MATH 350 - NOTES

Contents

Chapter 1

The Completeness of the Real Numbers

1.1 Ordered Fields

A field is a set $\mathcal{F}$ together with two binary operations, which we denote by $\oplus$ and $\otimes$ for the time being and the following properties:

1. $\mathcal{F}$ is closed under both operations.

2. $\mathcal{F}$ forms an abelian group under $\oplus$.

3. Let 0 denote the identy in $\mathcal{F}$, i.e. $a \oplus 0 = a$ for all $a \in \mathcal{F}$. Then $\mathcal{F} \setminus \{0\}$ forms an abelian group under $\otimes$.

4. For all $a, b, c \in \mathcal{F}$ we have $a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c)$. In order to save on notation we accept an order of operations in which $\otimes$ precedes $\oplus$, i.e. we may omit the parentheses in $(a \otimes b) \oplus (a \otimes c)$.

We will denote the identity under $\otimes$ by 1. The rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$, and the complex numbers $\mathbb{C}$, form a field with the usual addition and multiplication. In addition, we have an order relation " $>$ " defined on $\mathbb{R}$ and $\mathbb{Q}$ as follows.

Definition 1 *A field $\mathcal{F}$ is an ordered field if there exists a non-empty subset $P \subset \mathcal{F}$ such that*

1. *if $a, b \in P$, then $a \oplus b \in P$.*

2. *if $a, b \in P$ then $a \otimes b \in P$.*

3. *for every $x \in \mathcal{F}$ either $x \in P$ or $\ominus x \in P$ or $x = 0$.*

We define an order relation " $>$ " on $\mathcal{F}$ by setting $x > y$ if and only if $x \ominus y \in P$

To simplify notation we will use the familiar notation for multiplication and addition. But we are mindful, that our field does not have to be one of the familar ones.

Proposition 1 *Let $\mathcal{F}$ be an ordered field. The order relation satisfies the following properties:*

1. *If $x, y \in \mathcal{F}$ then either $x > y$, or $y > x$, or $x = y$.*
2. *If $y > x$ and $z > y$ then $z > x$.*
3. *If $x > y$ and $a \in P$ then $ax > ay$.*
4. *If $x > y$ then $-y > -x$.*

Proof: For the first assertion observe that if $x, y \in \mathcal{F}$ and $x \neq y$ then $x - y \neq 0$, and so $x - y \in P$ or $x - y \in -P$, thus $x.y$ or $y > x$. For the second assertion observe that $z - x = z - y + y - x$, and since $z - y \in P$ and $y - x \in P$ the result follows. The last two assertions have simlarly easy proofs.
For the real and rational numbers the set P is given by the set of all positive (real or rational) numbers. There is an important distinction between real numbers and raional numbers. We all know that $\sqrt{2}$ for example is a real number, but not a rational number. The irrational numbers are not limit to algebraic combinations of roots of positive reals, or algebraic numbers, they also include transcendental numbers such as π or e, which cannot not be obtained as zeros of polynomials with rational coefficients. Now in grade school, we learned that every real number has a decimal expansion, and irrational numbers have infinite, non-repeating expansions. If we look at the real number π we can construct a sequence of approximations as follows:

$$p_0 = 3,\ p - 1 = 3.1,\ p_2 = 3.14,\ p_3 = 3.141, \ldots$$

We can think of π being the limit of this sequence of rational numbers. It is the first order of business in any analysis course to make sense of this limit. And we will do so in the next section.

1.2 Bounded Subsets of the Real Numbers

Definition 2 *Let $S \subset \mathbb{R}$. We say that S is bounded above, if there exists a number $M \in \mathbb{R}$ such that $x \leq M$ for all $x \in S$. M is called an upper bound for S. Similarly, we say that S is bounded below if there exists a number $m \in \mathbb{R}$ such that $m \leq x$ for all $x \in S$. m is called a lower bound for S. If S is both bounded above and bounded below, we say that S is bounded.*

To illuminate this definition we look at a few examples. First let us consider the set of all integers $\mathbb{Z}$. This is clearly a subset of $\mathbb{R}$, but it is neither bounded above nor bounded below. Infact this is an important property called the Archimedean Principle, which we will prove later. On the other hand the set

$$\left\{x \in \mathbb{Q} : x^2 < 2\right\}$$

is bounded. The number 2 would be an upper bound for this set, as would be the number 3. This example shows that neither upper bounds nor lower bounds are unique. This brings us to the idea of looking for the "best possible" upper bound or lower bound. Now the properties of this "best possible" upper bound sould come quiet natural, first of all this number must be an upper bound itself, and secondly, every other upperbound must be bigger than the "best possible". We formalize these ideas in the following definition.

Definition 3 *Let S be a set that is bounded above. We say that M is the least upper bound or supremum of S, if M is an upper bound of S and if $M \leq M'$ for all upper bounds m' of S. We write*

$$M = \sup S$$

Similarly, we say that L is the greatest lower bound or infimum of a set S, that is bounded below, if L is a lower bound of S and if $L \geq L'$ for all lower bounds L' of S. We write

$$L = \inf S$$

Examples:

1. Consider the set $S = \{x \in \mathbb{Q} : x^2 < 2\}$. We know from algebra that if $x^2 < 2$ then $x < \sqrt{2}$, so $\sqrt{2}$ is clearly an upper bound. Now suppose that M is any upper bound of this set, if $M < \sqrt{2}$, then there exists a rational number q such that $M < q < \sqrt{2}$ and thus $M^2 < q^2 < 2$. So $q \in S$, which clearly contradicts that M is an upper bound. Thus $M \geq \sqrt{2}$.

2. Consider the set $\{x \in \mathbb{R} : x^2 + x + 1 \leq 2\}$. Then

$$-\frac{1}{2} - \frac{\sqrt{5}}{2} = \inf S, \qquad -\frac{1}{2} + \frac{\sqrt{5}}{2} = \sup S.$$

   Observe, that both the supremum and infimum are themselves elements of S, unlike the situation in the first example.

3. Modify the first example to $S = \{x \in \mathbb{Q} : x^2 \leq 2\}$. Is $\sup S \in S$? of course not since $\sup S = \sqrt{2}$ and S contains only rational numbers that is not the case.

There are some important questions to be asked. First, is the least upper bound of a set unique? Secondly, do bounded sets always have least upper bounds? The first question will be answered affirmatively in the next proposition. The second question is more tricky, however. It goes to the very definition of the real numbers. As you may have noticed by now, we actually have not defined what a real number is. We have given many examples of real numbers and have a sort of soft definition of a real number as a number with a decimal expansion. But we really do not have a definition. The least upper bound property is actually one way of defining the reals.

Definition 4 *The set $\mathbb{R}$ of real numbers is the smallest ordered field that contains $\mathbb{Q}$ and for which every non-empty subset A that is bounded above has a least upper bound.*

It is clear from this definition that for every non-negative $x \in \mathbb{R}$ (or $\mathbb{Q}$), $\sqrt{x} \in \mathbb{R}$. Moreover, any decimal expansion

$$a_N = \sum_{n=-k}^{N} \frac{a_n}{10^n}, \qquad a_n \in \{0, 1, \ldots, 9\}$$

has a limit in $\mathbb{R}$.

Proposition 2 *Let $A \subset \mathbb{R}$ be bounded above. Then $M = \sup A$ is unique.*

Proof: As in many uniqueness proofs we do this one by contradiction. Assume that M_1 and M_2 are to least upper bounds if $M_1 \neq m_2$ then either $M_1 > M_2$ which would make M_1 not the least uper bound, or $M_2 > M_1$ which makes M_2 not the least upper bound. Therefore, $M_1 = M_2$.

In analysis you will come to love a particular Greek character ε. We will introduce you to the use of this in the next proposition.

Proposition 3 *Let $A \subset \mathbb{R}$ be bounded above. Then the following statements are equivalent:*

1. *$M = \sup A$.*

2. *M is an upper bound of A and for every $\varepsilon > 0$ there is an $a \in A$ such that $M - \varepsilon < a \leq M$.*

Proof: Suppose $M = \sup A$, and suppose that there is a number $\varepsilon_0 > 0$ such that $a \leq M - \varepsilon_0$ for all $a \in A$. Then $M - \varepsilon_0$ is an upper bound of A which is smaller than M. Thus for every $\varepsilon > 0$ there is an $a \in A$ such that $M - \varepsilon < a \leq M$. On the other hand, suppose that the second statement holds. Let M' be an upper such that $M' < M$. Let $\varepsilon_0 = M - M'$ then $a \leq M - \varepsilon_0$ for all $a \in A$ which contradicts the second statement.

We can actually improve this Proposition as follows.

Corollary 1 *Let $A \subset \mathbb{R}$ be bounded above with least upper bound M. Then, if $M \notin A$, we have that*

1. *A is an infinite set.*

2. *For every $\varepsilon > 0$ there exist infinitely many $a \in A$ such that $M - \varepsilon < a < M$.*

Proof: If A is a finite set then $A = \{a_1, \ldots, a_n\}$ and it has a largest element a_k. Clearly a_k is the least upper bound of A and so $\sup A \in A$. For the second assertion assume that there is an $\varepsilon_0 > 0$ such that there is only finitely may elements $a_1, \ldots a_n$ which satisfy $M - \varepsilon_0 < a_k < M$. Assume that a_j is the largest of these. Then $a \leq a_j$ for all $a \in A$, and a_j is an upper bound of A which is smaller than M.

1.3 Problems:

1. Show that $\sqrt{2}$ is irrational.

2. Let x and y be irrational numbers with $x > y$. Prove that there is a rational number q such that $x > q > Y$.

3. Let p and q be rational numbers with $p > q$. Prove that there is an irrational number x such that $p > x > q$.

4. Show that the set $A = \{y \in \mathbb{R} : y = 2x^2 - x^4, \quad x \in \mathbb{R}\}$ is bounded above and compute its least upper bound.

5. For $A \subset \mathbb{R}$ define $-A = \{x \in \mathbb{R} : -x \in A\}$. Show that $m = \sup A$ if and only if $-m = \inf(-A)$.

6. Let $A \subset \mathbb{R}$ be non-empty and bounded above and define $B = \{x \in \mathbb{R} : x \geq a \forall a \in A\}$. Prove that $\sup A = \inf B$.

7. Let A and B be bounded non-empty subsets of $\mathbb{R}$ with $A \subset B$. Prove that $\sup A \leq \sup B$ and $\inf A \geq \inf B$.

1.4 Countability

We have a good idea what finite sets are. In a finite set we can list all elements explicitly. However, infinite sets can be different. We know the infinite set of positive integers and for this set we can also list the elements, but we will never be able to finish the list. Infinite sets of this type are called **countable** sets. Sets which are neither finite nor countable are called uncountable. The sets $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ are all infinite, but are they countable. Moreover, the question is are there any uncountable sets or tdid we just introduce a new confusion.

Definition 5 *Let A and B be sets. The cardinality of the set is the number of its elements, if the set is finite. Otherwise the cardinality is either countably infinite or uncountably infinite. A and B have the same cardinality if there is a one to one and onto function $\phi : A \to B$.*

Proposition 4 *Let A and B be sets and $\phi : A \to B$. Then the cardinality of A is*

1. *greater than or equal to the cardinality of B, if ϕ is onto.*

2. *less than or equal to the cardinality of B, if ϕ is one-to-one.*

Moreover, if $A \subset B$ then the cardinality of B is greater than or equal to the cardinality of A

Proof: Left as a homework assignment.

Proposition 5 *If A and B are countable then so is $A \times B = \{(a, b) : a \in A, \quad b \in B\}$.*

Proof: Since A and B are countable there exist bijections $\phi : A \to \mathbb{N}$ and $\psi : B \to \mathbb{N}$. Define the function

$$\Phi : A \times B \to \mathbb{N}, \qquad \Phi : (a, b) \mapsto 2^{\phi(a)} 3^{\psi(b)}$$

This function is one-to-one. Thus the cardinality of $A \times B$ is at most the same as the cardinality of $\mathbb{N}$, and it is thus countable or finite. But if either A or B is infinite then so is $A \times B$.

This has immediate consequences.

Corollary 2 *1. $\mathbb{Z}$ is countable.*

2. $\mathbb{Q}$ is countable.

Proof: For the first assertion we use the map ϕ which does $\phi(n) = 2n$ for $n \geq 0$, and $phi(n) = 2(-n) - 1$ for $n < 0$. this is cleearly a bijection and thus $\mathbb{Z}$ is countable. For the second assertion observe tah $\mathbb{Z} \times \mathbb{N}$ is countable and use the map $\phi(m, n) = \frac{m}{n}$. This function is onto $\mathbb{Q}$ and thus $\mathbb{Q}$ is countable as well.
The question remains whether ther actually is an uncountable set. Well, we look at he set $[0, 1] = \{x \in \mathbb{R} : 0 \leq x \leq 1\}$. First, we observe that this set contains all numbers x with binary expansions

$$x = \sum_{n=1}^{\infty} x_n 2^{-n}, \qquad x_n \in \{0, 1\}.$$

Let us assume that the set of numbers with such binary expansions is countable. then we can make a complete list of them as follows:

$$\begin{aligned} x_1 &= \sum_{j=1}^{\infty} x_{1j} 2^{-j} \\ x_2 &= \sum_{j=1}^{\infty} x_{2j} 2^{-j} \\ x_3 &= \sum_{j=1}^{\infty} x_{3j} 2^{-j} \\ \vdots &= \vdots \end{aligned}$$

In other words the kth element in this list has the expansion

$$x_k = \sum_{j=1}^{\infty} x_{kj} 2^{-j}$$

Now construct the binary expansion

$$y = \sum_{j=1}^{\infty} y_j 2^{-j}$$

where $y_j \neq x_{jj}$. Then for any k we have $y \neq x_k$ since it differs from x_k in the kth term. Thus y is not in the list and the list was not complete. This shows that the the set of numbers with such binary expansions is not countable and the immediate consequence is

Proposition 6 *The real numbers are not countable.*

1.5 Problems:

1. Prove Proposition 5.

2. Show that, if A and B are countable then so is $A \cup B$.

3. For every $n \in \mathbb{N}$ let $A - n$ be a countable set. Prove that $\bigcup_{n \in \mathbb{N}} A_n = A_1 \cup A_2 \cup \ldots$ is also countable.

Chapter 2

Sequences and Series of Real Numbers

2.1 Limits of Sequences

The most central concept of Analysis is the concept of limit. In this part of the course we will develop that concept. Unfortunately, most calculus courses give a rather unsatisfactory definition of the limit of a function, long before they develop the limit of a series or sequence. The section is called sequences and series of real numbers, but there really is no reason to distinguish between them, as any sequence

$$a_1, a_2, a_3, \ldots$$

can be made into a series by the following process

$$b_1 = a_1, \qquad b_2 = a_2 - b_1, \qquad b_3 = a_3 - b_2, \ldots$$

and

$$a_n = \sum_{k=1}^{n} b_k$$

So we will really only look at sequences. However we will first need a formal definition.

Definition 6 *A sequence* $\{f_n\}$ *is a function* $f : \mathbb{N} \to \mathbb{R}$ *which maps* $n \mapsto f_n$.

Now it is important to distinguish, between the sequence and the image of the sequence. The sequence given by $a_n = 1$ is an infinite sequence even though it takes on only a single value.
We next introduce the idea of the limit of a sequence.

Definition 7 *Let* $\{a_n\}$ *be a sequence of real numbers¿ We say that this sequence converges to a number* a*, if* **for every** *positive number* ε **there exists a** *positive integer* N *such that*

$$|a_n - a| < \varepsilon, \tag{2.1}$$

for all $n \geq N$. *If* $\{a_n\}$ *converges to* a *we write*

$$\lim_{n\to\infty} a_n = a \qquad \textit{or} \qquad \{a_n\} \to a. \tag{2.2}$$

Before giving some examples we want to paraphrase this definition. The definition basically says that $\{a_n\} \to a$ if and only if for any $\varepsilon > 0$ (no matter is chosen) **all but finitely many terms** of the sequence satisfy the inequality

$$a - \varepsilon < a_n < a + \varepsilon.$$

Or we can say that all but finitely many terms will lie in the interval $(a - \varepsilon, a + \varepsilon)$.

Examples on how to use this definition:

To use the definition we need to think in around about way.

1. Let $a_n = \frac{1}{n}$ then $\{a_n\} \to 0$. Indeed, let $\varepsilon > 0$ be given. We see that $\frac{1}{n} < \varepsilon$ if and only if $n > \frac{1}{\varepsilon}$. So if we choose $N = \inf\{n \in \mathbb{N} : n > \frac{1}{\varepsilon}$ we get:
$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} < \varepsilon$$
for all $n \geq N$. So what we need to compute is the value of N for the given ε.

2. Let $b_n = \frac{n^2+2n+1}{3n^2}$. We first need to guess a value for the limit. My guess is that the limit equals $\frac{1}{3}$ Next we look at
$$\left|\frac{n^2+2n+1}{3n^2} - \frac{1}{3}\right| = \left|\frac{n^2+2n+1-n^2}{3n^2}\right| = \frac{2n+1}{3n^2} \leq \frac{2n+n}{3n^2} = \frac{1}{n}.$$
So if we choose the same N as in the previous example we are doing fine.

Now, this procedure is rather awkward, since we first need to know the limit and then go through the procedure to prove that it is the limit. In order to streamline the process we will first prove a few results.

Lemma 1 *Let* a *be a non-negative number. And suppose that* $a < \varepsilon$ *for every* $\varepsilon > 0$*, then* $a = 0$*.*

Proof: Let $A = \{\varepsilon \in \mathbb{R} : \varepsilon > 0\}$, then $a \leq \inf A$, but $\inf A = 0$, and thus $a \leq 0$. Now since $a \geq 0$ it follows that $a = 0$.

Definition 8 *A sequence* $\{a_n\}$ *is bounded, if there exists a number* $M > 0$ *such that* $|a_n| < M$ *for all* $n \in \mathbb{N}$*.*

Proposition 7 *Convergent sequences are bounded.*

Proof: Let $\{a_n\}$ be a convergent sequence with limit a. Choose $\varepsilon = 1$, then there exists a number N such that

$$a - 1 < a_n < a + 1$$

for all $n \geq \mathbb{N}$. Next we look at the set

$$A = \{|a_1|, |a_2|, \ldots, |a_{N-1}|, |a - 1|, |a + 1|\},$$

which is a finite set and therefore has a larges element. Let M be the largest element, then

$$|a_n| \leq M$$

for all $n \in \mathbb{N}$. Thus the sequence is bounded.

Observe that the converse of this is not true. For example the sequence $a_n = (-1)^n$ is certainly bounded, but does not converge. Up to now we said that a sequence converge to the limit a. Our use of the definite article was a little premature. We first need to show that the limit of a convergent sequence is unique.

Proposition 8 *A sequence of real numbers can have at most one limit.*

Proof: If the sequence does not converge it has no limits and the statement is true. Now let $\{a_n\}$ be a sequence such that $\{a_n\} \to a$ and $\{a_n\} \to a'$. Let $\varepsilon > 0$, then there exist numbers N_1 and N_2 such that

$$\begin{aligned} |a_n - a| &< \frac{\varepsilon}{2}, \qquad \text{for all} \quad n \geq N_1 \\ |a_n - a'| &< \frac{\varepsilon}{2}, \qquad \text{for all} \quad n \geq N_2 \end{aligned}$$

Then

$$|a - a'| = |a - a_n + a_n - a'| \leq |a - a_n| + |a_n - a'| < frac\varepsilon 2 + \frac{\varepsilon}{2}$$

Now the result follows from the previous Lemma.
The next theorem will give us a lot of new tools to compute limits of more complex sequences.

Theorem 1 *Let $\{a_n\}$ and $\{b_n\}$ be two convergent sequences with limits a and b, respectively. Then*

1. $\lim_{n\to\infty} ca_n = ca$ *for every* $c \in \mathbb{R}$
2. $\lim_{n\to\infty}(a_n + b_n) = a + b$
3. $\lim_{n\to\infty} a_n b_n = ab$
4. *If* $b_n \neq 0$ *for all* $n \in \mathbb{N}$ *and* $b \neq 0$*, we also have*$\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}$

Proof: To prove the first statement we may assume that $c \neq 0$, since if $c = 0$ the sequence $\{ca_n\}$ is identically equal to zero and hence converges to zero. Now let $\varepsilon > 0$ be given. Since $\{a_n\}$ converges, there exists a N such that

$$|a_n - a| < \frac{\varepsilon}{|c|},$$

for all $n \geq N$. Therefore we have

$$|ca_n - ca| = |c|\,|a_n - a| < |c|\frac{\varepsilon}{|c|} = \varepsilon$$

for all $n \geq N$. For the second assertion let $\varepsilon > 0$ be given. Now, since $\{a_n\}$ converges there exist a number N_1 such that

$$|a_n - a| < \frac{\varepsilon}{2}$$

for all $n \geq N_1$. And since $\{b_n\}$ converges, there exists a number N_2 such that

$$|b_n - b| < \frac{\varepsilon}{2}$$

for all $n \geq N_2$. Let N be the larger one of N_1 and N_2. We have by the triangle inequality:

$$|a_n + b_n - (a+b)| = |a_n - a + b_n - b| \leq |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for all $n \geq N$, and we are done.

Things get a little more complicated for the last two assertions. First, observe that both sequences are bounded. There exists a number $M > 0$ such that $|a_n| < M$ and $|b_n| < M$ for all $n \in \mathbb{N}$, and $|a| < M$ and $|b| < M$. Now, we have

$$|a_nb_n - ab| = |a_nb_n - ab_n + ab_n - ab| = |(a_n - a)b_n + a(b_n - b)| \leq |b_n||a_n - a| + |a||b_n - b| < M|a_n - a| + M|b_n - b|.$$

Mow we can find for any $\varepsilon > 0$ numbers N_1 and N_2 such that

$$\begin{aligned} |a_n - a| &< \frac{\varepsilon}{2M}, \qquad \text{for all} \quad n \geq N_1 \\ |b_n - b| &< \frac{\varepsilon}{2M}, \qquad \text{for all} \quad n \geq N_2 \end{aligned}$$

Let $N = \max\{N_1, N_2\}$ and the result follows in the same way as the previous one.

For the last assertion we need to to some preparatory work. To begin, $b_n \neq 0$ and $b \neq 0$, there exists a number N_1 such that

$$-\frac{|b|}{2} < b_n - b < \frac{|b|}{2}$$

for all $n \geq N_1$. Subtracting b and some algebra implies now

$$-\frac{|b|}{2} < b_n < \frac{|b|}{2}$$

for all $n \geq N_1$. Next observe that

$$\left|\frac{a_n}{b_n} - \frac{a}{b}\right| = \left|\frac{a_nb - ab_n}{bb_n}\right| = \left|\frac{a_nb - ab + ab - ab_n}{bb_n}\right| \leq \frac{|b||a_n - a|}{|bb_n|} + \frac{|a||b_n - b|}{|bb_n|} < \frac{2|a_n - a|}{|b|} + \frac{|a|2|b_n - b|}{|b|^2},$$

for all $n \geq N_1$. We may assume that $|a| > 0$, since otherwise the second term equals 0. So for a given $\varepsilon > 0$ there exist numbers N_2 and N_3 such that

$$\begin{aligned} |a_n - a| &< \frac{\varepsilon|b|}{4}, \qquad \text{for all} \quad n \geq N_2 \\ |b_n - b| &< \frac{\varepsilon|b|^2}{2|a|}, \qquad \text{for all} \quad n \geq N_3 \end{aligned}$$

and the result follows by setting $N = \max\{N_1, N_2, N_3\}$.

Example: Consider the sequence $a_n = \frac{n^2+2n+3}{2(n+1)^2}$. Now we cannot use the fourth assertion of the Theorem, since neither the numerator nor the denominator are convergent. However, we can write

$$a_n = \frac{n^2+2n+3}{2(n+1)^2} = \frac{(n=1)^2+2}{2(n+1)^2} = \frac{1}{2} + \frac{1}{(n+1)^2},$$

and we have the sum of two convergent sequences. Thus $\lim_{n\to\infty} a_n = \frac{1}{2}$.
Sometimes sequences are defined by recursive relations. For example the sequence

$$x_1 = 2, \qquad x_{n+1} = \frac{x_n}{2} + \frac{1}{x_n}$$

converges, but it is very hard to use any of the previous methods to either show that it converges or to find its limit. However, If it converges we must have

$$x = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty}\left(\frac{x_n}{2} + \frac{1}{x_n}\right) = \frac{x}{2} + \frac{1}{x},$$

and thus

$$x^2 = 2.$$

So we are left to prove that the sequence actually converges. To this end we will use the monotone convergence theorem, which will be proved next.

Definition 9 *A sequence $\{a_n\}$ is eventually monotone inreasing (decreasing) if there is a number N such that $a_{n+1} \geq a_n$ ($a_{n+1} \leq a_n$) for all $n \geq N$.*

Theorem 2 *Let $\{a_n\}$ be a sequence that is eventually monotne increasing and bounded above. Then $\{a_n\}$ converges and*

$$\lim_{n\to\infty} a_n = \sup\{a_n : n \geq N\},$$

where N is a number such that $a_{n+1} \geq a_n$ for all $n \geq N$. The analogous result holds for decreasing sequences that are bounded below.

Proof: The set $\{a_n : n \geq N\}$ is nonempty and bouded above thus it has a least upper bound a. Now for every $\varepsilon > 0$ there exists an element a_M in this set such that

$$a - \varepsilon < a_M \leq a < a + \varepsilon$$

Moreover, since the sequence is increasing for $n \geq M$ we have

$$a_M \leq a_n \leq a < a + \varepsilon$$

for all $n \geq M$. Combining these statements proves the theorem. The proof for the analogous result for decreasing sequences follows the same line of reasoning.

We can now return to the above example. We need to show two things. First that the sequence x_n is either bounded above or below, and second that the sequence is monotone. Consider the function

$$f(x) = \frac{x}{2} - \frac{1}{x}$$

First if $x > 0$ then $f(x) > 0$, and since $x_1 = 2 > 0$ we have $x_n > 0$ for all n. Moreover, from Calculus we know that $f(x) \geq \sqrt{2}$ for all $x > 0$, and $f(\sqrt{2}) = \sqrt{2}$. So $\inf\{x_n : n \in \mathbb{N}\} = \sqrt{2}$. Now,

$$x_{n+1} - x_n = \frac{x_n}{2} + \frac{1}{x_n} - x_n = \frac{1}{x_n} - \frac{x_n}{2} \leq \frac{1}{\sqrt{2}} - \frac{\sqrt{2}}{2} = 0,$$

and thus the sequence is decreasing. The theorem applies and $\lim_{n\to\infty} x_n = \sqrt{2}$.

2.2 Problems

1. Let $0 < a < 1$. Prove that $\lim_{n\to\infty} a^n = 0$ by using the definition of convergence.

2. Let $0 < a < 1$. Show that the sequence $\{a^n\}$ is bounded and decreasing, and use this to show $\lim a^n = 0$.

3. Let x_n be a sequence of positive numbers, such that

   $$\lim_{n\to\infty} \frac{x_{n+1}}{x_n} = L.$$

   Prove that $\lim_{n\to\infty} x_n = 0$, if $L < 1$, and that x_n diverges if $L > 1$.

4. Give an example of a convergent sequence $\{x_n\}$ such that $\lim_{n\to\infty} \frac{x_{n+1}}{x_n} = 1$.

5. Find a sequence $\{x_n\}$ that does not converge, but $\lim_{n\to\infty} \frac{x_{n+1}}{x_n} = 1$.

2.3 Subsequences and the Bolzano-Weierstrass Theorem

In the previous section we explored convergent sequences. In particular we saw that all convergent sequences are bounded, but that the converse of this is not true. This secton will explore a partial converse of this result. To begin we need to say something about subsequences. Losely speaking a subsequence of a sequence is obtained by leaving out some terms of this sequence. For example the sequence $\{(-1)^n/n\}$ has a subsequence of the form $\{1/2n\}$ which is obtained by ignoring all the odd terms.

Definition 10 *Let $\{a_n\}$ be a sequence of numbers, and $\{n_k\}$ be a strictly increasing sequence of positive integers. Then $\{a_{n_k}\}$ is a subsequence of $\{a - n\}$.*

Of particular interest are convergent subsequences of sequences. If the sequence it self converges, then every subsequence of it will converge to the same limit as the sequence. This statement will be proved in an exercise. More interesting is the situation when the sequence itself doesn't converge.

Definition 11 *Let $\{a_n\}$ be a sequence of real numbers. A number a is a subsequential limit point of the sequence, if for every $\varepsilon > 0$ infinitely many terms of the sequence satisfy $a - \varepsilon < a_n < a + \varepsilon$.*

Now it is important to see the difference between this and the definition of the limit of a sequence. In the earlier definition we require all but finitely many terms to satisfy this inequality. In this definition there may infinitely many terms which do not satisfy the definition. However, it is clear that the limit of a sequence (if it exists) is also a subsequential limit point. A sequence may have no subsequential limit points at all, for example the sequence

$$\{a_n\} = \{n\}.$$

Convergent sequences have exactly one subsequential limit point. The sequence

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

has two subsequential limit points. The number of subsequential limit points can infinite and even uncountable. Indeed, since the rational are countable they theoretically form a sequence of real numbers, and for every $a \in \mathbb{R}$ and every $\varepsilon > 0$ the interval $(a - \varepsilon, a + \varepsilon$ contains infinitely many rational numbers, making a a subsequential limit point of $\mathbb{Q}$.

Lemma 2 *Let $\{a_n\}$ be a sequence that is bounded above and and has at least one subsequential limit point. Let A denote the set of all subsequential limit points of $\{a_n\}$ and*

$$L = \sup A$$

Then L is a subsequential limit point itself which we denote by

$$L = \lim \sup_{n\to\infty} a_n$$

Proof: A is non-empty and bounded above and has there for a least upper bound L. Now let $\varepsilon > 0$ then there is at least one $x \in A$ such that $L - \varepsilon/2 < x \leq L$. x is a sequential limit point of a_n, and therefore infinitely many terms of the sequence satisfy $x - \varepsilon/2 < a_n < x + \varepsilon/2$. Combining the two inequalities we get

$$L - \varepsilon < x - \frac{\varepsilon}{2} < a_n < x + \frac{\varepsilon}{2} < L + \varepsilon,$$

and hence L is itself a subsequential limit point.
For sequences bounded below we can define

$$\lim \inf_{n\to\infty} a_n = \inf A$$

where A is again the set of subsequential limit points.

Proposition 9 *Let $\{a_n\}$ be a sequence. Then L is a susequential limit poit of $\{a_n\}$ if and only if there exists a subsequence $\{a_{n_k}\}$ such that*

$$\lim_{k\to\infty} a_{n_k} = L$$

Proof: Let L be a subsequential limit point and for every $k \in \mathbb{N}$ let $\varepsilon_k = \frac{1}{k}$. Now ffor $k = 1$ there exist infinitely many terms of the sequence which satisfy

$$L - 1 < a_n < L + 1.$$

Pick of these terms and call it a_{n_1}. To continue, for every $k > 1$ there are infinitely many term sof the sequence qhich satisfy

$$L - \varepsilon_k < a_n < L + \varepsilon_k, \quad \textbf{and} \qquad n > n_{k-1}$$

pick one and call it a_{n_k}. Now for any $\varepsilon > 0$ there is a number K such that $\frac{1}{k} < \varepsilon$ for all $k \geq K$. It follows that

$$|a_{n_k} - L| < \frac{1}{k} < \varepsilon,$$

for all $k \geq K$, and thus

$$\lim_{k\to\infty} a_{n_k} = L.$$

For the converse, if a subsequence $\{a_{n_k}\}$ converges to L, then for every $\varepsilon > 0$ there are infinitely many terms of the subsequence in the interval $(L - \varepsilon, L + \varepsilon)$, but then this interval contains infinitely terms of the original sequence (since every term of the subsequence is a term of the original sequence as well) and thus L is a subsequential limit point.

We are now ready to state the main result of this section.

Theorem 3 BOLZANO-WEIERSTRASS. *Every bounded sequence has has a convergent subsequence.*

Before proving this theorem we need an other result

Proposition 10 *Let $\{a_n\}$, $\{b_n\}$, and $\{c_n\}$ be sequences such that there is a $N \in \mathbb{N}$ such that $a_n \leq b_n \leq c_n$ for all $n \geq N$ and $\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n$. The $\{b_n\}$ converges to the same limit.*

Proof: Let $\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = L$ and let $\varepsilon > 0$. Then there exist numbers N_1 and N_2 such that

$$L - \varepsilon < a_n \qquad \text{for all} \qquad n \geq N_1$$
$$c_n < L + \varepsilon \qquad \text{for all} \qquad n \geq N_2$$

Thus we have

$$L - \varepsilon < a_n \leq b_n \leq c_n < L + \varepsilon$$

for all $n \geq \max\{N, N_1, N_2\}$, and $\{b_n\}$ converges to L.

Proof of the Bolzano-Weierstrass Theorem: Let $\{a_n\}$ be a bounded sequence. Then there exists a number M such that $-M \leq a_n \leq M$ for all $n \in \mathbb{N}$. We now costruct two new sequences as follows. Let $x_0 =_M$ and $y_0 = M$. For next term we look at the two regions $x_0 \leq x \leq \frac{x_0+y_0}{2}$ and $\frac{x_0+y_0}{2} \leq x \leq y_0$. At least one of these regions contains infinitely many many terms we set

$$x_1 = x_0, \quad y_1 = \frac{x_0 + y_0}{2} \qquad \text{if infinitely many terms satisfy} \qquad x_0 \le a_k \le \frac{x_0 + y_0}{2}$$

and

$$x_1 = \frac{x_0 + y_0}{2}, \quad y_1 = y_0 \qquad \text{if finitely many terms satisfy} \qquad x_0 \le a_k \le \frac{x_0 + y_0}{2}$$

We continue to construct higher terms in the same way

$$x_{n+1} = x_n, \quad y_{n+1} = \frac{x_n + y_n}{2} \qquad \text{if infinitely many terms satisfy} \qquad x_n \le a_k \le \frac{x_n + y_n}{2}$$

and

$$x_{n+1} = \frac{x_n + y_n}{2}, \quad y_{n+1} = y_n \qquad \text{if finitely many terms satisfy} \qquad x_n \le a_k \le \frac{x_n + y_n}{2}$$

These two sequences satisfy the following properties:

$$x_{n+1} \ge x_n \quad \text{and} \quad x_n \le M \quad \text{for all} \quad n \in \mathbb{N}$$

and

$$y_{n+1} \le y_n \quad \text{and} \quad y_n \ge -M \quad \text{for all} \quad n \in \mathbb{N}$$

and

$$y_{n+1} - x_{n+1} = \frac{1}{2}(y_n - x_n) = \frac{1}{2}\frac{M}{2^n}, \qquad \text{for all} \quad n \in \mathbb{N}$$

The first property implies that the sequence $\{x_n\}$ converges to some limit L_1, the second implies that $\{y_n\}$ converges to some L_2. The last property implies that $L_2 - L_1 = 0$. So we may call the limit L. Now let $\varepsilon > 0$ the there are numbers N_1 and N_2 such that

$$L - \varepsilon < x_n \le L \qquad \text{for all} \qquad n \ge N_1$$
$$L \le y_n < L + \varepsilon \qquad \text{for all} \qquad n \ge N_2$$

Now let $N = \max\{N_1, N_2\}$ and we have

$$L - \varepsilon < x_N < y_N < L + \varepsilon$$

By the construction of the sequence infinitely may terms of $\{a_n\}$ lie between x_N and y_N, thus

$$L - \varepsilon < a_n < L + \varepsilon$$

for infinitely many terms of the sequence, i.e. L is a subsequential limit point of the sequence. Therefore, there is a subsequence which converge to L.

We finish this section with some more results about subsequences and limits.

Proposition 11 *Let $\{a_n\}$ be a bounded sequence. Then $\{a_n\}$ converges if and only if*

$$\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n$$

Proof: Suppose $a_n \to L$. Then for every $\varepsilon > 0$ there is only finitely many terms of the sequence which lie outside the interval $(L-\varepsilon, L+\varepsilon)$. In particular, there are only finitely many terms of the sequence which satisfy $a_n \geq L+\varepsilon$. Thus,

$$\limsup_{n\to\infty} a_n \leq L.$$

On the other hand there is only finitely may terms which satisfy $a_n \leq L-\varepsilon$ and therefore

$$\liminf_{n\to\infty} a_n \geq L.$$

But by definition

$$\liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n,$$

which implies

$$\liminf_{n\to\infty} a_n = L = \limsup_{n\to\infty} a_n.$$

Conversely, if

$$\liminf_{n\to\infty} a_n = L = \limsup_{n\to\infty} a_n,$$

the set of limit points of the sequence contains only a single point L, which satisfies that for every $\varepsilon > 0$ there are only finitely many terms such that

$$a_n \geq L+\varepsilon,$$

and finitily many terms such that

$$a_n \leq L-\varepsilon.$$

Thus, all but finitely many terms satisfy

$$L-\varepsilon < a_n < L+\varepsilon,$$

and the sequence converges to L.

Proposition 12 *Let $\{a_n\}$ and $\{b_n\}$ be sequences such that there is a number N such that*

$$a_n < b_n,$$

for all $n \geq N$. Then

$$\limsup_{n\to\infty} a_n \leq \limsup_{n\to\infty} b_n$$

and

$$\liminf_{n\to\infty} a_n \leq \liminf_{n\to\infty} b_n$$

The proof of this is left as a homework assignment.

2.4 Problems

1. Prove the last proposition.

2. Let $\{a_n\}$ and $\{b_n\}$ be to bounded sequences. Prove that $\limsup_{n\to\infty}(a_n + b_n) \leq \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n$.

3. Find an example of sequences $\{a_n\}$ and $\{b_n\}$ such that $a_n < b_n$ for all n, but $\limsup_{n\to\infty} a_n = \limsup_{n\to\infty} b_n$

4. Let $a_n = \sin n$. Find $\limsup_{n\to\infty} a_n$ and $\liminf_{n\to\infty} a_n$. Moreover show that the sequence has a subsequence with subsequential limit point $\frac{1}{2}$

2.5 Cauchy Sequences

In this section we explore another aspect of sequences of real numbers.

Definition 12 *A sequence of real numbers $\{a_n\}$ is called a Cauchy sequence if for every $\varepsilon >)$ there exists a number $N \in \mathbb{N}$ such that*

$$|a_n - a_m| < \varepsilon, \qquad \textit{for all} \quad m, n \geq N.$$

Losely speaking a Cauchy sequence is a sequence whose term get closer and closer together. Now such sequences seem to be good candidates for converging sequences, in particular we have

Proposition 13 *Every converging sequence is a Cauchy sequence.*

Proof: Let $\{a_n\}$ be a converging sequence with limit L. Then for every $\varepsilon > 0$ there is a number $N \in \mathbb{N}$ such that

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

for all $n \geq N$. Now we have

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

for all $n, m \geq N$.Thus the sequence is a Cauchy sequence.

Our objective is to prove the converse of this. This will provide us with an effective tool to check for the convergence of a sequence without the explicit knowledge of the limit. We start wit an easy step.

Lemma 3 *Every Cauchy sequence is bounded.*

Proof: The proof of this lemma is almost identical to the proof that every converging sequence is bounded. We start by letting $\varepsilon = 1$ for simplicity. The there is a number N such that

$$|a_n - a_m| < 1,$$

for all $n, m \geq N$. In particular we have

$$a_N - 1 < a_n < a_N + 1,$$

for all $n \geq N$. Let

$$M = \max\{|a_1|, |a_2|, \ldots, |a_{N-1}|, |a_N - 1|, |a_N + 1|\},$$

and we see that $|a_n| \leq M$ for all $n \in \mathbb{N}$ and therefore the sequence is bounded.

Now that Cauchy sequences are bounded they have a subsequential limit point by the Bolzano-Weierstrass Theorem. It remains to be shown that the whole sequence converges to that limit point.

Lemma 4 *Let $\{a_n\}$ be a Cauchy sequence and L be a subsequential limit point of this sequence. Then the sequence converges to L.*

Proof: Since L is a subsequential limit point of the sequence there exists a subsequence $\{a_{n_k}\}$ that converges to L. Now let $\varepsilon > 0$, then there exists a number K such that

$$|a_{n_k} - L| < \frac{\varepsilon}{2},$$

for all $k \geq K$. Observe that $n_k \geq k$. Since $\{a_n\}$ is a Cauchy sequence there is a number N such that

$$|a_n - a_m| < \frac{\varepsilon}{2},$$

for all $n, m \geq N$. Let $\hat{N} = \max\{N, K\}$ and pick a term of the subsequence such that $n_k \geq \hat{N}$, then

$$|a_n - L| \leq |a_n - a_{n_k}| + |a_{n_k} - L| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2},$$

for all $n \geq \hat{N}$. Thus the sequence converges.

We combine these results into

Theorem 4 *Every Cauchy sequence of real numbers converges.*

This concept of Cauchy sequences is far more important than it seems at first glance. Recall that we defined completeness of the real numbers by looking at least upper bounds. To do this it is necessary to have an order relation defined on the real numbers. If we want to do Calculus in several variables, we can't dot this anymore since there exist no meaningful order relation on $\mathbb{R}^n$ for $n \geq 2$. And this is where Cauchy sequences come in. Sequences and convergence can be define in rather abstract spaces, so called metric spaces. Completeness can thus be also defined in abstract metric spaces as follows.

Definition 13 *A metric space X is complete, if every Cauchy sequence in X converges to a limit in X.*

WE could have therefore constructed the real numbers as follows. Start with the rationals and look at all Cauchy sequences of rational numbers. Now some of them will converge to rational limits, others will have no limit inthe rationals. Next you define an equivalence relation between two Cauchy sequences, by saying the sequences are equivalent, if there difference converges to zero. Then you look at the set of all Cauchy sequences of rational numbers, they still form a field with equality being replace by the equivalence relation. You identify each sequence with a limit. The set of limits is the real numbers. You need to finish the construction by showing that this set has all the properties that we expect from the real numbers.

Chapter 3

The Topology of the Real Numbers

3.1 Open and Closed Subsets of the Real Numbers

We all are familiar with open and closed intervals since secondary schools. Topology expands these concepts to more general sets. To begin lets define intervals. We say a subset $A \subset \mathbb{R}$ is an interval if for any $a, b \in A$ with $a < b$ we have

$$t \in A \qquad \text{for all} \qquad a < t < b. \tag{3.1}$$

Definition 14 *A subset $O \subset \mathbb{R}$ is called open if for every $x \in O$ there exists an $\varepsilon > 0$ such that the interval $(x - \varepsilon, x + \varepsilon)$ is entirely contained in O, in other words $(x - \varepsilon, x + \varepsilon) \subset O$.*

Using this definition we can prove that the open interval (a, b) is opne in the sense of this definition as well. Let $x \in (a, b)$, then $a < x < b$. Let $\varepsilon = \min\{(x-a)/2, (b-x)/2\}$, then $x - \varepsilon \in (a, b)$ and $x + \varepsilon \in (a, b)$ and therefore $(x - \varepsilon, x + \varepsilon) \subset (a, b)$. Moreover, the set $\mathbb{R}$ is clearly open. For a more interesting example we look at the empty set. I claim that for every $x \in \emptyset$ the interval $(x - 1, x + 1) \subset \emptyset$. While this statement sounds ludricous at first it is definitely a true statement, since there are no such x's, and nonexisting entities have any property we want them to have.

We next look at collections of sets. WE start with an index set $\mathcal{J}$. This may be a finite, countable or uncountable set. For each $j \in \mathcal{J}$ we have a subset $O_j \subset \mathbb{R}$. Then the collection $\mathcal{O}$ of sets is defined as

$$\mathcal{O} = \{O_j : j \in \mathcal{J}\}.$$

Proposition 14 *1. $\mathbb{R}$ and $\emptyset$ are open sets.*

2. Let O_1 and O_2 be open, then $O_1 \cap O_2$ is open.

3. *Let $\mathcal{O}$ be an arbitrary collection of open sets, then*

$$\bigcup_{O \in \mathcal{O}} O$$

is open.

Proof: The first item was shown before. For item 2, if the intersection is empty there is nothing to show. If not let $x \in O_1 \cap O_2$, then there exist $\varepsilon_1, \varepsilon_2 > 0$ such that $(x - \varepsilon_1, x + \varepsilon_1) \subset O_1$ and $(x - \varepsilon_, x + \varepsilon_2) \subset O_2$. Let $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\}$, then $(x - \varepsilon, x + \varepsilon) \subset)_1$ and $(x - \varepsilon, x + \varepsilon) \subset O_2$ and thus

$$(x - \varepsilon, x + \varepsilon) \subset O_1 \cap O_2.$$

For the last assertion let $x \bigcup_{O \in \mathcal{O}} O$. Then there exists a $O_0 \in \mathcal{O}$ such that $x \in O_0$. Since O_0 is open there exists an $\varepsilon > 0$ suct that

$$(x - \varepsilon, x + \varepsilon) \subset O_0 \subset \bigcup_{O \in \mathcal{O}} O.$$

Therefore, the union of an arbitrary collection of open sets is open.

The collection of all open subsets of $\mathbb{R}$ is called the (usual) topology of $\mathbb{R}$. Any collection of subsets which contains $\mathbb{R}$ and the empty set, is closed under finite intersections, and closed under arbitrary union is called a topology on $\mathbb{R}$. We will not use any more abstract topologies in this course.

Definition 15 *A subset $F \subset \mathbb{R}$ is closed if its complement is an open set.*

Proposition 15 1. *$\mathbb{R}$ and $\emptyset$ are closed sets.*

2. *Let F_1 and F_2 be closed, then $F_1 \cup F_2$ is closed.*

3. *Let $\mathcal{F}$ be an arbitrary collection of closed sets, then*

$$\bigcap_{F \in \mathcal{F}} F$$

is closed.

Proof: See the assignment.

Definition 16 *Let A be an arbitrary subset of $\mathbb{R}$.*

1. *The interior of A is the largest open subset O of $\mathbb{R}$ such that $O \subset A$. We write $O = A^0$.*

2. *The closure of A is the smallest closed subset F of $\mathbb{R}$ such that $A \subset F$. We write $F = \overline{A}$.*

3. *The boundary of A is $\partial A = \overline{A} \cap \overline{A^c}$.*

If we look for example at the set $A = [0,1) \cup \{3\}$, then $A^0 = (0,1)$, $\overline{A} = [0,1] \cup 3$, and $\partial A = \{0,1,3\}$. More interesting cases are if $A = \mathbb{Q}$. In this case $\mathbb{Q}^0 = \emptyset$, $\overline{\mathbb{Q}} = \mathbb{R}$, and $\partial\mathbb{Q} = \mathbb{R}$.

Proposition 16 *Let A be an arbitrary subset of $\mathbb{R}$. Then*

1.
$$A^0 = \bigcup \mathcal{O} \qquad \textit{where} \qquad \mathcal{O} = \{O : O \quad \textit{open and} \quad O \subset A\}$$

2.
$$\overline{A} = \bigcap \mathcal{F} \qquad \textit{where} \qquad \mathcal{F} = \{F : F \quad \textit{closed and} \quad A \subset F\}$$

Proof: To show the first assertion, clearly

$$\bigcup \mathcal{O} \qquad \text{where} \qquad \mathcal{O} = \{O : O \quad \text{open and} \quad O \subset A\} \subset A^0,$$

since all the sets in this collection are subsets. On the other hand A^0 itself is an open subset of A and thus part of the collection and

$$A^0 \subset \bigcup \mathcal{O} \qquad \text{where} \qquad \mathcal{O} = \{O : O \quad \text{open and} \quad O \subset A\}.$$

The second assertion is left as an exercise.

Definition 17 *Let A be a subset of $\mathbb{R}$. Then*

1. *$x \in A$ is an interior point of A if there exists an $\varepsilon > 0$ such that $(x-\varepsilon, x+\varepsilon) \subset A$.*
2. *$x \in \mathbb{R}$ is a limit point of A if for every $\varepsilon > 0$ we have $((x-\varepsilon, x) \cup (x, x+\varepsilon)) \cap A \neq \emptyset$.*
3. *$x \in \mathbb{R}$ is a boundary point of A, if for every $\varepsilon > 0$ we have $(x-\varepsilon, x+\varepsilon) \cap A \neq \emptyset$ and $(x-\varepsilon, x+\varepsilon) \cap A^c \neq \emptyset$.*
4. *$x \in A$ is an isolated point of A if there exists an $\varepsilon > 0$ such that $(x-\varepsilon, x+\varepsilon) \cup A \setminus \{x\} = \emptyset$.*

This definition gives local properties of the concepts of interior and boundary. It is clear that a boundary point is either a limit point or an isolated point. It is also clear that every interior point must be a limit point. Moreover, the definition of isolated point implies that

$$(x-\varepsilon, x+\varepsilon) \cup A^c \neq,$$

and since $x \notin A^c$ this implies that an isolated point is a limit point of A^c. Boundary points are clearly limit points of either A or A^c.

Proposition 17 *Let A be a subset of $\mathbb{R}$. Then*

1. *The interior of A is the set of all interior points of A*

2. *The closure of A consist of A together with all the limit points of A.*

3. *The boundary of A is the set of all the boundary points of A*

Proof: Let $x \in A^0$, since A^0 is open there exists an $\varepsilon > 0$ such that $(x-\varepsilon, x+\varepsilon) \subset A^0 \subset A$, thus x is an interior point. On the other hand if x is an interior point then the $(x-\varepsilon, x+\varepsilon)$ is an open subset of A and thus in the union of all open subsets of A which is A^0. Let $x \in \overline{A}$ and suppose that $x \notin A$. We need to show that x is a limit point of A. Suppose x is not a limit point of A then there exists an $\varepsilon > 0$ such that $(x-\varepsilon, x+\varepsilon) \cap A = \emptyset$, and thus $(x-\varepsilon, x+\varepsilon) \subset A^c$. Let $F = \mathbb{R} \setminus (x-\varepsilon, x+\varepsilon)$, then F is closed and $A \subset F$, and therefore $\overline{A} \subset F$. But $x \notin F$ and hence $x \notin \overline{A}$. Finally, let x be a boundary point which is not in A, then x is alimit poit of both A and A^c and $x \in \overline{A} \cap \overline{A^c}$ and thus in the boundary. If $x \in A$, but not a limit point, then x must be an isolated point and thus a limit point of A^c.

We use these concepts to prove an important result.

Theorem 5 *Every bounded infinite subset of $\mathbb{R}$ has at least one limit point.*

Proof: A quick and easy proof of this is an application of the Bolzano-Weierstrass Theorem. Let A be an infinite bounded subset of $\mathbb{R}$. Create a sequence $\{x_n\}$ as follows:

$$\begin{aligned}
&x_1 \in A \\
&x_2 \in A \setminus \{x_1\} \\
&\vdots \\
&x_n \in A \setminus \{x_1, \ldots, x_{n-1}\}
\end{aligned}$$

This sequence is a bounded sequence and has a subsequential limit point L. No two terms of the sequence are equal. Now for every $\varepsilon > 0$ the interval $(L-\varepsilon, L+\varepsilon)$ contains infinitely may different terms of the sequence and thus infinitely many elements of A. Hence L is a limit point of A.

3.2 Problems

1. Show that all elements of a finite set are isolated points.

2. Let $A \subset \mathbb{R}$ and x be a limit point of A. Show that for every $\varepsilon > 0$ the set $(x-\varepsilon, x+\varepsilon)$ contains infinitely many elements of A.

3. Show that $\overline{A} = ((A^c)^0)^c$ and $A^0 = (\overline{A^c})^c$.

4. Let p be a prime number and
$$P = \left\{ \frac{k}{p^n} : k \in \mathbb{Z}, n \in \mathbb{N} \cup \{0\} \right\}.$$
Show that $\overline{P} = \mathbb{R}$.

5. Prove Proposition 15

3.3 Compactness and the Heine Borel Theorem

Compactness is one of the most important concepts in topology. It is unfortunately not easy to understand for beginnig students. The basic idea is that it allows us to reduce a complex infinite structure by a more managable finite one. We start by getting the stage set to introduce this concept.

Definition 18 *Let $C \subset \mathbb{R}$. An open covering of C is a collection $\mathcal{O}$ of open sets such that*

$$C \subset \bigcup_{O \in \mathcal{O}} . \tag{3.2}$$

A subcovering $\mathcal{O}'$ of $\mathcal{O}$ is a subcollection of $\mathcal{O}$ that still satisfies the covering property (3.2).

If we consider $C = [0, 1]$ we have the following open coverings.

1. $\mathcal{O}_1 = \{(-1/2, 1/2), (0, 1), (1/2, 3/2)\}$
2. $\mathcal{O}_2 = \{(r - 1/2, r + 1/2) : r \in \mathbb{Q} \cap [0, 1]\}$
3. $\mathcal{O}_3 = \{(r - q, r + q) : r \in \mathbb{Q} \cap [0, 1], q \in \mathbb{Q} \cap (0, 1)\}$
4. $\mathcal{O}_4 = \{x - \varepsilon, x + \varepsilon : x \in [0, 1], \varepsilon > 0\}$

Observe that $\mathcal{O}_1$ has finitely many elements, $\mathcal{O}_2$ and $\mathcal{O}_3$ have countably many elements, and $\mathcal{O}_4$ has uncountably many elements.

Definition 19 *A set $C \subset \mathbb{R}$ is compact if every open covering of C has a finite subcovering.*

Now it is sort of easier to understand what a non-compact set looks like. For example, the set $(0, 1]$ is not compact. To see this we have to find an open covering that does not have a finite subcovering. So for any $n \in \mathbb{N}$ let $O_n = (1/n, 3/2)$ and let

$$\mathcal{O} = \{O_n : n \in \mathbb{N}\}$$

Clearly if $x \in (0, 1]$ there is a number N such that $x \in O_N$ and therefore

$$(0, 1] \subset \bigcup_{n \in \mathbb{N}} O_n,$$

so $\mathcal{O}$ is an open covering of $(0, 1]$. Now take any finite subcollection of $\mathcal{O}$, say $\{O_{n_1}, \ldots, O_{n_k}\}$ and let $N = \max\{n_1, \ldots, n_k\}$. Clearly there is an $x \in (0, 1]$ such that $x < 1/N$, and thus $x \notin O_{n_1} \cup \ldots \cup O_{n_k}$, so the finite subcollection is not a subcovering. Since this is true for any finite subcollection, the open covering does not have a finite subcovering.

On the other hand it is clear that finite sets are compact. Indeed let $C = \{x_1, x_2, \ldots, x_n\}$ and $\mathcal{O}$ be an arbitrary open covering. Then there exist open sets $O_1, O_2, \ldots, O_n$ such that $x_j \in O_j$ for $j = 1, \ldots, n$. Hence, $C \subset O_1 \cup \ldots \cup O_n$ and we have thus constructed a finite subcover.

Proposition 18 *Let C be a compact subset of $\mathbb{R}$. Then C is bounded.*

Proof: Consider the open covering $\mathcal{O} = \{O_x : x \in C\}$ where $O_x = (x-1, x+1)$. Since C is compact this must have a finite subcover, i.e. there are numbers $x_1, \ldots, x_n$ such that $C \subset O_{x_1} \cup \ldots \cup O_{x_n}$. Now let $L = \min\{x_1, \ldots, x_n\}$ and $M = \max\{x_1, \ldots, x_n\}$. Then $L - 1 \le x \le M + 1$ for all $x \in C$, i.e. C is bounded.

Proposition 19 *Let C be a compact subset of $\mathbb{R}$. Then C is closed.*

Proof: Let $x \notin C$. Consider the open covering $\mathcal{O} = \{O_n : n \in \mathbb{N}\}$, where $O_n = (-\infty, x - 1/n) \cup (x + 1/n, \infty)$. Since C is compact there is numbers $N_1, \ldots, N_k$ such that

$$C \subset O_{N_1} \cup \ldots \cup O_{N_k}.$$

Let $N = \max\{N_1, \ldots, N_k\}$. Then

$$(x - \frac{1}{2N}, x + \frac{1}{2N}) \cap C \subset (x - \frac{1}{2N}, x + \frac{1}{2N}) \cap (O_{N_1} \cup \ldots \cup O_{N_k}) = \emptyset,$$

which means that x cannot be a limit point of C. Thus C contains all its limit points an is thus closed.

The next result will give us a better handle on things.

Proposition 20 *Let C be compact and $B \subset C$ be a closed set. Then B is compact.*

Proof: Let $\mathcal{O}$ be an arbitrary open cover of B. The idea of the proof is to construct an open cover of C using the open cover of B and then use the compactness of C. To this end let

$$\mathcal{O}' = \mathcal{O} \cup \{B^c\}.$$

This is clearly an open covering of C, as

$$\bigcup_{O \in \mathcal{O}'} O = \mathbb{R}.$$

Thus there exist $O_1, \ldots, O_N$ such that

$$C \subset O_1 \cup \ldots \cup O_N,$$

Now if B^c is not one of these sets we already have a finite subcover of $\mathcal{O}$. Otherwise say $B^c = O_1$, then

$$B \subset O_2 \cup \ldots \cup O_N,$$

and we thus have a finite subcover of $\mathcal{O}$.

A big advantage of the real numbers is that we have a countable dense subset of $\mathbb{R}$, that is a countable subset whose closure is all of $\mathbb{R}$. This subset is the rational numbers. We can use this fact to show that every open covering has a countable subcovering.

Proposition 21 *Let $A \subset \mathbb{R}$, and $\mathcal{O}$ be an open covering of A. Then $\mathcal{O}$ has a countable subcovering.*

Sketch of the Proof: Every $O \in \mathcal{O}$ contains at least one rational number q. If $q \in O$, we label this set by O_q, and ignore all other $O \in \mathcal{O}$ that may contain q. Now let$x \in O$, since O is open, there exists an $\varepsilon > 0$ such that $x - \varepsilon, x + \varepsilon) \subset O$. However, the interval $(x - \varepsilon/2, x + \varepsilon/2)$ contains a rational number q. If $O = O_q$ we are done, if not let replace O_q by O. Afeter doing this for all $x \in A$ we have

$$A \subset \bigcup_{q \in \mathbb{Q}} O_q.$$

We are now ready to prove the main result of this section. We will prove that every closed and bounded subset of $\mathbb{R}$ is compact. At first glance this seems like a nice result becacause the complicated concept of compactness is replaced by two much simpler concepts. Unfortunately, this replacement obscures the true nature of the beast, and this replacement is not valid in arbitray spaces, only in $\mathbb{R}^n$ for $n \geq 1$. This is a rather difficult pedagogical issue, as students will almost certainly fall into the trap of identifying compactness and closed and boundedness, thus forgetting about compactness.

Theorem 6 *(HEINE-BOREL-THEOREM) Let $C \subset \mathbb{R}$ be closed and bounded, then C is compact.*

Proof: Since C is bounded there exists a closed interval $[A, B]$ that contains C. We will show that this interval is compact. Then C is also compact as a result of Proposition 20. Let $\mathcal{O}$ be an aritrary open cover of $[A, B]$. Proposition 21 implies that it has a countable subcover, and we may thus assume that $\mathcal{O}$ itself is countable, so

$$\mathcal{O} = \{O_n : n \in \mathbb{N}\}.$$

Let us now assume that $\mathcal{O}$ has no finite subcovering. That means that for very $N \in \mathbb{N}$ there exists a $x_N \in [A, B]$ such that

$$x_N \notin O_1 \cup \ldots \cup O_N.$$

In this way we construct a sequence $\{x_n\}$ with the following properties:

$$A \leq x_n \leq B$$

for all $n \in \mathbb{N}$. As this is a bounded infinite sequence the BOLZANO-WEIERSTRASS-THEOREM implies that it has a convergent subsequece $\{x_{n_k}\}$with subsequential limit point L. Moreover, since $[A, B]$ is closed $L \in [A, B]$. Thus there exists a number N such that $L \in O_N$. Since O_N is open there is an $\varepsilon > 0$ such that

$$(L - \varepsilon, L + \varepsilon) \subset O_N,$$

and finally there is a $K \in \mathbb{N}$ such that

$$x_{n_k} \in (L - \varepsilon, L + \varepsilon) \subset O_N,$$

for all $k \geq K$. Thus there exists a $k \geq K$ such that $n_k \geq N$ and $x_{n_k} \in O_N$, which contradicts the construction of the sequence. Thus the above construction must terminate at a finite value $M \in \mathbb{N}$. But then

$$[A, B] \subset O_1 \cup \ldots \cup O_M,$$

and the open cover has a finite subcover. Thus $[A, B]$ and as a direct consequence C are compact.

The proof of this theorem indicates that compactness is closely related to the Bolzano-Weierstrass Property. This is indeed the case. For $\mathbb{R}^n$ and other complete metric spaces we have the following statements about an infinite set C that are equivalent:

- C is compact.
- C has at least one limit point in C. (C is limit point compact.)
- Every sequence in C has a convergent subsequence with subsequential limit in C. (C is sequentially compact)

3.4 Problems

1. Let $C = (0, 1)$. Find an open cover of C that doesn't have a finite subcover.
2. Let $\{A_n : n \in \mathbb{N}\}$ be a countable collection of non-empty closed and bounded subsets of $\mathbb{R}$ such that $A_{n+1} \subset A_n$ for all $n \in \mathbb{N}$. Show that $\bigcap n \in \mathbb{N} a_n \neq \emptyset$. (Hint: Construct a sequence and apply the BOLZANO-WEIERSTRASS THEOREM.)

Chapter 4

Continuous Functions

4.1 Basic Properties - Continuity and Limits

Before defining continuity we review a few properties of functions and sets. Let $A \subset \mathbb{R}$ and $f : A \to \mathbb{R}$ a function. For $B \subset A$ and $C \subset \mathbb{R}$ we define:

$$f(B) = \{y \in \mathbb{R} : y = f(x) \quad \text{for some} \quad x \in B\}$$

and

$$f^{-1}(C) = \{x \in A : f(x) \in C\}$$

Observe that we do not require f to be invertible for the latter definition. We call the first set the image of B and the second set the preimage or inverse image of C. We have the following properties.

$$f^{-1}(B \cup C) = f^{-1}(B) \cup f^{-1}(C) \tag{4.1}$$

$$f^{-1}(B \cap C) = f^{-1}(B) \cap f^{-1}(C) \tag{4.2}$$

$$f(B \cup C) = f(B) \cup f(C) \tag{4.3}$$

However, the statement for intersections does not hold for the image.

Definition 20 *Let $f : A \to \mathbb{R}$ be a function, and $x_0 \in A$. WE say that f is continuous at x_0 if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $|f(x) - f(x_0)| < \varepsilon$ for all $x \in A$ such that $|x - x_0| < \delta$.*

Before exploring this definition further we want to point out that continuity as such is a proprty of a function at a single point in its domain. Thismeans that a statement of continuity should always include the location where the statement holds. We often say sloppily that a function is continuous, instead of properlt saying that the function is contiuous at every point in its domain. We say that a function is continuous on a set A if is continuous at every $x \in A$.

The definition can be rewritten in terms of sets as follows. f is continuous at $x_0 \in A$ if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$f(x) \in (f(x_0) - \varepsilon, f(x_0) + \varepsilon) \quad \text{for all} x \in (x_0 - \delta, x_0 + \delta) \cap A, \tag{4.4}$$

or

$$f((x_0 - \delta, x_0 + \delta) \cap A) \subset (f(x_0) - \varepsilon, f(x_0) + \varepsilon), \tag{4.5}$$

or

$$(x_0 - \delta, x_0 + \delta) \cap A \subset f^{-1}((f(x_0) + \varepsilon, f(x_0) - \varepsilon)). \tag{4.6}$$

It is this last version which will turn out to be very useful later on. Related to the definition is the concept of the limit of a function.

Definition 21 *Let $f : A \to \mathbb{R}$ be a function and $x_0 \in \overline{A}$. We say that*

$$L = \lim_{x \to x_0} f(x),$$

if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$f(x) \in (L - \varepsilon, L + \varepsilon) \quad \textit{for all} \quad x \in (x_0 - \delta, x_0 + \delta) \cap A.$$

Observe that in this case f does not have to be defined at x_0. A consequence is the characterization of continuity at a point using this concept. Clearly f is continuous at x_0 if it is defined at x_0 and

$$\lim_{x \to x_0} f(x) = f(x_0).$$

Continuity and limits of functions can also be expressed using the concept of limits of sequences.

Proposition 22 *Let $f : A \to \mathbb{R}$ be a function, and $x_0 \in A$. Then f is continuous at x_0 if and only if*

$$\lim_{n \to \infty} f(x_n) = f(x_0)$$

for all sequences $\{x_n\}$ which converge to x_0.

Proof: Let f be continuous at x_0 and $\{x_n\}$ be a sequence that converges to x_0. Then for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $f(x) \in (f(x_0) - \varepsilon, f(x_0) + \varepsilon)$ for all $x \in (x_0 - \delta, x_0 + \delta) \cap A$. For this δ there exists an $N \in \mathbb{N}$ such that $x_n \in (x_o - \delta, x_0 + \delta)$ for all $n \geq N$. Combining the two statements yields the first part of the result. Conversely, assume that f is not continuous at x_0. Then there exists an $\varepsilon_0 > 0$ such that for every $n \in \mathbb{N}$ there is a $x_n \in (x_0 - 1/n, x_0 + 1/n) \cap A$ with

$$f(x_n) \notin (f(x_0) - \varepsilon_0, f(x_0) + \varepsilon_0)$$

thus $f(x_n)$ does not converge to $f(x_0)$, but $x_n \to x_0$.

As in the prove of the last proposition it is more important to understand what happens if f is not continuous at a point x_0. Consider the function

$$f(x) = \begin{cases} \sin\frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases} \tag{4.7}$$

This function is not continuous at 0. For example the sequence $x_n = 1/n\pi$ clearly converges to 0 and so does $f(x_n)$. However, the sequence $y_n = 2/(2n+1)\pi$ also converges to zero, but $f(y_n)$ does not converge at all.

Proposition 23 *Let g, f be functions such that $\lim_{x\to x_0} f(x) = F$ and $\lim_{x\to x_0} g(x) = G$. Then*

1. $\lim_{x\to x_0}(af(x) + bg(x)) = aF + bG$ *for all* $a, b \in \mathbb{R}$
2. $\lim_{x\to x_0} f(x)g(x) = FG$.
3. *If* $G \neq 0$ *we also have* $\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{F}{G}$

Proof: The proof of these results follows directly from the proof of the corresponding results for sequences.

Theorem 7 *$f : A \to \mathbb{R}$ is continuous on A if and only if for any open $u \subset \mathbb{R}$ $f^{-1}(U) = A \cap O$ for some open set $O \subset \mathbb{R}$.*

Proof: Let f be continuous and U be open. Then let $x \in f^{-1}(U)$, this implies $f(x) \in U$ and since U is open there exists an $\varepsilon_x > 0$ such that $(f(x) - \varepsilon_x, f(x) + \varepsilon_x) \subset U$. Since f is continuous there exists a $\delta_x > 0$ such that

$$(x - \delta_x, x + \delta_x) \cap A \subset f^{-1}((f(x) - \varepsilon_x, f(x) + \varepsilon_x)) \subset f^{-1}(U)$$

. Now let $O = \bigcup x \in f^{-1}(U)(x - \delta_x, x + \delta_x)$. Then O is open and

$$A \cap O \subset \bigcup_{x\in f^{-1}(U)} (x - \delta_x, x + \delta_x) \subset f^{-1}(U),$$

and $f^{-1}(U) \subset A \cap O$ by the construction of O.
Conversely, let $x \in A$, then for every $\varepsilon > 0$ the set $(f(x)-\varepsilon, f(x)+\varepsilon)$ is open and thus $f^{-1}((f(x)-\varepsilon, f(x)+\varepsilon)) = A \cap O$. Since x is an interior point of O there is a $\delta > 0$ such that

$$(x - \delta, x + \delta) \cap A \subset A \cap O = f^{-1}((f(x) - \varepsilon, f(x) + \varepsilon)),$$

which is the formulation (4.6) of continuity.

4.2 Problems

1. Find examples of two functions that are not fontinuous at some point x_0, but

(a) their sum is continuous.

(b) their product is continuous.

(c) their quotient is continuous.

2. Let $f : A \to \mathbb{R}$ be continuous at x_0 and $f(x_0) = y_0 \in B$. Let $g : B \to \mathbb{R}$ be continuous at y_0. Prove that the composition $g \circ f$ is continuous at x_0.

3. Let $f : A \to \mathbb{R}$ be be continuous on A and $C \subset A$ be closed. Prove that $f(C)$ is closed.

4. Find a continuous function f and an open subset O such that $f(O)$ is not open.

4.3 Properties of Continuous Functions

The main idea behind continuous functions is that they preserve certain topological properties. The most notable properties are compactness and connectedness.

Theorem 8 *Let $f : A \to \mathbb{R}$ be a continuous function and $C \subset A$ a compact set. Then $f(C)$ is compact.*

Proof: Let $\{O_\alpha : \alpha \in \mathcal{A}\}$ be an open cover of $f(C)$. Then for every $\alpha \in \mathcal{A}$ there exists a an open U_α such that $f^{-1}(O_\alpha) = U_\alpha \cap A$. Clearly,

$$C \subset \bigcup_{\alpha \in \mathcal{A}} U_\alpha \cap A,$$

and thus the collection $\{U_\alpha : \alpha \in \mathcal{A}\}$ is an open cover of C. Thus there exists a finite subcollection

$$\{U_{\alpha_1}, \ldots, U_{\alpha_n}\},$$

such that

$$C \subset U_{\alpha_1} \cup \cdots \cup U_{\alpha_n}$$

and thus

$$f(C) \subset O_{\alpha_1} \cup \cdots \cup O_{\alpha_n}.$$

Hence, $f(C)$ is compact.

An important consequence is that continuous functions attain there maxima and minima on compact sets.

Corollary 3 *Let $f : C \to \mathbb{R}$ be continuous and C compact. Then there exist $x, y \in C$ such that*

$$f(x) \le f(t) \le f(y),$$

for all $t \in C$.

Proof: $f(C)$ is compact and thus closed an bounded. Let $A = \inf f(C)$, since $f(C)$ is closed, we have $A \in f(C)$ and thus there is a number $x \in C$ such that $f(x) \leq f(t)$ for all $t \in C$. The analogous argument works for the upper bound.

An other important consequence is the intermediate value theorem.

Definition 22 *A function $f : [a,b] \to \mathbb{R}$ has the intermediate value property if for any $x_1, x-2 \in [a,b]$, with $f(x_1) \neq f(x_2)$ and any value y between $f(x_1)$ and $f(x_2)$ there is a ξ between x_1 and x_2 such that $f(\xi) = y$.*

Theorem 9 *Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$. Then f has the intermediate value property.*

Proof: We will first present a classical proof and then a more modern one. Let $x_1, x_2 \in [a,b]$ and assume $x_1 < x_2$. Moreover, we may assume that $f(x_1) < y < f(x_2)$. Let

$$M = \{x \in [x_1, x_2] : f(x) \leq y\}.$$

Then M is a nonempty bounded set and has a least upper bound $x_0 = \sup M$. We claim that $f(x_0) = y$. To prove this let $\{x_n\}$ be a sequence in M such that $x_n \to x_0$, thus $f(x_n) \to f(x_0) \leq y$. Next suppose that $f(x_0) \neq y$, then $y - f(x) = F(X)$ is a continuous function and $F(x_0) > 0$. Pick $\varepsilon = F(x_0)/2 > 0$. Then by the continuity of F there exists a $\delta > 0$ such that $F(x) > \varepsilon/2$ for all $x \in (x_0 - \delta, x_0 + \delta) \cap [a,b]$ In particular, there are value x such that $x_0 < x < x_0 + \delta$ such that $F(x) > 0$ and in turn $f(x) < y$. This contradicts that $x_0 = \sup M$.

The statement and the proof of the Intermediate Value Theorem relies heavily on the properties of $\mathbb{R}$ as a complete oredered field. It can therefore not easily be extended to higher dimension. However, we can extend by introducing a new topological concept, the concept of connected ness.

Definition 23 *Let $A \subset \mathbb{R}$. A is said to be connected, if there are no two nonempty open sets U and v such that.*

1. *$U \cap V = \emptyset$*
2. *$U \cap A \neq \emptyset$ and $V \cap A \neq \emptyset$*
3. *$A \subset U \cup V$.*

This is a rather akward looking definition. For the real numbers it comes to the fact that connected sets are intervals. Let A be a set that is not an interval. Then from the definition of intervals (3.1) there exist $a, b \in A$ with $a < b$ and a $t \in \mathbb{R}$ with $a < t < b$ such that $t \notin A$. Consider $U = (-\infty, t)$ and $V = (t, \infty)$ which satisfy all the properties of the sets in the definition. Hence A is not connected.

Proposition 24 *Let $f : A \to \mathbb{R}$ be continuous and A a connected subset. Then $f(A)$ is connected.*

Proof: Assume $f(A)$ is not connected, then there exist open sets U and V with the properties listed in the definition. Now

$$f^{-1}(U) = A \cap O, \quad \text{and} \quad f^{-1}(V) = A \cap W,$$

for some open sets O and W. First O and W are nonempty, and there intersection must lie in A^c. By removing the closure of there intersection we get open sets O' and W' which are nonempty, disjoint and there union contains all of A, hence A is not connected.

While this concept can now be extended to higher dimensions for the reals it basically means that the image of an interval is an interval.

Example: Let $f : [0,1] \to [0,1]$ be continuous, then there exists s number $\xi \in [0,1]$ such that $f(\xi) = \xi$. Indeed if you consider the function $g(x) = f(x) - x$, you may assume that $g(0) \neq 0$ and $g(1) \neq 0$, since otherwise 0 or 1 are the values for ξ. But then $g(0) > 0$ and $g(1) < 0$ and by the intermediate value theorem there exists a ξ such that $g(\xi) = 0$ and hence $f(\xi) = \xi$.

It is important to notice that a function may satisfy the intermediate value property, but is not continuous. For example the function in equation (4.7) satisfies the intermediate value property on $[-1,1]$, but it is not continuous.

Example: Consider the function $f(x) = \frac{1}{x}$ on the interval $(0,1)$. And let $x_0 \in (0,1)$ and $\varepsilon > 0$. Then

$$\left|\frac{1}{x} - \frac{1}{x_0}\right| = \frac{|x - x_0|}{|xx_0|} \leq \frac{2|x - x_0|}{|x_0^2|} < \varepsilon$$

for all x such that $|x - x_0| < \delta = \min\{\frac{|x_0|}{2}, \frac{\varepsilon|x_0^2|}{2}\}$. We see that δ depends very strongly not only on ε but also on the value of x_0. This example motivates the concept of uniform continuity.

Definition 24 *A function $f : A \to \mathbb{R}$ is uniformly continuous on A, if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that*

$$|f(x) - f(y)| < \varepsilon, \quad \textit{for all} \quad x, y \in A \quad \textit{which satisfy} \quad |x - y| < \delta.$$

Example: The function $f(x) = \sin x$ is uniformly continuous on $\mathbb{R}$. As we will see later, this function always satisfies $|\sin x - \sin y| \leq |x - y|$ and the uniform continuity follows immediately from this.

While even very simple functions are often not uniformly continuou on their domains. Every function can be made uniformly continuous by appropriately restricting its domain.

Proposition 25 *Let $f : C \to \mathbb{R}$ be continuous and C a compact interval. Then f is uniformly continuous on C.*

Proof: Let $\varepsilon > 0$, then for every $x \in C$ there exists a δ_x such that $|f(x) - f(t)| < \frac{\varepsilon}{3}$ for all $t \in (x - 3\delta_x, x + 3\delta_x) \cap C$. Now the collection $X = \{(x - \delta_x, x + \delta_x) : x \in C\}$ is an open covering of C. Since C is compact there exist finitely many values $x_1, x_2, \ldots, x_n$ such that

$$C \subset (x_1 - \delta_{x_1}, x_1 + \delta_{x_2}) \cup \cdots \cup (x_n - \delta_{x_n}, x_n + \delta_{x_n})$$

Define $\delta = \min\{\delta_1, \ldots, \delta_n\}$, and let $x, y \in C$ such that $|x - y| < \delta$. There exist numbers x_j, x_k such that $x \in (x_j - \delta_j, xj + \delta_j)$ and $y \in (x_k - \delta_k, x_k + \delta_k)$. Now

$$|x_k - x_j| \leq |x_k - y| + |y - x| + |x - x_j| \, \delta_{x_k} + \delta + \delta_{x_j}.$$

Thus either $x_j \in (x_k - 3\delta_{x_k}, x_k + 3\delta_{x_k})$ or $x_k \in (x_j - 3\delta_{x_j}, x_j + 3\delta_{x_j})$. Either way we have

$$|f(x_k) - f(x_j)| < \frac{\varepsilon}{3}.$$

Finally, we get

$$|f(x) - f(y)| \leq |f(x) - f(x_j)| + |f(x_j) - f(x_k)| + |f(x_k) - f(y)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \varepsilon,$$

and the Proposition is proved.

4.4 Monotone Functions and Types of Discontinuities

Definition 25 *Let $f : [a, b] \to \mathbb{R}$. We say that f is (strictly) monotonically increasing on $[a, b]$ if for all $s, t \in [a, b]$ $s < t$ implies $f(s)(<) \leq f(t)$.*
Analogously, we say that f is (strictly) monotonically decreasing on $[a, b]$ if for all $s, t \in [a, b]$ $s < t$ implies $f(s)(>) \geq f(t)$.

An immediate consequence of this definition is that strictly monotonically increasing (decreasing) functions on a set A are also one-to-one functions.
Example: The function $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = x^3$ is strictly monotonically increasing. To see this let $s < t$. Then

$$f(t) - f(s) = t^3 - s^3 = (t - s)(t^2 + st + s^2) = (t - s)\left(\frac{t^2}{2} + \frac{s^2}{2} + \frac{(t + s)^2}{2}\right).$$

Observe, that the second factor in the last expression is always positive. The first factor is positive since $t > s$, thus

$$f(t) - f(s) > 0,$$

and the function is srictly increasing.

The technique of investigating the difference function values – as used in this example – is common to prove the monotonicity of functions.

Definition 26 *Let $f : A \to \mathbb{R}$ be a function, and x_0 be a limit point of A. We say that L is the lower limit of f t x_0 if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that*

$$|f(x) - L| < \varepsilon \qquad \textit{for all} \qquad x \in (x_0 - \delta, x_0) \cap A.$$

We write

$$\lim_{x \to x_0^-} f(x) = L.$$

We say that L is the upper limit of f t x_0 if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - L| < \varepsilon \qquad \textit{for all} \qquad x \in (x_0, x_0 + \delta) \cap A.$$

We write

$$\lim_{x \to x_0^+} f(x) = L.$$

We see that this is sort of half the definition of limit. Indeed we have

Proposition 26 *Let $f : A \to \mathbb{R}$ be a function and x_0 be alimit point of A then*

$$\lim_{x \to x_0} f(x) = L$$

if and only if

$$\lim_{x \to x_0^-} f(x) = L \qquad \textit{and} \qquad \lim_{x \to x_0^+} f(x) = L.$$

Proof: If the limit exists then clearly the upper and lower limits exist. On the othe hand, suppose that the lower and upper limit exist and are equal. Let $\varepsilon > 0$ then there exist a $\delta^- > 0$ such that

$$|f(x) - L| < \varepsilon \qquad \text{for all} \qquad x \in (x_0 - \delta^-, x_0) \cap A.$$

and a $\delta^+ > 0$ such that

$$|f(x) - L| < \varepsilon \qquad \text{for all} \qquad x \in (x_0, x_0 + \delta^+) \cap A.$$

Hence for $\delta = \min\{\delta^-, \delta^+\}$ we have

$$|f(x) - L| < \varepsilon \qquad \text{for all} \qquad x \in (x_0 - \delta, x_0 + \delta) \cap A \setminus \{x_0\}.$$

In the section on sequences we learned that bounded monotonic sequences always have a limit. A similar statement also holds for monotonic functions.

Proposition 27 *Let $f : A \to \mathbb{R}$ be monotonically increasing (or decreasing) and bounded, and x_0 a limitpoint of A. then both the upper and lower limits at x_0 exist unless $x_0 \geq x$ or $x_0 \leq x$ for all $x \in A$.*

Proof: Without loss of generality we may assume that f is increasing. Consider the set

$$M = \{f(x) : x \in (-\infty, x_0) \cap A\}.$$

This is a set that is bounded above and thus we have $L = \sup M$. We claim that

$$L = \lim_{x \to x_0^-} f(x).$$

Suppose this is not the case. Then there exists an $\varepsilon_0 > 0$ such that for all $\delta > 0$ there is a $t \in (x_0 - \delta, x_0) \cap A$ such that

$$L - f(t) \geq \varepsilon_0 \qquad \text{or} \qquad f(t) \leq L - \varepsilon_0$$

Since f is increasing it follows that

$$f(s) \leq f(t) \leq L - \varepsilon_0 \qquad \text{for all} s \in (-\infty, x_0 - \delta] \cap A.$$

Since this holds for all $\delta > 0$ it follows that

$$f(s) \leq L - \varepsilon_0 \qquad forall \qquad s \in (-\infty, x_0) \cap A,$$

which contradicts the fact that $L = \sup M$. For the lower limit we use the analogous argument.

.

Observe that the Proposition doesn't guarantee the equality of upper and lower limits only their existence. Now we can start classifying types of discontinuities. The first case is that $\lim_{x \to x_0} f(x)$ actually exists, but it is not equal to $f(x_0)$. In this case we can make a continuous function out of f using the following construction

$$g(x) = \begin{cases} f(x) & x \neq x_0 \\ \lim_{x \to x_0} f(x) & x = x_0 \end{cases}$$

Then the function g is continuous at x_0 such an discontinuity is called **removeable**. Now in the seconcase we hav

$$\lim_{x \to x_0^-} f(x) = L, \qquad \lim_{x \to x_0^+} f(x) = K, \qquad \text{but} \qquad L \neq K.$$

In this case we say that f has a **jump discontinuity** at x_0. Finally there is the possibility that either the upper limit or the lower limit or both do not exist. This is what happens for example with the function

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

In this case we say that the function ha a **discontinuity of the third type**.
From our proposition it follows that monotone functions can only have removeable or jump discontinuities. And this also limits the number of discontinuities.

Theorem 10 *Let $f : A \to \mathbb{R}$ be a monotonic function. Then f has at most countably many jump discontinuities.*

Proof: Let S be the set of jump discontinuities of f, and suppose f is increasing. Then for every $s \in S$ define

$$L_s = \lim_{x \to s^-} f(x) \qquad \text{and} \qquad U_s = \lim_{x \to s^+} f(x).$$

Since f is increasing we have $L_s < U_s$, and for $s, t \in S$ with $s < t$ we have $U_s \leq L_t$. Thus for $s \neq t$ $(L_s, U_s) \cap (L_t, U_t) = \emptyset$ Since these intervals are open there exists a rational number $q_s \in (L_s, U_s)$, and for $s \neq t$ we have $q_s \neq q_t$. The mapping

$$\phi : S \to \mathbb{Q} \qquad \text{given by} \qquad \phi : s \mapsto q_s,$$

is a one to one map from S to a countable set. Therefore S is countable.
Strictly monotonic functions are invertible. However, if f is continuous and inverible its inverse is not necessarily continuous as the following example shows.
Example: Let

$$f(x) = \begin{cases} x & x \in (0,1] \\ x-1 & x \in (2,3) \end{cases}$$

then f is continuous and invertible., buf f^{-1} is not continuous. The problem here is that the domain of f is not connected. If we restrict ourselves to connected domains we have:

Proposition 28 *Let f be continuous and strictly monotonic on $[a, b]$. Then f^{-1} is continuous and strictly monotonic on its domain.*

Proof: Let us assume again that f is strictly increasing. Then for nay $x \in (a, b)$ we have $f(a) < f(x) < f(b)$, so $[f(a), f(b)]$ is the domain of f^{-1}. Let $y_1, y_2 \in [f(a), f(b)]$ with $y_1 < y_2$, then $x_1 = f^{-1}(y_1)$ and $x_2 = f^{-1}(y_2)$, and $x_1 < x_2$, since otherwise $f(x_1) \geq f(x_2)$ which would contradict the assumption $y_1 < y_2$. For cotinuity, we show that if $s, t \in [a, b]$ and $s < t$then $f((s, t)) = (f(s), f(t))$, in other words the function maps open intervals into open intervals. Let $x \in (s, t)$ then $f(s) < f(x) < f(t)$ which implies that $f((s, t)) \subset (f(s), f(t))$. Conversely let $y \in (f(s), f(t))$ since f is invertible there exists an $x \in [a, b]$ such that $y = f(x)$. Now since f^{-1} is increasing, we have $s < f^{-1}(y) = x < t$, and thus $(f(s), f(t)) \subset f((s, t))$. Since every open subset of $[a, b]$ is a union of open subintervals of $[a, b]$, f maps open subsets of $[a, b]$ to open subsets of $[f(a), f(b)]$. A similar argument goes for subsets of the form $[a, s)$ or $(s, b]$. Let O be an open subset of $\mathbb{R}$, then $f(O) = [a, b] \cap U$ for some open subset $U \subset \mathbb{R}$. The image of open subsets of $\mathbb{R}$ under f is relatively open. But the image under f is the same as the preimage under f^{-1}, thus f^{-1} is continuous.

Remark: Functions which map open sets to relatively open sets are called open functions. Strictly monotonic functions are open. Continuous functions which have continuous inverses are also called **homeomorphisms**.

Chapter 5

Differentiability on $\mathbb{R}$

5.1 Definition and Elementary Properties

We all recall the definition of differentiability of a function from Calculus. More importantly we should recall its geometric interpretation as the slope of the tangent line to a graph of the function. I.e. the derivative is used to create a local linear approximation of a function.

Definition 27 *Let $f : A \to \mathbb{R}$ be a function and x_0 be an interior point of A We say that f is differetiable at x_0 if*

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

exists. $f'(x_0)$ is called the derivative of f at x_0. We say that f is differentiable on an open set O if it is differentiable at x for every $x \in O$.

This definition slightly differs from the one we remember. In calculus the words interior point and open were missing. However, it is critical to have these words in there, since otherwise we cannot define the limit. This definition has some major short comings. The most important one is that it contains one of dreadful objects in mathematics, a fraction. On top of this, both numerator and denominator will converge to zero as $x \to x_0$. Fractions have the serious problem that they rely on the underying set to be a field and can thus not be readily extended to vectorspaces without this structure. For an alternate definition we return to the geometric interpretation of the derivative. The objective is to find a linear function that best approximates the given function near the point x_0, this means we want to find a number a such that

$$f(x) = f(x_0) + a(x - x_0) + E(|x - x_0|),$$

where E is an error term which should vanish as $x \to x_0$. Now if, $x \neq x_0$ we can trans form the above equation to

$$\frac{f(x) - f(x_0)}{x - x_0} = a + \frac{E(|x - x_0|}{x - x_0},$$

and we see that $a = f'(x_0)$, and

$$\lim_{x \to x_0} \frac{E(|x - x_0|)}{x - x_0} = 0.$$

This gets us to the following alternate definition.

Definition 28 *Let $f : A \to \mathbb{R}$ be a function and x_0 be an interior point of A We say that f is differetiable at x_0 if there exist numbers $f'(x_0)$ and $\delta > 0$ and a function $E : [0, \delta) \to \mathbb{R}$ such that*

$$|f(x) - f(x_0) - f'(x_0)(x - x_0)| \leq |E(|x - x_0|)|\,, \tag{5.1}$$

and

$$\lim_{t \to 0^+} \frac{|E(t)|}{t} = 0. \tag{5.2}$$

We see that this definition can be readily extended to vector spaces by changing the "number $f'(x_0)$ to the "linear operator $f'(x_0)$". It still involves a fraction, but this fraction only involves positive real numbers.

Example: Let $f(x) = x^3$ and $x_0 \in \mathbb{R}$. Observe that

$$x^3 - {}^3_0 = (x - x_0)(x^2 + xx_0 + x_0^2) = 3x_0^2(x - x_0) + (x - x_0)(x^2 - x_0^2 + xx_0 - x_0^2) = 3x_0^2(x - x_0) + (x - x_0)^2(x + 2x_0),$$

so $f'(x_0) = 3x_0^2$, $\delta > 0$ and $E(t) = At^2$ where the positive constat A is chosen such that $|x + 2x_0| \leq A$ for $x \in (x_0 - \delta, x_0 + \delta)$.

One of the first thing to investigate is the connection between differentiability and continuity. We start by the following

Example: Let $f(x) = |x|$ and $x_0 = 0$. We know that this function is continuous at the origin. But

$$\lim_{x \to 0^+} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0^+} \frac{x}{x} = 1,$$

and

$$\lim_{x \to 0^-} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0^-} \frac{-x}{x} = -1,$$

and so the function is not differentiable at 0. We see that differentiability is not an immediate consequence of continuity. In fact there exist functions which are uniformly continuous on $\mathbb{R}$, but they are nowhere differentiable. However continuiyy is an immediate consequence of differentiability.

Proposition 29 *Let f be differentiable at x_0, then f is continuous at x_0.*

Proof: From the above discussion of an alterate definition we have:

$$\lim_{x \to x_0} f(x) = f(x_0) + f'(x_0) \lim_{x \to x_0} (x - x_0) + \lim_{x \to x_0} E(|x - x_0|),$$

which immediately implies this result.

However, differentiability by itself doesn't imply uniform continuity. We haven't yet derived any rules of differentiation. But since the basic step in differentiation is a limit, we expect that similar rules apply.

Proposition 30 *Let X be set of all functions that are differentiable at a point x_0. Then X is a real vector space.*

Proof: We need to show that $af(x) + bg(x)$ is differentiable at x_0, if f and g are. This follows immediately from the same property of limits.

Proposition 31 *Let f and g be differentiable at x_0. Then fg is differentiable at x_0 and*

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0).$$

If in addition $g(x_0) \neq 0$, then $\frac{f}{g}$ is differetiable at x_0 and

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

Proof: See the next assignment at the end of this section.

Finally, we also have a result for the composition of functions known as the chain rule.

Proposition 32 *Let f be a function that is differentiable at x_0 and g a function that is differentiable at $y_0 = f(x_0)$. The $g \circ f$ is differentiable at x_0, and*

$$(g \circ f)'(x_0) = g'(y_0)f'(x_0).$$

Proof: See the next assignment at the end of this section.

Now if f is differentiable on (a, b) then $f' : (a, b) \to \mathbb{R}$. A priori we do not know any of the properties of this function. However, if this function is itself differentiable, we can compute its derivative $f" = (f')'$ which is called the econd derivative of f. Accordingly one can compute higher derivatives.

Definition 29 *A function $f : A \to \mathbb{R}$ has a local or relative maximum at x_0 if there exists a $\delta > 0$ such that*

$$f(x_0) \geq g(x) \qquad \textit{for all} \qquad x \in (x_0 - \delta, x_0 + \delta) \cap A.$$

A function $f : A \to \mathbb{R}$ has a local or relative minimum at x_0 if there exists a $\delta > 0$ such that

$$f(x_0) \leq g(x) \qquad \textit{for all} \qquad x \in (x_0 - \delta, x_0 + \delta) \cap A.$$

In either case we sy that f has a local extremum at x_0.

There is an important result for functions which are differentiable at a local extremum.

Proposition 33 *Let f be differentiable at x_0 and assume that f has a local extremum at x_0 Then*

$$f'(x_0) = 0.$$

Proof: With out loss of generality we may assume that f has a local maximum at x_0. Then

$$\frac{f(x)-f(x_0)}{x-x_0} \leq 0 \qquad \text{for all} \qquad x \in (x_0, x_0+\delta),$$

and thus

$$\lim_{x \to x_0^+} \frac{f(x)-f(x_0)}{x-x_0} \leq 0.$$

On the other hand

$$\frac{f(x)-f(x_0)}{x-x_0} \geq 0 \qquad \text{for all} \qquad x \in (x_0-\delta, x_0),$$

and thus

$$\lim_{x \to x_0^+} \frac{f(x)-f(x_0)}{x-x_0} \geq 0.$$

But since f is differentiable these two limits must be the same and thus $f'(x_0) = 0$. The proof for minima is analogous.

5.2 Exercises

1. Let f and g be functions that are differentiable at x_0. Prove that fg is differentiable at x_0 and that

$$(fg)'(x_0) = f'(x_0)g(x_0) - f(x_0)g'(x_0)$$

2. Let f and g be functions that are differentiable at x_0 with $g'(x_0) \neq 0$. Prove that $\frac{f}{g}$ is differentiable at x_0 and that

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

3. Let f be a function that is differentiable at x_0 and g a function that is differentiable at $y_0 = f(x_0)$. Prove that $g \circ f$ is differentable at x_0 and compute its derivative.

4. Give an example of two functions f and g that are not differentiable , but $f \circ g$ is differentiable.

5.3 Mean Value Theorems

The most useful results in Differential Calculus are Mean Value Theorems. As it turns out they are also some of the easiest theorems to prove. The basis of all mean value theorems is Rolle's Theorem

Theorem 11 Rolle's Theorem: *Let $f : (a,b) \to \mathbb{R}$ be differentiable on (a,b) and continuous on $[a,b]$ with $f(a) = f(b)$. Then there exists a $\xi \in (a,b)$ such that $f'(\xi) = 0$.*

Proof: With out loss of generality we may assume that $f(a) = f(b) = 0$. Then either $f(x) = 0$ for all $x \in (a, b)$, or it must have either a maximum or a minimum on $[a, b]$ since it is continuous. If f is identically equal to zero, then so is f' and there is nothing to show. If not the extremum in $[a, b]$ must be an extremum in (a, b). Let $\xi \in (a, b)$ such that $f(\xi)$ is an extreme value, then $f(\xi) = 0$.

While this theorem is a simple consequence of the properties of local extrema it has some important consequences.

Theorem 12 Mean Value Theorem: *Let $f : (a, b) \to \mathbb{R}$ be differentiable on (a, b) and continuous on $[a, b]$. Then there exists a $\xi \in (a, b)$ such that*

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

Proof: Define

$$F(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Then $F(a) = F(b)$ and F satisfies the hypotheses of Rolle's Theorem. Therefore there exists a $\xi \in (a, b)$ such that

$$0 = F'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a}.$$

This theorem has an important consequence. It is easy to see that if f is a constant function, then $f'(x) = 0$ for all x in the domain of f. Now, we can show that the converse is also true.

Corollary 4 *Suppose $f : [a, b] \to \mathbb{R}$ is continuous and difeferentiable on (a, b) and $f'(x) = 0$. Then f is constant on $[a, b]$ with $f(x) = f(a)$.*

Proof: We have for all $x \in (a, b]$

$$f(x) - f(a) = f'(\xi)(x - a) = 0.$$

The mean valur theorem can be further generalized to

Theorem 13 Cauchy's Mean Value Theorem: *Let $f, g : (a, b) \to \mathbb{R}$ be differentiable functions on (a, b) and continuous on $[a, b]$, and $g(a) \neq g(b)$. Then there exists a $\xi \in (a, b)$ such that*

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

Proof: Define

$$G(x) = (f(b) - f(a))g(x) - (g(b) - g(a))f(x).$$

Then

$$G(a) = f(b)g(a) - f(a)g(a) - g(b)f(a) + g(a)f(a) = f(b)g(a) - g(b)f(a),$$

and

$$G(b) = f(b)g(b) - f(a)g(b) - g(b)f(b) + g(a)f(b) = f(b)g(a) - g(b)f(a) = G(a).$$

And surprisingly this function satisfies the hypothesis of Rolle's Theorem. Therefore there exist a $\xi \in (a, b)$ such that

$$0 = G'(\xi) = (f(b) - f(a))g'(\xi) - (g(b) - g(a))f'(\xi),$$

which is the desired result.

This last result has some interesting consequences, one of them is the probably most abused result in calculus. It is known as l'Hospital's rule, even though it was probably proven long befor l'Hospital published the first textbook on calculus.

Corollary 5 l'Hospitals Rule: *Let $f, g : [a, b] \to \mathbb{R}$ be differentiable and continuous, $x_0 \in (a, b)$, and $f(x_0) = g(x_0) = 0$. Furthermore, assume 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 we may assume that there is a $\delta > 0$ such that $g'(x) \neq 0$ for all $x \in (x_0 - \delta, x_0 + \delta) \setminus \{x\}$. Observe that for $x \in (x_0, x_0 + \delta) \setminus \{x\}$ we have

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(x_0)}{g(x0 - g(x_0)} = \frac{f'(\xi)}{g'(\xi)},$$

for some $\xi \in (x_0, x)$. Moreover, $\xi \to x_0^+$ as $x \to x_0^+$ and thus.

$$\lim_{x \to x_0^+} \frac{f(x)}{g(x)} = \lim_{x \to x_0^+} \frac{f'(\xi)}{g'(\xi)} = L.$$

A similar argument can be made for the left side limit.

A more important consequence is the following.

Theorem 14 *Let f be a function such that $f, f', \ldots, f^{(n)}$ all exist and are continuous on some open interval containing x_0 and x. Moreover assume that $f^{(n+1)}$ exists ot thainterval. Then there exists a ξ between x and x_0 such that*

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \cdots + f^{(n)}(x_0)\frac{(x - x_0)^n}{n!} + f^{(n+1)}(\xi)\frac{(x - x_0)^{n+1}}{(n+1)!}$$

Proof: The proof is again just a clever application of the mean value theorem. To do this we assume that $x > x_0$. Then there is a number K such that

$$\frac{(x - x_0)^{n+1}}{(n+1)!}K = f(x) - \left(f(x_0) + \sum_{k=1}^{n} f^{(k)}(x_0)\frac{(x - x_0)^k}{k!} \right).$$

Next define

$$F(t) = f(x) - \left(f(t) + \sum_{k=1}^{n} f^{(k)}(x_0) \frac{(x-t)^k}{k!} + \frac{(x-t)^{n+1}}{(n+1)!} K \right).$$

Observe that $F(x_0) = F(x) = 0$. Thereforee ,there exists a $\xi \in (x_0, x)$ such that $F'(\xi) = 0$, or

$$0 = F'(\xi)) = f^{(n+1)}(\xi) \frac{(x-\xi)^{n=1}}{(n+1)!} - K \frac{(x-\xi)^{n=1}}{(n+1)!},$$

and since $x - \xi \neq 0$ the result follows.