) = h(b)$, so by the standard Mean Value Theorem, there is a $c \in (a, b)$ with $h'(c) = 0$.

Theorem 5.18 (L'Hôpital's Rule) *Suppose that $f, g : [a, b] \to \mathbb{R}$ are continous on $[a, b]$ and differentiable on $(a, b) \setminus \{x_0\}$ where $x_0 \in [a, b]$. Suppose*

that $f(x_0) = g(x_0) = 0$, *that* $g'(x) \neq 0$ *for all* $x \in [a,b] \setminus \{x_0\}$ *and that* $\lim_{x\to x_0} \frac{f'(x)}{g'(x)} = L$. *Then* $\lim_{x\to x_0} \frac{f(x)}{g(x)} = L$.

Proof: First notice that $g(x) \neq 0$ for $x \in [a,b] \setminus \{x_0\}$ by use of the Mean Value Theorem since $g(x_0) = 0$ and $g'(x) \neq 0$.

Let $\varepsilon > 0$ and choose $\delta > 0$ so that $x \in [a,b] \setminus \{x_0\}$ with $|x - x_0| < \delta$ implies that $|\frac{f'(x)}{g'(x)} - L| < \varepsilon$. Then for every $x \in [a,b] \setminus \{x_0\}$, Let I be the interval $[x, x_0]$ if $x < x_0$ and $[x_0, x]$ otherwise. Then f and g are continuous on I and differentiable except, perhaps, at the endpoints. So the generalized Mean Value Theorem gives $c \in I$ so that $f'(c)(g(x) - g(x_0)) = g'(c)(f(x) - f(x_0))$. Since $f(x_0) = g(x_0) = 0$ and $g'(c), g(x) \neq 0$, this gives that $|\frac{f(x)}{g(x)} - L| = |\frac{f'(c)}{g'(c)} - L| < \varepsilon$ since we have $|c - x_0| < |x - x_0| < \varepsilon$. This proves the result.

Exercises:

1. Suppose that $f : (a,b) \to \mathbb{R}$ is continuous on (a,b) and differentiable except, possibly at $x_0 \in (a,b)$. Suppose that $\lim_{x\to x_0} f'(x)$ exists. Show that $f'(x_0)$ exists and f' is continuous at x_0.

2. Suppose that $f : [a,b] \to \mathbb{R}$ is such that f' exists and is continuous $[a,b]$ and is differentiable on (a,b). Suppose that $f(a) = f'(a) = f(b) = 0$. Show that there is a $c \in (a,b)$ with $f''(c) = 0$.

Chapter 6

The Riemann Integral

Definition 6.1 *Let $a < b$. A parition of $[a, b]$ is a finite subset P of $[a, b]$ with $a, b \in P$. We will index the elements of a parition in such a way that $a = x_0 < x_1 < x_2 < \cdots < x_n = b$ and will often write $P = \{x_0 < x_1 < x_2 < \cdots < x_n\}$ for convenience.*

Suppose that $f : [a, b] \to \mathbb{R}$ is a bounded function and that $P = \{x_0 < \cdots x_n\}$ is a partition of $[a, b]$. Then f is also bounded on each subinterval $[x_{i-1}, x_i]$ for $i = 0, \cdots, n$, so the infima and suprema $m_i(f, P) = \inf\{f(x) : x \in [x_{i-1}, x_i]\}$ and $M_i(f, P) = \sup\{f(x) : x \in [x_{i-1}, x_i]\}$ are defined.

We define the lower and upper sums for f on the partition P by

$$L(f, P) = \sum_{i=1}^{n} m_i(f, P)(x_i - x_{i-1})$$

and

$$U(f, P) = \sum_{i=1}^{n} M_i(f, P)(x_i - x_{i-1}).$$

Alternatively, suppose that $g : [a, b] \to \mathbb{R}$ is a non-decreasing function. We can define the lower sum of f on P with respect to g by $U(f; g, P) = \sum_{i=1}^{n} m_i(f, P)(g(x_i) - g(x_{i-1}))$. A similar definition is given for the upper sum. The ordinary integrals are obtained when $g(x) = x$.

For example, let $f : [0, 2] \to \mathbb{R}$ be defined by $f(x) = x^2$. Then f is a bounded function. Let $P = \{0 < \frac{1}{2} < \frac{3}{2} < 2\}$, so $x_0 = 0, x_1 = \frac{1}{2}, x_2 = \frac{3}{2}$, and $x_3 = 2$.

Then $m_1(f,P) = 0, m_2(f,P) = \frac{1}{4}, m_3(f,P) = \frac{9}{4}, M_1(f,P) = \frac{1}{4}, M_2(f,P) = \frac{9}{4}$, and $M_3(f,P) = 4$. This gives $L(f,P) = 0 \cdot (\frac{1}{2} - 0) + \frac{1}{4} \cdot (\frac{3}{2} - \frac{1}{2}) + \frac{9}{4} \cdot (2 - \frac{3}{2}) = \frac{11}{8}$ while $U(f,P) = \frac{35}{8}$.

Proposition 6.2 *Suppose that $m \leq f(x) \leq M$ for $x \in [a,b]$. Then, for every partition P, we have $m(b-a) \leq L(f,P) \leq U(f,P) \leq M(b-a)$.*

Proof: In fact, if $P = \{x_0 < \cdots < x_n\}$, then for each $1 \leq i \leq n$, we have $m \leq m_i(f,P) \leq M_i(f,P) \leq M$. This gives

$$\sum_{i=1}^{n} m(x_i - x_{i-1}) \leq \sum_{i=1}^{n} m_i(f,P)(x_i - x_{i-1}) \leq \sum_{i=1}^{n} M_i(f,P)(x_i - x_{i-1}) \leq \sum_{i=1}^{n} M(x_i - x_{i-1}).$$

Now use the fact that $\sum_{i=1}^{n}(x_i - x_{i-1}) = (x_1 - x_0) + (x_2 - x_1) + \cdots (x_n - x_{n-1}) = x_n - x_0 = b - a$ to get the result.

Alternatively, we see that $m(g(b) - g(a)) \leq L(f;g,P) \leq U(f;g,P) \leq M(g(b) - g(a))$.

6.1 Integrability

Definition 6.3 *Suppose that $f : [a,b] \to \mathbb{R}$ is bounded. Define the upper and lower integrals of f on $[a,b]$ by*

$$\underline{\int_a^b} f(x)\, dx = \sup\{L(f,P) : P \text{ a partition of } [a,b]\}$$

$$\overline{\int_a^b} f(x)\, dx = \inf\{U(f,P) : P \text{ a partition of } [a,b]\}$$

If it happens that $\int_a^b f(x)\, dx = \int_a^b f(x)\, dx$, we say that f is (Riemann) integrable on $[a,b]$ and let $\int_a^b f(x)\, dx = \underline{\int_a^b} f(x)\, dx$, the common value. We write $f \in R[a,b]$.

The alternative definitions for $\int_a^b f(x)\, dg(x)$ and $\overline{\int_a^b} f(x)\, dg(x)$ are clear. In the case where the upper and lower integrals are equal, we say that f is (Riemann-Steiltjes) integrable with respect to g and write $f \in R([a,b], g)$.

As an example, if $f(x) = k$ is a constant function, then every $m_i(f,P) = M_i(f,P) = k$, so we have $L(f,P) = k(b-a)$ and $U(f,P) = k(b-a)$, so the upper

and lower integrals are both $k(b-a)$, so f is integrable and $\int_a^b f(x)\,dx = k(b-a)$. Similarly, $\int_a^b f(x)dg(x) = k(g(b) - g(a))$.

To proceed, we need to know some more about upper and lower sums.

Lemma 6.4 *Suppose that $Q = P \cup \{y\}$ is a partition of $[a, b]$ with $y \neq a, b$. Let f be a bounded function on $[a, b]$. Then $L(f, P) \leq L(f, Q) \leq U(f, Q) \leq U(f, P)$.*

Proof: Write $P = \{x_0 < x_1 < \cdots x_n\}$ and suppose that $y \in (x_{k-1}, x_k)$. Hence, $Q = \{x_0 < \cdots x_{k-1} < y < x_k \cdots x_n\}$. Then, temporarily writing $m_k^{'} = \inf\{f(x) : x \in [x_{k-1}, y]\}$ and $m_k^{''} = \inf\{f(x) : x \in [y, x_k]\}$ We have that $m_k(f, P) \leq m_k^{'}, m_k^{''}$, so that $m_k(f, P)(x_k - x_{k-1}) = m_k(f, P)(y - x_{k-1}) + m_k(f, P)(x_k - y) \leq m_k^{'}(y - x_{k-1}) + m_k^{''}(x_k - y)$.

But now,

$$\begin{aligned} L(f,P) &= \sum_{i=1}^{n} m_i(f,P)(x_i - x_{i-1}) \\ &= \sum_{i<k} m_i(f,P)(x_i - x_{i-1}) + m_k(f,P)(x_k - x_{k-1}) + \sum_{i>k} m_i(f,P)(x_i - x_{i-1}) \\ &\leq \sum_{i<k} m_i(f,P)(x_i - x_{i-1}) + m_k^{'}(y - x_{k-1}) + m_k^{''}(x_k - y) + \sum_{i>k} m_i(f,P)(x_i - x_{i-1}) \\ &= L(f,Q) \end{aligned}$$

A similar argument shows that $U(f, Q) \leq U(f, P)$. Since we always have $L(f, Q) \leq U(f, Q)$, we are done.

Proposition 6.5 *Suppose that P and Q are partitions of $[a, b]$ and f is a bounded function on $[a, b]$. Then*

a) If $P \subseteq Q$, then $L(f, P) \leq L(f, Q) \leq U(f, Q) \leq U(f, P)$,

b) In any case $L(f, P) \leq U(f, Q)$.

c) We always have

$$L(f,P) \leq \underline{\int_a^b} f(x)\,dx \leq \overline{\int_a^b} f(x)\,dx \leq U(f,Q)$$

.

Proof:

a) Suppose that $Q = P \cup \{y_1, \cdots y_m\}$. Then $L(f,P) \leq L(f, P \cup \{y_1\}) \leq L(f, P \cup \{y_1, y_2\}) \cdots \leq L(f,Q) \leq U(f,Q) \leq \cdots \leq U(f, P \cup \{y_1\}) \leq U(f,P)$.

b) In fact, we have $L(f,P) \leq L(f, P \cup Q) \leq U(f, P \cup Q) \leq U(f,Q)$.

c) Fix any partition Q. Then since $L(f,P) \leq U(f,Q)$ for *every* partition P, we have that $U(f,Q)$ is an upper bound for the collection of lower sums. By definition, this shows that

$$\underline{\int_a^b} f(x)\ dx \leq U(f,Q).$$

But now, this holds for *every* partition Q, so the lower integral is a lower bound for the uppser sums. Again, by definition, this shows that

$$\underline{\int_a^b} f(x)\ dx \leq \overline{\int_a^b} f(x)\ dx,$$

as desired.

The proof is the same for the Riemann-Steiltjes Integrals.

The following gives a quite usable test for when a function is integrable.

Theorem 6.6 *Suppose that f is a bounded function on $[a,b]$. Then f is integrable if and only if for each $\varepsilon > 0$, there is a partition P of $[a,b]$ such that $U(f,P) - L(f,P) < \varepsilon$.*

Proof: First, suppose that f is integrable and $\varepsilon > 0$. Then, there are partitions P and Q such that

$$\underline{\int_a^b} f(x)\,dx - \varepsilon/2 < L(f,P) \leq \underline{\int_a^b} f(x)\,dx \leq \overline{\int_a^b} f(x)\,dx \leq U(f,Q) < \overline{\int_a^b} f(x)\,dx + \varepsilon/2.$$

Also, $L(f,P) \leq L(f, P \cup Q) \leq U(f, P \cup Q) \leq U(f,Q)$. Since the upper and lower integrals are equal, this means that $U(f, P \cup Q) - L(f, P \cup Q) < \varepsilon/2 + \varepsilon/2 = \varepsilon$.

On the other hand, since

$$\overline{\int_a^b} f(x)\ dx - \underline{\int_a^b} f(x)\ dx \leq U(f,P) - L(f,P)$$

for every partition P, and since this last can be made as small as we like if the condition holds, we must have equality of the upper and lower integrals.

This is actually a very powerful result for showing integrability of a function. We apply it immediately to show that continuous functions are always integrable.

Theorem 6.7 *Suppose that $f : [a, b] \to \mathbb{R}$ is continuous. Then $f \in R[a, b]$.*

Proof: Let $\varepsilon > 0$. Since f is continuous and $[a, b]$ is compact, f is uniformly continuous. Hence, there is a $\delta > 0$ such that whenever $x, y \in [a, b]$ with $|x - y| < \delta$, we have that $|f(x) - f(y)| < \frac{\varepsilon}{b-a}$. Choose $n \in \mathbb{N}$ so that $\frac{b-a}{n} < \delta$ and let $x_k = a + k\frac{b-a}{n}$. Then $x_0 = a$, $x_n = b$, $x_0 < x_1 < \cdots < x_n$ and $x_i - x_{i-1} = \frac{b-a}{n}$ for all $1 \leq i \leq n$. Let $P = \{x_0 < \cdots < x_n\}$.

Now, for each $1 \leq i \leq n$, f is continuous on the compact set $[x_{i-1}, x_i]$, so it attains its maximum and minimum values, say $m_i(f, P) = f(y_i)$ and $M_i(f, P) = f(z_i)$ where $y_i, z_i \in [x_{i-1}, x_i]$. But then, $M_i(f, P) - m_i(f, P) = f(z_i) - f(y_i) < \frac{\varepsilon}{b-a}$ since $|z_i - y_i| < x_i - x_{i-1}| < \delta$.

Thus, we have

$$\begin{aligned}
U(f, P) - L(f, P) &= \sum_{i=1}^{n} M_i(f, P)(x_i - x_{i-1}) - \sum_{i=1}^{n} m_i(f, P)(x_i - x_{i-1}) \\
&= \sum_{i=1}^{n} \left(M_i(f, P) - m_i(f, P)\right)(x_i - x_{i-1}) \\
&< \sum_{i=1}^{n} \frac{\varepsilon}{b-a}(x_i - x_{i-1}) \\
&= \frac{\varepsilon}{b-a} \sum_{i=1}^{n} (x_i - x_{i-1}) \\
&= \frac{\varepsilon}{b-a}(b - a) = \varepsilon.
\end{aligned}$$

Hence, by our theorem, f is integrable.

In a similar way, we have that every non-decreasing function is integrable (this fails for Steiltjes integrals).

Theorem 6.8 *Suppose that $f : [a, b] \to \mathbb{R}$ is non-decreasing. Then $f \in R[a, b]$.*

Proof: If f is constant, we already know that f is integrable. Otherwise, we have $f(b) - f(a) > 0$. Let $\varepsilon > 0$ and choose $n \in \mathbb{N}$ so that $\frac{b-a}{n} < \frac{\varepsilon}{f(b)-f(a)}$. Again, choose $x_k = a + k\frac{b-a}{n}$ and $P = \{x_0 < x_1 < \cdots < x_n\}$. Since f is non-decreasing, we have that $m_i(f, P) = f(x_{i-1})$ and $M_i(f, P) = f(x_i)$.

Now, we have, like before,

$$\begin{aligned} U(f,P) - L(f,P) &= \sum_{i=1}^{n} (M_i(f,P) - m_i(f,P))\,(x_i - x_{i-1}) \\ &< \sum_{i=1}^{n} (M_i(f,P) - m_i(f,P)) \cdot \frac{\varepsilon}{f(b) - f(a)} \\ &= \frac{\varepsilon}{f(b) - f(a)} \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) \\ &= \frac{\varepsilon}{f(b) - f(a)} (f(x_n) - f(x_0) = \varepsilon. \end{aligned}$$

So, again, f is integrable.

6.2 Arithmetic of Integrals

Theorem 6.9 *Suppose that $f, g \in R[a, b]$ and $c \in \mathbb{R}$. Then*

a) $(f + g) \in R[a, b]$ and

$$\int_a^b (f+g)(x)\ dx = \int_a^b f(x)\ dx + \int_a^b g(x)\ dx.$$

b) $cf \in R[a, b]$ and

$$\int_a^b (cf)(x)\ dx = c \int_a^b f(x)\ dx.$$

Proof:

a) Let $P = \{x_0 < \cdots < x_n\}$ be any partition of $[a, b]$. For each $x \in [x_{i-1}, x_i]$, we have that $f(x) + g(x) \leq \sup\{f(x) : x \in [x_{i-1}, x_i]\} + \sup\{g(x) : x \in [x_{i-1}, x_i]\} = M_i(f, P) + M_i(g, P)$. Hence, $M_i(f + g, P) \leq M_i(f, P) + M_i(g, P)$. Similarly, $m_i(f, P) + m_i(g, P) \leq m_i(f + g, P)$. It follows that $L(f, P) + L(g, P) \leq L(f + g, P) \leq U(f + g, P) \leq U(f, P) + U(g, P)$.

Now, let $\varepsilon > 0$ and choose partitions P_1 and P_2 such that $U(f, P_1) - L(f, P_1) < \varepsilon/2$ and $U(g, P_2) - L(g, P_2) < \varepsilon/2$. Let $P = P_1 \cup P_2$. Then $U(f, P) - L(f, P), U(g, P) - L(g, P) < \varepsilon/2$. By the above paragraph, we have that $U(f + g, P) - L(f + g, P) \leq U(f, P) - L(f, P) + U(g, P) - L(g, P) < \varepsilon$. This shows that $f + g$ is an integrable function.

Furthermore, we have that

$$L(f + g, P) \leq \int_a^b (f + g)(x)\ dx \leq U(f + g, P)$$

for every partition P. Hence,

$$L(f, P) + L(g, P) \leq \int_a^b (f + g)(x)\ dx \leq U(f, P) + U(g, P)$$

$$L(f, P) + L(g, P) \leq \int_a^b f(x)\ dx + \int_a^b g(x)\ dx \leq U(f, P) + U(g, P)$$

for every partition P. Since, as the last paragraph showed, the outside expressions can be made as close as we wish, we get the desired result.

b) Case I: $c =$. Then both sides are 0.

Case II: $c > 0$. Then $m_i(cf, P) = c \cdot m_i(f, P)$ and $M_i(cf, P) = c \cdot M_i(f, P)$, so $L(cf, P) = cL(f, P)$ and $U(cf, P) = cU(f, P)$. This gives the result in this case.

Case III: $c < 0$. In this case, $m_i(cf, P) = cM_i(f, P)$ and $M_i(cf, P) = cm_i(f, P)$. Hence, $L(cf, P) = cU(f, P)$ and $U(cf, P) = cL(f, P)$, giving the result.

Proposition 6.10 *Suppose that $a < c < b$ and $f : [a, b] \to \mathbb{R}$. Then f is integrable on $[a, b]$ if and only if it is integrable on both $[a, c]$ and $[c, b]$. Furthermore, if this is the case,*

$$\int_a^b f(x)\ dx = \int_a^c f(x)\ dx + \int_c^b f(x)\ dx.$$

Proof: Suppose that P is a partition of $[a, b]$ with $c \in P$. Let $P_1 = \{x_i \in P : x_i \leq c\}$ and $P_2 = \{x_i \in P : c \leq x_i\}$. Then P_1 is a partition of $[a, c]$ and P_2

is a partition of $[c, b]$ such that $L(f, P) = L(f, P_1) + L(f, P_2)$ and $U(f, P) = U(f, P_1) + U(f, P_2)$. Thus, $U(f, P) - L(f, P) = U(f, P_1) - L(f, P_1) + U(f, P_2) - L(f, P_2)$.

Now, suppose that f is integrable on $[a, b]$ and let $\varepsilon > 0$. We can choose a partition P of $[a, b]$ such that $U(f, P) - L(f, P) < \varepsilon$. Furthermore, by possibly adding c to P, we can assume that $c \in P$. Then $U(f, P_1) - L(f, P_1) < \varepsilon$, showing that f is integrable on $[a, c]$. Similarly, f is integrable on $[c, b]$.

Conversely, if f is integrable on both $[a, c]$ and on $[c, b]$, then there are partitions P_1, P_2 of $[a, c], [c, b]$, resp. such that $U(f, P_1) - L(f, P_1) < \varepsilon/2$ and $U(f, P_2) - L(f, P_2) < \varepsilon/2$. Now, let $P = P_1 \cup P_2$. Then, as above, $U(f, P) - L(f, P) < \varepsilon/2 + \varepsilon/2 = \varepsilon$, so f is integrable on $[a, b]$.

Finally, we have both

$$L(f, P_1) + L(f, P_2) \leq \int_a^b f(x)\ dx \leq U(f, P_1) + U(f, P_2)$$

and

$$L(f, P_1) + L(f, P_2) \leq \int_a^c f(x)\ dx + \int_c^b f(x)\ dx \leq U(f, P_1) + U(f, P_2).$$

Since we can make the differences between the outside terms as small as we want, the two intermediate values are equal.

Proposition 6.11 *Suppose that $f, g : [a, b] \to \mathbb{R}$ are integrable. Then, if $f(x) \leq g(x)$ for all $x \in [a, b]$,*

$$\int_a^b f(x)\ dx \leq \int_a^b g(x)\ dx.$$

Proof: Let P be any partition of $[a, b]$. Then we have $m_i(f, P) = \inf\{(x) : x \in [x_{i-1}, x_i]\} \leq \inf\{g(x) : x \in [x_{i-1}, x_i]\} = m_i(g, P)$. Thus,

$$L(f, P) \leq L(g, P) \leq \int_a^b g(x)\ dx.$$

This shows that the integraal of g is an upper bound for the lower sums of f, givng the result.

Proposition 6.12 *Suppose that $f : [a, b] \to \mathbb{R}$ is integrable. Then $|f|$ is integrable and*

$$\left| \int_a^b f(x)\ dx \right| \leq \int_a^b |f(x)|\ dx.$$

Proof: Let $\varepsilon > 0$ and choose a partition P of $[a, b]$ with $U(f, P) - L(f, P) < \varepsilon$. Then, $M_i(|f|, P) - m_i(|f|, P) \leq M_i(f, P) - m_i(f, P)$ for every $1 \leq i \leq n$ as can be seen by considering the cases given by the signs of $M_i(f, P)$ and $m_i(f, P)$. But then $U(|f|, P) - L(|f|, P) \leq U(f, P) - L(f, P) < \varepsilon$, which shows that $|f$ is integrable. Now use the fact that $-|f(x)| \leq f(x) \leq |f(x)|$ for all x and integrate to obtain the result.

Notice that all the results of this section are also valid for the Steiltjes integrals.

6.3 The Fundamental Theorem

Theorem 6.13 (Fundamentqal Theorem of Calculus Ver I) *Suppose that $f : [a, b] \to \mathbb{R}$ is differentiable. Suppose also that f' is integrable. Then*

$$\int_a^b f'(x)\ dx = f(b) - f(a).$$

Proof: Suppose that $P = \{x_0 < x_1 < \cdots < x_n\}$ is a partition of $[a, b]$. The mean value theorem applies on each interval $[x_{i-1}, x_i]$ to give a $c_i \in (x_{i-1}, x_i)$ such that $f(x_i) - f(x_{i-1}) = f'(c_i)(x_i - x_{i-1})$. But, we then have $m_i(f', P) \leq f'(c_i) \leq M_i(f', P)$ so

$$\begin{aligned} L(f', P) &= \sum_{i=1}^{n} m_i(f', P)(x_i - x_{i-1}) \\ &\leq \sum_{i=1}^{n} f'(c_i)(x_i - x_{i-1}) \\ &= \sum_{i=1}^{n} f(x_i) - f(x_{i-1}) \\ &= f(b) - f(a) \end{aligned}$$

And similarly, $f(b) - f(a) \leq U(f', P)$. Since this is true for every partition P and since f' is integrable, we have the result.

The question now arises as to whether there is always an anti-derivative for a given integrable function. That is is not the case can be seen easily since derivatives satisfy the Intermediate Value Theorem and not all integrable functions do. However, it is the case that every continuous function has an anti-derivative. This is the essence of the second version of the fundamental theorem. To get to that, though, we need the following:

Theorem 6.14 (Mean value Theorem for Integrals) *Suppose that* $f : [a, b] \to \mathbb{R}$ *is continuous. Then, there is a* $c \in (a, b)$ *such that*

$$f(c) = \frac{1}{b-a} \int_a^b f(x)\ dx.$$

Proof: Let $m = \inf\{f(x) : x \in [a, b]\}$ and $M = \sup\{f(x) : x \in [a, b]\}$. Then $m \leq f(x) \leq M$ for all x, giving $m(b-a) \leq \int_a^b f(x)\ dx \leq M(b-a)$. But this says that

$$m \leq \frac{1}{b-a} \int_a^b f(x)\ dx \leq M$$

. The Intermediate Value Theorem now gives the result.

We now extend the definition of integrals just slightly. In the case where $b < a$, we define

$$\int_a^b f(x)\ dx = -\int_b^a f(x)\ dx$$

. We also define $\int_a^a f(x)\ dx = 0$. It is then an easy exercise in cases to show that

$$\int_a^b f(x)\ dx = \int_a^c f(x)\ dx + \int_c^b f(x)\ dx$$

no matter how a, b, and c are arranged. In this case, the mean value theorem for integrals says that for a continuous f, there is a c between a and b such that $f(c)(b-a) = \int_a^b f(x)\ dx$. This is true whether $a < b$ or not.

Theorem 6.15 (Fundamental Theorem of Calculus ver II) *Suppose that* $f : [a, b] \to \mathbb{R}$ *is integrable. Define*

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

for all $x \in [a,b]$. Then, F is a continuous function on $[a,b]$. Furthermore, if f is continuous at $x_0 \in [a,b]$, F is differentiable at x_0 and $F'(x_0) = f(x_0)$. In particular, if f is continuous, F is an anti-derivative for f.

Proof: Since f is integrable, it is bounded, so there is an M with $|f(x)| \leq M$ for all $x \in [a,b]$. Then, if $x < y$,

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

Thus, $|F(y) - F(x)| \leq M|y-x|$ for all x, y, showing that F is continuous.

Now suppose that f is continuous at x_0. We need to show that $\lim_{x\to x_0} \frac{F(x)-F(x_0)}{x-x_0} = f(x_0)$. So, let $\varepsilon > 0$. Choose $\delta > 0$ so that $x \in [a,b]$ and $|x - x_0| < \delta$ implies that $|f(x) - f(x_0)| < \varepsilon$. Let $x \in [a,b]$ with $|x - x_0| < \delta$. Then, by the Mean Value Theorem for Integrals, there is a c between x and x_0 such that $f(c)(x - x_0) = \int_{x_0}^x f(t)\ dt = \int_a^x f(t)\ dt - \int_a^{x_0} f(t)\ dt = F(x) - F(x_0)$. But, since c is between x and x_0, we have $|c - x_0| < |x - x_0| < \delta$, so

$$\left|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)\right| = |f(c) - f(x_0)| < \varepsilon.$$

This shows the required limit exists and equals $f(x_0)$.