Advanced Calculus I
MATH 409
Summer 2019 Lecture Notes

F. Baudier (Texas A&M University)

June 20, 2019

Contents

1 Introduction **5**
1.1 Functions . . . 5
1.2 Definition and Basic Properties . . . 5
1.3 Composition of Functions . . . 7
1.4 Surjective and Injective Functions . . . 8
1.4.1 Definitions and examples . . . 8
1.4.2 Injectivity, surjectivity and composition . . . 9
1.5 Invertible Functions . . . 10
1.6 Functions and Sets . . . 13
1.7 Principle of Mathematical Induction . . . 16
1.8 The Well-Ordering Principle . . . 17

2 The Real Numbers **19**
2.1 The absolute value . . . 21
2.2 The Least Upper Bound Property . . . 23

3 Sequences of Real Numbers **29**
3.1 Convergence of a sequence . . . 29
3.1.1 Definition and basic properties . . . 29
3.1.2 The Monotone Convergence Theorem . . . 31
3.2 Manipulations of limits . . . 32
3.3 Extraction of subsequences . . . 35
3.3.1 Bolzano-Weierstrass Theorem . . . 36
3.3.2 Limit supremum and limit infimum . . . 37

4 Introduction to Metric Topology **39**
4.1 Completeness . . . 39
4.1.1 Cauchy sequences . . . 39
4.2 Divergence to $\pm\infty$ . . . 41
4.3 Sequential Heine-Borel theorem . . . 41

5 Continuity **45**
5.1 Definition and basic properties . . . 45
5.2 The Intermediate Value Theorem . . . 50
5.3 The Extreme Value Theorem . . . 53
5.4 Uniform Continuity . . . 54
5.5 Continuity of inverse functions . . . 56

6 Differentiation 59
6.1 Definition and basic properties 59
6.2 Rules of differentiation 61
6.2.1 Basic Rules 61
6.2.2 Product Rule 63
6.2.3 Quotient Rule 63
6.2.4 Chain Rule 64
6.3 The Mean Value Theorem and its applications 65
6.3.1 Rolle's Theorem 65
6.3.2 The Mean Value Theorem 66
6.3.3 Cauchy's Mean Value Theorem 66
6.3.4 Applications 67
6.3.4.1 L'Hôpital's Rules 67
6.3.4.2 Taylor's Theorem 69
6.4 Monotone Functions and Derivatives 70
6.4.1 Various tests 70
6.4.2 Differentiablity of inverse functions 71

7 Integration 73
7.1 Definition of the Riemann Integral 73
7.1.1 Riemann Sums 73
7.1.2 Riemann integrable functions 74
7.2 Basic properties of the Riemann integral 75
7.2.1 Riemann integrability criteria 75
7.2.2 Algebraic and order properties 77
7.2.3 Integration by parts 78
7.3 The Fundamental Theorem of Calculus 78

Chapter 1

Introduction

In this introductory chapter we review the concept of function, the principle of mathematical induction, and the well-ordering principle. We assume that the reader is familiar with basic logic and set theory material from MATH 220. We also take the opportunity to prove a few statements and thus review a few proof techniques (double inclusion proofs, proof by contradiction...).

1.1 Functions

We recall some important definitions about functions.

1.2 Definition and Basic Properties

A function between two sets is a correspondence between elements of these two sets that enjoy some special properties.

Definition 1: Functions

Let X and Y be nonempty sets. A *function* from X to Y is a correspondence that assigns to *every* element in X *one and only one* element in Y. Formally, a function from X to Y is a subset $F \subseteq X \times Y$ such that

$$[(\forall x \in X)(\exists! y \in Y)\ (x, y) \in F].$$

Note that the logical formula $[(\forall x \in X)(\exists! y \in Y)\ (x, y) \in F]$ is equivalent to the logical formula

$$[(\forall x \in X)(\exists y \in Y)\ (x, y) \in F]$$

$$\wedge$$

$$[(\forall x \in X)[((x, y_1) \in F) \wedge ((x, y_2) \in F)] \implies (y_1 = y_2)]].$$

Remark 1

Since functions play a central role in set theory and in mathematics in general we use a specific terminology. A function is usually denoted by f (instead of F) and we write $f\colon X \to Y$ to say that f is a function from X to Y (instead of $F \subseteq X \times Y$). Since for every $x \in X$ there is a unique element $y \in Y$ such that $(x, y) \in F$, we prefer a much more convenient functional notation. Therefore, we will denote by $f(x)$ the unique element that is in correspondence with x. If $f(x) = y$ we say that y is the image of x or that x is the preimage of y. We call X the domain of f and Y the codomain.

To show that a correspondence $f\colon X \to Y$ is a function we must check that

$$(\forall x \in X)(\exists y \in Y)[f(x) = y]$$

and

$$(\forall x_1 \in X)(\forall x_2 \in X)[(x_1 = x_2) \implies (f(x_1) = f(x_2))].$$

We now define what it means for two functions to be equal.

Definition 2: Equality for functions

Two functions $f_1\colon X_1 \to Y_1$ and $f_2\colon X_2 \to Y_2$ are equal, denoted $f_1 = f_2$, if they have the *same* domain, the *same* codomain and their actions on elements in X are the same. Formally,

$$f_1 = f_2 \iff (X_1 = X_2) \wedge (Y_1 = Y_2) \wedge ((\forall z \in X_1)[f_1(z) = f_2(z)]).$$

The next definition introduces the concept of image, or range, of a function.

Definition 3: Image (or range) of a function

Let $f\colon X \to Y$ be a function. The image (or the range) of the function f is the set, denoted $\mathrm{Im}(f)$, of all elements in the codomain that are the image of an element in the domain. Formally,

$$\mathrm{Im}(f) = \{f(x) \mid x \in X\} = \{y \in Y \mid (\exists x \in X)[y = f(x)]\}.$$

Remark 2

The image of a function is a subset of the *codomain* of the function. It follows from the definition that

$$y \in \mathrm{Im}(f) \iff (\exists x \in X)[y = f(x)].$$

The next definition introduces the concept of graph of a function.

Definition 4: Graph of a function

Let X and Y be nonempty sets and $f\colon X \to Y$ be a function. The graph of the function f is the set, denoted G_f, of all ordered pairs (x, y) of elements $x \in X$ and $y \in Y$ that are in correspondence. Formally,

$$G_f = \{(x, y) \in X \times Y \mid y = f(x)\}.$$

Remark 3

The graph of a function is a subset of the Cartesian product of its domain with its codomain. It follows from the definition that

$$z \in G_f \iff (\exists x \in X)[z = (x, f(x))].$$

Remark 4

Let X and Y be nonempty sets. We denote $F(X, Y) = \{f \mid f\colon X \to Y\}$, the set of all functions from X to Y. If $X = Y$, we simply write $F(X)$.

1.3 Composition of Functions

Assume we are given two functions f and g. If the codomain of f coincides with the domain of g then it is make sense to look at what element is obtained if we first apply f and then g to an element in the domain of f. This procedure gives a function from the domain of f in the codomain of g.

Definition 5: Composition of functions

Let X, Y, Z be nonempty sets, and let $f\colon X \to Y$, $g\colon Y \to Z$. We define a function $g \circ f\colon X \to Z$, called the composition of f and g, by $g \circ f(x) = g(f(x))$, $\forall x \in X$.

Note that for the composition to be defined we just need the image of f to be a subset of the domain of g.

Remark 5

In general, $g \circ f \neq f \circ g$ and the composition is not a commutative operation! Indeed, consider the function $f\colon \mathbb{R} \to \mathbb{R}$ defined for all $x \in \mathbb{R}$ by $f(x) = 3x$ and the function $g\colon \mathbb{R} \to \mathbb{R}$ defined for all $x \in \mathbb{R}$ by $g(x) = x^2$. It is easy to see that $g \circ f$ and $f \circ g$ have the same domain and codomain, but for instance $g \circ f(1) = 9 \neq 3 = f \circ g(1)$.

The composition operation is associative.

Proposition 1: Associativity of the composition

Let W, X, Y, Z be nonempty sets. Let $f\colon W \to X$, $g\colon X \to Y$, and $h\colon Y \to Z$. Then, $(h \circ g) \circ f = h \circ (g \circ f)$.

Proof. Observe that W is the domain of both $(h \circ g) \circ f$ and $h \circ (g \circ f)$, and that Z is the codomain of both $(h \circ g) \circ f$ and $h \circ (g \circ f)$. It remains to show that for all $w \in W$, $((h \circ g) \circ f)(w) = (h \circ (g \circ f))(w)$. By definition of the composition operation it follows that if $x \in X$ then

$$((h \circ g) \circ f)(w) = (h \circ g)(f(w)) = h(g(f(w)))$$

and

$$(h \circ (g \circ f))(w) = h((g \circ f)(w)) = h(g(f(w))).$$

Therefore, $((h \circ g) \circ f)(w) = h(g(f(w))) = (h \circ (g \circ f))(w)$ and the two functions are equal. $\square$

1.4 Surjective and Injective Functions

1.4.1 Definitions and examples

A surjective function (or onto function) is a function whose image fills in completely the codomain.

Definition 6: Surjective function

Let X and Y be nonempty sets. A function $f\colon X \to Y$ is surjective (or onto, or a surjection) if *every* element in the codomain of f admits a preimage in the domain of f. Formally,

$$f\colon X \to Y \text{ is surjective} \iff (\forall y \in Y)(\exists x \in X)[y = f(x)].$$

The following proposition is a characterization of surjectivity in terms of the image of the function.

Proposition 2: Characterization of surjectivity in terms of the image

Let X and Y be nonempty sets. Let $f\colon X \to Y$ be a function. Then, f is surjective if and only if $\operatorname{Im}(f) = Y$.

Proof. We know that $\operatorname{Im}(f) \subseteq Y$ always holds, but the definition of injectivity says that $Y \subseteq \operatorname{Im}(f)$. Therefore $Y = \operatorname{Im}(f)$. $\square$

A function is injective (or one-to-one often abbreviated as $1-1$) if no two distinct elements in the domain are assigned the same element in the codomain.

Definition 7: Injective function

Let X and Y be nonempty sets. A function $f\colon X \to Y$ is injective (or one-to-one, or an injection) if *every two distinct* elements in the domain have *distinct* images in the codomain. Formally,

$$f\colon X \to Y \text{ is injective}$$
$$\iff$$
$$(\forall x_1 \in X)(\forall x_2 \in X)[\neg(x_1 = x_2) \implies \neg(f(x_1) = f(x_2))].$$

Remark 6

In practice, to show that a function is injective we need to prove *either* one of the following two logically equivalent statements (the second statement is the contrapositive of the first statement.):

- for all $x_1, x_2 \in X$ if $x_1 \neq x_2$ then $f(x_1) \neq f(x_2)$.
- for all $x_1, x_2 \in X$ if $f(x_1) = f(x_2)$ then $x_1 = x_2$.

Definition 8: Bijective function

Let X and Y be nonempty sets. Let $f\colon X \to Y$ be a function. Then f is bijective (or a bijection) if f is *both* injective and surjective. In the case where $X = Y$ a bijection is simply called a permutation.

1.4.2 Injectivity, surjectivity and composition

In this section we show that injectivity, surjectivity, and bijectivity are stable under composition.

Proposition 3: Stability of injectivity under composition

Let W, X, Y be nonempty sets. Let $f\colon W \to X$, $g\colon X \to Y$. If f and g are injective, then $g \circ f$ is also injective.

Proof. Assume that f and g are injective. Let $w_1, w_2 \in W$ such that $g \circ f(w_1) = g \circ f(w_2)$, then $g(f(w_1)) = g(f(w_2))$ (by definition of the composition) and $f(w_1) = f(w_2)$ (by injectivity of g). Now it follows from the injectivity of f that $w_1 = w_2$, and $g \circ f$ is injective.

□

Proposition 4: Stability of surjectivity under composition

Let W, X, Y be nonempty sets. Let $f\colon W \to X$, $g\colon X \to Y$. If f and g are surjective, then $g \circ f$ is also surjective.

Proof. Assume that f and g are surjective. Let $y \in Y$, then there exists $x \in X$ such that $g(x) = y$ (by surjectivity of g). Since $x \in X$, there exists $w \in W$ such that $x = f(w)$ (by surjectivity of f). And hence, $y = g(x) = g(f(w)) = g \circ f(w)$ (by definition of the composition). We have just shown that for every $y \in Y$ there exists $w \in W$ such that $y = g \circ f(w)$, which means that $g \circ f$ is surjective. $\square$

Proposition 5: Stability of bijectivity under composition

Let W, X, Y be nonempty sets. Let $f: W \to X$, $g: X \to Y$. If f and g are bijective, then $g \circ f$ is also bijective.

Proof. Assume that f and g are bijective, then in particular they are both injective . By Theorem 15, $g \circ f$ is then injective. A similar reasoning using Theorem 16 will show that $g \circ f$ is surjective, and hence $g \circ f$ is bijective. $\square$

1.5 Invertible Functions

In this section we take a look at those functions whose actions can be "undone".

Definition 9: Invertibility

Let X, Y be nonempty sets. Let $f: X \to Y$ be a function. We say that f is invertible (or admits an inverse) if there exists a function $g: Y \to X$ such that $f \circ g = i_Y$ and $g \circ f = i_X$.

Being invertible is closely connected to being bijective. Indeed, as we will see shortly invertibility and bijectivity are actually equivalent notions! The goal of this section is to prove this equivalence.

Theorem 1

Let X, Y be nonempty sets. Let $f: X \to Y$ be a function. If f is invertible then f is injective.

Proof. Assume that f is invertible. Then there is a function $g : Y \to X$ such that $g \circ f = i_X$ and $f \circ g = i_Y$. If $x_1, x_2 \in X$ and $f(x_1) = f(x_2)$, then $x_1 = g(f(x_1)) = g(f(x_2)) = x_2$. Thus f is injective. $\square$

It follows from the injectivity of invertible functions that the inverse of an invertible function is uniquely determined.

Proposition 6: Uniqueness of the inverse

Let X and Y be nonempty sets. Let $f: X \to Y$ be a function. If f is invertible then f has a unique inverse.

Proof. Let $f\colon X \to Y$ be a function. Our goal is to show that if there are two functions $g_1, g_2\colon Y \to X$ such that $f \circ g_1 = i_Y$, $g_1 \circ f = i_X$, $f \circ g_2 = i_Y$, and $g_2 \circ f = i_X$, then $g_1 = g_2$. Let $y \in Y$ then $(f \circ g_1)(y) = i_Y(y) = y$ and $(f \circ g_2)(y) = i_Y(y) = y$, thus $(f \circ g_1)(y) = (f \circ g_2)(y)$. It follows from the definition of the composition that $f(g_1(y)) = f(g_2(y))$ and since f is invertible, f is injective (Theorem 15) and hence $g_1(y) = g_2(y)$. Therefore, $g_1 = g_2$. □

Remark 7

If f is invertible, by Proposition 20 the unique function satisfying the conditions of the definition is called the inverse of f and is denoted f^{-1}.

Proposition 7: Stability of invertibility under composition

Let X, Y, Z be nonempty sets. Let $f\colon X \to Y$ and $g\colon Y \to Z$ be invertible functions. Then $g \circ f\colon X \to Z$ is invertible and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Proof. Let g^{-1} and f^{-1} be the inverses of g and f respectively. It follows from the associativity of the composition operation that,

$$\begin{aligned}(g \circ f) \circ (f^{-1} \circ g^{-1}) &= g \circ (f \circ f^{-1}) \circ g^{-1} \\ &= g \circ i_Y \circ g^{-1} \\ &= g \circ g^{-1} \\ &= i_Z\end{aligned}$$

and similarly,

$$\begin{aligned}(f^{-1} \circ g^{-1}) \circ (g \circ f) &= f^{-1} \circ (g^{-1} \circ g) \circ f) \\ &= f^{-1} \circ i_Y \circ f) \\ &= f^{-1} \circ f \\ &= i_X.\end{aligned}$$

Therefore, $g \circ f$ is invertible and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

□

Theorem 2

Let X, Y be nonempty sets. Let $f\colon X \to Y$ be a function. If f is invertible then f is surjective.

Proof. Assume that f is invertible. Then there is a function $g : Y \to X$ such that $g \circ f = i_X$ and $f \circ g = i_Y$. Let $y \in Y$, and put $x = g(y)$. Then by definition of g, one has $x \in X$, and thus

$$\begin{aligned} f(x) &= f(g(y)) \text{ (because } f \text{ is a function)} \\ &= (f \circ g)(y) \text{ (by definition of the composition)} \\ &= i_Y(y) \text{ (since } f \circ g(y) = i_Y(y) \text{ by our assumption)} \\ &= y \text{ (by definition of the identity function on } Y.) \end{aligned}$$

Therefore f is surjective. □

Theorem 3

Let X, Y be nonempty sets. Let $f\colon X \to Y$ be a function. If f is bijective then f is invertible.

Proof. Assume f is bijective. Given $y \in Y$, since f is surjective there is some $x \in X$ such that $y = f(x)$, and since f is injective this x is unique. Indeed if there are $x_1, x_2 \in X$ such that $f(x_1) = y = f(x_2)$, then $x_1 = x_2$ by injectivity of f. So for every $y \in Y$ there is a unique $x_y \in X$ such that $y = f(x_y)$. We will define a function $g : Y \to X$ by assigning to every element $y \in Y$ to the unique element $x_y \in X$ such that $f(x_y) = y$, i.e. $g(y) = x_y$. By uniqueness of x_y, g is a function.

Given $y \in Y$, then $g(y) = x_y$ where $f(x_y) = y$, and thus $f(g(y)) = f(x_y) = y$ (since f is a function). It follows from the definition of the composition that $(f \circ g)(y) = y$, and by definition of the identity function that $(f \circ g)(y) = i_Y(y)$. Since $y \in Y$ was arbitrary, one has $f \circ g = i_Y$.

It remains to show that $(g \circ f) = i_X$. Now given $x \in X$, $g(f(x))$ is the element $x_0 \in X$ such that $f(x_0) = f(x)$. That is, $g(f(x)) = x_0 = x$, since f is injective. Thus $g \circ f = i_X$, and therefore f is invertible. □

Combining the last three theorems we obtain the following corollary.

Corollary 1

Let X and Y be nonempty sets. Let $f\colon X \to Y$. Then,

$$f \text{ is invertible if and only if } f \text{ is bijective.}$$

Proof. Assume that f is invertible, then it follows by Theorem 15 that f is injectve and by Theorem 16 that f is surjective. Therefore, f is bijective. The converse is Theorem 17. □

We can also define the notion of right/left-inverse of a function.

Definition 10: Right-inverse

Let X, Y be nonempty sets. Let $f\colon X \to Y$ be a function. We say that f is right-invertible (or admits a right-inverse) if there exists a function $g\colon Y \to X$ such that $f \circ g = i_Y$.

Definition 11: Left-inverse

Let X, Y be nonempty sets. Let $f: X \to Y$ be a function. We say that f is left-invertible (or admits a left-inverse) if there exists a function $g: Y \to X$ such that $g \circ f = i_X$.

1.6 Functions and Sets

Recall that the image of a function $f: X \to Y$ is the set $\text{Im}(f) = \{y \in Y \mid (\exists x \in X)[y = f(x)]\}$. We generalize this concept in the following definition.

Definition 12: Direct image of a set

Let X, Y be nonempty sets. Let $f: X \to Y$ be a function. If $Z \subseteq X$, the image of Z under f is the set, denoted $f(Z)$, of all elements in the codomain that are the image of at least one element in Z. Formally,

$$f(Z) = \{y \in Y \mid (\exists z \in Z)[y = f(z)]\}.$$

Remark 8

- Note that $f(X)$ is simply the image of f, *i.e.*, $\text{Im}(f) = f(X)$.
- It follows from the definition that

$$v \in f(Z) \iff (\exists z \in Z)\ [v = f(z)].$$

The following proposition states that inclusion is preserved under taking direct images.

Proposition 8

Let X, Y be nonempty sets. Let $f: X \to Y$ be a function. Let W and Z be subsets of X. If $W \subseteq Z$, then $f(W) \subseteq f(Z)$

Proof. If $f(W)$ is empty then the conclusion holds. Otherwise, let $v \in f(W)$ then there exists $w \in W$ such that $v = f(w)$ (by definition of the direct image). But since $W \subseteq Z$ it follows that $w \in Z$ and thus $v \in f(Z)$ (by definition of the direct image). Therefore, $f(W) \subseteq f(Z)$. □

The following proposition states that the direct image of an union is the union of the direct images.

Proposition 9

Let X, Y be nonempty sets. Let $f: X \to Y$ be a function and W and Z be subsets of X. Then, $f(W \cup Z) = f(W) \cup f(Z)$

Proof. The proof is a classical double inclusion argument (and we do not include below the trivial cases when the sets are empty).

- We first show that $f(W \cup Z) \subseteq f(W) \cup f^{-1}(Z)$. Let $y \in f(W \cup Z)$, then there exists $x \in W \cup Z$ such that $y = f(x)$ (by definition of the image) thus $y = f(x)$ for some $x \in W$ or $y = f(x)$ for some $x \in Z$ (by definition of the union) and hence $y \in f(W)$ or $y \in f(Z)$ (by definition of the image) and $y \in f(W) \cup f(Z)$ (by definition of the union). Therefore $f(W \cup Z) \subseteq f(W) \cup f(Z)$.

- Now we show that $f(W) \cup f(Z) \subseteq f(W \cup Z)]$. Let $y \in f(W) \cup f(Z)$, then $y \in f(W)$ or $y \in f(Z)$ (by definition of the union) thus $y = f(x)$ for some $x \in W$ or $y = f(x)$ for some $x \in Z$ (by definition of the image) and $y = f(x)$ for some $x \in W \cup Z$ (by definition of the union) thus $y \in f(W \cup Z)$ (by definition of the inverse image). Therefore $f(W) \cup f(Z) \subseteq f(W \cup Z)$.

□

The situation is slightly different as far as intersection is concerned.

Proposition 10

Let X, Y be nonempty sets. Let $f\colon X \to Y$ be a function and W and Z be subsets of X. Then,

$$f(W \cap Z) \subseteq f(W) \cap f(Z).$$

Proof. Let $y \in f(W \cap Z)$, then there exists $x \in W \cap Z$ such that $y = f(x)$ (by definition of the image), thus $y = f(x)$ for some $x \in W$ and $y = f(x)$ for some $x \in Z$ (by definition of the intersection), and hence $y \in f(W)$ and $y \in f(Z)$ (by definition of the image), and $y \in f(W) \cap f(Z)$ (by definition of the intersection). Therefore $f(W \cap Z) \subseteq f(W) \cap f(Z)$. □

Definition 13: Inverse image of a set

Let X and Y be nonempty sets and let $f\colon X \to Y$ be a function. Let Z be a subset of Y. Then the inverse image of Z with respect to the function f, denoted $f^{-1}(Z)$, is the set of all elements in X that have their image in Z. Formally,

$$f^{-1}(Z) := \{x \in X \mid f(x) \in Z\}.$$

Remark 9

- In this context the symbol f^{-1} does not refer to the inverse of the function f (which might not exist in the first place).
- If follows from the definition that $v \in f^{-1}(Z) \iff f(v) \in Z$.

The following proposition states that inclusion is preserved under taking inverse images.

Proposition 11

Let X, Y be nonempty sets. Let $f\colon X \to Y$ be a function. Let W and Z be subsets of Y. If $W \subseteq Z$, then $f^{-1}(W) \subseteq f^{-1}(Z)$

Proof. If $f^{-1}(W)$ is empty then the conclusion holds. Otherwise, let $v \in f^{-1}(W)$ then $f(v) \in W$ and it follows from $W \subseteq Z$ that $f(v) \in Z$, and hence $v \in f^{-1}(Z)$. Therefore, $f^{-1}(W) \subseteq f^{-1}(Z)$. □

The following proposition states that the inverse image of an union is the union of the inverse images.

Proposition 12

Let X and Y be nonempty sets and let $f\colon X \to Y$ be a function. Let W and Z be subsets of Y. Then,

$$f^{-1}(W \cup Z) = f^{-1}(W) \cup f^{-1}(Z).$$

Proof. The proof is a classical double inclusion argument.

- We first show the inclusion $f^{-1}(W \cup Z) \subseteq f^{-1}(W) \cup f^{-1}(Z)$. Let $x \in f^{-1}(W \cup Z)$, then $f(x) \in W \cup Z$ (by definition of the inverse image) thus $f(x) \in W$ or $f(x) \in Z$ (by definition of the union) and hence $x \in f^{-1}(W)$ or $x \in f^{-1}(Z)$ (by definition of the inverse image) and $x \in f^{-1}(W) \cup f^{-1}(Z)$ (by definition of the union). Therefore $f^{-1}(W \cup Z) \subseteq f^{-1}(W) \cup f^{-1}(Z)$.

- Then we show that $f^{-1}(W) \cup f^{-1}(Z) \subseteq f^{-1}(W \cup Z)]$. Let $x \in f^{-1}(W) \cup f^{-1}(Z)$, then $x \in f^{-1}(W)$ or $x \in f^{-1}(Z)$ (by definition of the union) and $f(x) \in W$ or $f(x) \in Z$ (by definition of the inverse image) and hence $f(x) \in W \cup Z$ (by definition of the union) thus $x \in f^{-1}(W \cup Z)$ (by definition of the inverse image). Therefore $f^{-1}(W) \cup f^{-1}(Z) \subseteq f^{-1}(W \cup Z)$.

□

The following proposition states that the inverse image of an intersection is the intersection of the inverses images.

Proposition 13

Let X and Y be nonempty sets and let $f\colon X \to Y$ be a function. Let W and Z be subsets of Y. Then,

$$f^{-1}(W \cap Z) = f^{-1}(W) \cap f^{-1}(Z).$$

Proof. The proof is a classical double inclusion argument.

- First the inclusion $f^{-1}(W \cap Z) \subseteq f^{-1}(W) \cap f^{-1}(Z)]$. Let $x \in f^{-1}(W \cap Z)$, then $f(x) \in W \cap Z$ (by definition of the inverse image) thus $f(x) \in W$ and $f(x) \in Z$ (by definition of the intersection) and hence $x \in f^{-1}(W)$ and $x \in f^{-1}(Z)$ (by definition of the inverse image) and $x \in f^{-1}(W) \cap f^{-1}(Z)$ (by definition of the intersection). Therefore $f^{-1}(W \cap Z) \subseteq f^{-1}(W) \cap f^{-1}(Z)$.

- Then, the inclusion $f^{-1}(W) \cap f^{-1}(Z) \subseteq f^{-1}(W \cap Z)]$. Let $x \in f^{-1}(W) \cap f^{-1}(Z)$, then $x \in f^{-1}(W)$ and $x \in f^{-1}(Z)$ (by definition of the intersection) and $f(x) \in W$ and $f(x) \in Z$ by definition of the inverse image, and hence $f(x) \in W \cap Z$ (by definition of the intersection) thus $x \in f^{-1}(W \cap Z)$ (by definition of the inverse image). Therefore $f^{-1}(W) \cap f^{-1}(Z) \subseteq f^{-1}(W \cap Z)$.

□

1.7 Principle of Mathematical Induction

The principle of mathematical induction is a very powerful tool to deal with infinite objects and to prove rigorously infinitely many (in the sense that they can be enumerated) statements. In a very general form the principle can be stated as follows.

Theorem 4: Principle of Mathematical Induction

Let $P(n)$ be a predicate where the variable takes integer values. Suppose that there exists $k_0 \in \mathbb{Z}$ such that

$P(k_0)$ is true (the base case)

and

for all $k \geq k_0$, $P(k+1)$ is true under the assumption that $P(k)$ is true (the induction step),

then for all $k \geq k_0$ $P(k)$ is true (the conclusion).

The principle of mathematical induction is most commonly used when $k_0 = 1$. Recall that $\mathbb{N} := \{1, 2, 3, \ldots, \}$ is the set of natural numbers.

Theorem 5: Strong Induction Theorem

If Y is a subset of $\mathbb{N}$ such that:

1. $1 \in Y$,
2. for all $k \in \mathbb{N}$, if $\{1, 2, \ldots, k\} \subseteq Y$ implies that $k + 1 \in Y$,

then $Y = \mathbb{N}$.

Hint. Apply the principle of mathematical induction to the statement

$$P(n)\colon \{1, 2, \ldots, n\} \subseteq Y.$$

□

1.8 The Well-Ordering Principle

The natural total order relation on $\mathbb{N}$ is denoted $\leq$.

Definition 14: Least element

Let X be a nonempty subset of $\mathbb{N}$. An element $m \in X$ is said to be a least element of X if for all $k \in X$, $m \leq k$.

Proposition 14

Let X be a nonempty subset of $\mathbb{N}$. If X has a least element, then this element is unique.

Hint. Use the antisymmetry property of the ordering. □

Theorem 6: Well-Ordering Principle

Every nonempty subset X of $\mathbb{N}$ has a least element.

Hint. Assume by contradiction that X does not have a least element and apply the Strong Induction Theorem to $Y := \mathbb{N} \setminus X$. □

Note that we deduced the Well-Ordering Principle from the Induction Axiom. The Well-Ordering Principle is actually logically equivalent to the Induction Axiom. The Well-Ordering Principle has the following important application.

Recall that we write $x < y$ if and only if $(x \leqslant y) \wedge (x \neq y)$.

Proposition 15

There is no natural number n such that $0 < n < 1$.

Hint. Argue by contradiction and apply the Well-Ordering Principle to the set $X = \{n \in \mathbb{N} \colon 0 < n < 1\}$. □

Proof. Consider the set $X = \{n \in \mathbb{N} \colon 0 < n < 1\}$. Assume by contradiction that X is non-empty. By the Well-Ordering Principle X has a least element $a \in X$. But $0 < a < 1$ and thus $0 < a^2 < a < 1$. Since a^2 is a natural number we get a contradiction with the minimality of a. □

Chapter 2

The Real Numbers

Since we will be doing real analysis we need to understand the main properties of the set of real numbers $\mathbb{R}$ and of its classical subsets $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$. The set of real numbers is equipped with an algebraic structure and a compatible order structure. We start by discussing its algebraic structure.

The set of real numbers $\mathbb{R}$ is equipped with two binary operations $+$ and $\cdot$ satisfying the following properties hold.

C1 For all $x, y \in \mathbb{R}$, $x + y \in \mathbb{R}$.

C2 For all $x, y \in \mathbb{R}$, $x \cdot \in \mathbb{R}$.

F1 (Associativity) For all $x, y, z \in \mathbb{R}$,

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

and

$$(x \cdot y) \cdot z = x \cdot (y \cdot z).$$

F2 (Commutativity) For all $x, y \in \mathbb{R}$,

$$x + y = y + x$$

and

$$x \cdot y = y \cdot z.$$

F3 (Distributivity) For all $x, y, z \in \mathbb{R}$,

$$x \cdot (y + z) = x \cdot y + x \cdot z.$$

F4 (Additive identity) There exists a unique element $0 \in \mathbb{R}$ such that for all $x \in \mathbb{R}$, $x + 0 = 0 + x$.

F5 (Multiplicative identity) There exists a unique element $1 \in \mathbb{R}$ such that $1 \neq 0$ and for all $x \in \mathbb{R}$, $x \cdot 1 = 1 \cdot x$.

F6 (Additive inverses) For all $x \in \mathbb{R}$, there exists a unique element $-x \in \mathbb{R}$ such that $x + (-x) = -x + x = 0$.

F7 (Multiplicative inverses) For all $x \in \mathbb{R} \setminus \{0\}$, there exists a unique element $x^{-1} \in \mathbb{R}$ such that $x \cdot x^{-1} = x^{-1} \cdot x = 1$.

The properties F1-F7 are sufficient to recover all the usual algebraic laws of real numbers that you have been using without any justification in previous math courses. Every set equipped with two binary operations satisfying similar properties is called a commutative field, and the study of such sets equipped with such structure belongs to the field of abstract algebra.

Definition 15: Commutative fields

A set $\mathbb{F}$ equipped with two binary operation $+$ and $\cdot$ is called a commutative field if the following properties hold.

F1 (Associativity) For all $x, y, z \in \mathbb{F}$,

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

and

$$(x \cdot y) \cdot z = x \cdot (y \cdot z).$$

F2 (Commutativity) For all $x, y \in \mathbb{F}$,

$$x + y = y + x$$

and

$$x \cdot y = y \cdot z.$$

F3 (Distributivity) For all $x, y, z \in \mathbb{F}$,

$$x \cdot (y + z) = x \cdot y + x \cdot z.$$

F4 (Additive identity) There exists a unique element $0_{\mathbb{F}} \in \mathbb{F}$ such that for all $x \in \mathbb{F}$, $x + 0_{\mathbb{F}} = 0_{\mathbb{F}} + x$.

F5 (Multiplicative identity) There exists a unique element $1_{\mathbb{F}} \in \mathbb{F}$ such that $1_{\mathbb{F}} \neq 0_{\mathbb{F}}$ and for all $x \in \mathbb{F}$, $x \cdot 1_{\mathbb{F}} = 1_{\mathbb{F}} \cdot x$.

F6 (Additive inverses) For all $x \in \mathbb{F}$, there exists a unique element $-x \in \mathbb{F}$ such that $x + (-x) = -x + x = 0_{\mathbb{F}}$.

F7 (Multiplicative inverses) For all $x \in \mathbb{F} \setminus \{0_{\mathbb{F}}\}$, there exists a unique element $x^{-1} \in \mathbb{F}$ such that $x \cdot x^{-1} = x^{-1} \cdot x = 1_{\mathbb{F}}$.

The set of real numbers also comes with a total order relation, denoted $\leq$, which satisfies the following properties.

OF1 (Additivity Property) For all $x, y, z \in \mathbb{R}$, such that $x \leq y$,

$$x + z \leq y + z.$$

OF2 (Multiplicative Property) For all $x, y, z \in \mathbb{R}$, such that $x \leq y$, $x \cdot z \leq y \cdot z$ provided $0 \leq z$ and $y \cdot z \leq x \cdot z$ provided $z \leq 0$

Every commutative field equipped with a total order relation satisfying similar properties is called a commutative ordered field.

Definition 16: Commutative ordered fields

A commutative order field is a commutative field $(\mathbb{F}, +, \cdot)$ endowed with a total ordering $\preceq$ such that for all $x, y, z \in \mathbb{F}$ the following two properties hold.

OF1 (Additivity Property) For all $x, y, z \in \mathbb{F}$, such that $x \preceq y$,

$$x + z \preceq y + z.$$

OF2 (Multiplicative Property) For all $x, y, z \in \mathbb{F}$, such that $x \preceq y$, $x \cdot z \preceq y \cdot z$ provided $0_{\mathbb{F}} \preceq z$ and $y \cdot z \preceq x \cdot z$ provided $z \preceq 0_{\mathbb{F}}$

Beside $\mathbb{R}$, the set of rational numbers $\mathbb{Q}$ with the natural addition, multiplication and order relation is also a commutative ordered field, and thus we cannot distinguish $\mathbb{R}$ and $\mathbb{Q}$ only by considering the algebraic and order structures. To distinguish $\mathbb{R}$ and $\mathbb{Q}$, and to do analysis we need a metric structure induced by the notion of absolute value.

2.1 The absolute value

Definition 17: Absolute value

The absolute value of a number $x \in \mathbb{R}$, denoted $|x|$, is defined as the (unique) real number

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

Theorem 7: Basic properties of the absolute value

The absolute value satisfies the following properties.

1. (Positive definitness) For all $x \in \mathbb{R}$, $|x| \geq 0$ with $|x| = 0$ if and only if $x = 0$.
2. (Symmetry) For all $x, y \in \mathbb{R}$, $|x - y| = |y - x|$.
3. (Multiplicativity) For all $x, y \in \mathbb{R}$, $|xy| = |x||y|$.
4. Let $x \in \mathbb{R}$ and $M \geq 0$. Then $|x| \leq M$ if and only if $-M \leq x \leq M$.
5. (Subadditivity) For all $x, y \in \mathbb{R}$, $|x + y| \leq |x| + |y|$.
6. (Triangle Inequality) For all $x, y, z \in \mathbb{R}$, $|x - y| \leq |x - z| + |y - z|$.
7. (Reverse Triangle Inequality) For all $x, y \in \mathbb{R}$, $\big||x| - |y|\big| \leq |x - y|$.

Proof.

1. Let $x \in \mathbb{R}$. If $x \geq 0$ then $|x| = x \geq 0$ and if $x < 0$ then $|x| = -x > 0$.

2. Let $x, y \in \mathbb{R}$. If $x - y \geq 0$ then $y - x \leq 0$ and $|x - y| = x - y$ and $|y - x| = -(y-x) = x-y$. If $x-y < 0$ then $y-x > 0$ and $|x-y| = -(x-y) = y-x$ and $|y - x| = y - x$.

3. The proof is based on a study of various cases as in the previous two items and we leave the details as an exercise.

4. Let $x \in \mathbb{R}$ and $M \geq 0$. If $|x| \leq M$ then if $x \geq 0$, then $x = |x|$ and $-M \leq 0 \leq x = |x| \leq M$. Otherwise, if $x < 0$ then $|x| = -x$ and $-M \leq 0 < -x = |x| < M$.

5. Let $x, y \in \mathbb{R}$. Since $-|x| \leq x \leq |x|$ and $-|y| \leq y \leq |y|$ then by adding up theses two inequalities $-(|y|+|x|) \leq x+y \leq |x|+|y|$ and by the previous item $|x + y| \leq |x| + |y|$.

6. Let $x, y, z \in \mathbb{R}$ and set $a = x-z$ and $b = z-y$. It follows from subadditivity that $|x - y| = |a + b| \leq |a| + |b| = |x - z| + |y - z|$.

7. Let $x, y \in \mathbb{R}$, then by subadditivity $|x| = |x - y + y| \leq |x - y| + |y|$, and $|y| = |y-x+x| \leq |y-x|+|x|$. Thus $|x|-|y| \leq |x-y|$ and $|y|-|x| \leq |x-y|$ and the conclusion follows from 4..

□

We will finish this section with a very useful lemma.

Lemma 1

Let $x, y \in \mathbb{R}$.

1. $x \leq y$ if and only if $x < y + \varepsilon$ for all $\varepsilon > 0$.

2. $x \geq y$ if and only if $x > y - \varepsilon$ for all $\varepsilon > 0$.

3. $x = 0$ if and only if $|x| < \varepsilon$ for all $\varepsilon > 0$.

Hint. Prove (1) and (2) by contradiction and use (1), (2) and the antisymmetry of the order relation for (3). □

Proof.

1. The necessary implication is relatively easy to prove. Indeed, if $x \leq y$ then for all $\varepsilon > 0$ one has $x < x + \varepsilon \leq y + \varepsilon$ (we implicitly use OF1 twice here) and the necessary implication follows. Assume now, that $x < y + \varepsilon$ for all $\varepsilon > 0$. Assume by contradiction that $x > y$ and let $\varepsilon_0 = x - y > 0$. By our assumption, $x < y + \varepsilon_0 = y + (x - y) = x$; a contradiction.

2. We only prove the most difficult implication here. The easy one is left to the reader. Assume that $x > y - \varepsilon$ for all $\varepsilon > 0$. Assume by contradiction that $y > x$ and let $\varepsilon_0 = y - x > 0$. By our assumption, $x > y - \varepsilon_0 = y - (y - x) = x$; a contradiction.

3. It is clear that if $x = 0$ then for all $\varepsilon > 0$, $-\varepsilon < 0 = x < \varepsilon$ and the forward implication holds. If $|x| < \varepsilon$ for all $\varepsilon > 0$ then $0 - \varepsilon < x < 0 + \varepsilon$ and we apply (1) and (2) to get $0 \leq x$ and $x \leq 0$. By antisymmetry of the order relation $x = 0$.

□

2.2 The Least Upper Bound Property

Definition 18: Upper bound

Let $X \subset \mathbb{R}$ be non-empty. A number $M \in \mathbb{R}$ (not necessarily in X) is an upper bound for X if for all $x \in X$, $x \leq M$. A set admitting an upper bound is said to be bounded above.

Keep in mind the following remark.

Remark 10

A set has either no upper bound or infinitely many upper bounds.

Definition 19: Supremum

Let $X \subset \mathbb{R}$ be non-empty. A number $s \in \mathbb{R}$ (not necessarily in X) is called a (finite) supremum of the set X if and only if s is an upper bound for X and $s \leq M$ for all upper bounds M of X.

Proposition 16: Uniqueness of the supremum

If a non-empty set $X \subseteq \mathbb{R}$ has a supremum, then it is has only one supremum, that we shall denote $\sup(X)$.

Hint. At some point invoke the antisymmetry of the order relation. □

Proof. Assume that X has two suprema s_1 and s_2. Then by definition s_1 is an upper bound but s_2 being a supremum we have $s_2 \leq s_1$. A similar argument, tells us that $s_1 \leq s_2$, and we conclude by antisymmetry of the order relation that $s_1 = s_2$. □

We will use the following lemma repeatedly.

Lemma 2: Approximation property for suprema

Assume that a non-empty subset X of $\mathbb{R}$ has a finite supremum. Then for every $\varepsilon > 0$ there exists $x_\varepsilon \in X$ such that

$$\sup(X) - \varepsilon < x_\varepsilon \leq \sup(X).$$

Hint. The upper bound holds by definition of the supremum. For the lower bound argue by contradiction. □

Proof. The right-hand side inequality holds for every element in X by definition of the supremum and only the left-hand side inequality requires a proof. Assume by contradiction that there exists $\varepsilon_0 > 0$ such that for all $x \in X$, $\sup(X) - \varepsilon_0 \geq x$, and thus $\sup(X) - \varepsilon_0$ is an upper bound for X that is strictly smaller than $\sup(X)$; a contradiction. □

Proposition 17: Suprema for subsets of the integers

Assume that a non-empty subset X of $\mathbb{Z}$ has a finite supremum. Then $\sup(X) \in X$.

Hint. Use the approximation property for suprema. □

Proof. By the approximation property for suprema, for every $\varepsilon > 0$ there exists $x_\varepsilon \in X$ such that $\sup(X) - \varepsilon < x_\varepsilon \leq \sup(X)$. In particular for $\varepsilon = 1$, there exists $x_1 \in X$ such that $\sup(X) - 1 < x_1 \leq \sup(X)$. If $x_1 = \sup(X)$ then $\sup(X) \in X$ and we are done. Otherwise $x_1 < \sup(X)$ and one can apply the approximation property for suprema one more time to show that there exists $x_2 \in X$ such that $x_1 < x_2 \leq \sup(X)$ (why?). It follows that $0 < x_2 - x_1 \leq \sup(X) - x_1 < 1$ and hence $0 < x_2 - x_1 < 1$. But $x_2 - x_1 \in \mathbb{N}$ since both x_1 and x_2 belong to X which is a subset of the integers; a contradiction (why?). □

The next theorem will be taken for granted and we will not provide a proof.

Theorem 8: The least upper bound property

If X is a non-empty subset of $\mathbb{R}$ that is bounded above, then X has a finite supremum in $\mathbb{R}$.

The least upper bound property is also referred to as "completeness" of $\mathbb{R}$ since the suprema of a non-empty subset of $\mathbb{R}$ that is bounded above belongs to $\mathbb{R}$. In some sense there is no "hole" in $\mathbb{R}$. The least upper bound property does not hold is we consider $\mathbb{Q}$ instead of $\mathbb{R}$. There are "holes" in $\mathbb{Q}$.

Proposition 18: Unboundedness of $\mathbb{N}$

$\mathbb{N}$ is not bounded above in $\mathbb{R}$.

Hint. Argue by contradiction and use the least upper bound property. □

Proof. Assume by contradiction that $\mathbb{N}$ is bounbed above in $\mathbb{R}$, i.e. there exists $M \in \mathbb{R}$ such that for all $n \in \mathbb{N}$, $n \leq M$. By the least upper bound property $\mathbb{N}$ has a finite supremum $s := \sup(\mathbb{N}) \in \mathbb{R}$. Since $s - 1 < s$, $s - 1$ cannot be an upper bound for $\mathbb{N}$ (why?) and there must be a natural number $n_0 \in \mathbb{N}$ such that $s - 1 < n_0$. Therefore, $s < n_0 + 1$ and since $n_0 + 1 \in \mathbb{N}$, s is not an upper bound either; a contradiction. □

Theorem 9: Archimedean Principle

For every $x, y \in \mathbb{R}$, with $x > 0$, there is an integer $n \in \mathbb{N}$ such that $y < nx$.

Hint. Argue by contradiction and use the fact that $\mathbb{N}$ is unbounded. □

Proof. Let $x, y \in \mathbb{R}$, with $x > 0$ and assume by contradiction that the conclusion does not hold, i.e. for all $n \in \mathbb{N}$, $y \geq nx$. Since $x > 0$, one has that for all $n \in \mathbb{N}$, $n \leq \frac{y}{x}$, which means that $\mathbb{N}$ is bounded above by $\frac{y}{x}$. This contradicts the fact that $\mathbb{N}$ is not bounded above. □

To prove that $\mathbb{Q}$ is dense in $\mathbb{R}$ we will need the following lemma.

Lemma 3: Floor Lemma

For every $x \in \mathbb{R}$ there exists a unique $m \in \mathbb{Z}$ such that $m \leq x < m + 1$.

Hint. Use the Archimedean Principle and Proposition 17. □

Proof. We start by proving the uniqueness. Assume for the sake of a contradiction that there are $m_1 \neq m_2$ satisfying the conclusion of the lemma, and without loss of generality that $m_1 < m_2$ (otherwise reorder them). Then $m_1 + 1 \leq m + 2$. However, $m_1 \leq x < m_1 + 1$, and thus $x < m_2$ but also $m_2 \leq x < m_2 + 1$, a contradiction. Assume that $x \in \mathbb{R}$ and consider the set $X := \{n \in \mathbb{N} \colon n \leq x\}$ which is bounded above by definition. We need to show that it is non-empty. Consider $-X = \{-x \colon x \in X\}$ then $-X$ is non-empty by the Archimedean Principle. Indeed there exists $n \in \mathbb{N}$ such that $-x \leq n$ and thus $-n \leq x$ with $-n \in \mathbb{Z}$. Consequently, X is non-empty as well. Proposition 17 tells us X has a supremum $k \in X$. Since $k + 1 \in \mathbb{Z}$ and $k + 1 > k$, it follows that $k + 1 \notin X$. Therefore, $x < k + 1$, and $k \leq x < k + 1$ and we simply take $m = k$. □

Theorem 10: Density of $\mathbb{Q}$ in $\mathbb{R}$

For all $x, y \in \mathbb{R}$, such that $x < y$, there is a rational $r \in \mathbb{Q}$ such that $x < r < y$.

Hint. Observe that it is sufficient to treat the case where $0 \leq x < y$. Then use the Archimedean Principle and the Floor Lemma. □

Proof. Assume that $x < y$. Since $y - x > 0$, by the Archimedean Principle there exists $n_0 \in \mathbb{N}$ such that $n_0(y - x) > 1$ and $n_0 x + 1 < n_0 y$. Since $n_0 x \in \mathbb{R}$, by Lemma 3 there exists $m \in \mathbb{Z}$ such that $m \leq n_0 x < m + 1$. Thus, $m + 1 \leq n_0 x + 1 < n_0 y$. It follows from $n_0 x < m + 1$ and $m + 1 < n_0 y$, that $x < \frac{m+1}{n_0} < y$ and the rational $r = \frac{m+1}{n_0}$ is the rational sought. □

There is a “dual” notion to the notion of supremum that we discuss now. Most the the results that we have proven about suprema have a dual version that can be proven using similar arguments.

We first define lower bounds.

Definition 20: Lower bound

Let $S \subset \mathbb{R}$ be non-empty. A number $m \in \mathbb{R}$ (not necessarily in S) is said to be a lower bound for S if for all $x \in S$, $x \geq m$. A set admitting a lower bound is said to be bounded below.

This leads to the notion of infimum.

Definition 21: Infimum

Let $S \subset \mathbb{R}$ be non-empty. A number $t \in \mathbb{R}$ (not necessarily in S) is called a (finite) infimum of the set S if and only if t is an lower bound for S and $t \geq m$ for all lower bounds m of S.

We can also prove that if a set admits an infimum then this infimum is unique.

Proposition 19: Uniqueness of the supremum

If a non-empty set $X \subseteq \mathbb{R}$ has a infimum, then it is has only one infimum, that we shall denote $\inf(X)$.

Proof. The left-hand side inequality holds for every element in X by definition of the infimum and only the right-hand side inequality requires a proof. Assume by contradiction that there exists $\varepsilon_0 > 0$ such that for all $x \in X$, $\inf(X)+\varepsilon_0 \leqslant x$, and thus $\inf(X) + \varepsilon_0$ is a lower bound for X that is strictly larger than $\inf(X)$; a contradiction. □

Lemma 4: Approximation property for infima

Assume that a non-empty subset X of $\mathbb{R}$ has a finite infimum. Show that for every $\varepsilon > 0$ there exists $x_\varepsilon \in X$ such that

$$\inf(X) \leqslant x_\varepsilon < \inf(X) + \varepsilon.$$

Proof. The left-hand side inequality holds for every element in X by definition of the infimum and only the right-hand side inequality requires a proof. Assume by contradiction that there exists $\varepsilon_0 > 0$ such that for all $x \in X$, $\inf(X)+\varepsilon_0 \leqslant x$, and thus $\inf(X) + \varepsilon_0$ is a lower bound for X that is strictly larger than $\inf(X)$; a contradiction. □

Definition 22: Boundedness

Let $S \subset \mathbb{R}$ be non-empty. S is said to be bounded if it is bounded above and below.

Recall that $-E := \{x \in \mathbb{R} \colon -x \in E\}$. The following results are usually convenient to convert a result about suprema into a result about infima and vice-versa.

Proposition 20: Reflection Principle

Let $E \subset \mathbb{R}$ be non-empty.

1. E has a supremum if and only if $-E$ has an infimum, in which case

$$\inf(-E) = -\sup(E).$$

2. E has a infimum if and only if $-E$ has an supremum, in which case

$$\sup(-E) = -\inf(E).$$

Proof. The proofs of the four implications are very similar and we only prove one implication and leave the details of the others as an exercise.

Assume that E has a supremum $s = \sup(E)$, we will show that $t = -s$ is the infimum of $-E$. Indeed t is a lower bound for E since for all $x \in E$, $-x \in -E$ and $t \leq x$ follows from $-x \leq -t = s$. Assume now that l is another lower bound for $-E$ then $-l$ is an upper bound for E and $s \leq -l$, and thus $l \leq -s = t$ and the implication follows.

□

Proposition 21: Monotone property for suprema/infima

Let $A \subseteq B$ be non-empty subsets of $\mathbb{R}$.

1. If B has a supremum, then A has a supremum and $\sup(A) \leq \sup(B)$.
2. If B has a infimum, then A has a infimum and $\inf(B) \leq \inf(A)$.

Proof.

1. If $s = \sup(B)$ then for all $b \in B$, $b \leq s$. Since $A \subset B$, for every $a \in A$, $a \leq s$ and s is an upper bound for A. By the least upper bound property $\sup(A)$ exists and $\sup(A) \leq s$ by the definition of the supremum.
2. A proof similar to the one above will work or you can use the reflection principle and (1). Indeed, If $A \subset B$ then $-A \subset -B$ (prove it!). If B has a infimum, then $-B$ has a supremum and $\inf(B) = -\sup(-B) \leq -\sup(-A) = \inf(A)$.

□

Theorem 11: The greatest lower bound property

If E is a non-empty subset of $\mathbb{R}$ that is bounded below, then E has a finite infimum in $\mathbb{R}$.

Proof. Assume that E is bounded below then $-E$ is bounded above and has a supremum by the least upper bound property. It follows from the reflection principle that E has an infimum. $\square$

Chapter 3

Sequences of Real Numbers

Let us first formally define what we actually mean by a sequence of real numbers.

Definition 23: Sequence of real numbers

A sequence of real numbers is a function $f : \mathbb{N} \to \mathbb{R}$. A sequence whose terms are $x_n := f(n)$ will usually be denoted by $(x_n)_{n=1}^{\infty}$ or $(x_n)_{n\in\mathbb{N}}$.

Remark 11

Do not confuse a sequence $(x_n)_{n=1}^{\infty}$, which is a function, with its image, which is the set $\{x_n : n \in \mathbb{N}\}$.

3.1 Convergence of a sequence

3.1.1 Definition and basic properties

One of our main concern is to understand whether or not the terms in a sequence converge to a certain value, and we need to formally define what we mean by converging.

Definition 24

A sequence of real numbers $(x_n)_{n=1}^{\infty}$ is said to converge to a real number $\ell \in \mathbb{R}$, if for every $\varepsilon > 0$ there exists $N := N(\varepsilon) \in \mathbb{N}$ such that for all $n \geq N$, $|x_n - \ell| < \varepsilon$.

If a sequence $(x_n)_{n=1}^{\infty}$ converges to $\ell \in \mathbb{R}$, we write $\lim_{n\to\infty} x_n = \ell$

Remark 12

There is actually some freedom in the above definition. Indeed, one can replace $|x_n - \ell| < \varepsilon$ in the definition by $|x_n - \ell| \leq \varepsilon$ or $|x_n - \ell| < 2\varepsilon$ or $|x_n - \ell| \leq 100\varepsilon$ for instance, and yet obtain an equivalent definition. This can be useful on occasion.

Proposition 22: Uniqueness of the limit

A sequence of real numbers $(x_n)_{n=1}^{\infty}$ has at most one limit.

Hint 1. Show using the triangle inequality that if $(x_n)_{n=1}^{\infty}$ converges to ℓ_1 and ℓ_2 then for every $\varepsilon > 0$, $|\ell_1 - \ell_2| < \varepsilon$ and conclude. □

Proof 1. Let $\varepsilon > 0$ and assume that $(x_n)_{n=1}^{\infty}$ converges to ℓ_1 and ℓ_2. Then, by definition of convergence, there exist $N_1 \in \mathbb{N}$ such that for all $n \geq N_1$, $|x_n - \ell_1| < \frac{\varepsilon}{2}$, and N_2 in $\mathbb{N}$ such that for all $n \geq N_2$, $|x_n - \ell_2| < \frac{\varepsilon}{2}$. Now, $|\ell_1 - \ell_2| = |\ell_1 - x_n + x_n - \ell_2| \leq |\ell_1 - x_n| + |x_n - \ell_2|$ by the triangle inequality, and if n is such that $n \geq \max\{N_1, N_2\}$ one has $|\ell_1 - \ell_2| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. We just showed that for every $\varepsilon > 0$, $|\ell_1 - \ell_2| < \varepsilon$ and by Lemma 1 we can conclude that $\ell_1 = \ell_2$. □

Hint 2. Argue by contradiction using the triangle inequality. □

Proof 2. Assume by contradiction that $(x_n)_{n=1}^{\infty}$ converges to ℓ_1 and ℓ_2 with $\ell_1 \neq \ell_2$. Without loss of generality assume that $\ell_1 > \ell_2$. Let $\varepsilon = \frac{\ell_1 - \ell_2}{2} > 0$. Then, by the definition of convergence, there exist $N_1 \in \mathbb{N}$ such that for all $n \geq N_1$, $|x_n - \ell_1| < \varepsilon$, and N_2 in $\mathbb{N}$ such that for all $n \geq N_2$, $|x_n - \ell_2| < \varepsilon$. Now, $\ell_1 - \ell_2 = |\ell_1 - \ell_2| = |\ell_1 - x_n + x_n - \ell_2| \leq |\ell_1 - x_n| + |x_n - \ell_2|$ by the triangle inequality, and if $n \geq \max\{N_1, N_2\}$ one has $\ell_1 - \ell_2 < \varepsilon + \varepsilon = \ell_1 - \ell_2$; a contradiction. □

Example 1.

1. Let $c \in \mathbb{R}$ and consider the sequence defined by $x_n = c$ for all $n \in \mathbb{N}$. Show that $\lim_{n\to\infty} x_n = c$.
2. Consider the sequence defined by $x_n = (-1)^n$ for all $n \in \mathbb{N}$. Show that $(x_n)_{n=1}^{\infty}$ is not convergent.
3. Consider the sequence defined by $x_n = \frac{1}{n}$ for all $n \in \mathbb{N}$. Show that $\lim_{n\to\infty} x_n = 0$.

Solutions.

1. Let $\varepsilon > 0$. Since for all $n \in \mathbb{N}$, $|x_n - c| = 0 < \varepsilon$ the conclusion follows.
2. Argue by contradiction that $\lim_{n\to\infty} x_n = \ell$ for some $\ell \in \mathbb{R}$. For $\varepsilon =$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $|x_n - \ell| = |(-1)^n - \ell| < 1$, and hence $\max\{|1-\ell|, |-1-\ell|\} < 1$. It follows from the triangle inequality that $2 = |1+1| = |1 - \ell + \ell + 1| \leq |1-\ell| + |1+\ell| < 1 + 1 = 2$; a contradiction.
3. It follows from the Archimedean Principle that for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $0 < \frac{1}{\varepsilon} < N$. For $n \geq N$, $|\frac{1}{n} - 0| = \frac{1}{n} \leq \frac{1}{N} < \varepsilon$ and $\lim_{n\to\infty} \frac{1}{n} = 0$.

□

Example 2. Consider the Fibonacci sequence recursively defined by

$$f_n = \begin{cases} 1 & \text{if } n = 1 \\ 1 & \text{if } n = 2 \\ f_{n-1} + f_{n-2} & \text{if } n \geq 3. \end{cases}$$

Is the sequence $(f_n)_{n=1}^{\infty}$ convergent?

3.1.2 The Monotone Convergence Theorem

in this section we will discover a sufficient condition that guarantees that a sequence is convergent. Our first result links the convergence of the sequence with its boundedness.

Definition 25: Boundedness

- A sequence of real numbers $(x_n)_{n=1}^{\infty}$ is bounded above if there exists $M \in \mathbb{R}$ such that for all $n \in \mathbb{N}$, $x_n \leq M$.
- A sequence of real numbers $(x_n)_{n=1}^{\infty}$ is bounded below if there exists $m \in \mathbb{R}$ such that for all $n \in \mathbb{N}$, $x_n \geq m$.
- A sequence of real numbers $(x_n)_{n=1}^{\infty}$ is bounded if there exist $m, M \in \mathbb{R}$ such that for all $n \in \mathbb{N}$, $m \leq x_n \leq M$.

Proposition 23: Convergence implies boundedness

Every convergent sequence is bounded.

Proof. Assume that $(x_n)_{n=1}^{\infty}$ converges to ℓ. Then for $\varepsilon = 1$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $|x_n - \ell| \leq 1$, and by the triangle inequality $|x_n| \leq 1 + |\ell|$. Let $M := \max\{|x_1|, |x_2|, \ldots, |x_{N-1}|, 1 + |\ell|\}$, then for all $n \in \mathbb{N}$, $|x_n| \leq M$ and $(x_n)_{n=1}^{\infty}$ is bounded. □

We will prove the Monotone Convergence Theorem which relates the convergence of a sequence with its monotonicity.

Definition 26: Monotonicity for sequences

- A sequence $(x_n)_{n=1}^{\infty}$ is increasing if $x_m \leq x_n$ whenever $m \leq n$. Formally, $(x_n)_{n=1}^{\infty}$ is increasing $\iff \forall m, n \in \mathbb{N}[m \leq n \implies x_m \leq x_n]$.
- A sequence $(x_n)_{n=1}^{\infty}$ is decreasing if $x_m \geq x_n$ whenever $m \leq n$. Formally, $(x_n)_{n=1}^{\infty}$ is decreasing $\iff \forall m, n \in \mathbb{N}[m \leq n \implies x_m \geq x_n]$.
- A sequence $(x_n)_{n=1}^{\infty}$ is monotone if it is either increasing or decreasing.

Theorem 12: Monotone Convergence Theorem (increasing version)

Let $(x_n)_{n=1}^{\infty}$ be a sequence of real numbers. If $(x_n)_{n=1}^{\infty}$ is increasing and bounded above then $(x_n)_{n=1}^{\infty}$ is convergent.

Hint. Use the approximation property of suprema to show that $(x_n)_{n=1}^{\infty}$ converges to $\sup\{x_n : n \in \mathbb{N}\}$. □

Proof. Assume that $(x_n)_{n=1}^\infty$ is bounded above and increasing. By the Least Upper Bound Property $s := \sup\{x_n : n \in \mathbb{N}\}$ is finite. We will show that $(x_n)_{n=1}^\infty$ converges to s. Let $\varepsilon > 0$, then by the approximation property for suprema there exists $n_0 \in \mathbb{N}$ such that $s - \varepsilon < x_{n_0} \leq s$, and thus $0 \leq s - x_{n_0} < \varepsilon$. For $n \geq n_0$, since $(x_n)_{n=1}^\infty$ is increasing, $x_n \geq x_{n_0}$ and $|x_n - s| = s - x_n \leq s - x_{n_0} < \varepsilon$. We just proved that for every $\varepsilon > 0$ there is $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $|x_n - s| < \varepsilon$ and hence $\lim_{n\to\infty} x_n = s$. □

Theorem 13: Monotone Convergence Theorem (decreasing version)

Let $(x_n)_{n=1}^\infty$ be a sequence of real numbers. If $(x_n)_{n=1}^\infty$ is decreasing and bounded below then $(x_n)_{n=1}^\infty$ is convergent.

Proof. If we consider the sequence $(y_n)_{n=1}^\infty = (-x_n)_{n=1}^\infty$, then $(y_n)_{n=1}^\infty$ is increasing and bounded above. By the Monotone Convergence Theorem in increasing version $(y_n)_{n=1}^\infty$ is convergent, and in turn $(x_n)_{n=1}^\infty$ is convergent (prove it if you are not convinced or glimpse at Proposition 24!). □

By combining Proposition 22, Theorem 12, and Theorem 13 one gets the following corollary.

Corollary 2: Monotone Convergence Theorem

Let $(x_n)_{n=1}^\infty$ be a monotone sequence. Then, $(x_n)_{n=1}^\infty$ is convergent if and only if $(x_n)_{n=1}^\infty$ is bounded.

Example 3. Is the sequence $(2^{-n})_{n=1}^\infty$ convergent? If yes what is its limit?

3.2 Manipulations of limits

The results in this section tell us how the operation of taking the limit behaves with respect to the usual algebraic operations (arithmetic of limits) and the order relation (comparison theorems). We will use repeatedly these elementary but crucial properties.

Proposition 24: Arithmetic of limits

Let $(x_n)_{n=1}^{\infty}$ and $(y_n)_{n=1}^{\infty}$ be convergent sequences, then

1. the sequence $(x_n + y_n)_{n=1}^{\infty}$ is convergent and
$$\lim_{n\to\infty}(x_n + y_n) = \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n,$$
2. if $\lambda \in \mathbb{R}$, the sequence $(\lambda x_n)_{n=1}^{\infty}$ is convergent and
$$\lim_{n\to\infty}(\lambda x_n) = \lambda \lim_{n\to\infty} x_n,$$
3. the sequence $(x_n \cdot y_n)_{n=1}^{\infty}$ is convergent and
$$\lim_{n\to\infty}(x_n \cdot y_n) = \lim_{n\to\infty} x_n \cdot \lim_{n\to\infty} y_n,$$
4. if $x_n \neq 0$ for all $n \geq 1$ and if $\lim_{n\to\infty} x_n \neq 0$, then the sequence $(\frac{1}{x_n})_{n=1}^{\infty}$ is well defined and convergent, and
$$\lim_{n\to\infty} \frac{1}{x_n} = \frac{1}{\lim\limits_{n\to\infty} x_n},$$
5. if $y_n \neq 0$ for all $n \geq 1$ and if $\lim_{n\to\infty} y_n \neq 0$, then the sequence $(\frac{x_n}{y_n})_{n=1}^{\infty}$ is well defined and convergent, and
$$\lim_{n\to\infty} \frac{x_n}{y_n} = \frac{\lim\limits_{n\to\infty} x_n}{\lim\limits_{n\to\infty} y_n}$$

Proof. Assume that $\lim_{n\to\infty} x_n = \ell_1 < \infty$ and $\lim_{n\to\infty} y_n = \ell_2 < \infty$.

1. Let $\varepsilon > 0$, then there exist $N_1, N_2 \in \mathbb{N}$ such that for all $n \geq N_1$, $|x_n - \ell_1| < \frac{\varepsilon}{2}$ and for all $n \geq N_2$, $|y_n - \ell_2| < \frac{\varepsilon}{2}$. If follows from the triangle inequality that $|x_n + y_n - (\ell_1 + \ell_2)| = |x_n - \ell_1 + y_n - \ell_2| \leq |x_n - \ell_1| + |y_n - \ell_2|$, and hence for $n \geq \max\{N_1, N_2\}$, $|x_n + y_n - (\ell_1 + \ell_2) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.

2. If $\lambda = 0$ the equality clearly holds. Otherwise, let $\varepsilon > 0$, then $\frac{\varepsilon}{|\lambda|} > 0$ and there exists $N_1 \in \mathbb{N}$ such that for all $n \geq N_1$, $|x_n - \ell_1| < \frac{\varepsilon}{|\lambda|}$ and simply remark that $|\lambda x_n - \lambda| = |\lambda||x_n - \ell_1| < |\lambda|\frac{\varepsilon}{|\lambda|} = \varepsilon$.

3. If follows from the triangle inequality that $|x_n \cdot y_n - (\ell_1 \cdot \ell_2)| = |(x_n - \ell_1)y_n + \ell_1(y_n - \ell_2)| \leq |x_n - \ell_1||y_n| + |y_n - \ell_2||\ell_1|$. Since $(y_n)_{n=1}^{\infty}$ is convergent, and thus bounded, there exists $M > 0$ such that for all $n \in \mathbb{N}$, $|y_n| \leq M$. Let $\varepsilon > 0$. If $|\ell_1| > 0$, then there exist $N_1, N_2 \in \mathbb{N}$ such that for all $n \geq N_1$, $|x_n - \ell_1| < \frac{\varepsilon}{2M}$ and for all $n \geq N_2$, $|y_n - \ell_2| < \frac{\varepsilon}{2|\ell_1|}$, and hence for $n \geq \max\{N_1, N_2\}$, $|x_n \cdot y_n - (\ell_1 \cdot \ell_2)| < \frac{\varepsilon}{2M}M + \frac{\varepsilon}{2|\ell_1|}|\ell_1| = \varepsilon$. If $|\ell_1| = 0$ then for $n \geq \max\{N_1, N_2\}$, $|x_n \cdot y_n| < \frac{\varepsilon}{2M}M < \varepsilon$, and the proof is complete.

4. Assume $\lim_{n\to\infty} x_n = \ell_1 \neq 0$, then for $\varepsilon = \frac{|\ell_1|}{2} > 0$ there exists $N_1 \in \mathbb{N}$ such that for all $n \geq N_1$, $|x_n - \ell_1| < \frac{|\ell_1|}{2}$. It follows from the reverse triangle inequality that $|x_n| > \frac{|\ell_1|}{2} > 0$ for $n \geq N_1$. Also for $\varepsilon > 0$ there exists $N_2 \in \mathbb{N}$ such that for all $n \geq N_2$, $|x_n - \ell_1| < \frac{\varepsilon|\ell_1|^2}{2}$ and for $n \geq \max\{N_1, N_2\}$, $|\frac{1}{x_n} - \frac{1}{\ell_1}| = |\frac{\ell_1 - x_n}{x_n \ell_1}| < \frac{2}{|\ell_1|}|\frac{x_n - \ell_1}{\ell_1}| < \varepsilon$.

5. Since $\frac{x_n}{y_n} = x_n \frac{1}{y_n}$ the result follows by combining (3) and (4).

□

We will show that taking limits, whenever they exist is a "monotone operation".

Theorem 14: Comparison Theorem I

Let $(x_n)_{n=1}^\infty$ and $(y_n)_{n=1}^\infty$ be two convergent sequences. If there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $x_n \leq y_n$ then $\lim_{n\to\infty} x_n \leq \lim_{n\to\infty} y_n$.

Hint. You could either argue by contradiction or show that $\forall \varepsilon > 0, \lim_{n\to\infty} x_n < \lim_{n\to\infty} y_n + \varepsilon$. □

Proof. Let $\ell_1 = \lim_{n\to\infty} x_n$ and $\ell_2 = \lim_{n\to\infty} y_n$. Assume by contradiction that $\ell_1 > \ell_2$. If $\varepsilon = \frac{\ell_1 - \ell_2}{2} > 0$ there exists $N_1 \in \mathbb{N}$ such that for all $n \geq N_1$, $|x_n - \ell_1| < \varepsilon$ and $N_2 \in \mathbb{N}$ such that for every $n \geq N_2$, $|y_n - \ell_2| < \varepsilon$. Then, if $n \geq \max\{N_1, N_2\}$, $\frac{\ell_1 + \ell_2}{2} = \ell_1 - \varepsilon < x_n$ and $y_n < \varepsilon + \ell_2 = \frac{\ell_1 + \ell_2}{2}$, and hence $y_n < \frac{\ell_1 + \ell_2}{2} < x_n$ which is impossible if $n \geq n_0$. □

Since if $x_n < y_n$ implies that $x_n \leq y_n$ we obtain the following corollary.

Corollary 3: Comparison Theorem II

Let $(x_n)_{n=1}^\infty$ and $(y_n)_{n=1}^\infty$ be two convergent sequences. If there exist $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $x_n < y_n$ then $\lim_{n\to\infty} x_n \leq \lim_{n\to\infty} y_n$.

Remark 13

The inequality in the conclusion of the Comparison Theorem II cannot be strict. Indeed, simply consider $x_n = \frac{1}{n}$ and $y_n = \frac{2}{n}$.

Theorem 15: Squeeze Theorem

Let $(x_n)_{n=1}^\infty, (y_n)_{n=1}^\infty, (z_n)_{n=1}^\infty$ be sequences of real numbers such that

- $x_n \leq y_n \leq z_n$, for all $n \geq n_0$ for some $n_0 \in \mathbb{N}$,
- $(x_n)_{n=1}^\infty$ and $(z_n)_{n=1}^\infty$ are convergent and

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} z_n = \ell,$$

then, $(y_n)_{n=1}^\infty$ is convergent and $\lim_{n\to\infty} y_n = \ell$.

Remark 14

We cannot use the comparison theorem twice to prove the squeeze theorem, since we do not know a priori that the middle sequence converges.

Proof. Let $\varepsilon > 0$, then there exist $N_1, N_2 \in \mathbb{N}$ such that for all $n \geq \max\{N_1, N_2\}$, $|x_n - \ell| < \varepsilon$ and $|z_n - \ell| < \varepsilon$. Thus, if $n \geq \max\{N_1, N_2, n_0\}$, $\ell - \varepsilon < x_n \leq y_n \leq z_n < \ell + \varepsilon$, and $-\varepsilon < z_n - \ell < \varepsilon$, which shows that $(y_n)_{n=1}^{\infty}$ is convergent to ℓ. □

3.3 Extraction of subsequences

Definition 27: Subsequences

A subsequence of a sequence $(x_n)_{n=1}^{\infty}$ is a sequence $(y_n)_{n=1}^{\infty}$ such that for all $k \in \mathbb{N}$, $y_k = x_{n_k}$ for some natural numbers $n_1 < n_2 < \cdots < n_k < \cdots$.

The indices of a subsequence have the following useful property.

Lemma 5

Let $(n_k)_{k=1}^{\infty}$ be a strictly increasing sequence of natural numbers. Then for all $k \in \mathbb{N}$, $n_k \geq k$.

Hint. You use an induction. □

Proof. Assume that $(n_k)_{k=1}^{\infty}$ is a strictly increasing sequence of natural numbers. It is clear that $n_1 \geq 1$ and assume that $n_k \geq k$. Then, $n_{k+1} > n_k \geq k$ and $n_{k+1} \geq k + 1$ since n_{k+1} is a natural number. □

Proposition 25

If a sequence converges to ℓ then all its subsequences are also convergent and they all converge to the same limit ℓ.

Hint. You could use the previous lemma. □

Proof. Assume that $(x_n)_{n=1}^{\infty}$ converges to ℓ, and let $(x_{n_k})_{k=1}^{\infty}$ be a subsequence of $(x_n)_{n=1}^{\infty}$. Let $\varepsilon > 0$, then there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $|x_n - \ell| < \varepsilon$. If $k \geq N$, then by Lemma 5 $n_k \geq N$ and $|x_{n_k} - \ell| < \varepsilon$ which shows that $(x_{n_k})_{k=1}^{\infty}$ converges to ℓ. □

Remark 15

A sequence can have converging subsequences without being itself convergent. Think about an example.

Exercise 1. Give a new proof of the fact that the sequence $((-1)^n)_{n=1}^{\infty}$ is not convergent.

The proof of the following proposition is left as an exercise.

Proposition 26

If a sequence is bounded then all its subsequences are also bounded

3.3.1 Bolzano-Weierstrass Theorem

To prove Bolzano-Weirstrass Theorem we need a preliminary lemma.

Lemma 6: Monotone subsequence lemma

Every sequence of real numbers has a monotone subsequence.

The lemma can be proved using the notion of peak point

Definition 28: Peak point

Let $(x_n)_{n=1}^{\infty}$ be a sequence. A peak point of the sequence is a term x_p of the sequence such that for all $n \geq p$, $x_n < x_p$.

Hint. Consider the following three cases: the sequence has infinitely many peak points, or finitely many peak points, or no peak points. □

Proof of the monotone subsequence lemma. Assume first that $(x_n)_{n=1}^{\infty}$ has no peak points. Let $k_1 = 1$. Since x_{k_1} is not a peak point there exists $k_2 > k_1$ such that $x_{k_2} \geq x_{k_1}$. But x_{k_2} is not a peak point either and there exists $k_3 > k_2 > k_1$ such that $x_{k_3} \geq x_{k_2} \geq x_{k_1}$. If we continue this process indefinitely we can construct recursively a subsequence $(x_{k_n})_{n=1}^{\infty}$ that is increasing. If $(x_n)_{n=1}^{\infty}$ has finitely many peak points let x_p the largest of those peak points. Let $k_1 = p + 1$, then x_{k_1} is not a peak point and hence there exists $k_2 > k_1$ such that $x_{k_2} \geq x_{k_1}$. Since x_{k_2} is not a peak point either there exists $k_3 > k_2 > k_1$ such that $x_{k_3} \geq x_{k_2} \geq x_{k_1}$, and we can construct recursively a subsequence $(x_{k_n})_{n=1}^{\infty}$ that is increasing. Now, assume that a sequence $(x_n)_{n=1}^{\infty}$ has infinitely many peak points then there exist $p_1 < p_2 < \cdots < p_k < \cdots$ such that for all $m \leq n$, $x_{p_m} > x_{p_n}$ and the subsequence $(x_{p_k})_{k=1}^{\infty}$ is strictly decreasing. In all three cases, we were able to show the existence of a monotone subsequence. □

Theorem 16: Bolzano-Weierstrass Theorem

Every bounded sequence of real numbers has a convergent subsequence.

Hint. Use the monotone subsequence lemma and the Monotone Convergence Theorem. □

Proof. Assume that $(x_n)_{n=1}^{\infty}$ is bounded. Then by the monotone subsequence lemma there is a subsequence that is monotone. But this subsequence is also clearly bounded and by the Monotone Convergence Theorem the subsequence is convergent. □

3.3.2 Limit supremum and limit infimum

The concept of limit supremum and limit infimum (also called limit superior and limit inferior) is extremely useful and provides interesting information about sequences. The definition of these notions require some justification. In this section, we study these notions in the context of bounded sequences. We start with a crucial lemma.

Lemma 7

Let $(x_n)_{n=1}^{\infty}$ be a bounded sequence of real numbers. For all $n \in \mathbb{N}$, let $s_n := \sup\{x_k \colon k \geq n\}$. Then the sequence $(s_n)_{n=1}^{\infty}$ is convergent.

Hint. Use the Monotone Convergent Theorem. □

Proof. Let $n \in \mathbb{N}$. Since $\{x_k \colon k \geq n\} \supset \{x_k \colon k \geq n+1\}$, $s_n = \sup\{x_k \colon k \geq n\} \geq \sup\{x_k \colon k \geq n+1\} = s_{n+1}$, and $(s_n)_{n=1}^{\infty}$ is decreasing. Since $(x_n)_{n=1}^{\infty}$ is bounded, $(s_n)_{n=1}^{\infty}$ is also bounded. By the Monotone Convergence Theorem $(s_n)_{n=1}^{\infty}$ is convergent. □

Lemma 7 provides a sound justification that the following definition is meaningful.

Definition 29: Limit supremum

Let $(x_n)_{n=1}^{\infty}$ be a bounded sequence of real numbers. The number $\lim\limits_{n\to\infty}(\sup\limits_{k\geq n} x_k)$ is called the limit supremum of the sequence $(x_n)_{n=1}^{\infty}$ and denoted $\limsup_{n\to\infty} x_n$ (or $\overline{\lim}_{n\to\infty} x_n$).

Example 4. Find $\limsup_{n\to\infty} x_n$ if $x_n = (-1)^n$, for $n \in \mathbb{N}$.

Hint. Exploit the definitions! □

Solution. The sequence $(x_n)_{n=1}^{\infty}$ is clearly upper bounded by 1 and for all $n \in \mathbb{N}$, the non-empty set $X_n := \{x_k \colon k \geq n\}$ has a supremum denoted $s_n \leq 1$. Note that $s_n = 1$ otherwise it will contradict the fact that s_n is an upper bound for $(x_n)_{n=1}^{\infty}$. Therefore, $\lim_{n\to\infty}(\sup_{k\geq n} x_k) = \lim_{n\to\infty} s_n = 1$. □

Theorem 17

Let $(x_n)_{n=1}^{\infty}$ be a bounded sequence of real numbers. Then, there exists a subsequence of $(x_n)_{n=1}^{\infty}$ that converges to $\limsup_{n\to\infty} x_n$.

Hint. Construct the subsequence recursively using the approximation property for suprema and conclude with the Squeeze Theorem. □

Proof. For all $n \in \mathbb{N}$ denote $y_n = \sup\{x_k : k \geq n\}$. The sequence $(y_n)_{n=1}^{\infty}$ is decreasing (see the proof of Lemma 7) and $\lim_{n\to\infty} y_n = s$ by definition of the limit supremum. If we can find a subsequence $(x_{n_k})_{k=1}^{\infty}$ so that $s - \frac{1}{k} < x_{n_k} < s + \frac{1}{k}$. then the conclusion follows from the Squeeze Theorem. We will construct this subsequence recursively. To do this rigorously we will prove by induction that for all $k \in \mathbb{N}$ the following statement $P(k)$ is true: there exist $n_1 < n_2 < \cdots < n_k$ and $x_{n_1}, \ldots, x_{n_k}$ elements of the sequence such that $s - \frac{1}{k} < x_{n_k} < s + \frac{1}{k}$.

For $k = 1$, since $(y_n)_{n=1}^{\infty}$ converges to s there exists $N \in \mathbb{N}$ such that $s \leq y_N < s + 1$ (here we use that $y_n \geq s$ for all $n \in \mathbb{N}$). However $s - 1 < s \leq y_N$ and by the approximation property for suprema there exists x_{n_1} such that $s - 1 < x_{n_1} < y_N < s + 1$ (here we use that $y_N = \sup\{x_n : n \geq N\}$), and $P(1)$ is true.

Assume now that the statement $P(k)$ is true, then we have natural numbers $n_1 < n_2 < \cdots < n_k$ and real numbers $x_{n_1}, \ldots, x_{n_k}$ such that $s - \frac{1}{k} < x_{n_k} < s + \frac{1}{k}$. Since $(y_n)_{n=1}^{\infty}$ converges to s there exists $N \in \mathbb{N}$, that can be chosen such that $N > n_k$, and so that $s \leq y_N < s + \frac{1}{k+1}$. However $s - \frac{1}{k+1} < s \leq y_N$ and by the approximation property for suprema there exists $x_{n_{k+1}}$ for some $n_{k+1} > n_k$ such that $s - \frac{1}{k+1} < x_{n_{k+1}} < y_N < s + \frac{1}{k+1}$ (here we use that $y_N = \sup\{x_n : n \geq N\}$ and $N > n_k$), and $P(k+1)$ is true. We conclude by invoking the Principle of Mathematical Induction. □

Using similar arguments we can define the notion of limit infimum.

Definition 30: Limit infimum

Let $(x_n)_{n=1}^{\infty}$ be a bounded sequence of real numbers. The number $\lim\limits_{n\to\infty} \left(\inf\limits_{k\geq n} x_k\right)$ is called the limit infimum of the sequence $(x_n)_{n=1}^{\infty}$ and denoted $\liminf_{n\to\infty} x_n$ (or $\underline{\lim}_{n\to\infty} x_n$).

Example 5. Find $\liminf_{n\to\infty} x_n$ if $x_n = (-1)^n$, for $n \in \mathbb{N}$.

Hint. Exploit the definitions! □

Solution. The sequence $(x_n)_{n=1}^{\infty}$ is clearly lower bounded by -1 and for all $n \in \mathbb{N}$, the non-empty set $X_n := \{x_k : k \geq n\}$ has a infimum denoted $t_n \geqslant -1$. Note that $t_n = -1$ otherwise it will contradict the fact that t_n is a lower bound for $(x_n)_{n=1}^{\infty}$. Therefore, $\lim_{n\to\infty}(\inf_{k\geq n} x_k) = \lim_{n\to\infty} t_n = -1$. □

As for the limit supremum, the limit infimum is the limit of a convergent subsequence.

Theorem 18

Let $(x_n)_{n=1}^{\infty}$ be a bounded sequence of real numbers. Then, there exists a subsequence of $(x_n)_{n=1}^{\infty}$ that converges to $\liminf_{n\to\infty} x_n$.

Hint. Construct the subsequence recursively using the approximation property for infima and conclude with the Squeeze Theorem. □

Chapter 4

Introduction to Metric Topology

General Topology is a branch of Mathematics whose goal is to understand properties of spaces that are invariant under continuous transformations. The theory deals with purely set-theoretic concepts and is usually called point-set topology. In the metric space setting ($\mathbb{R}$ equipped with its absolute value will be our prototypical example of a metric space) the topology of the space can be conveniently studied using sequences of elements. This chapter can be seen as a light introduction to general topology in the metric space framework.

4.1 Completeness

A central notion in topology of metric spaces is the notion of completeness. We shall study it for the set of real numbers equipped with its absolute value. To notion of completeness relies on the notion of Cauchy sequences.

4.1.1 Cauchy sequences

Definition 31: Cauchy sequences

A Cauchy sequence of real numbers is a sequence of real numbers $(x_n)_{n=1}^{\infty}$ such that for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ so that for all $n \geq N$ and all $k \geq N$, $|x_n - x_k| < \varepsilon$.

The notion of Cauchy sequence is a weakening of the notion of converging sequence as shown in the next proposition.

Proposition 27

Every convergent sequence is a Cauchy sequence.

Hint. It follows from the Triangle Inequality and the definitions. □

Proof. Assume that a sequence $(x_n)_{n=1}^{\infty}$ converges to ℓ. Let $\varepsilon > 0$ then there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $|x_n - \ell| < \varepsilon$. Therefore, if $n, k \geq N$ $|x_n - x_k| \leq |x_n - \ell| + |x_k - \ell| < 2\varepsilon$, which shows that $(x_n)_{n=1}^{\infty}$ is Cauchy. □

So far when showing that a sequence was convergent we usually had to guess beforehand what would the potential limit be. For sequences of real numbers we will prove the remarkable fact that every Cauchy sequence is convergent. This is extremely useful since in order to check whether or not a sequence is convergent we only need to check whether the sequence is Cauchy, and this does not require guessing the eventual limit! We need a few more facts about Cauchy sequences before proving our main theorem.

Proposition 28

A Cauchy sequence is bounded.

Hint. Use an argument similar to the proof of boundedness of convergent sequence. □

Unlike arbitrary sequences, a Cauchy sequence cannot have converging subsequences without being itself convergent.

Proposition 29

If a Cauchy sequence $(x_n)_{n=1}^{\infty}$ has a convergent subsequence then $(x_n)_{n=1}^{\infty}$ is convergent.

Hint. It follows from the triangle inequality. □

Proof. Assume that $(x_{n_k})_{k=1}^{\infty}$ is a convergent subsequence of $(x_n)_{n=1}^{\infty}$ and denote ℓ its limit. Let $\varepsilon > 0$, then there exists $K \in \mathbb{N}$ such that for all $k \geq K$, $|x_{n_k} - \ell| < \varepsilon$. Since $(x_n)_{n=1}^{\infty}$ is Cauchy there exists $N \in \mathbb{N}$ such that for all $n, m \geq N$, $|x_n - x_m| < \varepsilon$. If $k \geq \max\{K, N\}$, then $n_k \geq N$ by Lemma 2 and $|x_k - \ell| \leq |x_k - x_{n_k}| + |x_{n_k} - \ell| < 2\varepsilon$. □

Theorem 19: Cauchy Completeness of $(\mathbb{R}, |\cdot|)$

Every Cauchy sequence of real numbers is convergent.

Hint. Use Proposition 28, Bolzano-Weierstrass Theorem, and Proposition 29. □

Proof. Every Cauchy sequence is bounded, and hence by Bolzano-Weierstrass Theorem it has a convergent subsequence, therefore the original sequence is convergent. □

Combining Proposition 27 and Theorem 19 one obtains the following corollary.

Corollary 4

Let $(x_n)_{n=1}^{\infty}$ be a sequence of real numbers. Then, $(x_n)_{n=1}^{\infty}$ is convergent if and only if $(x_n)_{n=1}^{\infty}$ is Cauchy.

4.2 Divergence to $\pm\infty$

Definition 32: Divergence to $+\infty$

A sequence of real numbers $(x_n)_{n=1}^{\infty}$ diverges to $+\infty$ if for all $M \in \mathbb{R}$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $x_n \geq M$.

Example 6. The sequence $(x_n)_{n=1}^{\infty}$ where $x_n = n$ diverges to $+\infty$.

Definition 33: Divergence to $-\infty$

A sequence of real numbers $(x_n)_{n=1}^{\infty}$ diverges to $-\infty$ if for all $m \in \mathbb{R}$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $x_n \leq m$.

Example 7. The sequence $(x_n)_{n=1}^{\infty}$ where $x_n = -n$ diverges to $-\infty$.

The following proposition can be proven using an argument similar to the proof of Proposition 25 and we leave the details to the reader.

Proposition 30

If a sequence $(x_n)_{n=1}^{\infty}$ diverges to $+\infty$ (resp. $-\infty$) then every subsequence of $(x_n)_{n=1}^{\infty}$ also diverges to $+\infty$ (resp. $-\infty$).

Analogues of the comparison theorem are valid in this context and we simply state those results and leave the proofs as exercises.

Proposition 31

Let $(x_n)_{n=1}^{\infty}$ and $(y_n)_{n=1}^{\infty}$ be sequences. If there exist $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $x_n \leq y_n$ and if $\lim_{n\to\infty} x_n = +\infty$, then $\lim_{n\to\infty} y_n = +\infty$.

Proposition 32

Let $(x_n)_{n=1}^{\infty}$ and $(y_n)_{n=1}^{\infty}$ be sequences. If there exist $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $x_n \leq y_n$ and if $\lim_{n\to\infty} y_n = -\infty$, then $\lim_{n\to\infty} x_n = -\infty$.

4.3 Sequential Heine-Borel theorem

The topological Heine-Borel theorem states that a subset of a Euclidean space is compact if and only if it is bounded and closed. The set of real numbers equipped with the absolute value is an example of an Euclidean space (of dimension 1) and the notion of boundedness is a property of the absolute value. Compactness is a topological notion, which can be defined in purely set-theoretic terms, and that expresses the fact of being "small". Closedness is also a topological notion, which can be defined in purely set-theoretic terms as well, and which has to do with a set containing its "boundary". In the presence of a "metric" (generated here by the absolute value) one can define analogous notions in terms of sequences. It can be shown that in the context of metric spaces the set-theoretic notions coincide with their sequential analogues. Our goal in this section is to prove the Sequential Heine-Borel Theorem for subsets of the real numbers. This theorem will be needed crucially in the sequel. We first define what is a sequentially closed set.

Definition 34: Sequential closedness

A non-empty set $X \subseteq \mathbb{R}$ is sequentially closed if the limit of every convergent sequence of elements of X belongs to X. i.e. if $(x_n)_{n=1}^{\infty}$ is a sequence in X such that $\lim_{n\to\infty} x_n = \ell$ then $\ell \in X$.

It is relatively easy to show that sequential closedness is implied the stronger notion of sequential compactness.

Definition 35: Sequential compactness

A non-empty set $X \subseteq \mathbb{R}$ is sequentially compact if every sequence in X has a subsequence that converges in X, i.e. for all sequence $(x_n)_{n=1}^{\infty}$ of elements in X there exists $(x_{n_k})_{k=1}^{\infty}$ so that $\lim_{k\to\infty} x_{n_k} \in X$.

Proposition 33: Sequential compactness implies sequential closedness

Let X be a non-empty subset of $\mathbb{R}$. If X is sequentially compact then X is sequentially closed.

Hint. Simply use the relevant definitions and Proposition 25. □

Proof. Assume that X is sequentially compact. Let $(x_n)_{n=1}^{\infty}$ be a sequence in X that is convergent to $\ell \in \mathbb{R}$ (not necessarily in X). By sequential compactness $(x_n)_{n=1}^{\infty}$ has a subsequence that is convergent in X. By Proposition 25 the limit of the subsequence is also ℓ and thus $\ell \in X$, and we can conclude that X is sequentially closed. □

We now show that sequential closedness and boundedness imply sequential compactness.

Proposition 34: Boundedness and sequential closedness imply sequential compactness

Let X be a non-empty subset of $\mathbb{R}$. If X is bounded and sequentially closed then X is sequentially compact.

Hint. Use Bolzano-Weierstrass Theorem and the relevant definitions. □

Proof. Assume that X is bounded and sequentially closed. Let $(x_n)_{n=1}^{\infty}$ be a sequence in X. Since X is bounded $(x_n)_{n=1}^{\infty}$ is also bounded and by Bolzano-Weierstrass Theorem $(x_n)_{n=1}^{\infty}$ has a subsequence that is convergent. The conclusion follows since by sequential closedness the limit is necessarily in X. □

To establish the link between sequential compactness and boundedness, we will need the following lemmas.

Lemma 8

Every non-empty subset $X \subset \mathbb{R}$ that is not bounded above admits a sequence $(x_n)_{n=1}^{\infty}$ diverging to $+\infty$.

Proof. Assume that $X \subset \mathbb{R}$ is not bounded above, then for every $n \in \mathbb{N}$ there exists $x_n \in X$ such that $x_n > n$. It follows from Proposition 33 that $(x_n)_{n=1}^{\infty}$ diverges to $+\infty$. □

If $X \subset \mathbb{R}$ is not bounded below we can argue similarly and appeal to Proposition 34 instead of Proposition 33.

Lemma 9

Every non-empty subset $X \subset \mathbb{R}$ that is not bounded below admits a sequence $(x_n)_{n=1}^{\infty}$ diverging to $-\infty$.

Proposition 35: Sequential compactness implies boundedness

Let X be a non-empty subset of $\mathbb{R}$. If X is sequentially compact then X is bounded.

Hint. Prove the contrapositive. □

Proof. Assume that X is not bounded. Then either X is not bounded above, and thus Lemma 8 tells us that there exists a sequence $(x_n)_{n=1}^{\infty}$ that diverges to $+\infty$. Therefore, every subsequence of $(x_n)_{n=1}^{\infty}$ is not convergent by Proposition 25, and X cannot be sequentially compact. Otherwise, X is not bounded below and we argue similarly. □

We now have all the ingredients to prove the Sequential Heine-Borel Theorem for subsets of real numbers.

Theorem 20: Sequential Heine-Borel Theorem

Let X be a non-empty subset of $\mathbb{R}$. Then, X is sequentially compact if and only if X is bounded and sequentially closed.

Proof. Assume that $X \subset \mathbb{R}$ is sequential compact. Then by Proposition 33 X is sequentially closed and by Proposition 35 X is bounded. Assume now that $X \subset \mathbb{R}$ is bounded and sequential closed, then by Proposition 34 X is sequentially compact. □

Recall that $[a, b] := \{x \in \mathbb{R} \colon a \leq x \leq b\}$. The following proposition will be used crucially in the sequel.

Proposition 36

Let $a \leq b$ two real numbers, then $[a, b]$ is sequentially compact.

Hint. Boundedness is clear. It remains to show sequential closedness. □

Chapter 5

Continuity

We recall the definitions of the various intervals.

Definition 36: Intervals

Let $a < b$ two real numbers. We define the following intervals:

- $(a, b) := \{x \in \mathbb{R} \colon a < x < b\}$
- $[a, b] := \{x \in \mathbb{R} \colon a \leq x \leq b\}$
- $(a, b] := \{x \in \mathbb{R} \colon a < x \leq b\}$
- $[a, b) := \{x \in \mathbb{R} \colon a \leq x < b\}$
- $(-\infty, b) := \{x \in \mathbb{R} \colon x < b\}$
- $(-\infty, b] := \{x \in \mathbb{R} \colon x \leq b\}$
- $(a, \infty) := \{x \in \mathbb{R} \colon a < x\}$
- $[a, \infty) := \{x \in \mathbb{R} \colon a \leq x\}$

5.1 Definition and basic properties

The notion of continuity of a function is based on the notion of limit. There are many variants of the notion of limit of a function. To define continuity of a function defined at an interior point of an interval of the form (a, b) we need to define what is a finite limit at a finite point.

Definition 37: Two-sided limit of a function

Let $x_0 \in (a, b)$ and $f \colon (a, b) \setminus \{x_0\} \to \mathbb{R}$. We say that f has a limit $\ell \in \mathbb{R}$ at x_0 if for every $\varepsilon > 0$ there exists $\delta := \delta(\varepsilon) > 0$ such that if $|x - x_0| < \delta$ and $x \neq x_0$, then $|f(x) - \ell| < \varepsilon$.

Remark 16

Using only logic symbols the definition of a two-sided limit is:

$$\forall \varepsilon > 0\, \exists \delta > 0\, \forall x\, [[x \in (a,b)] \wedge [x \neq x_0] \wedge [|x - x_0| < \delta]] \implies |f(x) - \ell| < \varepsilon$$

Proposition 37: Uniqueness of the limit at a point

Let $x_0 \in (a,b)$ and $f\colon (a,b) \setminus \{x_0\} \to \mathbb{R}$. If f has a limit at x_0 then this limit is unique.

Proof. Assume by contradiction that f has two limits $\ell_1 \neq \ell_2$ at x_0. Without loss of generality assume that $\ell_1 > \ell_2$. By definition, for $\varepsilon = \frac{\ell_1 - \ell_2}{2} > 0$ there exist $\delta_1 > 0$ such that if $|x - x_0| < \delta_1$ and $x \neq x_0$, then $|f(x) - \ell_1| < \varepsilon$, and $\delta_2 > 0$ so that if $|x - x_0| < \delta_2$ and $x \neq x_0$, then $|f(x) - \ell_2| < \varepsilon$. Then, $\ell_1 - \ell_2 = |\ell_1 - \ell_2| \leq |\ell_1 - f(x)| + |f(x) - \ell_2| < \varepsilon + \varepsilon = \ell_1 - \ell_2$ whenever $x \neq x_0$ and $|x - x_0| < \min\{\delta_1, \delta_2\}$; a contradiction. □

If f has a limit ℓ at x_0, we write $\lim_{x \to x_0} f(x) = \ell$.

Example 8. Show that $\lim_{x \to 2} x^2 = 4$.

The following proposition is very useful. It allows us to use all the results that we have proven for sequences to study limits of functions.

Proposition 38: Sequential characterization of limits

Let $\ell \in \mathbb{R}$, $x_0 \in (a,b)$, and $f\colon (a,b) \setminus \{x_0\} \to \mathbb{R}$. Then, $\lim_{x \to x_0} f(x) = \ell$ if and only if for every sequence $(z_n)_{n=1}^{\infty}$ of elements in $(a,b) \setminus \{x_0\}$ which converges to x_0 one has $\lim_{n \to \infty} f(z_n) = \ell$.

Proof. Assume that $\lim_{x \to x_0} f(x) = \ell$. Let $\varepsilon > 0$, then there exists $\delta > 0$ such that if $|x - x_0| < \delta$ and $x \neq x_0$, then $|f(x) - \ell| < \varepsilon$. Let $(z_n)_{n=1}^{\infty}$ be a sequence of elements in $(a,b) \setminus \{x_0\}$ which converges to x_0. Then, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $|z_n - x_0| < \delta$, and since $z_n \neq x_0$ by assumption one has that $|f(z_n) - \ell| < \varepsilon$. Therefore, $(f(z_n))_{n=1}^{\infty}$ converges to ℓ.

To prove the other implication we will prove the contrapositive. Assume that ℓ is not the limit of f at x_0, then there exists $\varepsilon_0 > 0$ such that for every $n \in \mathbb{N}$ there exists $z_n \in (a,b) \setminus \{x_0\}$ so that $|z_n - x_0| < \frac{1}{n}$ and $|f(z_n) - \ell| \geq \varepsilon_0$. By the Squeeze Theorem $(z_n)_{n=1}^{\infty}$ is a sequence of elements in $(a,b) \setminus \{x_0\}$ which converges to x_0 but $(f(z_n))_{n=1}^{\infty}$ does not converge to ℓ since $|f(z_n) - \ell| \geq \varepsilon_0 > 0$. □

Proposition 39: Limit arithmetic for functions

et $x_0 \in (a,b)$ and $f,g\colon (a,b)\setminus\{x_0\} \to \mathbb{R}$. Let $\lambda \in \mathbb{R}$. If f and g have limits at x_0 then,

1. $f+g$ has a limit at x_0 and $\lim_{x\to x_0}(f+g)(x) = \lim_{x\to x_0} f(x) + \lim_{x\to x_0} g(x)$,
2. $\lambda \cdot f$ has a limit and $\lim_{x\to x_0}(\lambda\cdot f)(x) = \lambda \lim_{x\to x_0} f(x)$
3. $f\cdot g$ has a limit at x_0 and $\lim_{x\to x_0}(f\cdot g)(x) = \lim_{x\to x_0} f(x) \cdot \lim_{x\to x_0} g(x)$.
4. If $f \neq 0$ on $(a,b)\setminus\{x_0\}$ and $\lim_{x\to x_0} f(x) \neq 0$ then $\frac{1}{f}$ is well defined on $(a,b)\setminus\{x_0\}$, has a limit at x_0, and
$$\lim_{x\to x_0}\left(\frac{1}{f}\right)(x) = \frac{1}{\lim\limits_{x\to x_0} f(x)}.$$
5. If $g \neq 0$ on $(a,b)\setminus\{x_0\}$ and $\lim_{x\to x_0} g(x) \neq 0$ then $\frac{f}{g}$ is well defined on $(a,b)\setminus\{x_0\}$, has a limit at x_0, and
$$\lim_{x\to x_0}\left(\frac{f}{g}\right)(x) = \frac{\lim\limits_{x\to x_0} g(x)}{\lim\limits_{x\to x_0} g(x)}.$$

Proof. It follows from the sequential characterization of limits and the analogue results for sequences.

1.

2.

3.

4. Let $(z_n)_{n=1}^{\infty}$ be a sequence such that for all $n \in \mathbb{N}$, $z_n \in (a,b)$, $z_n \neq x_0$ and $\lim_{n\to\infty} z_n = x_0$. By the sequential characterization of limits the sequence $(y_n)_{n=1}^{\infty} = (f(z_n))_{n=1}^{\infty}$ converges to $\lim_{x\to x_0} f(x) \neq 0$. By Proposition 18, it follows that $\lim_{n\to\infty} \frac{1}{y_n} = \frac{1}{\lim_{n\to\infty} y_n}$, i.e. $\lim_{n\to\infty} \frac{1}{f(z_n)} = \frac{1}{\lim_{n\to\infty} f(z_n)} = \frac{1}{\lim_{x\to x_0} f(x)}$. Since $(z_n)_{n=1}^{\infty}$ was fixed but arbitrary, by the sequential characterization of limits we can conclude that the function $\frac{1}{f}$ has a limit and $\lim_{x\to x_0} \frac{1}{f(x)} = \frac{1}{\lim\limits_{x\to x_0} f(x)}$.

5.

□

Theorem 21: Comparison Theorem for functions I

Let $x_0 \in (a,b)$ and $f,g\colon (a,b)\setminus\{x_0\} \to \mathbb{R}$.
If

1. there exists $0 < \delta_0 < \min\{b-x_0, x_0-a\}$ such that $f(x) \le g(x)$ for all $x \in (x_0-\delta_0, x_0+\delta_0)\setminus\{x_0\}$.
2. f and g have limits at x_0

then,

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

Proof. Let $\ell_1 = \lim_{x\to x_0} f(x)$ and $\ell_2 = \lim_{x\to x_0} g(x)$. Assume by contradiction that $\ell_1 > \ell_2$. Let $\varepsilon = \frac{\ell_1-\ell_2}{2} > 0$, then there exist $\delta_1 > 0$ such that for all $x \in (x_0-\delta_1, x_0+\delta_1)$, $|f(x)-\ell_1| < \varepsilon$ and $\delta_2 > 0$ such that for all $x \in (x_0-\delta_2, x_0+\delta_2)$, $|f(x)-\ell_2| < \varepsilon$. If $\delta = \min\{\delta_0,\delta_1,\delta_2\} > 0$, and if $x \in (x_0-\delta, x_0+\delta)$ then $\ell_1-\varepsilon < f(x)$ and $g(x) < \varepsilon+\ell_2$, and hence $g(x) < \varepsilon+\ell_2 = \frac{\ell_1+\ell_2}{2} = \ell_1-\varepsilon < f(x)$; a contradiction. $\square$

Remark 17

We could also have given a proof that follows from the sequential characterization of limits and the comparison theorem for sequences.

Corollary 5: Comparison Theorem for functions II

Let $x_0 \in (a,b)$ and $f,g\colon (a,b)\setminus\{x_0\} \to \mathbb{R}$.
If

1. $f(x) < g(x)$ for all $x \in (a,b)\setminus\{x_0\}$.
2. f and g have limits at x_0

then,

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

Remark 18

The inequality in the conclusion of the Comparison Theorem II cannot be strict. Indeed, simply consider $f(x) = \frac{1}{x}$ and $g(x) = \frac{2}{x}$.

Theorem 22: Squeeze Theorem for functions

Let $x_0 \in (a, b)$ and $f, g, h\colon (a, b) \setminus \{x_0\} \to \mathbb{R}$.
If

1. $f(x) \le g(x) \le h(x)$ for all $x \in (a, b) \setminus \{x_0\}$,
2. f and h have limits at x_0 and $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} h(x)$,

then,
g has a limit at x_0 and $\lim_{x\to x_0} g(x) = \lim_{x\to x_0} f(x) = \lim_{x\to x_0} h(x)$.

Proof. It follows from the sequential characterization of limits and the squeeze theorem for sequences. □

Definition 38: Local Continuity

Let $x_0 \in (a, b)$ and $f\colon (a, b) \to \mathbb{R}$. We say that f is continuous at x_0 if for every $\varepsilon > 0$ there exists $\delta := \delta(\varepsilon) > 0$ such that if $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \varepsilon$.

The following lemma follows easily form the definitions.

Lemma 10

Let $x_0 \in (a, b)$ and $f\colon (a, b) \to \mathbb{R}$. Then, f is continuous at x_0 if and only if f has a limit at x_0 and $\lim_{x\to x_0} f(x) = f(x_0)$.

An immediate consequence of Lemma 10 and the sequential characterization of limits is a sequential characterization of continuity at a point.

Proposition 40: Sequential characterization of local continuity

Let $x_0 \in (a, b)$ and $f\colon (a, b) \to \mathbb{R}$. Then, f is continuous at x_0 if and only if for every sequence $(z_n)_{n=1}^{\infty}$ which converges to x_0 one has $\lim_{n\to\infty} f(z_n) = f(x_0)$.

Proof. Assume that f is continuous at x_0. Then by Lemma 10, $\lim_{x\to x_0} f(x) = f(x_0)$. If $(z_n)_{n=1}^{\infty}$ is a sequence in (a, b) that converges to x_0, without loss of generality one can assume that $z_n \neq x_0$, and by the sequential characterization of limit $\lim_{n\to\infty} f(z_n) = f(x_0)$.

For the other implication, assume that for every sequence $(z_n)_{n=1}^{\infty}$ which converges to x_0 one has $\lim_{n\to\infty} f(z_n) = f(x_0)$. By sequential characterization of limit, f has a limit at x_0 and $\lim_{x\to x_0} f(x) = f(x_0)$ and it follows from Lemma 10 that f is continuous at x_0.

□

The following propositions can be proven using Proposition 40 and Lemma 10, and the proofs are left to the reader.

Proposition 41: Continuity and algebraic operations

Let $x_0 \in (a,b)$ and assume that $f, g\colon (a,b) \to \mathbb{R}$ are continuous at x_0. Let $\lambda \in \mathbb{R}$, then,

1. $f + g$ is continuous at x_0,
2. $\lambda \cdot f$ is continuous at x_0,
3. $f \cdot g$ is continuous at x_0.
4. If $f \neq 0$ on (a,b) then $\frac{1}{f}$ is well defined on (a,b) and continuous at x_0.
5. If If $g \neq 0$ on (a,b) then $\frac{f}{g}$ is well defined on (a,b) and continuous at x_0.

Proposition 42: Continuity and composition

1. Let $x_0 \in \mathbb{R}$ and assume that $f\colon (a,b)\backslash\{x_0\} \to (c,d)$ with $x_0 \in (a,b)$ and $g\colon (c,d) \to \mathbb{R}$. If f has a limit at x_0 and $\lim_{x\to x_0} f(x) := \ell \in (c,d)$ and if g is continuous at ℓ then $g \circ f$ has a limit at x_0 and $\lim_{x\to x_0} g \circ f(x) := g(\lim_{x\to x_0} f(x))$.
2. Let $x_0 \in \mathbb{R}$ and assume that $f\colon (a,b) \to (c,d)$ with $x_0 \in (a,b)$ and $g\colon (c,d) \to \mathbb{R}$. If f is continuous at x_0 and if g is continuous at $f(x_0)$ then $g \circ f$ is continuous at x_0.

5.2 The Intermediate Value Theorem

Continuity is a local property but we can easily define what it means for a function to be continuous globally, e.g. on an interval of the form (a,b).

Definition 39: Global Continuity on an open bounded interval

Let $f\colon (a,b) \to \mathbb{R}$. We say that f is continuous on (a,b) if f is continuous at every point $x_0 \in (a,b)$.

In the sequel we will also consider function on closed interval of the form $[a,b]$ and we need to define rigorously what it means to be continuous at the endpoint. For this purpose we need to define one-sided limits.

Definition 40: Left-sided limits

Let $f\colon (a,b) \to \mathbb{R}$. We say that f has a left-sided limit $\ell \in \mathbb{R}$ at b if for every $\varepsilon > 0$ there exists $\delta := \delta(\varepsilon) > 0$ such that if $b - \delta < x < b$, then $|f(x) - \ell| < \varepsilon$. In this case we write $\lim_{x\to a^-} f(x) = \ell$.

Definition 41: Right-sided limits

Let $f\colon (a,b) \to \mathbb{R}$. We say that f has a right-sided limit $\ell \in \mathbb{R}$ at a if for every $\varepsilon > 0$ there exists $\delta := \delta(\varepsilon) > 0$ such that if $a < x < a + \delta$, then $|f(x) - \ell| < \varepsilon$. In this case we write $\lim_{x\to a^+} f(x) = \ell$.

Lemma 11

Let $f\colon (a,b) \to \mathbb{R}$, $x_0 \in (a,b)$ and $\ell \in \mathbb{R}$. Then,

$\lim_{x\to x_0} f(x) = \ell$ if and only if $\lim_{x\to x_0^+} f(x) = \lim_{x\to x_0^-} f(x) = \ell$.

Proof. cf homework. □

Definition 42: Global Continuity on a closed bounded interval

Let $f\colon [a,b] \to \mathbb{R}$. We say that f is continuous on $[a,b]$ if f is continuous on (a,b), $\lim_{x\to a^+} f(x) = f(a)$ and $\lim_{x\to b^-} f(x) = f(b)$.

The Intermediate Value Theorem espresses the idea that the graph of a continuous function is a continuous curve without gaps, holes or jumps.

Theorem 23: Intermediate Value Theorem

Assume that $f\colon [a,b] \to \mathbb{R}$ is continuous such that $f(a) < f(b)$ (resp. $f(b) < f(a)$). Let $y_0 \in \mathbb{R}$ such that $f(a) < y_0 < f(b)$ (resp. $f(b) < y_0 < f(a)$), then there exists $x_0 \in (a,b)$ such that $f(x_0) = y_0$.

Hint: Use the Least Upper Bound Property. □

We will give two proofs in the case $f(a) < y_0 < f(b)$, the other case being completely similar. We first give a proof using sequences, the approximation property for suprema, the Squeeze Theorem for sequences, the comparison theorem for sequences, and the sequential characterization of continuity. Both proofs start with an argument using the Least Upper Bound Property and only the second part of the proofs differ.

Proof using sequences. Consider the set $X = \{x \in [a,b] \colon f(x) < y_0\}$. X is nonempty since $a \in X$ and upper bounded (actually bounded) by definition. By the Least Upper Bound Property $s = \sup(X)$ exists. First observe that $s \neq b$ and that $s \neq a$ as well. Indeed, let $\varepsilon = y_0 - f(a) > 0$, then by right-continuity of f at a there exists $\delta > 0$ such that whenever $a < x < a+\delta$ one has $|f(x) - f(a)| < \varepsilon$, and hence $f(x) < f(a) + \varepsilon = y_0$. This implies that $s > a$, otherwise assuming $s = a$, there would be a point, say $z = a + \frac{\delta}{2}$, such that $z \in X$ but $z > s$, a contradiction. Similarly, let $\varepsilon = f(b) - y_0 > 0$, then by left-continuity of f at b there exists $\delta > 0$ such that whenever $b - \delta < x < b$ one has $|f(x) - f(b)| < \varepsilon$, and hence $f(x) > f(b) - \varepsilon = y_0$. This implies that $s < b$, otherwise assuming $s = b$, there would be a point, say $z = b - \frac{\delta}{2}$, such that $z \in X$ but $z > s$, a contradiction. So $s \in (a, b)$, and it remains to prove that $f(s) = y_0$ and thus we can take $x_0 = s$.

By the approximation property of suprema there exists $(z_n)_{n=1}^{\infty}$ a sequence in X such that $\lim_{n\to\infty} z_n = s$. By by continuity of f at s and sequential characterization of continuity, $\lim_{n\to\infty} f(z_n) = f(s)$, but for all $n \in \mathbb{N}$, $f(z_n) < y_0$ and it follows from the comparison theorem for sequences that $f(s) \leq y_0$. Since $s < b$, for every $n \in \mathbb{N}$ there exists $z_n \in [a, b]$ such that $s < z_n < s + \frac{1}{n}$ (simply take $z_n =$). Since $z_n > s$, $z_n \notin X$ and for all $n \in \mathbb{N}$, $f(z_n) \geq y_0$. By the Squeeze Theorem $\lim_{n\to\infty} z_n = s$, and by continuity of f at s, sequential characterization of continuity, and the comparison theorem for sequences, $\lim_{n\to\infty} f(z_n) = f(s) \geq y_0$. Therefore, $f(s) = y_0$ and the proof is complete. □

We now give a proof that does not use sequences.

Proof not using sequences. Consider the set $X = \{x \in [a, b) : f(x) < y_0\}$. X is non-empty since $a \in X$ and upper bounded (actually bounded) by definition. By the Least Upper Bound Property $s = \sup(X)$ exists. First observe that $s \neq b$ by definition and that $s \neq b$ as well. Indeed, let $\varepsilon = y_0 - f(a) > 0$, then by right-continuity of f at a there exists $\delta > 0$ such that whenever $a < x < a + \delta$ one has $|f(x) - f(a)| < \varepsilon$, and hence $f(x) < f(a) + \varepsilon = y_0$. This implies that $s > a$, otherwise assuming $s = a$, there would be a point, say $z = a + \frac{\delta}{2}$, such that $z \in X$ but $z > s$, a contradiction. Similarly, let $\varepsilon = f(b) - y_0 > 0$, then by left-continuity of f at b there exists $\delta > 0$ such that whenever $b - \delta < x < b$ one has $|f(x) - f(b)| < \varepsilon$, and hence $f(x) > f(b) - \varepsilon = y_0$. This implies that $s < b$, otherwise assuming $s = b$, there would be a point, say $z = b - \frac{\delta}{2}$, such that $z \in X$ but $z > s$, a contradiction. So $s \in (a, b)$, and it remains to prove that $f(s) = y_0$ and thus we can take $x_0 = s$.

Assume by contradiction that $f(s) > y_0$. Let $\varepsilon = f(s) - y_0 > 0$. By continuity of f at s there is $\delta > 0$ such that if $|x - s| < \delta$ then $|f(x) - f(s)| < \varepsilon$. Thus, for all $x \in (s - \delta, s + \delta)$, one has $f(x) > f(s) - \varepsilon = f(s) - (f(s) - y_0) = y_0$. Since s is the supremum of X and $\delta > 0$, by the approximation property for suprema there exists $x_\delta \in X$ such that $s - \delta < x_\delta \leq s$, and hence there exists $x_\delta \in (s - \delta, s]$ such that $f(x_\delta) < y_0$ and $f(x_\delta) > y_0$; contradiction.
Assume by contradiction that $f(s) < y_0$. Let $\varepsilon = y_0 - f(s) > 0$. By continuity of f at s there is $\delta > 0$ such that if $|x - s| < \delta$ then $|f(x) - f(s)| < \varepsilon$. Thus, for all $x \in (s - \delta, s + \delta)$, one has $f(x) < f(s) + \varepsilon = f(s) - (y_0 - f(s)) = y_0$. But if $x \in (s, s + \delta)$ then $x > s$ and $f(x) < y_0$ which contradicts the fact that s is the supremum of X. We proved that $f(s) \leq y_0$ and $f(s) \geq y_0$ and thus $f(s) = y_0$. □

Remark 19

We could have given a proof of the Intermediate Value Theorem by considering the set $Y = \{x \in [a, b] : f(x) > y_0\}$ and using the greatest upper bound property.

Remark 20

The point x_0, whose existence is guarranted in the Intermediate Value Theorem, does not have to be unique.

The converse of the Intermediate Value Theorem is not true in general. Indeed, the function defined by $f(x) = \sin(\frac{1}{x})$ if $x \neq 0$ and $f(0) = 0$ is not continuous (prove it!) but satisfies the conclusion of the Intermediate Value Theorem. However for monotone function the converse holds.

Definition 43: Monotonicity for functions

A function $f\colon \mathbb{R} \to \mathbb{R}$ is monotone if it is either increasing (i.e. if $x \leq y$ then $f(x) \leq f(y)$) or decreasing (i.e. if $x \leq y$ then $f(x) \geq f(y)$).

Theorem 24: Converse of the Intermediate Value Theorem for monotone functions

Let $f\colon [a,b] \to \mathbb{R}$ be a monotone function such that $f(a) < f(b)$ (resp. $f(b) < f(b)$). If whenever $f(a) < y_0 < f(b)$ (resp. $f(b) < y_0 < f(a)$) there exists $x_0 \in (a,b)$ such that $f(x_0) = y_0$, then f is continuous on $[a,b]$.

5.3 The Extreme Value Theorem

Definition 44: Boundedness for functions

Let X be a non-empty subset of $\mathbb{R}$. A function $f\colon X \to \mathbb{R}$ is said to be bounded on X if and only if there exists $M \geq 0$ such that for all $x \in X$, $|f(x)| \leq M$.

Lemma 12

If a function $f\colon X \to \mathbb{R}$ is not bounded, then there exists a sequence $(z_n)_{n=1}^{\infty}$ in X such that $|f(z_n)| > n$ for all $n \geq 1$.

Proof. It follows simply by negating what it means for f to be bounded. If f is unbounded then for all $M > 0$ there exists $x_M \in X$ such that $|f(x_M)| > M$, and in particular for all $n \geq 1$ there exists $x_n \in X$ such that $|f(x_n)| > n$. □

Theorem 25: Extreme Value Theorem

Let $f\colon [a,b] \to \mathbb{R}$ be continuous. Then f is bounded and moreover there exist $x_i, x_s \in [a,b]$ such that $f(x_i) = \inf_{x\in[a,b]} f(x)$ and $f(x_s) = \sup_{x\in[a,b]} f(x)$.

Hint: Use the sequential characterization of continuity and sequential compactness to show that f is bounded. Then, use the Least Upper Bound Property, the Squeeze Theorem, and the approximation property for suprema and infima to show that the infimum and the supremum are attained. □

Proof. We first prove that f is bounded. Assume by contradiction that f is not bounded. Then there exists a sequence $(z_n)_{n=1}^{\infty}$ in $[a,b]$ such that $|f(z_n)| > n$ (why?). Since $[a,b]$ is sequentially compact, there is a subsequence $(y_n)_{n=1}^{\infty}$ of $(z_n)_{n=1}^{\infty}$ such that $\lim_{n\to\infty} y_n = \ell \in [a,b]$. By continuity of f, $\lim_{n\to\infty} f(y_n) = f(\ell)$ and $(f(y_n))_{n=1}^{\infty}$ is convergent; a contradiction.

Since f is bounded by the Greatest Lower Bound Property (resp. the Least Upper Bound Property) $m := \inf_{x\in[a,b]} f(x)$ (resp. $M := \sup_{x\in[a,b]} f(x)$) exists. It remains to show that these two extrema are attained. We start with the infimum.

By the approximation property for infima, for every $n \in \mathbb{N}$ there exists $z_n \in [a,b]$ such that $m \leq f(z_n) < m + \frac{1}{n}$. By the Squeeze Theorem, $\lim_{n\to\infty} f(z_n) = m$ and by sequential compactness of $[a,b]$, there is a subsequence $(y_n)_{n=1}^{\infty}$ of $(z_n)_{n=1}^{\infty}$ that converges to $\ell_m \in [a,b]$. By continuity of f, $m = \lim_{n\to\infty} f(z_n) = \lim_{n\to\infty} f(y_n) = f(\ell_m)$, and we conclude by setting $x_i = \ell_m$.

As for the supremum, by the approximation property for suprema, for every $n \in \mathbb{N}$ there exists $z_n \in [a,b]$ such that $M - \frac{1}{n} < f(z_n) \leq M$. By the Squeeze Theorem, $\lim_{n\to\infty} f(z_n) = M$ and by sequential compactness of $[a,b]$, there is a subsequence $(y_n)_{n=1}^{\infty}$ of $(z_n)_{n=1}^{\infty}$ that converges to $\ell_M \in [a,b]$. By continuity of f, $M = \lim_{n\to\infty} f(z_n) = \lim_{n\to\infty} f(y_n) = f(\ell_M)$, and we conclude by setting $x_s = \ell_M$.

□

5.4 Uniform Continuity

Some properties of sequences are preserved by taking the image of the sequence by a continuous function. For instance, if a sequence is bounded, say in $[a,b]$, then the sequence obtained by taking the images of the terms by a continuous function on $[a,b]$ is also bounded. However continuity is not strong enough to preserve certain properties of sequences. Here is an example. Consider the sequence $(\frac{1}{n})_{n=1}^{\infty}$ and the function $f(x) = \frac{1}{x}$ which is continuous on $(0,1)$, then $(\frac{1}{n})_{n=1}^{\infty}$ is a Cauchy sequence but $(f(\frac{1}{n}))_{n=1}^{\infty} = (n)_{n=1}^{\infty}$ is not Cauchy. We will study a very useful notion which is stronger than continuity.

Definition 45: Uniform Continuity

Let I be a subset of $\mathbb{R}$ and $f\colon I \to \mathbb{R}$. We say that f is uniformly continuous on I if for every $\varepsilon > 0$ there exists $\delta := \delta(\varepsilon) > 0$ such that if $x, y \in I$ and $|x - y| < \delta$, then $|f(x) - f(y)| < \varepsilon$.

The fact that uniform continuity is stronger than continuity follows clearly form the definitions.

Proposition 43: Uniform continuity implies continuity

Let $f\colon I \to \mathbb{R}$. If f is uniformly continuous on I then f is continuous on I.

Proof. Assume that f is uniformly continuous and let $x_0 \in I$ and $\varepsilon > 0$. Since f is uniformly continuous, there exists $\delta := \delta(\varepsilon) > 0$ such that if $x, y \in I$ and

$|x - y| < \delta$, then $|f(x) - f(y)| < \varepsilon$, and by taking $y = x_0$, one has that for all $x \in I$ if $|x - x_0| < \delta$, then $|f(x) - f(x_0)| < \varepsilon$ and the proof is complete. □

Remark 1. Observe that to show that a function is not uniformly continuous on I it is sufficient to show that there exists $\varepsilon > 0$ such that for all $n \geq 1$ there exist sequence $(x_n)_{n=1}^{\infty}$ and $(y_n)_{n=1}^{\infty}$ in I such that $\lim_{n\to\infty} |x_n - y_n| = 0$ and for all $n \geq 1$, $|f(x_n) - f(y_n)| \geq \varepsilon$.

Example 9. 1. Show that $x \mapsto \frac{1}{x}$ is uniformly continuous on $(1, 2)$.

2. Show that $x \mapsto \frac{1}{x}$ is not uniformly continuous on $(0, 1)$.

The next proposition says that Cauchy sequences are preserved by uniformly continuous functions.

Proposition 44: Preservation of Cauchy sequences by uniformly continuous functions

Let $f\colon (a, b) \to \mathbb{R}$ and $(x_n)_{n=1}^{\infty}$ a Cauchy sequence of element in (a, b). Then, $(f(x_n))_{n=1}^{\infty}$ is a Cauchy sequence.

Proof. Let $(x_n)_{n=1}^{\infty}$ be a Cauchy sequence of element in (a, b) and $\varepsilon > 0$. By uniform continuity of f, there exists $\delta > 0$ such that for all $x, y \in I$, if $|x-y| < \delta$ then $|f(x) - f(y)| < \varepsilon$. Since $(x_n)_{n=1}^{\infty}$ is Cauchy there exists $N \in \mathbb{N}$ such that for all $n, m \geq N$, $|x_n - x_m| < \delta$ and thus $|f(x_n) - f(x_m)| < \varepsilon$. □

A continuous function on a sequentially compact interval $[a, b]$ has a remarkable property since it is automatically uniformly continuous.

Theorem 26: Continuity on compact intervals implies uniform continuity

Let f be a real-valued continuous function on a closed interval $[a, b]$. Then, f is uniformly continuous on $[a, b]$.

Proof. Assume by contradiction that f is continuous but not uniformly continuous on $[a, b]$. Then, there exists $\varepsilon_0 > 0$ such that for all $n \in \mathbb{N}$, there is $x_n, y_n \in [a, b]$ with $|x_n - y_n| < \frac{1}{n}$ and $|f(x_n) - f(y_n)| \geq \varepsilon_0$. By sequential compactness of $[a, b]$ there is a subsequence $(x_{n_k})_{k=1}^{\infty}$ of $(x_n)_{n=1}^{\infty}$ that converges to some $\ell_1 \in [a, b]$. Consider the subsequence $(y_{n_k})_{k=1}^{\infty}$ of $(y_n)_{n=1}^{\infty}$, then by sequential compactness again, the sequence $(y_{n_k})_{k=1}^{\infty}$ has a subsequence $(y_{m_k})_{k=1}^{\infty}$ that converges to some $\ell_2 \in [a, b]$. Note that the subsequence $(x_{m_k})_{k=1}^{\infty}$ of $(x_{n_k})_{k=1}^{\infty}$ still converges to ℓ_1. By the Squeeze Theorem the sequence $(x_n - y_n)_{n=1}^{\infty}$ converges to 0 and hence all its subsequences, in particular $\lim_{k\to\infty}(x_{m_k} - y_{m_k}) = 0$. Therefore, $\ell_1 - \ell_2 = 0$ and $\ell_1 = \ell_2$. By continuity of f the sequence $(f(x_{m_k}) - f(y_{m_k}))_{k=1}^{\infty}$ converges to $f(\ell_1) - f(\ell_2) = 0$. This is impossible since for all $n \in \mathbb{N}$, $|f(x_n) - f(y_n)| \geq \varepsilon_0 > 0$. □

Theorem 27: Continuous extendability

Let $f\colon (a,b) \to \mathbb{R}$. Then, f is uniformly continuous if and only if there is a continuous function $g\colon [a,b] \to \mathbb{R}$ which extends f, i.e. g satisfies $g(x) = f(x)$ for all $x \in (a,b)$.

Proof. Assume that there is a continuous function $g\colon [a,b] \to \mathbb{R}$ which extends f. Then g is uniformly continuous on $[a,b]$ and thus on (a,b) and f being the restriction of g on (a,b) it is also uniformly continuous on (a,b). Assume now that f is uniformly continuous on (a,b). Define $g\colon (a,b) \to \mathbb{R}$ by $g(x) = f(x)$. The function g is clearly continuous on (a,b). It remains to show that $\lim_{x\to a^+} f(x)$ and $\lim_{x\to b^-} f(x)$ exist and are finite and set $g(a) = \lim_{x\to a^+} f(x)$ and $g(b) = \lim_{x\to b^-} f(x)$ to complete the proof. Let $(z_n)_{n=1}^{\infty}$ be a sequence in (a,b) that is convergent to a. Then $(z_n)_{n=1}^{\infty}$ is Cauchy and by the previous exercise $(f(z_n))_{n=1}^{\infty}$ is also Cauchy. Since every Cauchy sequence of real numbers is convergent $(f(z_n))_{n=1}^{\infty}$ converges to some real number ℓ_z. At this point we still need to justify that the limit does not depend on the sequence. Let $(z_n)_{n=1}^{\infty}$ and $(y_n)_{n=1}^{\infty}$ be sequences in (a,b) that converge to a and such that $(f(z_n))_{n=1}^{\infty}$ converges to some real number ℓ_z and $(f(y_n))_{n=1}^{\infty}$ converges to some real number ℓ_y. Let $\varepsilon > 0$, and note that $|\ell_z - \ell_y| \leq |\ell_z - f(z_n)| + |f(z_n) - f(y_n)| + |f(y_n) - \ell_y|$. But there exists $N_1 \in \mathbb{N}$ such that for all $n \geq N_1$, $|\ell_z - f(z_n)| < \frac{\varepsilon}{3}$, $N_2 \in \mathbb{N}$ such that for all $n \geq N_2$, $|\ell_y - f(y_n)| < \frac{\varepsilon}{3}$. There is also $N_3 \in \mathbb{N}$ such that for all $n \geq N_3$, $|f(z_n) - f(y_n)| < \frac{\varepsilon}{3}$. Indeed, since f is uniformly continuous on (a,b) there exists $\delta > 0$ such that if $|x-y| < \delta$ then $|f(x) - f(y)| < \frac{\varepsilon}{3}$. Let $K_1 \geq 1$ such that for all $n \geq K_1$, $|x_n - a| < \frac{\delta}{2}$ and $K_2 \geq 1$ such that $|y_n - a| < \frac{\delta}{2}$ then for $n \geq \max\{K_1, K_2\}$ $|x_n - y_n| < \delta$ and thus $|f(x_n) - f(y_n)| < \frac{\varepsilon}{3}$. So if $N_3 = \max\{K_1, K_2\}$ then for all $n \geq N_3$, $|f(x_n) - f(y_n)| < \frac{\varepsilon}{3}$. Therefore, if $n \geq \max\{N_1, N_2, N_3\}$, $|\ell_z - \ell_y| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$. We just proved that for all $\varepsilon > 0$, $|\ell_z - \ell_y| < \varepsilon$ which implies that $\ell_z = \ell_y$. By sequential characterization of limits, $\lim_{x\to a^+} f(x)$ exists and is finite and we set $g(a) = \lim_{x\to a^+} f(x)$. The case of b is identical. □

5.5 Continuity of inverse functions

Proposition 45: Continuity and injectivity implies strict monotonicity

Let I be an interval and let $f\colon I \to \mathbb{R}$. If f is continuous and injective on I, then either f is strictly increasing or strictly decreasing on I.

Proof. Assume that f is continuous and injective on I. Assume by contradiction that f is neither strictly increasing nor strictly decreasing then there exist $x_1 < x_2 < x_3$ in I such that $f(x_1) \leq f(x_2)$ and $f(x_3) \leq f(x_2)$ (or $f(x_1) \geq f(x_2)$ and $f(x_3) \geq f(x_2)$) (why?). Since the proof for the latter case is similar to the proof of the former case we only treat the case where $f(x_1) \leq f(x_2)$ and $f(x_3) \leq f(x_2)$. Since f is injective $f(x_1) < f(x_2)$ and $f(x_3) < f(x_2)$. Let α such that $f(x_1) < \alpha < f(x_2)$ and $f(x_3) < \alpha < f(x_2)$ (why such an α exists?).

Since f is continuous on $[x_1, x_2]$, the IVT implies that there exists $c \in (x_1, x_2)$ such that $f(c) = \alpha$. Similarly, since f is continuous on $[x_2, x_3]$, the IVT implies that there exists $d \in (x_2, x_3)$ such that $f(d) = \alpha$ and thus $f(c) = f(d)$ for some $x_1 < c < x_2 < d < x_3$ which contradicts the injectivity of f. Therefore f is either strictly increasing or strictly decreasing. $\square$

Theorem 28: Continuity of inverse functions on compact intervals

Let $f : [a, b] \to \mathbb{R}$ be continuous and injective on $[a, b]$, then $f([a, b])$ is a closed interval and the inverse of f onto its image, $f^{-1} : f([a, b]) \to [a, b]$, is continuous.

Proof. Note f is either strictly increasing or strictly decreasing by the previous proposition. We will assume that f is strictly increasing as the proof that f is strictly decreasing will follow by similar arguments or by considering $g = -f$. Since f is strictly increasing, we obtain $f(a) < f(b)$. Since f is continuous, we obtain by the Intermediate Value Theorem that $f([a, b]) = [f(a), f(b)]$. We claim that f^{-1} is strictly increasing. To see this, suppose $y_1, y_2 \in f([a, b])$ are such that $y_1 < y_2$. Choose $x_1, x_2 \in [a, b]$ such that $f(x_1) = y_1$ and $f(x_2) = y_2$. Since $f(x_1) < f(x_2)$, it must be the case that $x_1 < x_2$ as f was strictly increasing. Hence $f^{-1}(y_1) = f^{-1}(f(x_1)) = x_1 < x_2 = f^{-1}(f(x_2)) = f^{-1}(y_2)$. Hence f^{-1} is strictly increasing. Therefore, $f^{-1} : [f(a), f(b)] \to [a, b]$ is a strictly increasing function such that $f^{-1}([f(a), f(b)]) = [a, b]$. Therefore f^{-1} is continuous by the converse of the Intermediate Value Theorem for monotone functions as f^{-1} satisfies the conclusions of the Intermediate Value Theorem. $\square$

Chapter 6

Differentiation

6.1 Definition and basic properties

Definition 46: Local differentiability

Let $f\colon (a,b) \to \mathbb{R}$ and $x_0 \in (a,b)$. We say that f is differentiable at x_0 if $\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}$ exists and is finite.

If f is differentiable at x_0 we write $f^{'}(x_0)$ the limit $\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}$, and if f is differentiable at every point $x_0 \in (a,b)$ the function $x \in (a,b) \mapsto f^{'}(x)$ is called the derivative of f and is denoted $f^{'}$.

Remark 21

It is often convenient to write the limit above as follows:

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

Example 10. Let $\alpha \in \mathbb{R}$. Where is the function $x \mapsto \alpha$ differentiable? Compute the derivative wherever the function is differentiable.

Example 11. Where is the function $x \mapsto |x|$ differentiable? Compute the derivative wherever the function is differentiable.

Example 12. Where is the function $x \mapsto \sqrt{x}$ differentiable? Compute the derivative wherever the function is differentiable.

We will now prove the simple but useful Carathéodory's Theorem.

Theorem 29: Carathéodory's Theorem

Let $x_0 \in (a,b)$ and $f\colon (a,b) \to \mathbb{R}$. Then, f is differentiable at x_0 if and only if there exists a function φ defined on (a,b) such that φ is continuous at x_0 and $f(x) = f(x_0) + \varphi(x)(x-x_0)$. Furthermore, $f^{'}(x_0) = \varphi(x_0)$.

Proof. First suppose that there is a function $\varphi\colon (a,b) \to \mathbb{R}$ such that φ is continuous at x_0 and $f(x) = f(x_0) + \varphi(x)(x-x_0)$. Note that for $x \neq x_0$,

$\frac{f(x)-f(x_0)}{x-x_0} = \frac{\varphi(x)(x-x_0)}{x-x_0} = \varphi(x)$. Therefore $\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}$ exists by continuity of φ and $f^{'}(x_0) = \lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0} = \lim_{x\to x_0}\varphi(x) = \varphi(x_0)$. Conversely, suppose that f is differentiable at x_0. Define $\varphi\colon (a,b) \to \mathbb{R}$ via $\varphi(x) = \begin{cases} f^{'}(x_0) \text{ if } x = x_0 \\ \frac{f(x)-f(x_0)}{x-x_0} \text{ if } x \neq x_0, \end{cases}$ for all $x \in (a,b)$. Clearly $f(x) = f(x_0) + \varphi(x)(x-x_0)$ for all $x \in (a,b)$. Furthermore, since $\lim_{x\to x_0}\varphi(x) = \lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0} = f^{'}(x_0) = \varphi(x_0)$, φ is continuous at x_0 as desired. □

Proposition 46: Differentiability implies continuity

Let $x_0 \in (a,b)$ and $f\colon (a,b) \to \mathbb{R}$. If f is differentiable at x_0 then f is continuous at x_0.

Proof. Suppose that f is differentiable at x_0. By Carathéodory's Theorem, there exists a function φ defined on (a,b) such that φ is continuous at x_0 and $f(x) = f(x_0)+\varphi(x)(x-x_0)$. Therefore, $\lim_{x\to x_0} f(x) = f(x_0)+\varphi(x_0)0 = f(x_0)$ by continuity of φ at x_0. and the conclusion follows. □

Definition 47: Local extrema

Let I be an interval and let $f\colon I \to \mathbb{R}$. We say that f has a local maximum at $c \in I$ if there exists an open interval $(a,b) \subset I$ such that $c \in (a,b)$ and $f(x) \leq f(c)$ for all $x \in (a,b)$, and a local minimum at $c \in I$ if $f(x) \geq f(c)$ for all $x \in (a,b)$. If f has a local maximum or a local minimum at c we simply say that f has a local extremum at c.

Proposition 47: Local maximum and derivatives

Let I be an interval and $f\colon I \to \mathbb{R}$.
If

1. f has a local maximum at $c \in I$
2. f is differentiable at c,

then $f^{'}(c) = 0$.

Proof. Let $c \in I$ be such that $f^{'}(c)$ exists and such that f has a local maximum at c. Since f has a local maximum at c, there exists an open interval $(a,b) \subset I$ such that $c \in (a,b)$ and $f(x) \leq f(c)$ for all $x \in (a,b)$. If $x \in (a,b)$ and $x > c$, then $\frac{f(x)-f(c)}{x-c} \leq 0$. Therefore as (a,b) is an open interval containing c, $\lim_{x\to c^+}\frac{f(x)-f(c)}{x-c} \leq 0$. Similarly, if $x \in (a,b)$ and $x < c$, then $\frac{f(x)-f(c)}{x-c} \geq 0$. Therefore as (a,b) is an open interval containing c, $\lim_{x\to c^-}\frac{f(x)-f(c)}{x-c} \geq 0$. Since $f^{'}(c)$ exists

$$\lim_{x\to c^-}\frac{f(x)-f(c)}{x-c} = \lim_{x\to c^+}\frac{f(x)-f(c)}{x-c} = f^{'}(c).$$

Hence the above inequalities show $0 \leq f^{'}(c) \leq 0$ and thus $f^{'}(c) = 0$. □

Proposition 48: Local minimun and derivatives

Let I be an interval and $f\colon I \to \mathbb{R}$.
If

1. f has a local minimum at $c \in I$
2. f is differentiable at c,

then $f^{'}(c) = 0$.

Proof. Let $c \in I$ be such that $f^{'}(c)$ exists and such that f has a local minimum at c. Since f has a local minimum at c, there exits an open interval $(a, b) \subset I$ such that $c \in (a, b)$ and $f(x) \geq f(c)$ for all $x \in (a, b)$. If $x \in (a, b)$ and $x > c$, then $\frac{f(x)-f(c)}{x-c} \geq 0$. Therefore as (a, b) is an open interval containing c, $\lim_{x\to c^+} \frac{f(x)-f(c)}{x-c} \geq 0$. Similarly, if $x \in (a, b)$ and $x < c$, then $\frac{f(x)-f(c)}{x-c} \leq 0$. Therefore as (a, b) is an open interval containing c, $\lim_{x\to c^-} \frac{f(x)-f(c)}{x-c} \leq 0$. Since $f^{'}(c)$ exists

$$\lim_{x\to c^-} \frac{f(x) - f(c)}{x - c} = \lim_{x\to c^+} \frac{f(x) - f(c)}{x - c} = f^{'}(c).$$

Hence the above inequalities show $0 \leq f^{'}(c) \leq 0$ and thus $f^{'}(c) = 0$. ☐

Combining the last two propositions we obtain Fermat's Theorem.

Corollary 6: Fermat's Theorem

Let I be an interval and $f\colon I \to \mathbb{R}$. If f has a local extremum at $c \in I$ and if f is differentiable at c, then $f^{'}(c) = 0$.

Definition 48: Global extrema

Let I be an interval and let $f\colon I \to \mathbb{R}$. It is said that f has a global maximum at $c \in I$ if $f(x) \leq f(c)$ for all $x \in I$, and a global minimum at $c \in I$ if $f(x) \geq f(c)$ for all $x \in I$. If f has a global maximum or a global minimum at c we simply say that f has a global extremum at c.

6.2 Rules of differentiation

In this section we present classical and fundamental rules of differentiation.

6.2.1 Basic Rules

Proposition 49: Addition and multiplication by a scalar

Let $\lambda \in \mathbb{R}$, $x_0 \in (a,b)$ and $f, g\colon (a,b) \to \mathbb{R}$ differentiable at x_0. Then,

1. $f+g$ is differentiable at x_0 and $(f+g)^{'}(x_0) = f^{'}(x_0) + g^{'}(x_0)$,
2. λf is differentiable at x_0 and $(\lambda f)^{'}(x_0) = \lambda f^{'}(x_0)$.

Proof. 1. Assume that f and g are differentiable at x_0. If $h \neq 0$, then

$$\begin{aligned}\frac{(f+g)(x_0+h)-(f+g)(x_0)}{h} &= \frac{f(x_0+h)+g(x_0+h)-(f(x_0)+g(x_0))}{h}\\ &= \frac{f(x_0+h)-f(x_0)+g(x_0+h)-g(x_0))}{h}\\ &= \frac{f(x_0+h)-f(x_0)}{h}+\frac{g(x_0+h)-g(x_0))}{h}.\end{aligned}$$

But $\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h} = f^{'}(x_0)$ and $\lim_{h\to 0}\frac{g(x_0+h)-g(x_0)}{h} = g^{'}(x_0)$ by assumption. Therefore, $\frac{(f+g)(x_0+h)-(f+g)(x_0)}{h}$ has a limit when h tends to 0 and $(f+g)$ is differentiable at x_0. Moreover,

$$\begin{aligned}(f+g)^{'}(x_0) &= \lim_{h\to 0}\frac{(f+g)(x_0+h)-(f+g)(x_0)}{h}\\ &= \lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h} + \lim_{h\to 0}\frac{g(x_0+h)-g(x_0)}{h}\\ &= f^{'}(x_0)+g^{'}(x_0).\end{aligned}$$

2. Assume that f is differentiable at x_0. If $h \neq 0$, then

$$\begin{aligned}\frac{(\lambda\cdot f)(x_0+h)-(\lambda\cdot f)(x_0)}{h} &= \frac{\lambda f(x_0+h)-\lambda f(x_0)}{h}\\ &= \lambda\frac{f(x_0+h)-f(x_0)}{h}.\end{aligned}$$

But $\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h} = f^{'}(x_0)$ by assumption. Therefore, $\frac{(\lambda\cdot f)(x_0+h)-(\lambda\cdot f)(x_0)}{h}$ has a limit when h tends to 0 and $(\lambda\cdot f)$ is differentiable at x_0. Moreover,

$$\begin{aligned}(\lambda\cdot f)^{'}(x_0) &= \lim_{h\to 0}\frac{(\lambda\cdot f)(x_0+h)-(\lambda\cdot f)(x_0)}{h}\\ &= \lim_{h\to 0}\lambda\frac{f(x_0+h)-f(x_0)}{h}\\ &= \lambda f^{'}(x_0).\end{aligned}$$

□

6.2.2 Product Rule

Proposition 50: Product rule

Let $x_0 \in (a,b)$ and $f,g\colon (a,b) \to \mathbb{R}$ differentiable at x_0. Then, $f \cdot g$ is differentiable at x_0 and $(f \cdot g)^{'}(x_0) = f^{'}(x_0)g(x_0) + f(x_0)g^{'}(x_0)$.

Proof. Assume that f and g are differentiable at x_0. If $h \neq 0$, then

$$\begin{aligned}\frac{(f\cdot g)(x_0+h)-(f\cdot g)(x_0)}{h} &= \frac{f(x_0+h)g(x_0+h)-(f(x_0)g(x_0))}{h}\\ &= \frac{(f(x_0+h)-f(x_0))g(x_0+h)+f(x_0)(g(x_0+h)-g(x_0))}{h}\\ &= \frac{f(x_0+h)-f(x_0)}{h}g(x_0+h)+f(x_0)\frac{g(x_0+h)-g(x_0)}{h}.\end{aligned}$$

But $\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h} = f^{'}(x_0)$, $\lim_{h\to 0}\frac{g(x_0+h)-g(x_0)}{h} = g^{'}(x_0)$ and $\lim_{h\to 0} g(x_0+h) = g(x_0)$ by assumption. Therefore, $\frac{(f\cdot g)(x_0+h)-(f\cdot g)(x_0)}{h}$ has a limit when h tends to 0 and $(f \cdot g)$ is differentiable at x_0. Moreover,

$$\begin{aligned}(f\cdot g)^{'}(x_0) &= \lim_{h\to 0}\frac{(f\cdot g)(x_0+h)-(f\cdot g)(x_0)}{h}\\ &= \lim_{h\to 0}[\frac{f(x_0+h)-f(x_0)}{h}g(x_0+h)] + f(x_0)\lim_{h\to 0}\frac{g(x_0+h)-g(x_0)}{h}\\ &= f^{'}(x_0)g(x_0)+f(x_0)g^{'}(x_0).\end{aligned}$$

□

6.2.3 Quotient Rule

Proposition 51: Quotient rule

Let $x_0 \in (a,b)$ and $f,g\colon (a,b) \to \mathbb{R}$ differentiable at x_0.

1. If $f(x_0) \neq 0$, then $\frac{1}{f}$ is well defined around x_0 and differentiable at x_0 and $\left(\frac{1}{f}\right)^{'}(x_0) = -\frac{f^{'}(x_0)}{(f(x_0))^2}$.
2. If $g(x_0) \neq 0$, then $\frac{f}{g}$ is well defined around x_0 and differentiable 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. 1. Assume that f is differentiable at x_0. If $h \neq 0$, since f is continuous and $f(x_0) \neq 0$ we can assume that if h is small enough that $f(x_0+h) \neq 0$ and then

$$\begin{aligned}\frac{\left(\frac{1}{f}\right)(x_0+h)-\left(\frac{1}{f}\right)(x_0)}{h} &= \frac{\frac{1}{f(x_0+h)}-\frac{1}{f(x_0)}}{h}\\ &= \frac{1}{h}\frac{f(x_0)-f(x_0+h)}{f(x_0+h)f(x_0)}\\ &= -\frac{f(x_0+h)-f(x_0)}{h}\frac{1}{f(x_0+h)f(x_0)}.\end{aligned}$$

But $\lim_{h\to 0}\frac{f(x_0+h)-f(x_0)}{h} = f^{'}(x_0)$ and $\lim_{h\to 0} f(x_0+h) = f(x_0)$ follow from the assumptions. Therefore, $\frac{\left(\frac{1}{f}\right)(x_0+h)-\left(\frac{1}{f}\right)(x_0)}{h}$ has a limit when h tends to 0 and $\frac{1}{f}$ is differentiable at x_0. Moreover,

$$\begin{aligned}\left(\frac{1}{f}\right)^{'}(x_0) &= \lim_{h\to 0}\frac{\left(\frac{1}{f}\right)(x_0+h)-\left(\frac{1}{f}\right)(x_0)}{h}\\ &= \lim_{h\to 0} -\frac{f(x_0+h)-f(x_0)}{h}\frac{1}{f(x_0+h)f(x_0)}\\ &= -\frac{f^{'}(x_0)}{\big(f(x_0)\big)^2}.\end{aligned}$$

2. Assume that f and g are differentiable at x_0 and that $g(x_0) \neq 0$. Remark that $\frac{f}{g} = f \cdot \frac{1}{g}$ and by the product rule and (1) $\frac{f}{g}$ is differentiable at x_0. Moreover,

$$\begin{aligned}\left(\frac{f}{g}\right)^{'}(x_0) &= \left(f\cdot\frac{1}{g}\right)^{'}(x_0)\\ &= f^{'}(x_0)\left(\frac{1}{g}\right)(x_0) + f(x_0)\left(\frac{1}{g}\right)^{'}(x_0)\\ &= f^{'}(x_0)\frac{1}{g(x_0)} - f(x_0)\frac{g^{'}(x_0)}{\big(g(x_0)\big)^2}\\ &= \frac{f^{'}(x_0)g(x_0)-f(x_0)g^{'}(x_0)}{\big(g(x_0)\big)^2}.\end{aligned}$$

□

6.2.4 Chain Rule

Proposition 52: Chain rule

Let $x_0 \in (a, b)$, $f\colon (a,b) \to \mathbb{R}$ such that $f(x_0) \in (c, d)$ and $g\colon (c,d) \to \mathbb{R}$. If f differentiable at x_0 and g is differentiable at $f(x_0)$ then, $g \circ f$ is differentiable at x_0 and $(g \circ f)^{'}(x_0) = g^{'}(f(x_0)) \cdot f^{'}(x_0)$.

Proof. Since f is differentiable at x_0 and g is differentiable at $f(x_0)$, by Caratheodory's Theorem there exist φ and ψ such that φ is continuous at x_0, $f(x) = f(x_0) + \varphi(x)(x - x_0)$ for all $x \in (a, b)$, $f^{'}(x_0) = \varphi(x_0)$ and ψ is continuous at $f(x_0)$, $g(x) = g(f(x_0)) + \psi(x)(x - f(x_0))$ for all $x \in (c, d)$, $g^{'}(f(x_0)) = \psi(f(x_0))$. Therefore, $g(f(x)) - g(f(x_0)) = \psi(f(x_0))(f(x) - f(x_0)) = \psi(f(x_0))(\varphi(x)(x - x_0))$, and $\frac{g(f(x))-g(f(x_0))}{x-x_0} = \psi(f(x_0))\varphi(x)$ if $x \neq x_0$. Since g is continuous at $f(x_0)$ and f is continuous at x_0 it follows that $g \circ f$ is differentiable at x_0 and $(g \circ f)^{'}(x_0) = \lim_{x \to x_0} \frac{g(f(x))-g(f(x_0))}{x-x_0} = \lim_{x \to x_0} \psi(f(x_0))\varphi(x) = g^{'}(f(x_0)) \cdot f^{'}(x_0)$ □

6.3 The Mean Value Theorem and its applications

6.3.1 Rolle's Theorem

Theorem 30: Rolle's Theorem

Let $f\colon [a, b] \to \mathbb{R}$. If,

(i) f is continuous on $[a, b]$,

(ii) f is differentiable on (a, b),

(iii) $f(a) = f(b)$,

then there exists $c \in (a, b)$ such that $f^{'}(c) = 0$.

Proof. By the EVT, f has a finite maximum M and a finite miminum m on $[a, b]$. The proof will be divided into three cases.

Case 1 If $M = m$, then f is constant on (a, b) and $f^{'}(x) = 0$ for all $x \in (a, b)$.

Case 2 Suppose that $M \neq m$. Since $f(a) = f(b)$, f must assume one of the values M or m at some point $c \in (a, b)$.

Case 2.a Consider the case where $f(c) = M$. Since M is the maximum of f on $[a, b]$ we have $f(c + h) - f(c) \leq 0$ for all h which satisfy $c + h \in (a, b)$. In the case $h > 0$ this implies $f^{'}(c) = \lim_{h \to 0^+} \frac{f(c+h)-f(c)}{h} \leq 0$, and in the case $h < 0$ this implies $f^{'}(c) = \lim_{h \to 0^-} \frac{f(c+h)-f(c)}{h} \geq 0$. It follows that $f^{'}(c) = 0$.

Case 2.b Finally, consider the case where $f(c) = m$. Since m is the minimum of f on $[a, b]$ we have $f(c+h) - f(c) \geq 0$ for all h which satisfy $c+h \in (a, b)$. In the case $h > 0$ this implies $f^{'}(c) = \lim_{h\to 0^+} \frac{f(c+h)-f(c)}{h} \geq 0$, and in the case $h < 0$ this implies $f^{'}(c) = \lim_{h\to 0^-} \frac{f(c+h)-f(c)}{h} \leq 0$. It follows that $f^{'}(c) = 0$.

□

We shall use Rolle's Theorem to obtain two important and very useful results: the Mean Value Theorem and Cauchy's Mean Value Theorem.

6.3.2 The Mean Value Theorem

The Mean Value Theorem has a geometric meaning when we look at the slopes of tangents to the graph of the function.

Theorem 31: Mean Value Theorem

Let $f: [a, b] \to \mathbb{R}$. If,

(i) f is continuous on $[a, b]$,

(ii) f is differentiable on (a, b),

then there exists $c \in (a, b)$ such that $f^{'}(c) = \frac{f(b)-f(a)}{b-a}$.

Proof. Let $h(x) = f(x)(b-a) - x(f(b) - f(a))$. By (i) and (ii), h is continuous on $[a, b]$ and differentiable on (a, b), with $h^{'}(x) = f^{'}(x)(b-a) - (f(b) - f(a))$. Since $h(a) = f(a)b - af(b) = h(b)$ it follows from Rolle's Theorem that there exists $c \in (a, b)$ such that $h^{'}(c) = f^{'}(c)(b-a) - (f(b) - f(a)) = 0$, i.e. $f^{'}(c) = \frac{f(b)-f(a)}{b-a}$. □

6.3.3 Cauchy's Mean Value Theorem

We will need a strengthening of the Mean Value Theorem to prove L'Hôpital's rules.

Theorem 32: Cauchy's Mean Value Theorem

Let $f, g: [a, b] \to \mathbb{R}$. If,

(i) f and g are continuous on $[a, b]$,

(ii) f is differentiable on (a, b),

(iii) g is differentiable on (a, b) with $g^{'}(x) \neq 0$ for every $x \in (a, b)$,

then there exists $c \in (a, b)$ such that $\frac{f^{'}(c)}{g^{'}(c)} = \frac{f(b)-f(a)}{g(b)-g(a)}$.

Proof. Observe first that $g(b) \neq g(a)$. Otherwise, by Rolle's Theorem there would exist $c \in (a, b)$ such that $g'(c) = 0$; contradiction with (iii). Let $h(x) =$

$f(x)(g(b)-g(a))-g(x)(f(b)-f(a))$. By (i) and (ii), h is continuous on $[a,b]$ and differentiable on (a,b), with $h^{'}(x)=f^{'}(x)(g(b)-g(a))-g^{'}(x)(f(b)-f(a))$. Since $h(a)=f(a)g(b)-g(a)f(b)=h(b)$ it follows from Rolle's Theorem that there exists $c\in(a,b)$ such that $h^{'}(c)=f^{'}(c)(g(b)-g(a))-g^{'}(c)(f(b)-f(a))=0$, i.e. $\frac{f^{'}(c)}{g^{'}(c)}=\frac{f(b)-f(a)}{g(b)-g(a)}$. □

6.3.4 Applications

6.3.4.1 L'Hôpital's Rules

There are many variations of L'Hôpital's rule and we shall present a proof a one of them and simply state a few others.

Theorem 33: L'Hôpital Rule (two-sided finite limits)

Let $x_0\in(a,b)$ and $f,g\colon(a,b)\setminus\{x_0\}\to\mathbb{R}$. If,

(i) f is differentiable on $(a,b)\setminus\{x_0\}$,

(ii) g is differentiable on $(a,b)\setminus\{x_0\}$ with $g^{'}(x)\neq 0$ for every $x\in(a,b)\setminus\{x_0\}$,

(iii) $\lim\limits_{x\to x_0} f(x)=\lim\limits_{x\to x_0} g(x)=0$

(iv) there exists $\ell\in\mathbb{R}$ such that $\lim\limits_{x\to x_0}\dfrac{f^{'}(x)}{g^{'}(x)}=\ell$

then $\lim\limits_{x\to x_0}\dfrac{f(x)}{g(x)}=\ell$.

Hint. First, show using the Mean Value Theorem that there is at most one point in (x_0,b) where g vanishes. Then show that $\lim_{x\to x_0^+}\frac{f(x)}{g(x)}=\ell$ using Cauchy's Mean Value Theorem. Adjust the arguments to prove that $\lim_{x\to x_0^-}\frac{f(x)}{g(x)}=\ell$ and conclude. □

Proof. We prove that both right-sided and left-sided limits exist and are the same.

Left-sided limit First, we show using the Mean Value Theorem that there is at most one point in (a,x_0) where g vanishes. Indeed, if $x_1<x_2$ are in (a,x_0) then g is continuous on $[x_1,x_2]$ and differentiable on (x_1,x_2). Hence, by the MVT there exists $c\in(x_1,x_2)$ such that $g^{'}(c)=\frac{g(x_2)-g(x_1)}{x_2-x_1}$. By assumption, $g^{'}(c)\neq 0$ and thus $g(x_1)\neq g(x_2)$. As this hold for all $x_1<x_2$ in (a,x_0), g is injective on (a,x_0) and there is at most one point, say $\gamma_1\in(a,x_0)$, where g vanishes. Assume now that $\lim\limits_{x\to x_0}\dfrac{f^{'}(x)}{g^{'}(x)}=\ell$ and let $\varepsilon>0$. Then there exists $\delta_1>0$ such that for all x such that $x_0>x>x_0-\delta_1>\gamma_1>a$ we have $|\frac{f^{'}(x)}{g^{'}(x)}-\ell|<\varepsilon$. Fix x such that $x_0>x>x_0-\delta_1$ and let $x_0>y>x$. Since f and g are continuous on $[x,y]$, differentiable on (x,y) and $g^{'}(t)\neq 0$ for all $t\in(x,y)$, Cauchy's MVT implies that there exists $c \in (x, y)$ such that $\frac{f'(c)}{g'(c)} = \frac{f(y)-f(x)}{g(y)-g(x)}$. Hence, as $c \in (x, y) \subseteq (x_0 - \delta_1, x_0)$, we obtain that $|\frac{f'(c)}{g'(c)} - \ell| = |\frac{f(y)-f(x)}{g(y)-g(x)} - \ell| < \varepsilon$. Since the above inequality holds for any y such that $x_0 > y > x > x_0 - \delta_1$, by taking the limit when y tends to x_0 from the left, we obtain $\frac{f(x)}{g(x)} = \lim_{y \to x_0^+} \frac{f(y)-f(x)}{g(y)-g(x)}$, indeed $\lim_{x \to x_0^-} f(x) = \lim_{x \to x_0^-} g(x) = 0$ by assumption and since $x \in (\gamma_1, x_0)$ one has $g(x) \neq 0$. Therefore, if $x_0 - \delta_1 < x < x_0$ one has $|\frac{f(x)}{g(x)} - \ell| \leq \varepsilon$. Finally, since x was fixed but arbitrary in $(x_0 - \delta_1, x_0)$ we have proven that for all $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x \in (x_0 - \delta_1, x_0)$, $|\frac{f(x)}{g(x)} - \ell| \leq \varepsilon$ and hence $\lim_{x \to x_0^-} \frac{f(x)}{g(x)} = \ell$.

Right-sided limit First, we show using the Mean Value Theorem that there is at most one point in (x_0, b) where g vanishes. Indeed, if $x_1 < x_2$ are in (x_0, b) then g is continuous on $[x_1, x_2]$ and differentiable on (x_1, x_2). Hence, by the MVT there exists $c \in (x_1, x_2)$ such that $g'(c) = \frac{g(x_2)-g(x_1)}{x_2-x_1}$. By assumption, $g'(c) \neq 0$ and thus $g(x_1) \neq g(x_2)$. As this hold for all $x_1 < x_2$ in (x_0, b), g is injective on (x_0, b) and there is at most one point, say $\gamma_2 \in (x_0, b)$, where g vanishes. Assume now that $\lim\limits_{x \to x_0} \dfrac{f'(x)}{g'(x)} = \ell$ and let $\varepsilon > 0$. Then there exists $\delta_2 > 0$ such that for all $x_0 < x < x_0 + \delta_2 < \gamma_2 < b$ such that $|\frac{f'(x)}{g'(x)} - \ell| < \varepsilon$. Fix x such that $x_0 < x < x_0 + \delta_2$ and let $x_0 < y < x$. Since f and g are continuous on $[y, x]$, differentiable on (y, x) and $g'(t) \neq 0$ for all $t \in (y, x)$, Cauchy's MVT implies that there exists $c \in (y, x)$ such that $\frac{f'(c)}{g'(c)} = \frac{f(x)-f(y)}{g(x)-g(y)}$. Hence, as $c \in (y, x) \subseteq (x_0, x_0 + \delta_2)$, we obtain that $|\frac{f'(c)}{g'(c)} - \ell| = |\frac{f(x)-f(y)}{g(x)-g(y)} - \ell| < \varepsilon$. Since the above inequality holds for any y such that $x_0 < y < x < x_0 + \delta$, by taking the limit when y tends to x_0 from the right, we obtain $\frac{f(x)}{g(x)} = \lim_{y \to x_0^+} \frac{f(x)-f(y)}{g(x)-g(y)}$, indeed since $x \in (x_0, \gamma)$ $g(x) \neq 0$, and $\lim_{x \to x_0^+} f(x) = \lim_{x \to x_0^+} g(x) = 0$ by assumption. Therefore, if $x_0 < x < x_0 + \delta_2$ one has $|\frac{f(x)}{g(x)} - \ell| \leq \varepsilon$. Finally, since x was fixed but arbitrary in $(x_0, x_0 + \delta_2)$ we have proven that for all $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x \in (x_0, x_0 + \delta_2)$, $|\frac{f(x)}{g(x)} - \ell| \leq \varepsilon$ and hence $\lim_{x \to x_0^+} \frac{f(x)}{g(x)} = \ell$.

Since $\frac{f}{g}$ has a right-sided limit and a left-sided limit at x_0, and both limits are equal to ℓ it follows that $\frac{f}{g}$ has a limit at x_0 which is ℓ. □

With similar proofs we can show various versions of L'Hôpital' rule. We state without proofs the most useful ones. The next rule is for right-sided limit and if we write b^- instead of a^+ we get the left-sided version.

Theorem 34: L'Hôpital Rule (one-sided finite limits)

Let $f, g\colon (a,b) \to \mathbb{R}$. If,

(i) f is differentiable on (a,b),

(ii) g is differentiable on (a,b) with $g^{'}(x) \neq 0$ for every $x \in (a,b)$,

(iii) $\lim_{x\to a^+} f(x) = \lim_{x\to a+} g(x) = 0$

(iv) there exists $\ell \in \mathbb{R}$ such that $\lim_{x\to a^+} \frac{f^{'}(x)}{g^{'}(x)} = \ell$

then $\lim_{x\to a^+} \frac{f(x)}{g(x)} = \ell$.

The next version is about two-sided infinite limits.

Theorem 35: L'Hôpital Rule (two-sided infinite limits)

Let $x_0 \in (a,b)$ and $f, g\colon (a,b)\setminus\{x_0\} \to \mathbb{R}$. If,

(i) f is differentiable on $(a,b)\setminus\{x_0\}$,

(ii) g is differentiable on $(a,b)\setminus\{x_0\}$ with $g^{'}(x) \neq 0$ for every $x \in (a,b)\setminus\{x_0\}$,

(iii) $\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = \pm\infty$

(iv) there exists $\ell \in \mathbb{R}$ such that $\lim_{x\to x_0} \frac{f^{'}(x)}{g^{'}(x)} = \pm\infty$

then $\lim_{x\to x_0} \frac{f(x)}{g(x)} = \pm\infty$.

6.3.4.2 Taylor's Theorem

Theorem 36: Taylor's Theorem

Let $x_0 \in (a,b)$ and $f\colon (a,b) \to \mathbb{R}$. If f is $n+1$ times differentiable on (a,b) and if $x \in (a,b)\setminus\{x_0\}$, then there exists $c_x \in (a,b)\setminus\{x_0\}$ such that

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{f^{(n+1)}(c_x)}{(n+1)!}(x-x_0)^{n+1}.$$

Proof. Consider the functions $g(t) = f(x) - f(t) - \sum_{k=1}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k$ and $h(t) = g(t) - (\frac{x-t}{x-x_0})^{n+1} g(x_0)$ and apply Rolle's Theorem. □

6.4 Monotone Functions and Derivatives

6.4.1 Various tests

Proposition 53: Increasing function test

Let $f\colon [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). If $f^{'}(x) \geq 0$ for all $x \in (a,b)$, then f is increasing on $[a,b]$. If $f^{'}(x) > 0$ for all $x \in (a,b)$, then f is strictly increasing on $[a,b]$.

Proof. Let $x_1 < x_2$ in $[a,b]$ then f is continuous on $[x_1,x_2]$ and differentiable on (x_1,x_2). By the MVT there exists $c \in (x_1,x_2)$ such that $f^{'}(c) = \frac{f(x_2)-f(x_1)}{x_2-x_1}$. But $f^{'}(c) \geq 0$ and thus $f(x_2) - f(x_1) \geq 0$. Therefore for every $x_1 < x_2$ in $[a,b]$ one has $f(x_1) \leq f(x_2)$ and f is increasing. In the case where $f^{'}(c) > 0$, then for every $x_1 < x_2$ in $[a,b]$ one has $f(x_1) < f(x_2)$ and f is strictly increasing. □

Proposition 54: Decreasing function test

Let $f\colon [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). If $f^{'}(x) \leq 0$ for all $x \in (a,b)$, then f is decreasing on $[a,b]$. If $f^{'}(x) < 0$ for all $x \in (a,b)$, then f is strictly decreasing on $[a,b]$.

Proof. Let $x_1 < x_2$ in $[a,b]$ then f is continuous on $[x_1,x_2]$ and differentiable on (x_1,x_2). By the MVT there exists $c \in (x_1,x_2)$ such that $f^{'}(c) = \frac{f(x_2)-f(x_1)}{x_2-x_1}$. But $f^{'}(c) \leq 0$ and thus $f(x_2) - f(x_1) \leq 0$. Therefore for every $x_1 < x_2$ in $[a,b]$ one has $f(x_1) \geq f(x_2)$ and f is decreasing. In the case where $f^{'}(c) < 0$, then for every $x_1 < x_2$ in $[a,b]$ one has $f(x_1) > f(x_2)$ and f is strictly decreasing. □

Proposition 55: First Derivative Test (existence of local minima)

Let $f\colon (a,b) \to \mathbb{R}$ be continuous on (a,b). Suppose $c \in (a,b)$ has the property that there exists $\delta > 0$ such that

1. $f^{'}(x)$ exists and $f^{'}(x) \leq 0$ for all $x \in (c-\delta, c) \subseteq (a,b)$, and
2. $f^{'}(x)$ exists and $f^{'}(x) \geq 0$ for all $x \in (c, c+\delta) \subseteq (a,b)$.

Then f has a local minimum at c.

Proof. Assume that f satisfies the above assumptions and let $x \in (c, c+\delta)$. Since f is continuous on $[c,x]$ and differentiable on (c,x), the MVT implies that there exists $d \in (c,x)$ such that $f^{'}(d) = \frac{f(x)-f(c)}{x-c}$. Since $d \in (c, c+\delta)$, $f^{'}(d) \geq 0$. Hence the above equation implies $f(x) \geq f(c)$ for all $x \in (c, c+\delta)$. Similarly, let $x \in (c-\delta, c)$. Since f is continuous on $[x,c]$ and differentiable on (x,c), the MVT implies that there exists $d \in (x, c)$ such that $f^{'}(d) = \frac{f(c)-f(x)}{c-x}$. Since $d \in (c-\delta, c)$, $f^{'}(d) \leq 0$. Hence the above equation implies $f(x) \geq f(c)$ for all $x \in (c-\delta, c)$. Therefore, f has a local minimum at c. □

Proposition 56: First Derivative Test (existence of local maxima)

Let $f\colon (a, b) \to \mathbb{R}$ be continuous on (a, b). Suppose $c \in (a, b)$ has the property that there exists $\delta > 0$ such that

1. $f^{'}(x)$ exists and $f^{'}(x) \geq 0$ for all $x \in (c-\delta, c) \subseteq (a, b)$, and
2. $f^{'}(x)$ exists and $f^{'}(x) \leq 0$ for all $x \in (c, c+\delta) \subseteq (a, b)$.

Then f has a local maximum at c.

Proof. Assume that f satisfies the above assumptions and let $x \in (c, c+\delta)$. Since f is continuous on $[c, x]$ and differentiable on (c, x), the MVT implies that there exists $d \in (c, x)$ such that $f^{'}(d) = \frac{f(x)-f(c)}{x-c}$. Since $d \in (c, c+\delta)$, $f^{'}(d) \leq 0$. Hence the above equation implies $f(x) \leq f(c)$ for all $x \in (c, c+\delta)$. Similarly, let $x \in (c-\delta, c)$. Since f is continuous on $[x, c]$ and differentiable on (x, c), the MVT implies that there exists $d \in (x, c)$ such that $f^{'}(d) = \frac{f(c)-f(x)}{c-x}$. Since $d \in (c-\delta, c)$, $f^{'}(d) \geq 0$. Hence the above equation implies $f(x) \leq f(c)$ for all $x \in (c-\delta, c)$. Therefore, f has a local maximum at c. □

6.4.2 Differentiablity of inverse functions

In this section we study the differentiability of the inverse (whenever it exists) of a differentiable function.

Theorem 37: Inverse Function Theorem

Let $f\colon [a, b] \to \mathbb{R}$ be continuous and injective on $[a, b]$. Let $g\colon f([a, b]) \to [a, b]$ be the inverse of f onto its image. If $x_0 \in (a, b)$ and f is differentiable at x_0 with $f^{'}(x_0) \neq 0$, then g is differentiable at $f(x_0)$ and $g^{'}(f(x_0)) = \frac{1}{f^{'}(x_0)}$.

Hint: Use the sequential characterization limit and the fact the g is continuous. □

Chapter 7

Integration

7.1 Definition of the Riemann Integral

7.1.1 Riemann Sums

We define upper and lower Riemann sums which are based on the notion of partition.

Definition 49: Partition

A partition of a closed interval $[a,b]$ is a finite list of real numbers $(t_k)_{k=0}^n$, where $n \geq 1$, such that $a = t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = b$.

Example 13. The partition $(t_k)_{k=0}^n$ of $[a,b]$ where $t_k = a + k\frac{b-a}{n}$ is called the regular partition.

Definition 50: Lower Riemann sum

Let $P = (t_k)_{k=0}^n$ be a partition of $[a,b]$ and let $f\colon [a,b] \to \mathbb{R}$ be bounded. The lower Riemann sum of f associated to P, denoted $L(f,P)$, is

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

where, for all $k \in \{1,\ldots,n\}$, $m_k = \inf\{f(x)\colon x \in [t_{k-1},t_k]\}$.

Definition 51: Upper Riemann sum

Let $P = (t_k)_{k=0}^n$ be a partition of $[a,b]$ and let $f\colon [a,b] \to \mathbb{R}$ be bounded. The upper Riemann sum of f associated to P, denoted $L(f,P)$, is

$$U(f,P) = \sum_{k=1}^{n} M_k(t_k - t_{k-1})$$

where, for all $k \in \{1,\ldots,n\}$, $M_k = \sup\{f(x)\colon x \in [t_{k-1},t_k]\}$.

Definition 52

Let P and Q be partitions of $[a, b]$. It is said that Q is a refinement of P if $P \subseteq Q$

Lemma 13

Let P and Q be partitions of $[a, b]$ and let $f: [a, b] \to \mathbb{R}$ be bounded. If Q is a refinement of P, then $L(f, P) \leq L(f, Q) \leq U(f, Q) \leq U(f, P)$. Therefore,

$$\sup\{L(f, P): P \text{ a partition of } [a, b]\} \leq \inf\{U(f, P): P \text{ a partition of } [a, b]\}.$$

7.1.2 Riemann integrable functions

Definition 53

Let $f: [a, b] \to \mathbb{R}$ be bounded. We say that f is Riemann integrable on $[a, b]$ if

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

If f is Riemann integrable on $[a, b]$ the Riemann integral of f from a to b, denoted $\int_a^b f(x)dx$, is defined as

$$\begin{aligned} \int_a^b f(x)dx &= \sup\{L(f, P): P \text{ a partition of } [a, b]\} \\ &= \inf\{U(f, P): P \text{ a partition of } [a, b]\}. \end{aligned}$$

Example 14. Show that the function $f: [a, b] \to \mathbb{R}$ defined by $f(x) = 2x$ for all $x \in [a, b]$ is Riemann integrable and that $\int_a^b f(x)dx = 2(b - a)$.

The definition of the Riemann integral suggests that we need to look at all the partitions, and this might be possible for very elementary functions as in the example above but in general it is often intractable (try for the function $f(x) = x^2$ on $[0, 1]$). The following characterization of Riemann integrability says that if we fix a given precision we only need to exhibit a partition such that the upper Riemann sum and the lower Riemann sum for this specific partition are almost equal up to this precision.

Proposition 57

Let $f: [a, b] \to \mathbb{R}$ be bounded. Then f is Riemann integrable on $[a, b]$ if and only if for every $\varepsilon > 0$ there exists a partition P of $[a, b]$ such that

$$0 \leq U(f, P) - L(f, P) < \varepsilon.$$

To show that f is Riemann integrable it follows from Proposition 57, that it is sufficient to find a sequence of partition $(P_n)_{n=1}^{\infty}$ such that $\lim_{n\to\infty}(U(f,P_n) - L(f,P_n)) = 0$. This is in general a much easier task than looking at all the partitions (try again the example $f(x) = x^2$ on $[0,1]$ with the regular partitions).

Definition 54: Riemann sum

Let $P = (t_k)_{k=0}^n$ be a partition of $[a,b]$ and let $f\colon [a,b] \to \mathbb{R}$ be bounded. For a finite sequence $(x_k)_{k=1}^n$ such that $x_k \in [t_k - 1, t_k]$ we define the Riemann sum, denoted $R(f,P,(x_k)_{k=1}^n)$, by

$$R(f,P,(x_k)_{k=1}^n) = \sum_{k=1}^{n} f(x_k)(t_k - t_{k-1})$$

7.2 Basic properties of the Riemann integral

7.2.1 Riemann integrability criteria

Theorem 38

Let $f\colon [a,b] \to \mathbb{R}$. If f is bounded and increasing on $[a,b]$, then f is Riemann integrable on $[a,b]$.

Proof. Consider a partition, to be specified later, $P = (t_k)_{k=0}^N$ of $[a,b]$. Since f is increasing $\inf\{f(x) : t_{k-1} \leqslant x \leqslant t_k\} \geq f(t_{k-1})$ and $\sup\{f(x) : t_{k-1} \leqslant x \leqslant t_k\} \leq f(t_k)$. Since,

$$L(f,P) = \sum_{k=1}^{N} m_k(t_k - t_{k-1})$$

where, for all $k \in \{1,\dots,N\}$, $m_k = \inf\{f(x)\colon x \in [t_{k-1}, t_k]\}$, we have

$$L(f,P) \geq \sum_{k=1}^{N} f(t_{k-1})(t_k - t_{k-1}).$$

Similarly,

$$U(f,P) = \sum_{k=1}^{N} M_k(t_k - t_{k-1})$$

where, for all $k \in \{1,\dots,N\}$, $M_k = \sup\{f(x)\colon x \in [t_{k-1}, t_k]\}$, and

$$U(f,P) \leq \sum_{k=1}^{N} f(t_k)(t_k - t_{k-1}).$$

Therefore,

$$\begin{aligned}0 \leq U(f,P) - L(f,P) &\leq \sum_{k=1}^{N} f(t_k)(t_k - t_{k-1}) - \sum_{k=1}^{N} f(t_{k-1})(t_k - t_{k-1}) \\ &= \sum_{k=1}^{N} (f(t_k) - f(t_{k-1}))(t_k - t_{k-1})\end{aligned}$$

Let $\varepsilon > 0$ and let $N \geq 1$ such that $\frac{b-a}{N}(f(b) - f(a)) < \varepsilon$. If $P = (t_k)_{k=0}^N$ was chosen to be the regular partition of $[a,b]$ where $t_k = a + k\frac{b-a}{N}$, one has

$$\begin{aligned}0 \leq U(f,P) - L(f,P) &\leq \sum_{k=1}^{N} (f(t_k) - f(t_{k-1}))(\frac{b-a}{N}) \\ &= \frac{b-a}{N} \sum_{k=1}^{N} (f(t_k) - f(t_{k-1})) \\ &= \frac{b-a}{N}(f(b) - f(a)) < \varepsilon.\end{aligned}$$

By Proposition 57 we conclude that f is Riemann integrable on $[a,b]$. □

A similar proof gives the following theorem.

Theorem 39

Let $f\colon [a,b] \to \mathbb{R}$. If f is bounded and decreasing on $[a,b]$, then f is Riemann integrable on $[a,b]$.

Theorem 40

Let $f\colon [a,b] \to \mathbb{R}$. If f is continuous, then f is Riemann integrable on $[a,b]$.

Proof. Since f is continuous on $[a,b]$, it is uniformly continuous on $[a,b]$. Hence, there exists $\delta > 0$, that for all $x, y \in [a,b]$ with $|x-y| < \delta$, one has $|f(x)-f(y)| < \varepsilon$. By the Archimedean Property, there exists $N \in \mathbb{N}$, such that $(b-a)/N < \delta$, and consider the regular partition $P_{reg} = (t_k)_{k=0}^N$ of $[a,b]$ where $t_k = a + k\frac{b-a}{N}$. Let $1 \leqslant k \leqslant N$. By the Extreme Value Theorem, there exist $x_k, y_k \in [t_{k-1}, t_k]$, so that $f(x_k) = \inf\{f(x) : t_{k-1} \leqslant x \leqslant t_k\}$ and $f(y_k) = \sup\{f(x) : t_{k-1} \leqslant x \leqslant t_k\}$. As $x_k, y_k \in [t_{k-1}, t_k]$, we have $|x_k - y_k| \leq \frac{b-a}{N} < \delta$ and therefore $|f(x_k) - f(y_k)| < \varepsilon$. Since,

$$L(f, P_{reg}) = \sum_{k=1}^{N} m_k(t_k - t_{k-1})$$

where, for all $k \in \{1, \ldots, N\}$, $m_k = \inf\{f(x)\colon x \in [t_{k-1}, t_k]\}$, we have

$$L(f, P_{reg}) = \sum_{k=1}^{N} m_k \frac{b-a}{N} = \frac{b-a}{N} \sum_{k=1}^{N} f(x_k).$$

Similarly,

$$U(f, P_{reg}) = \sum_{k=1}^{N} M_k(t_k - t_{k-1})$$

where, for all $k \in \{1, \ldots, N\}$, $M_k = \sup\{f(x) : x \in [t_{k-1}, t_k]\}$, and

$$U(f, P_{reg}) = \sum_{k=1}^{N} M_k \frac{b-a}{N} = \frac{b-a}{N} \sum_{k=1}^{N} f(y_k).$$

Therefore,

$$\begin{aligned} 0 \leq U(f, P_{reg}) - L(f, P_{reg}) &= \frac{b-a}{N} \sum_{k=1}^{N} f(y_k) - \frac{b-a}{N} \sum_{k=1}^{N} f(x_k) \\ &= \frac{b-a}{N} \sum_{k=1}^{N} (f(y_k) - f(x_k)) \\ &\leq \frac{b-a}{N} \sum_{k=1}^{N} |f(y_k) - f(x_k)| \\ &< \frac{b-a}{N} \sum_{k=1}^{N} \varepsilon = (b-a)\varepsilon. \end{aligned}$$

By Proposition 57 we conclude that f is Riemann integrable on $[a, b]$. □

7.2.2 Algebraic and order properties

Proposition 58: Linearity

Let $f, g \colon [a, b] \to \mathbb{R}$ be Riemann integrable and $\lambda \in \mathbb{R}$. Then,

1. the function $(\lambda \cdot f)$ is Riemann integrable and

$$\int_a^b (\lambda \cdot f)(x)dx = \lambda \int_a^b f(x)dx.$$

2. the function $(f + g)$ is Riemann integrable and

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

Proposition 59: Chasles' relation

Let $f \colon [a, b] \to \mathbb{R}$ be Riemann integrable and $c \in [a, b]$. Then, f is Riemann integrable on $[a, c]$ and on $[c, b]$, and

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

Proposition 60: Preservation of the order relation

Let $f, g\colon [a,b] \to \mathbb{R}$ be Riemann integrable.

1. If $f \leq g$ on $[a,b]$ then $\int_a^b f(x)dx \leq \int_a^b g(x)dx$.
2. If there exist $m, M \in \mathbb{R}$ such that for all $x \in [a,b]$, $m \leq f(x) \leq M$ then $m(b-a) \leq \int_a^b f(x)dx \leq M(b-a)$.

Proposition 61: Triangle Inequality

Let $f\colon [a,b] \to \mathbb{R}$ be Riemann integrable. Then, the function $|f|$ is Riemann integrable on $[a,b]$ and

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

Proposition 62: Integrability of products

Let $f, g\colon [a,b] \to \mathbb{R}$ be Riemann integrable. Then, the function $(f \cdot g)$ is Riemann integrable on $[a,b]$.

7.2.3 Integration by parts

Theorem 41: Integration by part

Let f, g be real valued functions on $[a,b]$ such that:

1. f and g are differentiable on $[a,b]$,
2. f' and g' are Riemann integrable on $[a,b]$.

Then,

$$\int_a^b f'(x)g(x)dx = f(b)g(b) - f(a)g(a) - \int_a^b f(x)g'(x)dx. \tag{7.1}$$

7.3 The Fundamental Theorem of Calculus

We start with a discussion of the notion of antiderivative.

Definition 55: Antiderivative

Let $f\colon (a,b) \to \mathbb{R}$. A function $F\colon (a,b) \to \mathbb{R}$ is said to be an antiderivative of f on (a,b), if F is differentiable on (a,b) with $F^{'}(x) = f(x)$ for all $x \in (a,b)$.

Lemma 14

Let $f\colon (a,b) \to \mathbb{R}$. If f is differentiable on (a,b) with $f^{'}(x) = 0$ for all $x \in (a,b)$, then there exists $\alpha \in \mathbb{R}$ such that $f(x) = \alpha$ for all $x \in (a,b)$.

Proof. Assume that $f\colon (a,b) \to \mathbb{R}$ is differentiable on (a,b) with $f^{'}(x) = 0$ for all $x \in (a,b)$. Let $x_1 < x_2$ in (a,b). Since f is continuous on $[x_1, x_2]$ and differentiable on (x_1, x_2), by the MVT, there exists $x_0 \in (x_1, x_2)$ such that $\frac{f(x_2)-f(x_1)}{x_2-x_1} = f^{'}(x_0) = 0$ and thus $f(x_1) = f(x_2)$. Pick $c \in (a,b)$ and let $x \in (a,b)$, $x \neq c$ then either $x > c$ or $x < c$. In either case the above argument shows that $f(x) = f(c)$. If we set $\alpha = f(c)$ then it follows that for all $x \in (a,b)$, $f(x) = \alpha$. □

Proposition 63

Let $f, g\colon (a,b) \to \mathbb{R}$. If f and g are differentiable on (a,b) with $f^{'}(x) = g^{'}(x)$ for all $x \in (a,b)$, then there exists $C \in \mathbb{R}$ such that $f(x) = g(x)+C$ for all $x \in (a,b)$.

Proof. Assume that $f, g\colon (a,b) \to \mathbb{R}$ are differentiable on (a,b) with $f^{'}(x) = g^{'}(x)$ for all $x \in (a,b)$. Let $x_1 < x_2$ in (a,b). If $h = f - g$ then h is continuous on $[x_1, x_2]$ and differentiable on (x_1, x_2). Therefore, by the MVT, there exists $x_0 \in (x_1, x_2)$ such that $\frac{h(x_2)-h(x_1)}{x_2-x_1} = h^{'}(x_0) = 0$ and thus $h(x_1) = h(x_2)$. Pick $c \in (a,b)$ and let $x \in (a,b)$, $x \neq c$ then either $x > c$ or $x < c$. In either case the above argument shows that $h(x) = h(c)$. If we set $\alpha = h(c)$ then it follows that for all $x \in (a,b)$, $f(x) = g(x) + \alpha$. □

Thus Proposition 63 implies that if F is an antiderivative of f, then the function G defined via $G\colon x \mapsto F(x) + C$, for some fixed constant $C \in \mathbb{R}$, is also an antiderivative of f.

The First Fundamental Theorem of Calculus states that every continuous function admits an antiderivative (and thus infinitely many).

Theorem 42: First Fundamental Theorem of Calculus (existence of derivatives for continuous functions)

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

1. for all $x \in [a,b]$ $F(x) = \int_a^x f(t)dt$ is well-defined,
2. the function $F: [a,b] \to \mathbb{R}$ is Lipschitz on $[a,b]$, in particular F is uniformly continuous on $[a,b]$, and F is differentiable on (a,b) with $F^{'}(x) = f(x)$ for all $x \in (a,b)$.

Proof. Since f is continuous on $[a,b]$, f is continuous on every interval of the form $[a,x]$ where $a \leqslant x \leqslant b$ and thus Riemann integrable on $[a,x]$ and $\int_a^x f(t)dt$ is a finite real number. For any $x \in [a,b]$, let $F(x) = \int_a^x f(t)dt$, then F is a well-defined function on $[a,b]$. For any $x, y \in [a,b]$, say $x < y$ one has

$$\begin{aligned}
|F(x) - F(y)| &= \left| \int_a^x f(t)dt - \int_a^y f(t)dt \right| \\
&= \left| \int_x^y f(t)dt \right| \\
&\leqslant \int_x^y |f(t)|dt \\
&\leqslant \int_x^y \sup\{|f(z)| \colon z \in [a,b]\}dt \\
&\leqslant \sup\{|f(z)| \colon z \in [a,b]\}|x-y|.
\end{aligned}$$

Therefore, F is M-Lipschitz with $M = \sup\{|f(z)| \colon z \in [a,b]\}$.

Let $x_0 \in (a,b)$ and $\varepsilon > 0$. By continuity of f at x_0, there exists $\delta > 0$ such that for all $x \in (x_0 - \delta, x_0 + \delta)$ if $|x - x_0| < \delta$ then $|f(x) - f(x_0)| < \varepsilon$. Assume that $0 < h < \delta$ then

$$\begin{aligned}
\left| \frac{F(x_0+h) - F(x_0)}{h} - f(x_0) \right| &= \left| \frac{1}{h}\int_a^{x_0+h} f(t)dt - \frac{1}{h}\int_a^{x_0} f(t)dt - f(x_0) \right| \\
&= \left| \frac{1}{h}\int_{x_0}^{x_0+h} f(t)dt - f(x_0) \right| \\
&= \left| \frac{1}{h}\int_{x_0}^{x_0+h} (f(t) - f(x_0))dt \right| \\
&\leqslant \frac{1}{h}\int_{x_0}^{x_0+h} |f(t) - f(x_0)|dt \\
&< \frac{1}{h}\int_{x_0}^{x_0+h} \varepsilon dt = \frac{h}{h}\varepsilon = \varepsilon.
\end{aligned}$$

Therefore, $\lim_{x \to x_0^+} \frac{F(x_0+h) - F(x_0)}{h} = f(x_0)$.

If we assume now that $-\delta < h < 0$ then

$$\begin{aligned}\left|\frac{F(x_0+h)-F(x_0)}{h} - f(x_0)\right| &= \left|\frac{1}{h}\int_a^{x_0+h} f(t)dt - \frac{1}{h}\int_a^{x_0} f(t)dt - f(x_0)\right| \\ &= \left|\frac{1}{h}\int_{x_0+h}^{x_0} f(t)dt - f(x_0)\right| \\ &= \left|\frac{1}{h}\int_{x_0+h}^{x_0} (f(t)-f(x_0))dt\right| \\ &\leqslant \frac{1}{|h|}\int_{x_0+h}^{x_0} |f(t)-f(x_0)|dt \\ &< \frac{1}{|h|}\int_{x_0+h}^{x_0} \varepsilon dt = \frac{-h}{|h|}\varepsilon = \varepsilon.\end{aligned}$$

Therefore, $\lim_{x\to x_0^-} \frac{F(x_0+h)-F(x_0)}{h} = f(x_0)$. Finally, we can conclude that $\lim_{x\to x_0} \frac{F(x_0+h)-F(x_0)}{h} = f(x_0)$ and thus F is differentiable at x_0, and furthermore $F'(x_0) = f(x_0)$.

□

Remark 2. If $f\colon [a,b] \to \mathbb{R}$ is Riemann integrable on $[a,b]$, then for all $x \in [a,b]$ $F(x) = \int_a^x f(t)dt$ is well-defined as well but the conclusion in (2) might not be true anymore. It can be shown that Thomae's function f_T is Riemann integrable on $[a,b]$ and that for all $x \in [a,b]$, $F(x) = \int_a^x f_T(t)dt = 0$ (and F is Lipschitz, uniformly continuous, and differentiable on $[a,b]$). However it follows from Darboux's Theorem that Thomae's function does not admit an antiderivative.

The Second Fundamental Theorem of Calculus is a very convenient tool to estimate the integral of a continuous function in terms of its antiderivatives. We can thus avoid the cumbersome and often intractable use of Riemann sums to compute integrals.

Theorem 43: Second Fundamental Theorem of Calculus (computation of the integral by means of an antiderivative)

Let $f\colon [a,b] \to \mathbb{R}$ be Riemann integrable on $[a,b]$ and assume that f admits an antiderivative F on (a,b) such that F is continuous on $[a,b]$. Then, $\int_a^b f(x)dx = F(b) - F(a)$.

Proof. Let $\varepsilon > 0$, then there exists a partition $P = (t_k)_{k=0}^N$ of $[a,b]$ such that $L(f,P) \leqslant \int_a^b f(x)dx \leqslant U(f,P) < L(f,P) + \varepsilon$. Indeed, since f is Riemann integrable on $[a,b]$ one has

$$\begin{aligned}\int_a^b f(x)dx &= \sup\{L(f,P)\colon P \text{ a partition of } [a,b]\} \\ &= \inf\{U(f,P)\colon P \text{ a partition of } [a,b]\}.\end{aligned}$$

By the approximation property of infima and suprema and the definition of the integral there exist partitions P_L and P_U such that

$$\int_a^b f(x)dx - \frac{\varepsilon}{2} < L(f, P_L) \leqslant \int_a^b f(x)dx$$

and

$$\int_a^b f(x)dx \leqslant U(f, P_U) < \int_a^b f(x)dx + \frac{\varepsilon}{2}.$$

Let P be the refinement of P_L and P_U, then it follows from Lemma 13 that

$$\int_a^b f(x)dx - \frac{\varepsilon}{2} < L(f, P) \leqslant \int_a^b f(x)dx$$

and

$$\int_a^b f(x)dx \leqslant U(f, P) < \int_a^b f(x)dx + \frac{\varepsilon}{2}.$$

Therefore, $U(f, P) < L(f, P) + \varepsilon$. For any $k \in \{1, 2, \ldots, N\}$ by the MVT there exists $c_k \in (t_{k-1}, t_k)$ such that $F(t_k) - F(t_{k-1}) = F'(c_k)(t_k - t_{k-1}) = f(c_k)(t_k - t_{k-1})$. Summing over k we get $F(b) - F(a) = \sum_{k=1}^N f(c_k)(t_k - t_{k-1})$. Observe that $L(f, P) \leqslant \sum_{k=1}^N f(c_k)(t_k - t_{k-1}) \leqslant U(f, P)$, and hence $L(f, P) \leqslant F(b) - F(a) < L(f, P) + \varepsilon$. Finally, one obtains that $-\varepsilon < \int_a^b f(x)dx - (F(b) - F(a)) < -\varepsilon$ and thus $\left|\int_a^b f(x)dx - (F(b) - F(a))\right| < \varepsilon$, for $\varepsilon > 0$ fixed but arbitrary. Therefore, $\int_a^b f(x)dx = F(b) - F(a)$.

□