Math 4031 - Advanced Calculus I

Instructor: Dr. Cristopher HERMOSILLA
Louisiana State University - Spring 2016

Week 1: Basic elements of Logic

The set of real numbers is the base of the Calculus. In these notes we will study the real numbers from an axiomatic point of view, that is, using some prescribed rules, called *axioms*, we will deduce several properties of the real numbers. To do so, one essential tool we require is the *Mathematical logic*, which will allow us to understand how a proof of a theorem works. Next week we will review another important tool, *Set theory*.

1.1 Statements and Truth values

The basic object in logic is called **statement**, which is a verbal assertions characterized by the fact that it has a unique *truth value*, that is, either *true* (T) or *false* (F). Statements are usually denoted by the letters p, q or r with or without subscripts.

Example 1.1. *We can use p to denote the statement "Paris is in France" and q for the statement "London is in Italy". In this example, p is true and q is false.*

Since each statement has a unique truth value, we can associate one with another statement that has the opposite truth value. This is called the **negation** of the statement and it is denoted by $\overline{p}$. This statement can be read as *it is false that* ... or simply *not p*. The truth value of the negation of a statement is described by the following table

p	$\overline{p}$
T	F
F	T

Example 1.2. *The negation of "Paris is in France" can be written as "It is false that Paris is in France" and, since the original statement is true, this negation is false.*

As may seem clear, the negation of the negation of a proposition has the same truth value as the proposition itself. To formally express this idea, we need to define first what we mean by claiming that two statements are the same. We say that two statements, p and q, are **equivalent** if they have the same truth values. In this case we write

$$p \Longleftrightarrow q.$$

Note that the equivalence of two statements is a statement too, it can be either true or false. Furthermore, the truth table that defines this relation is given by

p	q	$p \iff q$
T	T	T
T	F	F
F	T	F
F	F	T

In particular, we can check the following equivalence

$$p \iff \overline{\overline{p}}. \tag{1.1}$$

Algebra of propositions

Simple statements can be used to create composite statements, which will be called **proposition**. The fundamental property of a composite statement is that its truth value is completely determined by the truth value of the statements that compose it and the connectors used to create it. As we will see later, theorems, lemmas and corollary are examples of proposition whose truth value is true.

Example 1.3. *Let us consider the composite statement "Paris is in France and London is in Italy". We understand that this statement is false, and it will continue being false as long as the false statement "London is in Italy" is part of the composite statement.*

In the preceding example we have used the verbal connector *and* to create a **conjunction** of two statements. Symbolically, the conjunction of two statements p and q is denoted by

$$p \wedge q$$

and its truth value is the true if and only if both statements are true as well. This property is summarized in the next table

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

On the other hand, if in the example above we would have used the verbal connector *or* instead of *and*, the truth value of the statement would have been true. Furthermore, the truth value of "Paris is in France or London is in Italy" is true and it is not going to change if we replace the statement "London is in Italy" by any other statement (true or false). A proposition made with the connector *or* is called a **disjunction** and is it denoted by

$$p \vee q.$$

The truth values of a disjunction are summarized in the following table

p	q	$p \vee q$
T	T	T
T	F	T
F	T	T
F	F	F

Remark 1.1. *Let us point out that in conjunctions and disjunctions the order in which the statements p and q appear is not relevant, that is,*

$$p \wedge q \Longleftrightarrow q \wedge p \qquad \text{and} \qquad p \vee q \Longleftrightarrow q \vee p.$$

We can see this fact directly from the truth tables. In this case, we say that these connectors are **commutative**.

It is not difficult to see from the last two truth tables that conjunctions and disjunctions are related by means of the negation operation (the negation of one provides a statement that looks like the other one). These relations are known as the **De Morgan's laws** and read as follows

$$\overline{p \vee q} \Longleftrightarrow \overline{p} \wedge \overline{q} \qquad \text{and} \qquad \overline{p \wedge q} \Longleftrightarrow \overline{p} \vee \overline{q}. \tag{1.2}$$

The proof of the first De Morgan's law follows from the next truth table (the other De Morgan's law can be proved in similar way and it is left as exercise for the reader).

p	q	$\overline{p \vee q}$	$\overline{p}$	$\overline{q}$	$\overline{p} \wedge \overline{q}$
T	T	F	F	F	F
T	F	F	F	T	F
F	T	F	T	F	F
F	F	T	T	T	T

For the purposes of this course, the most important propositions are the **conditionals** or **implications**, which are of the form *if p then q*. This kind of statements allows us to decide whether a deduction is correct or not. This is essentially the nature of any theorem in mathematics, assume that a statement p is true, and then deduce after a sequence of logical steps, that another statement q, is also true. Symbolically, the implication is denoted by

$$p \Longrightarrow q$$

and it is defined by the truth table

p	q	$p \Longrightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

As mentioned above, a theorem (or lemma or corollary), corresponds to the first line in the truth table. In other words, when the *hypothesis* p is satisfied (the statement p is true), an implication is true if and only if the *consequence* q is true as well; if the hypothesis is false, the implication is true regardless the truth value of the consequence q.

There is an alternative way to define a logical implication. If we look at its truth table, we can see that it has only one case when it can be false, same as the the disjunction. It turns out that the logical implication is equivalent to the following disjunction

$$\overline{p} \vee q. \tag{1.3}$$

To see this, it is enough to check its truth table, which is

p	q	$\overline{p}$	$\overline{p} \vee q$
T	T	F	T
T	F	F	F
F	T	T	T
F	F	T	T

Tautologies and proofs

The implication and the equivalence can be compared through the following proposition that is always true (we leave the proof of this as exercise for the reader)

$$(p \Longrightarrow q) \wedge (q \Longrightarrow p) \Longleftrightarrow (p \Longleftrightarrow q).$$

A proposition that is always true is called a **tautology**.

Tautologies provides a way to write a particular statement in several different ways. This is very useful when trying to prove a theorem. Indeed, proving directly that $p \Longrightarrow q$ is true can be very hard and we might need to explore other strategies to do it.

Proofs by contradiction

A proof by contradiction consists in assuming that the hypothesis p is true and that the negation of the consequence q is true as well, and then reach another consequence r that is false. The underlying idea behind this kind of proofs is to show that the statement

$$p \wedge \overline{q}$$

cannot be true. The logical explanation for this fact is based on the following tautology

$$\overline{p \Longrightarrow q} \Longleftrightarrow p \wedge \overline{q}. \tag{1.4}$$

This tautology can be proved with truth tables, but we are now going to show an alternative way to do it, using a sequence of equivalences that we have already presented. The proof of the tautology (1.4) is as follows:

$$\begin{aligned}
\overline{p \Longrightarrow q} &\Longleftrightarrow \overline{\overline{p} \vee q} && \text{(alternative definition of an implication (1.3))}\\
&\Longleftrightarrow \overline{\overline{p}} \wedge \overline{q} && \text{(De Morgan's law (1.2))}\\
&\Longleftrightarrow p \wedge \overline{q} && \text{(negation of a negation (1.1))}
\end{aligned}$$

Proof by contraposition

The contrapositive of an implication $p \Longrightarrow q$ is the following proposition

$$\overline{q} \Longrightarrow \overline{p}.$$

As we will see shortly, the truth values of the contrapositive and the implication are the same, and so they are equivalent as statements. Hence, a proof by contraposition consists in assuming that the negation of the consequence q is true and then prove that the negation of the hypothesis is also true.

To show that the implication is equivalent to its contrapositive we follow the next steps

$$\begin{aligned}
p \Longrightarrow q &\Longleftrightarrow \overline{p} \vee q && \text{(alternative definition of an implication (1.3))} \\
&\Longleftrightarrow q \vee \overline{p} && \text{(Commutativity of the connector } \textit{or}\text{, Remark 1.1)} \\
&\Longleftrightarrow \overline{\overline{q}} \vee \overline{p} && \text{(negation of a negation (1.1))} \\
&\Longleftrightarrow \overline{q} \Longrightarrow \overline{p} && \text{(alternative definition of an implication (1.3))}
\end{aligned}$$

1.2 Quantifiers

There are several situations in mathematics where the truth value of a statement depends on some variables. For example, the equation $x + 1 = 0$ is true if and only if $x = -1$, and in any other situation, the equation if false.

We define a **propositional function** $p(x)$ as an undetermined statement that assumes a truth value whenever the variable x is fixed. The variable x might be understood as a generic parameter that belong to some collection of options.

Example 1.4. *Let us consider the propositional function $p(x)$ given by "x is in France". The truth value of $p(x)$ depends on x and also on the collection where x is assumed to belong. Indeed, if we suppose that x is part of the collection of all national capitals, we have that $p(x)$ is true if and only if x is "Paris", otherwise, it is false. However, if assume that x belongs to the collection of all continental french cities, $p(x)$ is always true.*

The usual way in which propositional functions are turned into statements is by means of quantifiers. On the one hand, in the first case in our example, the propositional function is turned into a true statement if before the propositional function we write *there exists at least one national capitals x such that ...*. In mathematical terms, it is written as

$$\exists x,\ p(x).$$

The symbol $\exists$ is called the **existential quantifier** and it is use to say a propositional function $p(x)$ is true for at least one element x.

On the other hand, in the second case in our example, by writing *for each continental french cities x ...* we turn the propositional function into a true statement. Symbolically, we write this as

$$\forall x,\ p(x).$$

The symbol $\forall$ is called the **universal quantifier** and it is use to express that a propositional function $p(x)$ is true for any possible choice we can make for x.

Remark 1.2. *When proving a proposition that includes the universal quantifier, we need to be careful and prove it for a generic x. It is a common mistake to prove that a statement $p(x)$ is true only for some instances (or even for only one!). The universal quantifier makes reference to any possible x, and thus, it needs to be proved for any arbitrary case (not for a particular one).*

A statement written with the universal quantifier is false if we can find at least one element x_0 for which $p(x_0)$ is false. If such element exists, it is usually called a **counterexample**. Following this reasoning, we can find a way to compute the negation of the universal quantifier. Indeed, we have that $\forall x,\ p(x)$ is false, this means that for some x, $\overline{p(x)}$ must be true. In other words, we have

$$\overline{\forall x,\ p(x)} \Longleftrightarrow \exists x,\ \overline{p(x)} \tag{1.5}$$

In a similar way, we have

$$\overline{\exists x,\ p(x)} \Longleftrightarrow \forall x,\ \overline{p(x)}. \tag{1.6}$$

Notice that we can prove the equivalence (1.6), by taking the negation in (1.5) and replacing $p(x)$ by $\overline{p(x)}$ when appropriate.

We finish by reviewing a last quantifier that is a composition of the universal and existential quantifier. We introduce the **uniqueness quantifier**, denoted by $\exists!$ to indicate that there is one and only one x for which a propositional function $p(x)$ is true. In mathematical terms, we write

$$\exists! x,\ p(x)$$

for the statement that is equivalent to

$$(\exists x,\ p(x)) \ \wedge \ (\forall x,\ \forall y,\ [(p(x) \wedge p(y)) \Longrightarrow x = y]).$$

In mathematics, when we want to prove that a statement $\exists! x,\ p(x)$ is true, we first prove the existence and then the uniqueness. Moreover, note that the negation of a uniqueness statement consists in two parts, either there is no x such that $p(x)$ is true or there are more than one instances for which $p(x)$ is true. Formally, we have

$$\overline{\exists! x,\ p(x)} \Longleftrightarrow (\forall x,\ \overline{p(x)}) \ \vee \ (\exists x,\ \exists y,\ [(p(x) \wedge p(y)) \wedge x \neq y]).$$

This tautology can be proved using the De Morgan's law (we leave it as exercise for the reader).

1.3 Exercises

1. Show using truth tables that the *and* and *or* connectors are **associative** with respect to each other, that is, for any statements p, q and r we have

   (a) $p \wedge (q \wedge r) \Longleftrightarrow (p \wedge q) \wedge r$

   (b) $p \vee (q \vee r) \Longleftrightarrow (p \vee q) \vee r$

2. Show using truth tables that the *and* and *or* connectors are **distributive** with respect to each other, that is, for any statements p, q and r we have

   (a) $p \wedge (q \vee r) \iff (p \wedge q) \vee (p \wedge r)$

   (b) $p \vee (q \wedge r) \iff (p \vee q) \wedge (p \vee r)$

3. Show that $(p \iff q) \iff ([\overline{p} \wedge \overline{q}] \vee [p \wedge q])$ is a tautology.

4. Show without using truth tables that the next proposition is a tautology

$$[(p \implies \overline{q}) \wedge (\overline{r} \vee q) \wedge r] \implies \overline{p}.$$

5. Determine the truth value of the statements p, q, r and s by knowing that the following proposition is true:

$$[s \implies (\overline{r} \vee r)] \implies [\overline{p \implies q} \wedge s \wedge \overline{r}].$$

6. Prove that if $\exists x,\ p(x) \implies \forall x, p(x)$ is true, then $p(x)$ is either true or false, regardless the value of x.

7. Let $p(x)$ and $q(x)$ be two propositional functions. Prove that if

$$(\exists! x,\ p(x)) \wedge (\exists! x,\ q(x))$$

   is true, then the following proposition is also true

$$(\exists x,\ p(x) \wedge q(x)) \implies (\exists! x,\ p(x) \wedge q(x)).$$

Math 4031 - Advanced Calculus I

Instructor: Dr. Cristopher HERMOSILLA
Louisiana State University - Spring 2016

Week 2: Basic elements of Set theory

As we mentioned last week, we need to review some mathematical tools in order to provide a self-contained exposition for the rest of the course. We have already studied basic notions of Logic, we now turn our attention into *Set theory*. This theory will allow us to set up the notion of *set* and also the symbolic language we are going to use along the course.

2.1 Basic definitions

The intuition tells us that a **set** is a collection of elements of some kind. For example, the set

$$D = \{0,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9\}$$

is a collection of symbols that we call digits. The set D is completely characterized by its elements, and so, we can determine the truth value of the propositional function

$$p(x) \iff x \textit{ is an element of } D.$$

Actually, the set D, can also be though as the collection of number that makes $p(x)$ true. In general, if we denote by A the set to be defined, the statement that describe the elements of the set A is written as

$$x \in A.$$

This is read as x *belongs to* A, and if its truth value is true, we implicitly understand that x is an element of A. The negation of $x \in A$, that is, $\overline{x \in A}$ is read as x *doesn't belong to* A and is denoted by

$$x \notin A.$$

In practice, when we write $x \in A$ or $x \notin A$ we are assuming that the corresponding statement is true. Therefore, a set A is the collection of all the elements x such that the statement $x \in A$ is true, and we say that we know the set A if we can determine all the elements that make the statement $x \in A$ true.

Remark 2.1. *Note that the definition of set that we have adopted doesn't take into account the order of the elements nor if an element has been written more than once. This means that, for us, the following are equivalent descriptions of the same set:*

$$\{0,\ 1\}, \quad \{1,\ 0\} \quad \textit{or} \quad \{0,\ 1,\ 0\}.$$

The usual way in which sets are denoted is through capital letters such as A, B or C (with or without subscripts), and sometimes, to denote especial sets, we use more sophisticated letters such as $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ or $\mathbb{R}$.

Remark 2.2. *At first glance, we might think that any definable proposition such as*

$$x \in A \iff p(x)$$

is enough to define a set A. This is not correct and a classical counterexample is the so-called **Russell's paradox**. *In simple words, this paradox considers the collection of sets x that don't belong to the themselves, that is,*

$$x \in R \iff x \notin x.$$

If the collection R is assumed to be a set, this readily leads to

$$R \in R \iff R \notin R,$$

which contradicts the very definition of R. In order to avoid this kind of paradox, mathematicians needed to introduce some rules to operate with sets, which are called the **Zermelo-Fraenkel axioms***; the underlying idea is that sets must be constructed in some way from previously constructed sets. We shall not discuss the Zermelo-Fraenkel Theory in details, but will mention it sometimes along these notes. The curious reader is referred to [1] for a concise exposition on the subject, or to [2, Chapter 7] for a shorter presentation.*

2.1.1 Some important sets

In these notes we are going to accept that the real, natural and integer numbers are sets, which we denote by $\mathbb{R}$, $\mathbb{N}$ and $\mathbb{Z}$, respectively; a rigorous proof of this fact goes far beyond the scope of this course. We may assume that $0 \in \mathbb{N}$, and so we can construct the set of rational numbers, denoted by $\mathbb{Q}$, as

$$\mathbb{Q} = \{x \in \mathbb{R} \mid \exists n \in \mathbb{Z},\ \exists m \in \mathbb{Z},\ m \neq 0\ \wedge mx = n\}.$$

Remark 2.3. *Let us point out that we have defined the rational number using the following scheme*

$$x \in \mathbb{Q} \iff [x \in \mathbb{R}\ \wedge\ p(x)].$$

Note that this differs from the way how the Russell's paradox has been stated. The essential point here is that $\mathbb{Q}$ is constructed as a subcollection of the set of real numbers and the Russell's paradox is done as a subcollection of some "universal set". We will see shortly that the notion of "set of all sets" is not suitable for a consistent mathematical theory.

We define the **empty set** as the set that contains no element at all. This set is denoted by $\emptyset$ and it admits several alternative definitions, although it is existence is accepted within the Zermelo-Fraenkel axioms. I f a set has at least an element, we say that it is **nonempty**.

2.1.2 Inclusion and equality of sets

Together with the notion of set comes that one of subset, which intuitively is a collection of elements that belong to the initial set and that satisfy some further properties; see for example the discussion in Remark 2.3. Formally, we say that a set A is a subset of B, we write $A \subseteq B$, if and only if the next proposition is true

$$\forall x, \ (x \in A \Longrightarrow x \in B). \tag{2.1}$$

Example 2.1. *Let D be the set of digits, and let A be the set of even digits, that is, $\{0, 2, 4, 6, 8\}$. We see that whenever the statement $x \in A$ is true, we also have that $x \in D$ is true, and thus, the proposition (2.1) is also true (for D in place of B).*

In what follows, when we write $A \subseteq B$ we are implicitly assuming that the statement (2.1) is true. Moreover, with a slight abuse of notation, we may use $A \subseteq B$ equivalently as (2.1).

There are two situations to be considered. Either A is strictly included in B or, A and B have the same elements (they are equal). In the first case we write $A \subsetneq B$ to indicate that

$$A \subseteq B \ \wedge \ [\exists x, \ (x \in B \wedge x \notin A)].$$

We say in this case that A is a **proper subset** of B.

Remark 2.4. *We always have that $\emptyset$ is a proper subset of any nonempty set A. This fact comes from the very definition. Indeed, the statement*

$$x \in \emptyset \ \Longrightarrow x \in A$$

is always true, regardless A, because the statement $x \in \emptyset$ is always false.

On the second case, we write $A = B$ to indicate that A and B are **equal**, that is, that the following proposition is true

$$\forall x, \ [(x \in A \ \Longrightarrow x \in B) \ \wedge \ (x \in B \ \Longrightarrow x \in A)].$$

Furthermore, from the definition of the inclusion we have the following characterization of the equality of set

$$A = B \ \Longleftrightarrow [(A \subseteq B) \ \wedge (B \subseteq A)]. \tag{2.2}$$

Remark 2.5. *The characterization (2.2) will play a fundamental role in the upcoming discussion. Indeed, this shows that to prove that two sets are the same, we need to prove that each one of them is a subset of the other. This is very similar to the fact that to prove $p \Longleftrightarrow q$, we prove $p \Longrightarrow q$ and $q \Longrightarrow p$ separately.*

Power set

Given a set A, we call the **power set** of A, written as $\mathcal{P}(A)$, to the collection of all the subsets of A, including the empty set. In this notes we assume that the power set is actually a set (this is one of the Zermelo-Fraenkel axioms).

Example 2.2. *Let $A = \{0,\ 1,\ 2\}$, then its power set is*

$$\mathcal{P}(A) = \{\emptyset,\ \{0\},\ \{1\},\ \{2\},\ \{0,\ 1\},\ \{0,\ 2\},\ \{1,\ 2\},\ \{0,\ 1,\ 2\}\}.$$

In the preceding example, since we can list all the elements of A, it is not difficult to provide an explicit expression of $\mathcal{P}(A)$. In other situations this is not possible and we might need to find alternative ways to do it. This is the case of $\mathcal{P}(\mathbb{N})$, the power set of the natural numbers. It is not possible to write down each of the elements of $\mathcal{P}(\mathbb{N})$. For the moment, we just relay on the formal definition to describe it, but we will show one way to express it later on.

2.1.3 Cardinality

We say that a set A is **finite** if it has a finite number of different elements. This means that we can count the elements of A finishing at some point. Furthermore, we can associate A with a natural number, called the **cardinality** of A, that corresponds to the number of (different) elements of A. We denote the cardinality of a finite set A by $|A|$.

Note that in Example 2.2 the set A has 3 elements and $\mathcal{P}(A)$ has $2^3 = 8$. This is because, A has a single subset with 0 elements (the empty set), 3 subsets with 1 element, 3 subsets having 2 elements and a single subset having 3 elements (the same set). This remark is not a coincidence and it is a general fact.

Theorem 2.1. *If A is a finite set, then $|\mathcal{P}(A)| = 2^{|A|}$.*

Proof. Let $n \in \mathbb{N}$ be the natural number that represents $|A|$. Let us first answer the question of how many subsets of A having k elements we can find. We claim that the answer is

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{2.3}$$

Note that there is a single subset of A that contains $k = 0$ and $k = n$, respectively. These are the empty set and the whole set A. Clearly, the claim holds in these cases.

On the other hand, suppose that we can assign an order to the elements of any subset of A that has k elements. For the first position we have n options, but for the second one we only have $n-1$ options (one the $n-1$ elements remaining). We can continue the process and see that for the third position we only have $n-2$ options and so on. We end up having only $n-k+1$ option for the k-th position. If we count all the possible combinations, we get

$$n(n-1)(n-2)\ldots(n-k+1) = \frac{n!}{(n-k)!}$$

possible subsets of A considering the order we have described above. However, subsets don't take into account the order of their elements; see Remark

Finally, the number of all possible subsets of A is

$$\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{k} + \ldots + \binom{n}{n-1} + \binom{n}{n} = \sum_{k=0}^{n} \binom{n}{k} = (1+1)^n = 2^n,$$

where we have use the so-called binomial theorem. $\square$

If a set is not finite, we say that it is **infinite**. Let us point out that the natural, integer, rational and real numbers are all examples of infinite sets. However, the kind of infinite that each of them represents is slightly different. The natural, integer and rational are examples **countable** sets. The idea of countability refers to the fact that we can count the elements of $\mathbb{N}$, $\mathbb{Z}$ or $\mathbb{Q}$ even though we will never finish doing so. It might seem counterintuitive, but it can be proved that the "number" of elements of $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$, which is also called cardinality, is the same for all these sets. The cardinality of $\mathbb{N}$ is denoted by the symbol $\aleph_0$, which of course, is not a number in the usual sense.

This topic (as well as a proper definition of the phrase *having the same cardinality*) requires further developments. We will stop the discussion shortly and resume it later on the course. But, in order to motivate the reader's curiosity we make the following remark.

Remark 2.6. *When A is a finite set, Theorem 2.1 is implicitly saying that $\mathcal{P}(A)$ has more elements than A. This fact is also true when dealing with infinite sets and it is called the* **Cantor's Theorem**. *This result has two important consequences:*

a) It can be proved that the cardinality of real numbers is the same as the cardinality of $\mathcal{P}(\mathbb{N})$; we will prove this later on the course. Hence, the cardinality of $\mathbb{R}$ is strictly bigger than $\aleph_0$, which yields to the idea that some infinities are bigger than others. Consequently, the set of real numbers is said to be **uncountable**.

b) Cantor's Theorem rules out the existence of a "set of all sets". Indeed, if such set exists, then its power set must have more elements than the "universal set" and so there is at least a set that doesn't belong to the "set of all sets". This contradicts the definition of the "universal set", and so, a "set of all sets" can't exist in our setting.

2.2 Basic set operations

In what follows, we assume that A and B are subsets of a given reference set X; the set X is sometimes called the **universal set** and it is supposed to be understood from the context one is working in; note that this notion differs from the idea of "set of all sets".

We define the **difference** between A and B, written as $A \setminus B$, as the set that contains all the elements of A that are not in B. Formally,

$$x \in A \setminus B \iff [(x \in A) \ \wedge (x \notin B)].$$

Example 2.3. *Let D be the set of digits, and E the set of even digits. Then $D \setminus E$ is the set of odd digits, that is,*

$$D \setminus E = \{1,\ 3,\ 5,\ 7,\ 9\}.$$

When X plays the role of A, the set $X \setminus B$ is called the **complement** of B (with respect to the universal set X), and it is denoted by B^c.

Example 2.4. *When taking the complement of a set, it is important to understand who is acting as the universal set X. For instance, in Example 2.3 if we take D as X, we have that the complement of E, the set of even digits, is the set of odd digits. However, if X is taken as $\mathbb{N}$, then E^c includes the set of odd digits as well as other natural numbers.*

Recall that the contrapositive is equivalent to the original statement. Hence by the definition of the inclusion we have

$$A \subseteq B \iff [\forall x, (x \notin B) \implies (x \notin A)].$$

Now, by the definition of the complement, we get the following identity:

$$A \subseteq B \iff B^c \subseteq A^c.$$

We define the **union** of A and B, as the set that joins all the elements of A and B. Symbolically, we write $A \cup B$ and its definition is

$$x \in A \cup B \iff [(x \in A) \vee (x \in B)].$$

Some essential properties of the union are listed below

1. $A \cup A = A$: this is clear from the definition.
2. $A \subseteq A \cup B$: this follows from the tautology

$$x \in A \implies (x \in A \vee x \in B).$$

3. If $A \subseteq C$ and $B \subseteq C$, then $A \cup B \subseteq C$: to prove this we follow the next logical steps:

$$\begin{aligned} A \cup B \subseteq C &\iff \forall x,\ (x \in A \vee x \in B) \implies x \in C \\ &\iff \forall x,\ \overline{(x \in A \vee x \in B)} \vee x \in C \\ &\iff \forall x,\ (x \notin A \wedge x \notin B) \vee x \in C \\ &\iff \forall x,\ (x \notin A \vee x \in C) \wedge (x \notin B \vee x \in C) \\ &\iff \forall x,\ (\overline{x \in A} \vee x \in C) \wedge (\overline{x \in B} \vee x \in C) \\ &\iff \forall x,\ (x \in A \implies x \in C) \wedge (x \in B \implies x \in C) \\ &\iff A \subseteq C \wedge B \subseteq C \end{aligned}$$

4. $A \cup A^c = X$: since $A \subseteq X$ and $A^c \subseteq X$, by property 3 we get $A \cup A^c \subseteq X$. Therefore, to prove the equality we only need to prove that $X \subseteq A \cup A^c$, but this is a consequence of the following tautology:

$$x \in X \implies (x \in A \vee x \notin A).$$

5. If $A \subseteq B$ and $C \in \mathcal{P}(X)$, then $A \cup C \subseteq B \cup C$: by property 2, we have $A \subseteq B \subseteq B \cup C$ and $C \subseteq B \cup C$. Then the result follows from property 3.
6. $A \cup \emptyset = A$: by property 2 we have $A \subseteq A \cup \emptyset$. Also, since $\emptyset \subseteq A$ and $A \subseteq A$, by property 3 we get $A \cup \emptyset \subseteq A$.
7. $A \cup X = X$: by property 2 we have $X \subseteq A \cup X$. Moreover, since $A \subseteq X$ and $X \subseteq X$, by property 3 the conclusion follows.

Also, we define the **intersection** of A and B, denoted by $A \cap B$, as the set that contains the common elements between A and B. In other words,

$$x \in A \cap B \iff [(x \in A) \ \wedge (x \in B)].$$

Some essential properties of the intersection are listed below, and their proofs are left as exercise for the reader.

1. $A \cap A = A$.

2. $A \cap B \subseteq A$.

3. If $C \subseteq A$ and $C \subseteq B$, then $C \subseteq A \cap B$.

4. $A \cap A^c = \emptyset$.

5. $A \subseteq B$ and $C \in \mathcal{P}(X)$, then $A \cap C \subseteq B \cap C$.

6. $A \cap \emptyset = \emptyset$.

7. $A \cap X = A$.

By the commutative, associative and distributive properties of the $\wedge$ and $\vee$ connectors, we can see that the union and intersection of sets satisfy similar properties, that is, suppose C is another subset of X, then

- **Commutativity:** $A \cup B = B \cup A$ and $A \cap B = B \cap A$.

- **Associativity:** $A \cup (B \cup C) = (A \cup B) \cup C$ and $A \cap (B \cap C) = (A \cap B) \cap C$.

- **Distributivity:** $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ and $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

In general, we can also consider families of subsets of X with arbitrary large number of elements, say $n \in \mathbb{N}$. In this case, we might use the notation $\{A_k\}_{k=0}^n$ to indicate the set of $\mathcal{P}(X)$ whose elements are A_0, A_1, $\ldots$, A_n. We can then define the union of $\{A_k\}_{k=0}^n$ as the following set

$$\bigcup_{k=0}^{n} A_k = \{x \in X \mid \ \exists k \in \{0, \ldots, n\}, \ x \in A_k\}.$$

In a similar way, we can define the intersection of the elements of $\{A_k\}_{k=0}^n$ as the following set

$$\bigcap_{k=0}^{n} A_k = \{x \in X \mid \ \forall k \in \{0, \ldots, n\}, \ x \in A_k\}.$$

If $A \cap B = \emptyset$ we say that they are **disjoint**, and if $\bigcap_{k=0}^n A_k = \emptyset$, we say that $\{A_k\}_{k=0}^n$ is a **pairwise disjoint** family.

Remark 2.7. *Note that, by the commutative and associative properties, it doesn't matter the order in which the union or intersection are taken, so the notation we have chosen is consistent. Furthermore, the definition of union and intersection of a finite family of sets can also be extended to any infinite family of sets* $\{A_k\}_{k\in K}$ *indexed by a set* K *(countable or not) in the following way*

$$\bigcup_{k\in K} A_k = \{x \in X \mid \exists k \in K,\ x \in A_k\} \quad \textit{and} \quad \bigcap_{k\in K} A_k = \{x \in X \mid \forall k \in K,\ x \in A_k\}.$$

By applying the De Morgan's laws to the definitions of the union and intersection, we get the following identities

$$(A \cap B)^c = A^c \cup B^c \quad \text{and} \quad (A \cup B)^c = A^c \cap B^c.$$

It is not difficult to see that, if $\{A_k\}_{k=0}^n$ is a finite family of sets, then the De Morgan's laws provides the following identities

$$\left(\bigcap_{k=0}^{n} A_k\right)^c = \bigcup_{k=0}^{n} A_k^c \quad \text{and} \quad \left(\bigcup_{k=0}^{n} A_k\right)^c = \bigcap_{k=0}^{n} A_k^c.$$

We left as an exercise for the reader to prove that this is also true for $\{A_k\}_{k\in K}$, an infinite family of subsets of X.

2.3 Exercises

1. Let X be a nonempty set and $E \subseteq X$. Suppose that the following proposition is true

   $$\forall A,\ B \in \mathcal{P}(X),\ (E \cup A = E \cup B \implies A = B).$$

   Then show that $E = \emptyset$.

2. Let A and B be subsets of a given set X. Show that

   $$A \cap B = \emptyset \iff \mathcal{P}(A) \cap \mathcal{P}(B) = \{\emptyset\}.$$

   Hint: Prove ($\Longrightarrow$) by contradiction and ($\Longleftarrow$) by contraposition.

3. Let A and B be subsets of a given set X.

   (a) Show that $\emptyset \notin \mathcal{P}(A) \setminus \mathcal{P}(B)$.

   (b) Prove that $\mathcal{P}(A \setminus B) \subseteq (\mathcal{P}(A) \setminus \mathcal{P}(B)) \cup \{\emptyset\}$.

   (c) Give an example of set A and B such that

   $$\mathcal{P}(A \setminus B) \neq (\mathcal{P}(A) \setminus \mathcal{P}(B)) \cup \{\emptyset\}$$

4. Let us consider the set operation $\circledast$ defined via

$$A \circledast B = A^c \cap B^c.$$

Let X be a nonempty set and $\mathcal{F} \subseteq \mathcal{P}(X)$ be a nonempty set such that

$$\forall A,\ B \in \mathcal{F}, A \circledast B \in \mathcal{F}.$$

Show that if $A,\ B \in \mathcal{F}$, then

(a) $A^c \in \mathcal{F}$.

(b) $A \cap B \in \mathcal{F}$.

(c) $A \cup B \in \mathcal{F}$.

Conclude that $\emptyset \in \mathcal{F}$ and $X \in \mathcal{F}$.

5. Let A and B be subsets of a given set X. Show by contraposition that

$$[(A^c \cap B) \cup (A \cap B^c)] = B \implies (A = \emptyset).$$

6. Let A, B and C be subsets of a given set X. Prove that

$$A \cap B \cap C = \emptyset \implies (A \setminus B) \cup (B \setminus C) \cup (C \setminus A) = A \cup B \cup C.$$

Bibliography

[1] Paul R. Halmos *Naive set theory.* Princeton, N.J., Van Nostrand, 1960.

[2] Robert R. Stoll *Set theory and Logic.* New York : Dover Publications, 1979.

Math 4031 - Advanced Calculus I

Instructor: Dr. Cristopher HERMOSILLA
Louisiana State University - Spring 2016

Week 3: The Real numbers

We start now the study of the set of Real numbers, which we have denoted by $\mathbb{R}$. As we said the first week, our exposition will be based on some prescribed rules, called *axioms*, that will allow us to provide a consistent description of the set of Real numbers. The axioms we will present can be divided into three groups: algebraic, order and the supremum axiom.

3.1 Algebraic properties of the Real numbers

First of all, we write $x = y$ to indicate that the Real numbers x and y are the same; otherwise we write $x \neq y$. Formally speaking, $=$ is a relation on $\mathbb{R}$ that can be defined using the equality of sets as below:

$$\forall x, y \in \mathbb{R},\ x = y \iff \{x\} = \{y\}.$$

We now consider two basic operations defined on $\mathbb{R}$, the **sum** $(+)$ and **product** $(\cdot)$ of Real numbers. We will write

$$x + y \qquad \text{and} \qquad x \cdot y$$

for the elements obtained respectively as the sum and product of two Real numbers x and y.

It may sound familiar to you that *the sum and product of Real numbers is a Real number*. We don't prove this statement and just accept it as a prescribed rule. Actually, this is the first axiom for Real numbers we are going to study and it is known as the **closure axiom**:

$$(A_0) \qquad \forall x, y \in \mathbb{R},\ x + y \in \mathbb{R}\ \wedge\ x \cdot y \in \mathbb{R}$$

Each of the operations, sum and product, has four additional axioms associated with it. These are the associative, the commutative, the additive/multiplicative identity and the additive/multiplicative inverse axioms. In mathematical terms, these rules are written respectively:

$$\begin{array}{llll}
(A_1) & \forall x, y, z \in \mathbb{R}, & (x + y) + z = x + (y + z) \quad \wedge & (x \cdot y) \cdot z = x \cdot (y \cdot z) \\
(A_2) & \forall x, y \in \mathbb{R}, & x + y = y + x \quad \wedge & x \cdot y = y \cdot x \\
(A_3) & \exists 0,\ 1 \in \mathbb{R},\ 0 \neq 1, & \forall x \in \mathbb{R},\ x + 0 = x \quad \wedge & \forall x \in \mathbb{R},\ x \cdot 1 = x \\
(A_4) & \forall x \in \mathbb{R} \setminus \{0\}, & \exists! y \in \mathbb{R},\ x + y = 0 \quad \wedge & \exists! z \in \mathbb{R},\ x \cdot z = 1
\end{array}$$

In the last axiom, the Real numbers y and z are called the **additive** and **multiplicative inverses**, respectively. Since for any $x \in \mathbb{R}$ there is a unique inverse, we normally use the notation $-x$ and x^{-1} to indicate such elements. Note that the additive inverse of 0 has not been defined in (A_4). However, it is not difficult to prove, using (A_3) and (A_4), that 0 is the (unique) additive inverse of 0; we leave this as exercise for the reader.

Remark 3.1. *Note that (A_3) provides the existence of at least two different Real numbers: 0 and 1. Therefore, using the sum we can construct other numbers, such as the Naturals via*

$$2 := 1 + 1, \quad 3 := 2 + 1, \quad 4 := 3 + 1, \ldots$$

Here the notation $a := b$ means that the symbol a is defined with the value b.

The sum and product of Real numbers are related via an algebraic relation called the **distributive** axiom, which can be expressed as follows

$$(A_5) \qquad \forall x, y, z \in \mathbb{R}, \qquad (x + y) \cdot z = (x \cdot z) + (y \cdot z) \ \wedge \ z \cdot (x + y) = (z \cdot x) + (z \cdot y).$$

The six axioms we have presented are enough to derive several properties of the Real numbers. For example, we can prove the so-called *square of a binomial formula*:

$$(a + b)^2 = a^2 + 2 \cdot (a \cdot b) + b^2,$$

where the notation x^2 stands for the product $x \cdot x$, which is called the **square of** x. Let us see a detailed proof of this formula:

$$\begin{aligned}
(a + b)^2 &= (a + b) \cdot (a + b) && \text{(Definition of the square of } (a + b)\text{)} \\
&= a \cdot a + b \cdot a + a \cdot b + b \cdot b && \text{(Axiom } (A_5) \text{ used twice)} \\
&= a \cdot a + a \cdot b + a \cdot b + b \cdot b && \text{(Axiom } (A_2)\text{)} \\
&= a^2 + a \cdot b + a \cdot b + b^2 && \text{(Definition of the square } a \text{ and } b\text{)} \\
&= a^2 + 1 \cdot (a \cdot b) + 1 \cdot (a \cdot b) + b^2 && \text{(Axiom } (A_3)\text{)} \\
&= a^2 + (1 + 1) \cdot (a \cdot b) + b^2 && \text{(Axiom } (A_5)\text{)} \\
&= a^2 + 2 \cdot (a \cdot b) + b^2 && \text{(Definition of 2)}
\end{aligned}$$

Other properties of Real numbers based on the axioms are:

(1) $x \cdot 0 = 0$: By (A_3) and (A_5), $x \cdot 0 = x \cdot (0 + 0) = x \cdot 0 + x \cdot 0$, the adding $-(x \cdot 0)$ in the equality, we get the result.

(2) $-(x + y) = (-x) + (-y)$: Remark that $(x + y) + (-x) + (-y) = 0$ This is a consequence of the associative and commutative axioms. Then, the conclusion follows from the uniqueness of the additive inverse of $x + y$.

(3) $x \cdot (-y) = -(x \cdot y)$: We need to show that $x \cdot y + x \cdot (-y) = 0$. Using ($A_5$), ($A_3$) and property (1), we get $x \cdot y + x \cdot (-y) = x \cdot (y + -y) = x \cdot 0 = 0$, which completes the proof.

(4) $x \cdot y = (-x) \cdot (-y)$: Using property (3) and (A_2) twice each, we get

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

It is not difficult to see from (A_4), that $-(-(x \cdot y)) = x \cdot y$, and so the conclusion follows.

From the sum and product we can define new operations on $\mathbb{R}$. For example, the **subtraction** (-) and **division** (/) which are defined via:

$$\forall x, y \in \mathbb{R},\ y \neq 0, \quad x - y := x + (-y)\ \wedge\ x/y := x \cdot y^{-1}.$$

Remark 3.2. *Since the additive inverse of* 0 *is well defined (it is* 0 *itself), we can extend the subtraction for the case* $y = 0$*. Under these circumstances, we just get*

$$x - 0 := x + 0 = x.$$

Nevertheless, it is not possible to do the same for the division without getting some inconsistency with the axioms; this is essentially because the multiplicative inverse of 0 *can not exist. For example, suppose that* 0^{-1} *exists and it is a Real number. Then, on the one hand we have*

$$0 \cdot 0^{-1} = 1.$$

But, on the other hand, by axioms (A_3), (A_4) *and* (A_5)

$$0 = 0 \cdot 0^{-1} - 0 \cdot 0^{-1} = (0 + 0) \cdot 0^{-1} - 0 \cdot 0^{-1} = 0 \cdot 0^{-1} + 0 \cdot 0^{-1} - 0 \cdot 0^{-1} = 0 \cdot 0^{-1} = 1.$$

This conclusion contradicts axiom (A_3).

3.2 Order axioms

So far, we know that Real numbers can be summed o multiplied, but we don't know how to compare them. To do so, we introduce an **order relation** ($\leq$), which satisfies the following axioms:

(O_1)	$\forall x \in \mathbb{R},$	$x \leq x$	(Reflexivity).
(O_2)	$\forall x, y \in \mathbb{R},$	$(x \leq y\ \wedge\ y \leq x) \implies x = y$	(Antisymmetry).
(O_3)	$\forall x, y, z \in \mathbb{R},$	$(x \leq y\ \wedge\ y \leq z) \implies x \leq z$	(Transitivity).
(O_4)	$\forall x, y \in \mathbb{R},$	$x \leq y\ \vee\ y \leq x$	(Comparability).

The statement $x \leq y$ is read as x *is less than or equal to* y and we say that $\mathbb{R}$ is a **totally ordered set**. We can also read $x \leq y$ as y *is greater than or equal to* x.

Remark 3.3. *Let us point out that by the Reflexive axiom, the converse of the Antisymmetric axiom holds true, that is*

$$\forall x, y \in \mathbb{R}, \qquad x = y \implies (x \leq y\ \wedge\ y \leq x).$$

This means in particular that

$$x \leq y\ \wedge\ x = y \iff x = y.$$

We leave the proof of the last equivalence as exercise for the reader.

From the order relation $\leq$ we can define other order relations. For example, we say that x *is less than* y, written as $x < y$, if and only if

$$x \leq y \ \wedge \ x \neq y.$$

This new order relation satisfies neither the axioms (O_1), (O_2) nor (O_4). However, its utility is reflected in the next trichotomy result.

Theorem 3.1. *For any $x, y \in \mathbb{R}$, only one of the following statements is true*

$$x < y, \qquad x = y \qquad \text{or} \qquad y < x.$$

Proof. Let us first check that the statements are mutually exclusive. Let $x, y \in \mathbb{R}$ be fixed but arbitrary.

- Suppose that $x = y$, then $x \neq y$ is false and thus, neither $x < y$ nor $y < x$ can be true.
- Suppose that $x \neq y$, then clearly $x = y$ is false. Furthermore, by the Comparability axiom (O_4) we have that either $x \leq y$ or $y \leq x$ is true. Note that $x \leq y$ and $y \leq x$ can not be both true at the same time, because otherwise the Antisymmetric axiom (O_2) would imply that $x = y$ which can not be. Therefore, by the definition of $<$, either $x < y$ o $y < x$, but not both at the same time.

In order to complete the proof, we need to show that at least one of the statements is true. By the Comparability axiom (O_4), we have that either $x \leq y$ or $y \leq x$ is true. Without loss of generality, we can assume that $x \leq y$; if $y \leq x$, we use the same arguments but switching the roles of x and y. We claim that

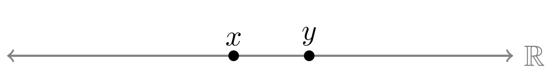
$$x \leq y \iff (x < y) \ \vee (x = y).$$

Indeed, this follows from the following reasoning

$$\begin{aligned} x \leq y &\iff x \leq y \ \wedge \ (x \neq y \ \vee x = y) \\ &\iff (x \leq y \ \wedge \ x \neq y) \ \vee (x \leq y \ \wedge \ x = y) \\ &\iff (x < y) \ \vee (x = y) \end{aligned}$$

This mean that either $x < y$ or $x = y$ is true, and so, at least one of the statements is true. $\square$

Theorem 3.1 allows us to represent the set of Real numbers as a horizontal line that extends infinitely in both directions. This means that each Real number can be associated with a point in the line, and for any $x, y \in \mathbb{R}$, if $x < y$ we have that x appears at the left of y. Otherwise, x appears at the right of y.



The order relation we have introduced can be related to the sum and product by means of the following Compatibility axioms:

(C_1) $\quad \forall x, y, z \in \mathbb{R}, \qquad x \leq y \implies x + z \leq y + z.$

(C_2) $\quad \forall x, y, z \in \mathbb{R}, \qquad x \leq y \ \wedge \ 0 \leq z \implies x \cdot z \leq y \cdot z.$

Moreover, the following properties can be derived from the preceding axioms:

(a) $x \leq y \implies -y \leq -x$: Since $x \leq y$, in the light of (C_1) we get $0 = x - x \leq y - x$. Hence, adding $-y$ to the last inequality and, using (C_1) and (A_2), we get the desired inequality.

(b) $x \leq y \ \wedge \ z \leq 0 \implies y \cdot z \leq x \cdot z$: By property (a), $0 \leq -z$, and so by (C_2), we get

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

The conclusion follows from property (a) and the fact that $a = -(-a)$.

(c) $0 \leq x^2$: If $0 \leq x$, the result is a direct consequence of (C_2). If on the contrary, $x \leq 0$, the result follows from property (b).

(d) $x < x + 1$: It is clear that $x \neq x + 1$, otherwise we get $0 = 1$. Furthermore, by property (c) we get that $0 \leq 1^2 = 1 \cdot 1 = 1$. Hence, by (C_1), we obtain $x = x + 0 \leq x + 1$.

(e) $1 \leq x \implies x \leq x^2$: Since $0 \leq 1$, by(O_3) we get in particular that $0 \leq x$, and so by (C_2) we get $x = 1 \cdot x \leq x \cdot x = x^2$.

(f) $x \leq 1 \ \wedge \ 0 \leq x \implies x^2 \leq x$: Similar as the proof above.

(g) $x \leq y \wedge 0 < x \implies y^{-1} \leq x^{-1}$: Let us first prove that $0 < x^{-1}$. Suppose by contradiction that $0 < x^{-1}$ is false. This means that $x^{-1} \leq 0$, and by Theorem 3.1, we must have that $x^{-1} < 0$. Hence, by property (b), we get $1 = x \cdot x^{-1} \leq 0 \cdot x^{-1} = 0$. Which is impossible, so we must have $0 < x^{-1}$. In the same way, we can prove that $0 < y^{-1}$. Therefore, multiplying the inequality $x \leq y$ first by x^{-1} and then y^{-1}, we get the result wanted.

3.2.1 Some particular subsets of $\mathbb{R}$

The order relations we have introduced are also helpful to define new subsets of $\mathbb{R}$, called **intervals**. For any $a, b \in \mathbb{R}$ with $a \leq b$ we define:

- The **closed bounded interval**: $[a, b] := \{x \in \mathbb{R} \mid \ a \leq x \ \wedge \ x \leq b\}$.
- The **open-closed bounded interval**: $(a, b] := \{x \in \mathbb{R} \mid \ a < x \ \wedge \ x \leq b\}$.
- The **closed-open bounded interval**: $[a, b) := \{x \in \mathbb{R} \mid \ a \leq x \ \wedge \ x < b\}$.
- The **open bounded interval**: $(a, b) := \{x \in \mathbb{R} \mid \ a < x \ \wedge \ x < b\}$.

Remark 3.4. *For the extreme case $a = b$ we have that $[a, b] = [a, a] = \{a\}$ but*

$$(a, b] = (a, a] = [a, b) = [a, a) = (a, b) = (a, a) = \emptyset.$$

We might also consider the following **unbounded intervals**:

- The **closed unbounded intervals**:

$$[a, +\infty) := \{x \in \mathbb{R} \mid \ a \leq x \ \} \quad \text{and} \quad (-\infty, b] := \{x \in \mathbb{R} \mid \ x \leq b\}.$$

- The **open unbounded interval**:

$$(a, +\infty) := \{x \in \mathbb{R} \mid \ a < x \ \} \quad \text{and} \quad (-\infty, b) := \{x \in \mathbb{R} \mid \ x < b\}.$$

Here the symbols $+\infty$ and $-\infty$ are only used for sake of notation, which means that they don't have a meaning by themselves. Furthermore, the intervals $(0, +\infty)$ and $(-\infty, 0)$ will be particularly interesting. We say that x is a **positive** Real number if $x \in (0, +\infty)$. Similarly, x is said to be a **negative** Real number if $x \in (-\infty, 0)$.

3.2.2 The absolute value

We define the absolute value of a Real number x, denoted by $|x|$, as the Real number that agrees with x, when the latter is positive and with $-x$ when x is negative. In other words,

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

Note that by definition, $|x| = |-x|$ for any $x \in \mathbb{R}$.

The absolute value will play a crucial role when studying sequences and convergence. For the moment, we restrict our attention to some of its basic properties:

(i) $x \leq |x|$: Note that we always have $0 \leq |x|$, so if $x \leq 0$, the result is straightforward. On the other hand, if $0 \leq x$, then $|x| = x$, so the conclusion follows as well.

(ii) (**triangular inequality**) $|x + y| \leq |x| + |y|$: By property (i) we have that $x \leq |x|$, $-x \leq |x|$, $y \leq |y|$ and $-y \leq |y|$. Hence. if on the one hand we have $0 \leq x + y$, then $|x+y| = x+y \leq |x|+|y|$. On the other hand, if $x+y \leq 0$, then $0 \leq -(x+y) = (-x)+(-y)$ and so $|x + y| = (-x) + (-y) \leq |x| + |y|$, and the proof is complete.

(iii) $|x \cdot y| = |x| \cdot |y|$: We might assume that $x \neq 0$ and $y \neq 0$, otherwise the result is straightforward. To complete the proof we need to put on cases. We are only going to do the case $x \leq 0$ and $0 \leq y$, the others remain as exercise for the reader. In the circumstances we described before, $|x| = -x$ and $|y| = y$, but $x \cdot y \leq 0$ (this is because (C_2) and property (b)), then $|x \cdot y| = -(x \cdot y) = (-x) \cdot y = |x| \cdot |y|$.

(iv) $||x| - |y|| \leq |x - y|$: Let us assume without loss of generality that $||x| - |y|| = |x| - |y|$. By (A_4) and the triangular inequality we get $|x| = |x - y + y| \leq |x - y| + |y|$. Finally, adding $-|y|$ in the last inequality we get the result wanted.

(v) For any $a \in (0, +\infty)$, $|x| \leq a$ if and only if $-a \leq x \ \wedge \ x \leq a$: By property (i) and ($O_3$) we immediately have $x \leq a$. Moreover, by property ((a)), we get that $-a \leq -|x|$ and $-|x| = -|-x| \leq -(-x) = x$. Thus, by ($O_3$) we get $-a \leq x$, which completes the proof.

3.3 The supremum axiom

Note that if we replace, in all the axiom we have presented, the set $\mathbb{R}$ by $\mathbb{Q}$, the axioms still make sense. This could wrongly yield to the idea that the set of Rational and Real numbers are the same. There is another axiom, called the **Supremum axiom**, that will allow us to distinguish $\mathbb{Q}$ from $\mathbb{R}$; see Exercise 7.

Let us now introduce some definitions. Given a subset $A \subseteq \mathbb{R}$, we say that it is **bounded above** if

$$\exists M \in \mathbb{R},\ \forall x \in A,\ x \leq M. \tag{3.1}$$

In a similar way, we say that A is **bounded below** if

$$\exists m \in \mathbb{R},\ \forall x \in A,\ m \leq x. \tag{3.2}$$

Any Real number that satisfy (3.1) or (3.2), is called **upper** or **lower bound of** A, respectively. If A is bounded below and above, we just say that it is a **bounded** set.

Example 3.1. *The definition of bounded set is consistent with the definition of bounded intervals we have done. Indeed, for any $a, b \in \mathbb{R}$ with $a \leq b$, the intervals*

$$[a, b], \quad [a, b), \quad (a, b] \quad \text{and} \quad (a, b)$$

are all examples of bounded set. Here a is a lower bound and b is an upper bound for any of those subsets. Notice as well that the unbounded intervals

$$[a, +\infty), \quad (a, +\infty), \quad (-\infty, b] \quad \text{and} \quad (-\infty, b)$$

are not bounded, but the first two are bounded below (a is a lower bound) and the last two are bounded above (b is an upper bound).

Note that in Example 3.1, any number greater than b is also an upper bound for any of the bounded intervals, but there are no other Real numbers, less than b than can be an upper bound of any of the bounded intervals. In other words, b is the least upper bound we can find of, for instance, $A = [a, b)$. In this case, we say that this upper bound is the **supremum of** A and we denote it by $\sup(A)$. In mathematical terms, $M = sup(A)$ if and only if M is an upper bound of A and

$$\forall N \in \mathbb{R},\ [(\forall x \in A,\ x \leq N) \implies M \leq N].$$

From this definition we can see that the supremum is uniquely determined (verify this, suppose there are two different supremum and then conclude using the Antisymmetric axiom).

Analogously, we define the **infimum of** A, denoted by $\inf(A)$, as the Real number m that is a lower bound of A and such that

$$\forall n \in \mathbb{R},\ [(\forall x \in A,\ n \leq x) \implies n \leq m].$$

Example 3.2. *We have that for any $a, b \in \mathbb{R}$ with $a \leq b$*

$$b = \sup([a, b]) = \sup([a, b)) = \sup((a, b]) = \sup((a, b)) = \sup((-\infty, b]) = \sup((-\infty, b))$$
$$a = \inf([a, b]) = \inf([a, b)) = \inf((a, b]) = \inf((a, b)) = \inf([a, +\infty)) = \inf((a, +\infty)).$$

In the special case that $\inf(A) \in A$, we say that $\inf(A)$ is the **minimum of** A. Analogously, when $\sup(A) \in A$ we say that $\sup(A)$ is the **maximum of** A. In this case, we change the notation, and we write $\min(A)$ and $\max(A)$, for the minimum and maximum of A, respectively.

Example 3.3. *We have that for any $a, b \in \mathbb{R}$ with $a \leq b$*

$$\begin{aligned} b &= \max([a,b]) = \sup((a,b]) = \sup((-\infty, b]) \\ a &= \min([a,b]) = \inf([a,b)) = \inf([a,+\infty)). \end{aligned}$$

We are now in position to state the Supremum axiom. This says that for any nonempty bounded above subset of $\mathbb{R}$ its supremum is a Real number, that is,

$$(S) \qquad \forall A \in \mathcal{P}(\mathbb{R}) \setminus \{\emptyset\},\ (A \text{ is bounded above } \implies \sup(A) \in \mathbb{R}).$$

As we mentioned earlier, this is what distinguish $\mathbb{R}$ from $\mathbb{Q}$, because, as you will see in Exercise 7, not every subset of $\mathbb{Q}$ has a supremum that is also a Rational number.

Remark 3.5. *The Supremum axiom has a counterpart, the Infimum axiom which says that for any nonempty bounded below subset of $\mathbb{R}$ its infimum is a Real number. We leave the proof that both axioms are equivalent as exercise for the reader.*

One of the main consequences of the Supremum axiom is that the set of Real numbers is **Archimedean**. This is summarized in the next theorem.

Theorem 3.2. $\mathbb{R}$ *is Archimedean, that is,*

$$\forall x, M \in (0,+\infty),\ \exists n \in \mathbb{N},\ M < n \cdot x. \tag{3.3}$$

Proof. We divide the proof into three steps:

1. Let us first prove that $\mathbb{N}$ is not bounded above. We argue by contradiction, that is, suppose that $\mathbb{N}$ is bounded above. Since it is also nonempty ($1 \in \mathbb{N}$ for instance), by the Supremum axiom, $\sup(\mathbb{N})$ is a Real number. In particular, $\sup(\mathbb{N}) - \frac{1}{2}$ is a Real number that can not be an upper bound of $\mathbb{N}$. So, there is $n \in \mathbb{N}$ such that
$$\sup(\mathbb{N}) - \frac{1}{2} < n.$$
Then, adding 1 in the inequality, we get
$$\sup(\mathbb{N}) + \frac{1}{2} < n + 1.$$
But, by definition, $n + 1 \leq \sup(\mathbb{N})$, because $n + 1 \in \mathbb{N}$. Thus, by (O_3)
$$\sup(\mathbb{N}) + \frac{1}{2} \leq \sup(\mathbb{N}),$$
which is not possible, and consequently, $\mathbb{N}$ is not bounded above.

2. To finish, let us also use a contradiction argument again. Suppose that (3.3) is false, then its negation is true, that is
$$\exists x, M \in (0,+\infty),\ \forall n \in \mathbb{N},\ n \cdot x \leq M.$$
Since, $x \neq 0$, we get that $M \cdot x^{-1} \in \mathbb{R}$ and it is an upper bound of $\mathbb{N}$, which contradicts what we proved in the first step. Therefore, the conclusion follows.

□

3.4 Exercises

1. Let $x, y, z, w \in \mathbb{R}$ with $y, w \neq 0$. Prove, mentioning each of the axioms and properties you are using, that

$$x \cdot y^{-1} + z \cdot w^{-1} = (x \cdot w + y \cdot z) \cdot (y \cdot w)^{-1}$$

2. Let $x, y, z, w \in \mathbb{R}$ such that $(x \cdot w) + (-(y \cdot z)) = 0$. Prove, mentioning each of the axioms and properties you are using, that

$$((x + y) \cdot w) + (-(z + w) \cdot y) = 0$$

3. Prove, detailing each step you do, the following properties:
   (a) $\forall x, y \in \mathbb{R},\ \ 2 \cdot (x \cdot y) \leq x^2 + y^2$.
   (b) $\forall x \in (0, +\infty),\ \ 2 \leq x + x^{-1}$.
   (c) $\forall x \in (0, +\infty),\ \ 1 + 2 \cdot x \leq (1 + x)^2$.
   (d) $\forall x, y \in (0, +\infty),\ \ 4 \cdot (x + y)^{-1} \leq x^{-1} + y^{-1}$.
   (e) $\forall x, y, z \in (0, +\infty),\ \ x \cdot y + x \cdot z + y \cdot z \leq x^2 + y^2 + z^2$.
   (f) $\forall x, y, z \in (0, +\infty),\ \ 8 \cdot x \cdot y \cdot z \leq (x + y) \cdot (y + z) \cdot (z + x)$.

4. Prove, detailing each step you do, the following properties:
   (a) $x + y + z = 1 \ \wedge\ x, y, z \neq 0 \implies 8 \leq (x^{-1} - 1) \cdot (y^{-1} - 1) \cdot (z^{-1} - 1)$.
   (b) $x \cdot y \cdot z = 1 \implies 3 \leq x + y + z$.
   (c) $\forall z \in (0, +\infty),\ x < y + z \implies x \leq y$.

5. Show that for any $x, y \in \mathbb{R}$,
   (a) if $x, y \neq 0$, then $|x \cdot y^{-1} - y \cdot x^{-1}| \leq \frac{1}{2} \cdot [(|x| + |x^{-1}|) \cdot |y - y^{-1}| + (|y| + |y^{-1}|) \cdot |x - x^{-1}|)]$.
   (b) $|x| \leq \max(\{|x - y|, |x + y|\})$.

   Hint: Argue by contradiction and show that for any $t \in [0, 1]$

$$|t(x + y) + (1 - t)(x - y)| < |x|.$$

6. Let $x \in \mathbb{R}$ and consider the set $A_x = \{n \in \mathbb{N} \mid \ n \leq x\}$.
   (a) Prove that $\sup(A_x)$ is a Real number. This number is called the **integer part of** x and it is denoted by $[x]$.
   (b) Prove now that $[x] \in A_x$. To do so, follow the next step:
      i. Prove that there is $n \in \mathbb{N}$ such that $[x] - \frac{1}{2} < n \ \wedge\ n \leq [x]$.
      ii. Show that for any $m \in \mathbb{N}$, if $n < m$ then $m \notin A_x$.
      iii. Prove that $[x]$ is the maximum of A_x and conclude.

7. Let us consider the set $Q = \{x \in \mathbb{Q} \mid x^2 < 2\}$. The aim of this problem is to show that the supremum of Q is not a Rational number, and so the Supremum axiom doesn't hold when replacing $\mathbb{Q}$ for $\mathbb{R}$. We divide the proof in several steps:

   (a) Prove that Q is nonempty and bounded above.

   (b) Prove that $\sup(Q)^2 = 2$. To do so, follow the next steps:

      i. Assume that $\sup(Q)^2 < 2$ and get a contradiction by proving that

$$(\sup(Q) + a)^2 < 2, \quad \text{where } a = \min\left(\left\{\frac{1}{2}, \frac{2 - \sup(Q)^2}{2\sup(Q) + 1}\right\}\right).$$

      ii. Suppose that $\sup(Q)^2 > 2$ and show that $s = \sup(Q) - \frac{\sup(Q)^2 - 2}{2\sup(Q)}$ is also an upper bound of Q. Conclude that the only possible option is that $\sup(Q)^2 = 2$.

   (c) Suppose that $\sup(Q) \in \mathbb{Q}$ and let $\frac{p}{q}$ be its lowest terms representation, that is, p and q are Natural numbers having no common factors.

      i. Prove that p is even if and only if p^2 is even.
      Hint: Recall that p is even if $p = 2 \cdot n$ for some $n \in \mathbb{N}$. To prove the implication $(\Longleftarrow)$ use the contrapositive and the fact that a number that is not even, must be odd, that is, it can be written as $2 \cdot n + 1$ for some $n \in \mathbb{N}$.

      ii. Prove that p and q must be both even, and then conclude.

Math 4031 - Advanced Calculus I

Instructor: Dr. Cristopher HERMOSILLA
Louisiana State University - Spring 2016

Week 4: Sequences of Real numbers

We continue the study of the set of Real numbers by introducing the notion of sequence. From now on, we assume that all the Real numbers axioms are well known and we use them without making any reference, except in particular situations.

Before going further let us introduce some notation. If A is a finite set, whose elements are $a_1,\ a_2,\ \ldots,\ a_k$, we write $\max\{a_1,\ a_2,\ \ldots,\ a_k\}$ to indicate $\max(A)$. Moreover, for any $n \in \mathbb{N}$, the notation x^n stands for the Real number defined recursively via

$$\forall x \in \mathbb{R}, \quad x^0 = 1 \quad \text{and} \quad x^n := x^{n-1} \cdot x, \quad n \in \mathbb{N} \setminus \{0\}.$$

4.1 Sequences and converge

Let us begin the exposition with an example. We consider the set of Real numbers defined via

$$A = \left\{ x \in \mathbb{R} \;\middle|\; \exists n \in \mathbb{N} \setminus \{0\},\ x = \frac{1}{n} \right\}$$

We know that the order in which we write the elements of A is irrelevant for its description. For instance,

$$A_1 := \left\{1,\ \frac{1}{2},\ \frac{1}{3},\ \frac{1}{4},\ \frac{1}{5},\ \frac{1}{6}, \ldots \right\} \quad \text{and} \quad A_2 := \left\{\frac{1}{2},\ 1,\ \frac{1}{4},\ \frac{1}{3},\ \frac{1}{6},\ \frac{1}{5},\ \ldots \right\}$$

are equivalent description of A. However, if we assign an order to each positions of A_1 and A_2, say from left to right, and we associate each blank with a positive Natural number (1 for the first position, 2 for the second one, and so on). We get that x_n and y_n, the n^{th} element of A_1 and A_2, can be described via

$$x_n = \frac{1}{n} \quad \text{and} \quad y_n = \begin{cases} \frac{1}{n+1} & n \text{ is odd} \\ \frac{1}{n-1} & n \text{ is even} \end{cases}. \tag{4.1}$$

We write $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ to indicate the objects created from the sets A_1 and A_2 with the order we have prescribed. This kind of objects is what we call a **sequence**, that is, a (countable) infinite set whose elements have an order associated with them.

Similarly as done above, we denote a generic sequence by $\{x_n\}_{n=1}^{\infty}$, and we may use also y_n or any other letter with the subscript n to indicate the n^{th} element of the sequence at hand.

The main interest in studying sequences is their relation with the idea of *approximation*. Intuitively, we see that the sequences given by (4.1) get closer and closer to 0 as n increases its value, until the point that for any error $\varepsilon \in (0, +\infty)$ we may take, we can pick an $n \in \mathbb{N}$ large enough so that $|x_n| \leq \varepsilon$. This yields to the concept of **convergence**.

Definition 4.1. *We say that a sequence $\{x_n\}_{n=1}^{\infty}$ converges if there exists $L \in \mathbb{R}$ such that*

$$\forall \varepsilon \in (0, +\infty),\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ (N \leq n \implies |x_n - L| \leq \varepsilon). \tag{4.2}$$

Under these circumstances, we write $x_n \to L$ and L is called the **limit** *of $\{x_n\}_{n=1}^{\infty}$.*

Example 4.1. *Let us prove that $x_n \to 0$, where x_n is given by (4.1). Since, we have a* **candidate** *to limit that has been prescribed, we just need to verify that (4.2) is satisfied. Let us exhibit one way to prove this:*

1. *We take any $\varepsilon \in (0, +\infty)$, that is, ε is a generic positive Real number.*
2. *We write the functional statement $|x_n - L| \leq \varepsilon$ as explicit as we can. In this case, since $x_n = \frac{1}{n}$ and $L = 0$, we get $\frac{1}{n} \leq \varepsilon$; here we have used $|\frac{1}{n}| = \frac{1}{n}$.*
3. *We assume there exists $N \in \mathbb{N}$ such that (4.2) holds true and find some condition over it. In this case, it will be that $1 \leq N \cdot \varepsilon$. Furthermore, note that by the Compatibility axiom (C_2), this condition is sufficient to prove the convergence. Indeed, since $N \leq n$ and $\varepsilon \in (0, +\infty)$, we get $N \cdot \varepsilon \leq n \cdot \varepsilon$. Consequently, if $1 \leq N \cdot \varepsilon$ then $1 \leq n \cdot \varepsilon$.*
4. *Find $N \in \mathbb{N}$ for which the condition found in the previous part is satisfied. To do this, we need to get the existence from another source; for instance any theorem that provides the existence of some Real number. In this case, we use the Archimedean property (Theorem 3.2) with $M = 1$ and $x = \varepsilon$. This theorem gives the existence of $N \in \mathbb{N}$ such that $1 = M < N \cdot x = N \cdot \varepsilon$. In particular, we also have $1 \leq N \cdot \varepsilon$, so the conclusion follows.*

Notice that in Definition 4.1, the Natural number N depends in general on ε, as in the preceding example. This means that, the N associated with $\varepsilon = 1$ may be different from the one associated with $\varepsilon = \frac{1}{2}$; normally, the latter will be greater than the first one.

Remark 4.1. *The limit of a sequence is uniquely determined. To see this, assume by contradiction that there are two limits, say L_1 and L_2. Since $L_1 - L_2 \neq 0$, we have that $\frac{1}{4}|L_1 - L_2|$ is a positive Real number, and so, it can be used as ε in (4.2). Furthermore, given that $\{x_n\}_{n=1}^{\infty}$ converges, for $\varepsilon = \frac{1}{4}|L_1 - L_2|$, we can find $N_1, N_2 \in \mathbb{N}$ such that*

$$\forall n \in \mathbb{N},\ [(N_1 \leq n \implies |x_n - L_1| \leq \varepsilon)\ \wedge\ (N_2 \leq n \implies |x_n - L_2| \leq \varepsilon)].$$

Let $n \in \mathbb{N}$ be such that $N_1 \leq n$ and $N_2 \leq n$ (for instance $n = \max\{N_1, N_2\}$). Hence,

$$|L_1 - L_2| = |L_1 - x_n + x_n - L_2| \leq |L_1 - x_n| + |x_n - L_2| \leq \varepsilon + \varepsilon = 2\varepsilon = \frac{1}{2} \cdot |L_1 - L_2|$$

which yields to a contradiction. So, we must have $L_1 = L_2$.

4.1.1 Some conditions for convergence

Recall that sequences can also be seen as (infinite) nonempty subsets of $\mathbb{R}$. Hence, we say that a sequence $\{x_n\}_{n=1}^{\infty}$ is **bounded below** if it is bounded below as a set. Since each of the elements of the sequence can be enumerated, we write this definition in the following way:

$$\exists m \in \mathbb{R},\ \forall n \in \mathbb{N},\ m \leq x_n.$$

In an analogous way, we say that $\{x_n\}_{n=1}^{\infty}$ is **bounded above** provided that

$$\exists M \in \mathbb{R},\ \forall n \in \mathbb{N},\ x_n \leq M.$$

We say that $\{x_n\}_{n=1}^{\infty}$ is **bounded** if it is bounded below and above at the same time. We can easily verify that the sequences given by (4.1) are both bounded; it is enough to take $m = 0$ and $M = 1$. Moreover, by the Supremum axion, $\sup(\{x_n\}_{n=1}^{\infty})$ is well defined whenever $\{x_n\}_{n=1}^{\infty}$ is bounded above. By using similar arguments, we also get that $\inf(\{x_n\}_{n=1}^{\infty})$ is well defined whenever $\{x_n\}_{n=1}^{\infty}$ is bounded below. Under theses circumstances, we might use the following notation

$$\sup\{x_n\} := \sup(\{x_n\}_{n=1}^{\infty}) \quad \wedge \quad \inf\{x_n\} := \inf(\{x_n\}_{n=1}^{\infty})$$

It turns out that any sequence that converges must be bounded.

Theorem 4.1. *Suppose that $\{x_n\}_{n=1}^{\infty}$ converges, the it is bounded.*

Proof. By the definition of convergence, we have that there is $L \in \mathbb{R}$ so that (4.1) holds true. In particular, let $N \in \mathbb{N}$ be the number associated with $\varepsilon = 1$. We then have that

$$\forall n \in \mathbb{N},\ N \leq n,\ |x_n| = |x_n - L + L| \leq |x_n - L| + |L| \leq 1 + |L|.$$

This means that the tail of the sequences is bounded, that is,

$$\forall n \in \mathbb{N},\ N \leq n,\ -(1 + |L|) \leq x_n \quad \wedge\ x_n \leq 1 + |L|.$$

On the other hand, the set $\{x_1,\ x_2,\ \ldots,\ x_{N-1}\}$ is a finite, so it is bounded. Let m and M be a lower and upper bound of this set. Then, the following is always true

$$\forall n \in \mathbb{N},\ \min(\{m, -(1 + |L|)\}) \leq x_n \quad \wedge\ x_n \leq \max(\{M, 1 + |L|\}.$$

Hence, $\{x_n\}_{n=1}^{\infty}$ is bounded and the proof is complete. □

The preceding theorem provides a necessary condition for a sequence to converge. However, this condition is not sufficient. For example, the sequence whose terms are given by

$$x_n = \begin{cases} 1 & \text{if } n \text{ is even} \\ -1 & \text{if } n \text{ is odd} \end{cases}$$

is bounded but it doesn't converge. One way to make boundedness a sufficient condition for convergence is to ask some additional structure to the sequence.

Note that the sequence $\{x_n\}_{n=1}^{\infty}$ given by (4.1) has the special feature that each of its elements is less than the preceding one. We call this property **monotonicity**.

Definition 4.2. *A sequence* $\{x_n\}_{n=1}^{\infty}$ *is called* **monotonic** *if one of the following holds true:*

- $\{x_n\}_{n=1}^{\infty}$ *is a* **decreasing sequence**, *that is,* $\forall n \in \mathbb{N},\ x_{n+1} \leq x_n$.
- $\{x_n\}_{n=1}^{\infty}$ *is a* **increasing sequence**, *that is,* $\forall n \in \mathbb{N},\ x_n \leq x_{n+1}$.

We now provide a criterion to determine that a sequence converges based on monotonicity.

Theorem 4.2. *Let* $\{x_n\}_{n=1}^{\infty}$ *be an increasing sequence bounded above, then it converges. Analogously, if* $\{x_n\}_{n=1}^{\infty}$ *is a decreasing sequence bounded below, then it converges.*

Proof. We only prove the case of increasing sequences, the other is similar and is left as exercise for the reader. Since $\{x_n\}_{n=1}^{\infty}$ is bounded above, by the Supremum axiom, $L := \sup\{x_n\} \in \mathbb{R}$. We claim that $x_n \to L$. Take $\varepsilon \in (0, +\infty)$ arbitrary, by the definition of the supremum, there is $N \in \mathbb{N}$ such that

$$L - \varepsilon \leq x_N.$$

By the monotonicity of the sequence and the Transitivity axiom, we get that

$$\forall n \in \mathbb{N},\ (N \leq n \implies L - \varepsilon \leq x_n).$$

Furthermore, by definition $0 \leq L - x_n = |L - x_n|$ for any $n \in \mathbb{N}$. Hence, after a few algebraic steps we get

$$\forall n \in \mathbb{N},\ (N \leq n \implies |L - x_n| \leq \varepsilon).$$

Therefore, since $\varepsilon \in (0, +\infty)$ is arbitrary, the conclusion follows. □

Example 4.2. *Let us consider the sequence given by* $x_n = \frac{n-1}{n}$. *It is not difficult to see that* $\{x_n\}_{n=1}^{\infty}$ *is bounded above by 1; actually,* $\sup\{x_n\} = 1$. *We claim that it is also increasing. Indeed, since* $(n-1) \cdot (n+1) = n^2 - 1 \leq n^2$ *we get that*

$$x_n = \frac{n-1}{n} \leq \frac{n^2}{n \cdot (n+1)} = \frac{n}{n+1} = x_{n+1}.$$

Thus, by Theorem 4.2 the sequence converges to $\sup\{x_n\} = 1$.

The next result is a very useful tool to prove the convergence of a sequence and it is called the **Squeeze Theorem**. This criterion doesn't require a monotone character on the sequence but that it belongs to an interval whose length is smaller and smaller as n increases its value.

Theorem 4.3. *Let* $\{x_n\}_{n=1}^{\infty}$ *and* $\{y_n\}_{n=1}^{\infty}$ *be two sequences. Suppose that* $x_n \to L$ *and* $y_n \to L$ *for some* $L \in \mathbb{R}$. *Let* $\{z_n\}_{n=1}^{\infty}$ *be another sequence that satisfies*

$$\forall n \in \mathbb{N},\ x_n \leq z_n \ \wedge \ z_n \leq y_n.$$

Then $\{z_n\}_{n=1}^{\infty}$ *converges to* L.

Proof. Let $\varepsilon \in (0, +\infty)$ and $N_1, N_2 \in \mathbb{N}$ given by the definition of convergence of $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$, respectively. Let $N = \max\{N_1, N_2\}$ and take any $n \in \mathbb{N}$ such that $N \leq n$, then using the properties of the absolute value

$$z_n - L \leq y_n - L \leq |y_n - L| \leq \varepsilon \quad \wedge -\varepsilon \leq -|x_n - L| \leq x_n - L \leq z_n - L.$$

This means that $z_n - L \leq \varepsilon$ and $-\varepsilon \leq z_n - L$, or in other words, $|z_n - L| \leq \varepsilon$. This completes the proof. □

Example 4.3. *We consider the sequence determined by*

$$z_n = \begin{cases} \frac{1}{n} & \text{if } n \text{ is even} \\ -\frac{1}{n} & \text{if } n \text{ is odd} \end{cases}$$

It is not difficult to see that, setting $x_n = -\frac{1}{n}$ and $y_n = \frac{1}{n}$, the hypothesis of Theorem 4.3 are satisfies with $L = 0$. Hence, we also have $z_n \to 0$.

4.1.2 Algebraic combination of sequences

We finish this section by showing that algebraic combinations of convergent sequences also converge and their limits can be obtained in terms of the initial sequences. This will help us to study the convergence of complicated sequence in terms of simple ones.

Theorem 4.4. *Let $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ be two sequence. Suppose that $x_n \to x$ and $y_n \to y$ for some $x, y \in \mathbb{R}$. Then*

1. *$\forall c \in \mathbb{R}$, the sequence $\{c \cdot x_n\}$ converges to $c \cdot x$.*
2. *the sequence $\{x_n + y_n\}$ converges to $x + y$.*
3. *the sequence $\{x_n \cdot y_n\}$ converges to $x \cdot y$.*
4. *if $x \neq 0$, then $\frac{1}{x_n} \to \frac{1}{x}$.*

Proof. Let $\varepsilon \in (0, +\infty)$ be fixed but arbitrary.

1. We can rule out the case $c = 0$, because the sequence defined by $z_n = 0$ converges to 0. Let $N \in \mathbb{N}$ be given by the definition of convergence of $\{x_n\}_{n=1}^{\infty}$ but associated with $\tilde{\varepsilon} = \frac{1}{|c|} \cdot \varepsilon$, that is,
$$\forall n \in \mathbb{N},\ N \leq n \implies |x_n - x| \leq \frac{1}{|c|} \cdot \varepsilon.$$
By the properties of the absolute value we have that for any $n \in \mathbb{N}$ with $N \leq n$ the following holds true:
$$|c \cdot x_n - c \cdot x| \leq |c| \cdot |x_n - x| \leq |c| \cdot \frac{1}{|c|} \cdot \varepsilon = \varepsilon.$$
This means that $c \cdot x_n \to c \cdot x$.

2. Let $N_1, N_2 \in \mathbb{N}$ given by the definition of convergence of $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$, respectively, but associated with $\tilde{\varepsilon} = \frac{1}{2} \cdot \varepsilon$. We set $N = \max\{N_1, N_2\}$, and in particular we have
$$\forall n \in \mathbb{N},\ N \leq n \implies \left(|x_n - x| \leq \frac{1}{2} \cdot \varepsilon \ \wedge \ |y_n - y| \leq \frac{1}{2} \cdot \varepsilon \right).$$
It follows that for any $n \in \mathbb{N}$ with $N \leq n$:
$$|(x_n + y_n) - (x + y)| \leq |x_n - x| + |y_n - y| \leq \frac{1}{2} \cdot \varepsilon + \frac{1}{2} \cdot \varepsilon = \varepsilon.$$
We conclude then that $x_n + y_n \to x + y$.

3. Since $\{x_n\}_{n=1}^{\infty}$ converges, it is bounded. In particular, its absolute value is bounded above by some $c \in (0, +\infty)$ (it is always bounded below by 0). Moreover, we can always assume that $|y| \leq c$, because c is only an upper bound and any number greater than c is also an upper bound for $\{|x_n|\}_{n=1}^{\infty}$.

   Using the same argument as above, we take $N_1, N_2 \in \mathbb{N}$ given by the definition of convergence of $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$, respectively, but associated with $\tilde{\varepsilon} = \frac{1}{2\cdot c} \cdot \varepsilon$ and we also set $N = \max\{N_1, N_2\}$. We take any $n \in \mathbb{N}$ with $N \leq n$ and obtain that

$$|x_n \cdot y_n - x \cdot y| = |x_n \cdot y_n - x_n \cdot y + x_n \cdot y - x \cdot y| \leq |x_n| \cdot |y_n - y| + |y| \cdot |x_n - x| \leq c \cdot |y_n - y| + c \cdot |x_n - x|.$$

   On the other hand, since then terms $|x_n - x|$ and $|y_n - y|$ are both bounded above by $\frac{1}{2\cdot c}|c| \cdot \varepsilon$, we finally obtain that $x_n \cdot y_n \to x \cdot y$, because

$$|x_n \cdot y_n - x \cdot y| \leq \frac{1}{2} \cdot \varepsilon + \frac{1}{2} \cdot \varepsilon = \varepsilon.$$

4. Let $N_1, N_2 \in \mathbb{N}$ be given by the definition of convergence of $\{x_n\}_{n=1}^{\infty}$ but associated with $\tilde{\varepsilon} = \frac{1}{2} \cdot |x|$ and $\tilde{\varepsilon} = \frac{1}{2} \cdot |x|^2 \cdot \varepsilon$. Therefore,

$$\forall n \in \mathbb{N},\ N_1 \leq n \implies |x_n - x| \leq \frac{1}{2} \cdot |x| \tag{4.3}$$

$$\forall n \in \mathbb{N},\ N_2 \leq n \implies |x_n - x| \leq \frac{1}{2} \cdot |x|^2 \cdot \varepsilon \tag{4.4}$$

   Let $N = \max\{N_1, N_2\}$. Note that (4.3) and (4.4) hold as well if we replace N_1 and N_2 with N, respectively.

   On the one hand, from (4.3) we get for any $n \in \mathbb{N}$ with $N \leq n$ that

$$|x| \leq |x - x_n + x_n| \leq |x - x_n| + |x_n| \leq \frac{1}{2} \cdot |x| + |x_n|.$$

   This means that $\frac{1}{2} \cdot |x| \leq |x_n|$ for $n \in \mathbb{N}$ appropriate. In particular, $0 < |x_n|$ and $\frac{1}{|x_n|} \leq 2 \cdot \frac{1}{|x|}$ for all $n \in \mathbb{N}$ that satisfies $N \leq n$.

   On the other hand, using (4.4) and the remark we did above, we can conclude, because

$$\forall n \in \mathbb{N},\ N \leq n \implies \left|\frac{1}{x_n} - \frac{1}{x}\right| = \frac{1}{|x_n| \cdot |x|} \cdot |x_n - x| \leq 2 \cdot \frac{1}{|x|^2} \cdot \frac{1}{2} \cdot |x|^2 \cdot \varepsilon = \varepsilon.$$

□

4.2 Completeness

We turn our attention into an important convergence criterion called **completeness**. This is a very powerful tool that allows us to determine whether a sequence has a limit by only studying the distances between its terms. To be more accurate, we introduce the following definition.

Definition 4.3. *We say that* $\{x_n\}_{n=1}^{\infty}$ *is a* **Cauchy sequence** *if for any* $\varepsilon \in (0, +\infty)$ *the following condition is met*

$$\exists N \in \mathbb{N},\ \forall n, p \in \mathbb{N},\ (N \leq n \implies |x_n - x_{n+p}| \leq \varepsilon). \tag{4.5}$$

Example 4.4. *Let us consider the sequence given by*

$$x_n = \sum_{k=1}^{n} \frac{1}{k^2} := 1 + \frac{1}{4} + \frac{1}{9} + \ldots + \frac{1}{n^2}.$$

This sequence is a Cauchy sequence. Indeed, since $k - 1 \leq k$ *for any* $k \in \mathbb{N}$, *we get*

$$|x_n - x_{n+p}| = \sum_{k=n+1}^{n+p} \frac{1}{k^2} \leq \sum_{k=n+1}^{n+p} \frac{1}{k(k-1)} = \sum_{k=n+1}^{n+p} \frac{1}{k-1} - \frac{1}{k} = \frac{1}{n} - \frac{1}{n+p} \leq \frac{1}{n}.$$

Hence, given $\varepsilon \in (0, +\infty)$ *we know by the Archimedean property that there is* $N \in \mathbb{N}$ *such that* $1 \leq \varepsilon \cdot N$, *and so,* (4.5) *holds too with the* N *we have chosen above.*

It turns out, as we will prove shortly, that a sequence converges if and only if it is a Cauchy sequences. In this case, we say that $\mathbb{R}$ is a **complete space**. This claim shows the utility of the notion of completeness. This criterion doesn't require a priori knowledge about the limit; it can be proven (by more sophisticated means) that the sequence in Example 4.4 converges to $\frac{1}{6} \cdot \pi^2$, which is not obvious from the definition of the sequences. Moreover, it is not difficult to see that the sequence in Example 4.4 is also increasing. However, it is not straightforward to prove that it is bounded above, which may turn Theorem 4.2 difficult to apply.

Theorem 4.5. *Let* $\{x_n\}_{n=1}^{\infty}$ *be a sequence of Real numbers, then* $\{x_n\}_{n=1}^{\infty}$ *converges if and only if it is a Cauchy sequence.*

Proof. The proof consists in two parts. We first prove that if a sequence converges, then it must also be of Cauchy type. Let $\varepsilon \in (0, +\infty)$ be fixed but arbitrary. By the definition of convergence, there is $N \in \mathbb{N}$ such that

$$\forall n \in \mathbb{N},\ \left(N \leq n \implies |x_n - L| \leq \frac{\varepsilon}{2}\right). \tag{4.6}$$

Let $n \in \mathbb{N}$ such that $N \leq n$. It is clear that $N \leq n + p$ for any $p \in \mathbb{N}$. In particular, by (4.6), we have that, if $n, p \in \mathbb{N}$ are as above, we get

$$|x_n - L| \leq \frac{\varepsilon}{2} \ \wedge\ |x_{n+p} - L| \leq \frac{\varepsilon}{2}.$$

Combining these two inequalities we get

$$|x_n - x_{n+p}| = |x_n - L + L - x_{n+p}| \leq |x_n - L| + |L - x_{n+p}| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

In other words, $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence and the first part of the proof is complete.

Let us now assume that $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence and prove that it converges. This part is considerably more difficult and for this reason we divide it into two steps: we first construct a candidate to limit (as the supremum of a particular set) and then we prove that it is actually the limit of the sequence.

1. By the definition of Cauchy sequence, we have that there is $N_1 \in \mathbb{N}$ such that
$$\forall n, p \in \mathbb{N}, \ \left(N_1 \leq n \implies |x_n - x_{n+p}| \leq \frac{1}{2} \right).$$
In a similar way, there is $N_2 \in \mathbb{N}$ such that
$$\forall n, p \in \mathbb{N}, \ \left(N_2 \leq n \implies |x_n - x_{n+p}| \leq \frac{1}{4} \right).$$
Note that we can also assume that $N_1 \leq N_2$ and the preceding statement is still valid. Hence, continuing the process we see that for any $k \in \mathbb{N} \setminus \{0, 1\}$ we can find $N_k \in \mathbb{N}$ such that $N_{k-1} \leq N_k$ and
$$\forall n, p \in \mathbb{N}, \ \left(N_k \leq n \implies |x_n - x_{n+p}| \leq \frac{1}{2^k} \right).$$
Let us now consider the set
$$A = \left\{ x \in \mathbb{R}^N \mid \ \exists k \in \mathbb{N}, \ x = x_{N_k} - \frac{1}{2^k} \right\}.$$
This set is nonempty. Furthermore, it is bounded above. Indeed, let $x \in A$. By definition, there is $k \in \mathbb{N}$ so that $x = x_{N_k} - \frac{1}{2^k}$. In particular,
$$x \leq x_{N_k} = x_{N_k} - x_{N_1} + x_{N_1} \leq |x_{N_k} - x_{N_1}| + x_{N_1} \leq \frac{1}{2} + x_{N_1}.$$
So, $M = \frac{1}{2} + x_{N_1}$ is an upper bound of A. Thus, by the Supremum axiom, $\sup(A) \in \mathbb{R}$. Note that if $\{x_n\}_{n=1}^{\infty}$ were convergent, then each $x_{N_k} - \frac{1}{2^k}$ should approach to the supremum. Consequently, we take $L = \sup(A)$ as our candidate to limit.

2. Given that $L = \sup(A)$, there is $x \in A$ such that $L - \frac{\varepsilon}{3} \leq x$. In other words, there is $k \in \mathbb{N}$ such that
$$x = x_{N_k} - \frac{1}{2^k} \quad \text{and} \quad L \leq x_{N_k} - \frac{1}{2^k} + \frac{\varepsilon}{3}.$$
Note that there are infinitely many $k \in \mathbb{N}$ that verifies the preceding condition; otherwise, L wouldn't be the supremum. The latter means that we can take $k \in \mathbb{N}$ as large as we want. In particular, we can assume that $3 \leq 2^k \cdot \varepsilon$. Therefore, we get that for any $n \in \mathbb{N}$, if $N_k \leq n$ then
$$|x_n - L| = |x_n - x + x - L| \leq |x_n - x| + |x - L| \leq |x_n - x_{N_k}| + \frac{1}{2^k} + |x - L|.$$
On the other hand, we have that $|x - L| = L - x \leq \frac{\varepsilon}{3}$ and by definition $|x_n - x_{N_k}| \leq \frac{1}{2^k}$. So finally, since $3 \leq 2^k \cdot \varepsilon$ we get
$$|x_n - L| \leq \frac{1}{2^k} + \frac{1}{2^k} + \frac{\varepsilon}{3} = \varepsilon,$$
which completes the proof.

$\square$

4.3 Exercises

1. Let $a \in [-1, 1]$ and consider the sequence $\{x_n\}_{n=1}^{\infty}$ defined via
$$x_n = 3 \cdot \frac{a^n}{n^2}.$$
Prove using the definition of convergence that $x_n \to 0$.

2. Let $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ be two sequences. Suppose that $x_n \to x$ and $y_n \to y$ for some $x, y \in \mathbb{R}$. Prove that if $x_n \leq y_n$ for all $n \in \mathbb{N}$, then $x \leq y$. Based on this, determine whether the limit of a sequence of positive numbers can be a negative number or not.

   Hint: Assume that $y < x$ and use $\varepsilon = x - y$ to get a contradiction.

3. Let $\{x_n\}_{n=1}^{\infty}$ be an increasing sequence and $\{y_n\}_{n=1}^{\infty}$ be a decreasing sequence such that $x_n - y_n \to 0$. Show that $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ converge and have the same limit.

4. Let $\{x_n\}_{n=1}^{\infty}$ be the sequence defined via
$$x_1 = 1, \quad x_{n+1} = \sqrt{\frac{9 + x_n^2}{2}}, \ n \in \mathbb{N} \setminus \{0\}.$$
Here $\sqrt{a}$ stands for the unique positive Real number x that verifies $a = x^2$.
   (a) Show that $\{x_n\}_{n=1}^{\infty}$ is bounded above by 3.
   (b) Show that $\{x_n\}_{n=1}^{\infty}$ is increasing and converges to 3.

5. Let $\{x_n\}_{n=1}^{\infty}$ be a Cauchy sequence. Consider another sequence $\{y_n\}_{n=1}^{\infty}$ such that
$$\forall n \in \mathbb{N} \setminus \{0\}, \ |x_n - y_n| \leq \frac{1}{n}.$$
   (a) Prove that $\{y_n\}_{n=1}^{\infty}$ is also a Cauchy sequence.
   (b) Prove that $\{x_n\}_{n=1}^{\infty}$ converges to $L \in \mathbb{R}$ if and only if $\{y_n\}_{n=1}^{\infty}$ converges to $L \in \mathbb{R}$.

6. Let $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ be two sequence that satisfy
   - $\{y_n\}_{n=1}^{\infty}$ is decreasing with $y_n \to 0$ and each y_n being positive.
   - The sequence $\{z_n\}_{n=1}^{\infty}$, is bounded, where $z_n := \sum_{k=1}^{n} x_k$.

   Show that the sequence $\{w_n\}_{n=1}^{\infty}$ converges, where $w_n := \sum_{k=1}^{n} x_k y_k$.

   Hint: Prove that $\{w_n\}_{n=1}^{\infty}$ is a Cauchy sequence. To do so, show that for any $n, p \in \mathbb{N}$ we have
$$w_{n+p} - w_n = y_{n+p} z_{n+p} + \sum_{k=n+1}^{n+p-1} (y_k - y_{k+1}) z_k - y_{n+1} z_n$$

Math 4031 - Advanced Calculus I

Instructor: Dr. Cristopher HERMOSILLA
Louisiana State University - Spring 2016

Week 5: Subsequences and Compactness

We turn our attention into a new mathematical concept called compactness, and we exhibit its relation with sequences of Real numbers. To do this, we need to introduce a new object called a subsequence.

5.1 Subsequences

Let us consider the sequence

$$\{x_n\}_{n=1}^{\infty} := \left\{\frac{1}{n}\right\}_{n=1}^{\infty} = \left\{1,\ \frac{1}{2},\ \frac{1}{3},\ \frac{1}{4},\ \frac{1}{5},\ \frac{1}{6},\ldots\right\}.$$

Note that we can construct another sequence from $\{x_n\}_{n=1}^{\infty}$ by, for example, taking only the terms associated with an even number

$$\{y_k\}_{k=1}^{\infty} := \left\{\frac{1}{2\cdot k}\right\}_{k=1}^{\infty} = \left\{\frac{1}{2},\ \frac{1}{4},\ \frac{1}{6},\ \frac{1}{8},\ \frac{1}{10},\ldots\right\}.$$

In this case, we have used the rule $n = 2\cdot k$ to construct the new sequence $\{y_k\}_{k=1}^{\infty}$, but we could have used any other, as for intance, $n = 2\cdot k+1$. Sequences constructed in this way are called **subsequences** of $\{x_n\}_{n=1}^{\infty}$.

Definition 5.1. *Let $\{x_n\}_{n=1}^{\infty}$ be a given sequence, we say that $\{y_k\}_{k=1}^{\infty}$ is a subsequence of $\{x_n\}_{n=1}^{\infty}$ if there is a strictly increasing sequence of positive Natural numbers $\{n_k\}_{k=1}^{\infty}$, that is, $n_k < n_{k+1}$ for any $k \in \mathbb{N}\setminus\{0\}$, such that*

$$\forall k \in \mathbb{N}\setminus\{0\}, \quad y_k = x_{n_k}. \tag{5.1}$$

We simply write $\{x_{n_k}\}_{k=1}^{\infty}$ to denote such subsequence.

Remark 5.1. *Note that in the preceding definition, since $\{n_k\}_{k=1}^{\infty}$ is an strictly increasing sequence of positive Natural numbers we must have*

$$\forall k \in \mathbb{N}\setminus\{0\}, \quad k \le n_k.$$

Since a subsequence is essentially a sequence, we can talk about its convergence. In particular, we say that a subsequence converges to some limit $L\in\mathbb{R}$ provided that

$$\forall \varepsilon \in (0,+\infty),\ \exists N\in\mathbb{N},\ \forall k\in\mathbb{N},\ (N \le k \implies |x_{n_k} - L| \le \varepsilon). \tag{5.2}$$

Note that (5.2) agrees with the usual definition of convergence but applied to the sequence $\{y_k\}_{k=1}^{\infty}$ that verifies (5.1). Moreover, the collection of all the limits of subsequences of $\{x_n\}_{n=1}^{\infty}$ are called the **accumulation points** of $\{x_n\}_{n=1}^{\infty}$.

Example 5.1. *Let us consider the sequence given by $x_n = (-1)^n$. We have already discussed that this sequence doesn't converge. However, it is not difficult to see that it has several subsequences that converge. For example, the subsequences given by the index $n = 2 \cdot k$ and $n = 2 \cdot k + 1$. In the first case we get the sequence whose elements are all identically 1 and in the other case, the sequence with all the elements being -1. It is clear that both sequences converge, to 1 and -1, respectively. Hence, 1 and -1 are accumulation points of $\{(-1)^n\}_{n=1}^{\infty}$. It can be proved that they are the only accumulation points of $\{(-1)^n\}_{n=1}^{\infty}$. We leave this as exercise for the reader.*

For the purposes of the course, the utility of the notion of subsequence is twofold. We describe them in the next subsections.

5.1.1 Convergence

Let us start with a question. Suppose that a sequence converges, what happens with any of its subsequences? Does it converge too?

It turns out that subsequences can be thought as particular selections of terms of the original sequences. Therefore, it may seem natural that if a sequence converges, then any subsequence must also approach to the limit.

Theorem 5.1. *Let $\{x_n\}_{n=1}^{\infty}$ be a given sequence. Suppose that $x_n \to L$, then any subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ converges to L.*

Proof. Note that by definition we have

$$\forall \varepsilon \in (0, +\infty), \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ (N \leq n \implies |x_n - L| \leq \varepsilon). \tag{5.3}$$

On the other hand, by Remark 5.1, we also have $k \leq n_k$ for any positive Natural number. Therefore, if $N \leq k$ for some $N \in \mathbb{N}$, then $N \leq n_k$, which means that if (5.3) is true, then so must be (5.2). Consequently, any subsequence of $\{x_n\}_{n=1}^{\infty}$ must converges to L as well. □

In practice, the theorem is usually used in the contrapositive form, that is, if you find two subsequences that converge to different limits or a subsequence that diverges, then you can be sure that the sequence you have been provided at the beginning doesn't converge.

Remark 5.2. *A direct consequence of Theorem 5.1 is that a sequence that converges must have a unique accumulation point. The converse of this affirmation doesn't hold in general, that is, a sequence that diverges can have a unique accumulation point. For example, let us consider the sequence determined by*

$$x_n = \begin{cases} 1 & \text{if } n \text{ is odd,} \\ n & \text{if } n \text{ is even.} \end{cases}$$

It is clear that the unique accumulation point of this sequence is 1, however the sequence is not bounded, so it can not converge.

5.1.2 Sequentially compactness

Remark 5.2 shows that there can be sequences having unique accumulation points that don't converge. To avoid this pathological cases we need to rule out the cases of unbounded sequences. For this purpose, we turn our attention into the relation between bounded sequence and subsequences. The most important result concerning this is called the **Bolzano-Weierstrass Theorem**.

Theorem 5.2. *Let $\{x_n\}_{n=1}^{\infty}$ be a given sequence. Suppose $\{x_n\}_{n=1}^{\infty}$ is bounded, then it has a subsequence that converges. In other words, the set of accumulation points of a bounded sequence is nonempty.*

Proof. Let $\{x_n\}_{n=1}^{\infty}$ be a bounded sequence. We are going to construct a subsequence that is also a Cauchy sequence, and the result will follow from the completeness of $\mathbb{R}$.

Let $a, b \in \mathbb{R}$ be a lower and upper bound of $\{x_n\}_{n=1}^{\infty}$, respectively. Let us consider the intervals

$$A_{1,1} = \left[a, \frac{a+b}{2}\right] \quad \text{and} \quad A_{1,2} = \left[\frac{a+b}{2}, b\right].$$

Since the sequence $\{x_n\}_{n=1}^{\infty}$ has infinitely many elements, either $A_{1,1}$ or $A_{1,2}$ contains infinitely many elements of $\{x_n\}_{n=1}^{\infty}$ (not necessarily all the elements of the sequence). We set

$$a_2 = a, \quad b_2 = \frac{a+b}{2}, \quad \text{and} \quad A_1 = A_{1,1}$$

if $A_{1,1}$ contains infinitely many elements of the sequence, otherwise we set

$$a_2 = \frac{a+b}{2}, \quad b_2 = b, \quad \text{and} \quad A_1 = A_{1,2}.$$

We take $n_1 := \min\{n \in \mathbb{N} \setminus \{0\} \mid x_n \in A_1\}$. Notice that $1 \leq n_1$.

We repeat the process but with a_2 and b_2, instead of a and b. Hence, consider the intervals

$$A_{2,1} = \left[a_2, \frac{a_2+b_2}{2}\right] \quad \text{and} \quad A_{2,2} = \left[\frac{a_2+b_2}{2}, b_2\right].$$

By the same arguments used above, either $A_{2,1}$ or $A_{2,2}$ contains infinitely many elements of $\{x_n\}_{n=1}^{\infty}$. We set

$$a_3 = a_2, \quad b_3 = \frac{a_2+b_2}{2}, \quad \text{and} \quad A_2 = A_{2,1}$$

if $A_{2,1}$ contains infinitely many elements of the sequence, otherwise we set

$$a_3 = \frac{a_2+b_2}{2}, \quad b_3 = b_2, \quad \text{and} \quad A_2 = A_{2,2}.$$

Then we take $n_2 := \min\{n \in \mathbb{N} \setminus \{0\} \mid n_1 < n \wedge x_n \in A_2\}$. Note that, since $A_2 \subseteq A_1$ we have that $x_{n_1}, x_{n_2} \in A_1$ and so

$$|x_{n_1} - x_{n_2}| \leq \frac{b-a}{2}.$$

Using a recursive argument we can construct a strictly increasing sequence of positive Natural numbers $\{n_k\}_{k=1}^n$ for which we have

$$|x_{n_k} - x_{n_{k+1}}| \leq \frac{b-a}{2^k}.$$

We claim that $\{x_{n_k}\}_{k=1}^\infty$ is a Cauchy sequence. Indeed, for any $k, p \in \mathbb{N} \setminus \{0\}$ we get

$$|x_{n_k} - x_{n_{k+p}}| \leq \sum_{i=k}^{k+p-1} |x_{n_i} - x_{n_{i+1}}| \leq \sum_{i=k}^{k+p-1} \frac{b-a}{2^i} = \frac{b-a}{2^k} \sum_{i=0}^{p-1} \frac{1}{2^i}.$$

Note that $\sum_{i=0}^{p-1} \frac{1}{2^i} \leq 2$ so we get that

$$\forall k, p \in \mathbb{N} \setminus \{0\}, \quad |x_{n_k} - x_{n_{k+p}}| \leq \frac{b-a}{2^{k-1}}.$$

Hence, given $\varepsilon \in (0, +\infty)$, by the Archimedean property of $\mathbb{R}$, we can find $N \in \mathbb{N}$ such that $\frac{b-a}{2^{N-1}} \leq \varepsilon$. Therefore, since for any $k \in \mathbb{N} \setminus \{0\}$ such that $N \leq k$ we have $2^{N-1} \leq 2^{k-1}$ (because $1 \leq 2$), we get that for any $k, p \in \mathbb{N} \setminus \{0\}$ we have

$$|x_{n_k} - x_{n_{k+p}}| \leq \frac{b-a}{2^{k-1}} \leq \frac{b-a}{2^{N-1}} \leq \varepsilon.$$

Therefore $\{x_{n_k}\}_{k=1}^\infty$ is a Cauchy sequence, and by the completeness of $\mathbb{R}$ it also converges. So, the conclusion follows. □

In view of the preceding theorem and Exercise 4.2 (Week 4 notes), we have that any interval of the form $[a, b]$ with $a, b \in \mathbb{R}$ and $a < b$ satisfies the following property:

Any sequence contained in $[a, b]$ has a subsequence that converges to some $L \in [a, b]$.

This property is known as **sequential compactness** and it can be defined for general sets in the following way.

Definition 5.2. *Let $A \subseteq \mathbb{R}$ be a given set. We say that A is* **sequentially compact** *provided that any sequence $\{x_n\}_{n=1}^\infty$ contained in A has a subsequence that converges to some $L \in A$. In other words, if $x_n \in A$ for any $n \in \mathbb{N}$, then there are $L \in A$ and $\{x_{n_k}\}_{k=1}^\infty$ such that $x_{n_k} \to L$.*

We will study a useful characterization of sequentially compact sets in Exercise 4.

The last result we present is in direct correlation with Remark 5.2. We postpone the proof of the theorem to the Exercises section.

Theorem 5.3. *Let $\{x_n\}_{n=1}^\infty$ be a given sequence. Suppose $\{x_n\}_{n=1}^\infty$ is bounded and has a unique accumulation point, $L \in \mathbb{R}$, then $x_n \to L$.*

5.2 Exercises

1. Find all the accumulations points of the sequence given by
$$x_n := (-1)^n \cdot \sin\left(n \cdot \frac{\pi}{4}\right)$$

2. Let $\{x_n\}_{n=1}^{\infty}$ be a Cauchy sequence. Suppose that $\{x_n\}_{n=1}^{\infty}$ has a subsequence that converge to $L \in \mathbb{R}$. Prove that $x_n \to L$.

3. Let $\{x_n\}_{n=1}^{\infty}$ be a given sequence. Suppose that $\{x_n\}_{n=1}^{\infty}$ is monotonic and has a unique accumulation point $L \in \mathbb{R}$. Show that $x_n \to L$.

4. We say that a set $A \subseteq \mathbb{R}$ is **closed** if the limit of any convergent sequence contained in A belongs to A, that is,
$$(\forall n \in \mathbb{N},\ x_n \in A)\ \wedge\ x_n \to L \implies L \in A.$$
The aim of this exercise is to prove that a set is sequentially compact if and only if it is bounded and closed. We divide the proof in several steps:
   (a) Suppose that A is bounded and closed. Show that A is sequentially compact.
   (b) Suppose that A is sequentially compact. Show that A is closed.
   (c) Suppose that A is sequentially compact. Show that A is bounded.
   Hint: Note that if the set A is unbounded, then there is a sequence contained in A such that $n \leq x_n$ for any $n \in \mathbb{N} \setminus \{0\}$.

5. The goal of this problem is to provide a proof of Theorem 5.3. The idea is to argue by contradiction in the following way.
   (a) Suppose $\{x_n\}_{n=1}^{\infty}$ doesn't converge to L and show that there are $\varepsilon \in (0, +\infty)$ and a sequence $\{y_N\}_{N=1}^{\infty}$ such that
$$\forall N \in \mathbb{N}, \quad (\varepsilon < |y_N - L|\ \wedge\ \exists n \in \mathbb{N},\ y_N = x_n).$$
   (b) Show that no subsequence of $\{y_N\}_{N=1}^{\infty}$ can converge to L.
   (c) Prove that $\{y_N\}_{N=1}^{\infty}$ has a subsequence that converges.
   (d) Prove that any subsequence of $\{y_N\}_{N=1}^{\infty}$ is a subsequence of $\{x_n\}_{n=1}^{\infty}$.
   (e) Find a contradiction and conclude the result.

6. Let $\{a_n\}_{n=1}^{\infty}$ be an increasing sequence and $\{b_n\}_{n=1}^{\infty}$ be decreasing sequence. Suppose that for any $n \in \mathbb{N} \setminus \{0\}$ we have $a_n < b_n$.
   (a) Show that $\bigcap_{n=1}^{\infty} [a_n, b_n]$ is nonempty.
   Hint: Construct a sequence $\{x_n\}_{n=1}^{\infty}$ with the property that $x_n \in [a_n, b_n]$ for any $n \in \mathbb{N}$. To conclude, recall that each interval $[a_n, b_n]$ is sequentially compact.
   (b) Suppose that $a_n - b_n \to 0$. Prove that $\bigcap_{n=1}^{\infty} [a_n, b_n] = \{x\}$ for some $x \in \mathbb{R}$.

Math 4031 - Advanced Calculus I

Instructor: Dr. Cristopher HERMOSILLA
Louisiana State University - Spring 2016

Week 6: Cardinality of some subsets of $\mathbb{R}$

We finish the study of Real numbers with the notion of Cardinality. We now present a formal exposition of the ideas we have described in the short overview of Set Theory. For this purpose, we might need to remember some notions of functions.

Recall that a function $f : A \to B$ is a rule that associates any $a \in A$ with a unique element in B, usually denoted by $f(a)$. Furthermore, we say that $f : A \to B$ is **bijective** if for any $b \in B$ there is a unique $a \in A$ such that $f(a) = b$. In practice, we divide the definition of a bijection in two part, the **injective** and **surjective** properties:

- $f : A \to B$ is injective provided that $f(a_1) = f(a_2) \implies a_1 = a_2$.
- $f : A \to B$ is surjective provided that for any $b \in B$ there is $a \in A$ such that $f(a) = b$.

For any bijective function $f : A \to B$, there is a unique function, $g : B \to A$ such that

$$\forall a \in A,\ g \circ f(a) := g(f(a)) = a \quad \wedge \quad \forall b \in B,\ f \circ g(b) := f(g(b)) = b.$$

The function $g : B \to A$ is called the **inverse function of** f, and it is denoted by f^{-1}.

6.1 Cardinality of a set

Recall that we have said that a set A is finite if there is $n \in \mathbb{N}$ such that A has exactly n (different) elements. The number n was called the **cardinality** of A and was denoted by $|A|$. In particular, this means that any finite set that has n elements has the same cardinality than the set

$$A_n := \{1,\ 2,\ \ldots,\ n\}.$$

If we look this from another point of view, we can also remark that a finite set A has n elements if and only if there is a one-to-one relation between A and the set A_n defined above. This one-to-one relation is the process of counting.

This idea yields to the following definition which applies to infinite sets too.

Definition 6.1. *Let A, B be two given sets, we say that A and B have the same cardinality if there is a bijective function $f : A \to B$. Under these circumstances, we just write $|A| = |B|$.*

Let us come back to the case of finite sets. Our intuition says that we can assign an order to the elements of A, say $a_1,\ a_2,\ \ldots,\ a_n$. Considering this order, we define the function $f : A \to A_n$ via

$$f(a_k) = k.$$

Let us check that $|A| = |A_n|$ according to Definition 6.1. On the one hand we see that $f : A \to A_n$ is injective, because the counting process has assigned a unique element to the k^{th} position. On the other hand, $f : A \to A_n$ is surjective, because any of the n possible positions in the order has been filled with an element of A.

6.2 Countable sets

We now introduce a formal definition of a **countably infinite** set.

Definition 6.2. *Let A be a given set, we say that A is countably infinite if $|A| = |\mathbb{N}|$, that is, there is a bijection between A and $\mathbb{N}$. In this case, we also write $|A| = \aleph_0$.*

Remark 6.1. *Note that according to Definition 6.2, any sequence is a countably infinite set. To see this, it is enough to use the function $f : \{x_n\}_{n=1}^{\infty} \to \mathbb{N}$ defined via*

$$f(x_n) = n - 1.$$

Clearly, this function is a bijection between $\{x_n\}_{n=1}^{\infty}$ and $\mathbb{N}$. Moreover, the converse is also true, that is, any countably infinite set can be described as a sequences. To see this, suppose that $f : A \to \mathbb{N}$ is a bijection, then it is enough to define

$$x_n = f^{-1}(n - 1).$$

We are now in position to prove that $\mathbb{Z}$ is countably infinite, that is, it has the same cardinality than $\mathbb{N}$.

Theorem 6.1. *The set of Integers $\mathbb{Z}$ is countably infinite.*

Proof. Let us consider the function $f : \mathbb{Z} \to \mathbb{N}$ defined via

$$f(z) = \begin{cases} 2 \cdot z & \text{if } 0 \leq z, \\ 2 \cdot (-z) - 1 & \text{if } z < 0. \end{cases}$$

Let us divide the proof into three steps:

1. We first need to check that $f : \mathbb{Z} \to \mathbb{N}$ is well-defined, that is, $f(z) \in \mathbb{N}$ for any $z \in \mathbb{Z}$. Let $z \in \mathbb{Z}$ and suppose that $0 \leq z$, then by definition $z \in \mathbb{N}$ and so $f(z) = 2 \cdot z \in \mathbb{N}$. Assume now that $z < 0$, then by definition $-z \in \mathbb{N}$ and $1 \leq -z$. This means that $2 \cdot (-z)$ is a Natural number greater than or equal to 2, and so $f(z) = 2 \cdot (-z) - 1$ is a Natural number greater than or equal to 1. Hence, the function is well-defined.

2. We now show that $f : \mathbb{Z} \to \mathbb{N}$ is injective. From the preceding part, we have that if $0 \leq z$, then $f(z)$ is even, and if $z < 0$, then $f(z)$ is odd. Hence, since any Natural number is either even or odd, if $f(z_1) = f(z_2)$ we get that $z_1 = z_2$, and so $f : \mathbb{Z} \to \mathbb{N}$ is injective.

3. We finally prove that $f : \mathbb{Z} \to \mathbb{N}$ is surjective. Let $n \in \mathbb{N}$ and suppose that n is even, that is, there is $k \in \mathbb{N}$ such that $n = 2 \cdot k$. But, since $\mathbb{N} \subseteq \mathbb{Z}$ we have that $f(k) = n$. In on the other hand, n is odd, there is $k \in \mathbb{N} \setminus \{0\}$ such that $n = 2 \cdot k - 1$, which means that $n = 2 \cdot k - 1 = f(-k)$. Since, $-k \in \mathbb{Z}$ we get that $f : \mathbb{Z} \to \mathbb{N}$ is surjective, and this completes the proof.

□

6.2.1 The Cantor-Schröeder-Bernstein Theorem and consequences

When two sets A and B are given, it may be too dificult to construct a bijection between both sets to prove that they have the same cardinality. Instead, it may be easier to contruct two injections, one from A into B and another from B into A. It turns out that there is a powerful result, called the **Cantor-Schröeder-Bernstein Theorem**, that says that both strategies are equivalent. We state the theorem, however we don't provide a proof; we refer the interested reader to [1, Theorem 3.1] for details.

Theorem 6.2. *Let A and B be two given sets. Suppose that there are two function, $f : A \to B$ and $g : B \to A$, both injective. Then, there is a bijective function $h : A \to B$.*

We have claimed that the set of Rational numbers is countably infinite. Nevertheless, it proof is slightly more complicated than for $\mathbb{Z}$ and it requires some intermediate results.

We introduce the Cartesian product $\mathbb{N}\times\mathbb{N}$ as the collection all elements that can we written as $\{\{m\}, \{m, n\}\}$ for some $n, m \in \mathbb{N}$. Formally speaking $\mathbb{N} \times \mathbb{N}$ is a subset of $\mathcal{P}(\mathbb{N})$, so it is a set in the sense we have defined at the beginning of the course. For sake of notation, we denote an element of $\mathbb{N} \times \mathbb{N}$ by (m, n).

Lemma 6.1. *The set $\mathbb{N} \times \mathbb{N}$ is countably infinite.*

Proof. By the Cantor-Schröeder-Bernstein Theorem (Theorem 6.2), we only need to exhibit the existence of an injective function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, because it is clear that the function $g : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ defined via $g(n) = (n, n)$ is an injection.

Let us consider the function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ defined via

$$f(m, n) = 2^m \cdot (2 \cdot n + 1) - 1.$$

It is clear that $f(m, n) \in \mathbb{N}$, so we need to prove that it is injective. Take two elements in $\mathbb{N} \times \mathbb{N}$, say (m_1, n_1) and (m_2, n_2), such that $(m_1, n_1) \neq (m_2, n_2)$, that is,

$$m_1 \neq m_2 \quad \vee \quad n_1 \neq n_2.$$

Suppose by contradiction that $f(m_1, n_1) = f(m_2, n_2)$, then

$$2^{m_1} \cdot (2 \cdot n_1 + 1) = 2^{m_2} \cdot (2 \cdot n_2 + 1). \tag{6.1}$$

Suppose that $m_1 = m_2$, then by (6.1) we get $2 \cdot n_1 + 1 = 2 \cdot n_2 + 1$ and so $n_1 = n_2$, which leads to a contradiction. Assume now that $m_1 \neq m_2$. Without loss of generality, we can suppose that $m_1 < m_2$. But, by (6.1) we get

$$2 \cdot n_1 + 1 = 2^{m_2 - m_1} \cdot (2 \cdot n_2 + 1),$$

which is impossible because the left handside is odd but the right handside is even because $2^{m_2 - m_1}$ is a multiple of 2. This case also leads to a contradiction, so $f(m_1, n_1) = f(m_2, n_2)$ can not be true, and $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ must be injective, which finishes the proof. ☐

Remark 6.2. *Lemma 6.1 can be proved without using the Cantor-Schröeder-Bernstein Theorem. It can be proved directly that the function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ given in the proof is also bijective. We leave this as exercise for the reader.*

We are now in position to prove that the set of Rational numbers is countable.

Theorem 6.3. *The set of Rational numbers $\mathbb{Q}$ is countably infinite.*

Proof. Assume that any Rational number can be written as $\frac{p}{q}$ where $p \in \mathbb{Z}$ and $q \in \mathbb{N} \setminus \{0\}$. Then we consider the function $f : \mathbb{Q} \to \mathbb{N} \times \mathbb{N}$ defined via

$$f\left(\frac{p}{q}\right) = \begin{cases} (2 \cdot p, q) & \text{if } 0 \leq p, \\ (2 \cdot (-p) - 1, q) & \text{if } p < 0. \end{cases}$$

It is clear that $f : \mathbb{Q} \to \mathbb{N} \times \mathbb{N}$ is a well-defined function, which in addition is injective. To conclude the proof, we need to show that there is an injection from $\mathbb{N} \times \mathbb{N}$ into $\mathbb{Q}$, then by Theorem 6.2 and Lemma 6.1 the result will follow.

On the one hand, by Lemma 6.1, there is a bijection between $\mathbb{N} \times \mathbb{N}$ and $\mathbb{N}$. On the other hand, there is a canonical injection from $\mathbb{N}$ into $\mathbb{Q}$, that is, $n \mapsto n$. Then, composing both functions we obtain an injection from $\mathbb{N} \times \mathbb{N}$ into $\mathbb{Q}$. Then, in the light of the Cantor-Schröeder-Bernstein Theorem (Theorem 6.2), we get that there is a bijection between $\mathbb{N} \times \mathbb{N}$ and $\mathbb{Q}$, and so $\mathbb{Q}$ must be countably infinite because by Lemma 6.1, $\mathbb{N} \times \mathbb{N}$ is countably infinite. $\square$

6.3 Uncountable sets

We say that a set is **uncountable** if it is infinite and not countably infinite. This yields to the idea that every uncountable set has, in some sense, more elements that $\mathbb{N}$. Recall that we have said that two sets have the same cardinality if there is a bijection between them. Consequently, if two sets have different cardinality, no injection from one of the set in the other can be surjective.

Definition 6.3. *Let A, B be two given sets, we say that the cardinality of A is less than the cardinality of B if there is an injective function $f : A \to B$, but there is no bijection between A and B. Under these circumstances, we just write $|A| < |B|$.*

We now evoke the fact that if A is a finite set, then $|\mathcal{P}(A)| = 2^{|A|}$, which means that $|A| < |\mathcal{P}(A)|$. This fact can be generalized to infinite set and it is known as the **Cantor's Theorem**. This results reads as follows.

Theorem 6.4. *Let A be a given, then $|A| < |\mathcal{P}(A)|$. In particular, $\mathcal{P}(\mathbb{N})$ doesn't have the same cardinality than $\mathbb{N}$ and it is an uncountable set.*

Proof. Let $f : A \to \mathcal{P}(A)$ defined via

$$f(x) = \{x\}.$$

It is clear that $f : A \to \mathcal{P}(A)$ is injective. So to conclude we need to show that there is no bijection between A and $\mathcal{P}(A)$. Suppose by contradiction that there is $g : A \to \mathcal{P}(A)$ bijective. Let us define

$$B = \{x \in A \mid \ x \notin g(x)\}.$$

It is clear that B is a subset of A. Moreover, since $g : A \to \mathcal{P}(A)$ is surjective, there is $b \in A$ such that $g(b) = B$. We now have two possibilities for b, either $b \in B$ or $b \notin B$.

On the one hand, if $b \in B$, then $b \notin g(b)$, but $g(b) = B$, so $b \notin B$, which leads to a contradiction. On the other hand, if $b \notin B$, we get that $b \notin g(b)$, which means that $b \in B$ and we get a contradiction too. Therefore, such function $g : A \to \mathcal{P}(A)$ can not exist and the conclusion follows. $\square$

6.3.1 Cardinality of the set of Real numbers

The set of Real numbers is one of the most important examples of uncountable sets. There are several ways to prove that this set in uncountable, but the one we present is based on the Cantor's Theorem.

We begin by noting that $\mathbb{R}$ has the same cardinality than the interval $(0, 1)$; too see this, it is enough to consider the bijection $f : \mathbb{R} \to (0, 1)$ defined via

$$f(x) = \frac{\exp(x)}{\exp(x) + 1}.$$

So, the fact that $|\mathbb{R}| = |\mathcal{P}(\mathbb{N})|$ will be a direct consequence of the following result.

Theorem 6.5. *$\mathcal{P}(\mathbb{N})$ has the same cardinality than the interval $(0, 1)$.*

Proof. Let us construct two injections, one from $\mathcal{P}(\mathbb{N})$ into $(0, 1)$ and another from $(0, 1)$ into $\mathcal{P}(\mathbb{N})$. Then the conclusion will follow from the Cantor-Schröeder-Bernstein Theorem.

1. Recall that each Real number $x \in (0, 1)$ has a unique decimal representation

$$x = 0.n_1 n_2 n_3 n_4 \ldots$$

   where each $n_k \in \{0,\ 1,\ \ldots,\ 8,\ 9\}$ and there is no repeating sequences of $9's$ at the end. We define $f : (0, 1) \to \mathcal{P}(\mathbb{N})$ via

$$f(x = 0.n_1 n_2 n_3 n_4 \ldots) = \{n_1 \cdot 10,\ n_2 \cdot 10^2,\ n_3 \cdot 10^3,\ \ldots\} = \bigcup_{k=1}^{\infty} \{n_k \cdot 10^k\}.$$

   It is easy to see that $f : (0, 1) \to \mathcal{P}(\mathbb{N})$ is well-defined, so we need to prove that it is injective. Let $x = 0.n_1 n_2 n_3 n_4 \ldots$ and $y = 0.m_1 m_2 m_3 m_4 \ldots$, both different elements of $(0, 1)$. Let $k \in \mathbb{N}$ be the first index such that $n_k \neq m_k$. This index exists, otherwise $x = y$. In particular, $n_k \cdot 10^k \notin f(y)$, so $f(x) \neq f(y)$, which means that $f : (0, 1) \to \mathcal{P}(\mathbb{N})$ is injective.

2. Let us now construct an injection from $\mathcal{P}(\mathbb{N})$ into $(0, 1)$. Let $A \subseteq \mathbb{N}$, we consider the sequence given by

$$x_n^A := \begin{cases} 1 & \text{if } n - 1 \in A, \\ 0 & \text{otherwise.} \end{cases}$$

We consider the function $g : \mathcal{P}(\mathbb{N}) \to (0,1)$ defined by

$$g(A) = 0.1x_1^A x_2^A x_3^A \ldots = 0.1 + \sum_{n=1}^{\infty} x_n^A \cdot 10^{-(n+1)}.$$

We readily see that $g : \mathcal{P}(\mathbb{N}) \to (0,1)$ is well-defined. Furthermore, it is also injective. Indeed, let $A, B \in \mathcal{P}(\mathbb{N})$ with $A \neq B$. Suppose that there is $p \in A \setminus B$. In particular, $x_{p+1}^A = 1$ and $x_{p+1}^B = 0$, so $g(A) \neq g(B)$, because each element in $(0,1)$ admits a unique decimal representation. Therefore, $g : \mathcal{P}(\mathbb{N}) \to (0,1)$ is injective and by the Cantor-Schröeder-Bernstein Theorem the conclusion follows.

□

6.4 Exercises

1. Prove that the set of even and odd Natural numbers are countably infinite.
2. Let $a, b \in \mathbb{R}$ with $a < b$. Prove that the intervals $[0,1]$ and $[a,b]$ have the same cardinality.
3. Prove that the intervals $[0,1]$, $[0,1)$, $(0,1]$ and $(0,1)$ have the same cardinality.
4. Prove that if A and B are countable sets, then $A \cup B$ is countably infinite. Can the set of Irrational numbers be countably infinite?
5. The aim of this problem is to prove that, if X is a countably infinite set and $A \subseteq X$ is infinite, then A is countably infinite. To do so, follow the next steps:
   (a) Show that it is enough to prove the result for the particular case $X = \mathbb{N}$.
   (b) Show that any nonempty subset B of $\mathbb{N}$ has a unique minimum.
   (c) Consider the function $f : \mathbb{N} \to A$ defined recursively via
   $$f(0) = \min(A) \quad \text{and} \quad f(n+1) = \min(A \setminus \{f(0),\ \ldots,\ f(n)\}).$$
   Show that the function is well-defined (the minimum are attained) and that
   $$\forall n \in \mathbb{N},\ f(n) < f(n+1)\ \wedge\ n \leq f(n).$$
   (d) Prove by contraposition that $f : \mathbb{N} \to A$ is injective.
   (e) Prove by contradiction that $f : \mathbb{N} \to A$ is surjective.
   Hint: Suppose first that $a < f(n)$ for some $n \in \mathbb{N}$ and consider the set
   $$B = \{n \in \mathbb{N} \mid\ a < f(n)\}.$$
   Show that $f(n_0) \leq a$ for $n_0 := \min(B)$ and get a contradiction. Finally, study the case in which $f(n) < a$ for any $n \in \mathbb{N}$.

Bibliography

[1] Robert R. Stoll *Set theory and Logic*. New York : Dover Publications, 1979.

Math 4031 - Advanced Calculus I

Instructor: Dr. Cristopher HERMOSILLA
Louisiana State University - Spring 2016

Week 7: Continuous Real-valued functions: Sequential definition

We now begin the study of continuous Real-valued function. We might focus, unless otherwise stated, on the case of functions defined on a bounded closed interval and whose values belong to $\mathbb{R}$. These functions will be generically denoted by $f : [a, b] \to \mathbb{R}$. Furthermore, if the label of the function is of little importance, we might also use the notation $x \mapsto f(x)$ to highlight the expression that defines the function.

Let us emphasize that several of the results discussed in the first part of the course will be applied, in particular those for sequence of Real numbers.

7.1 Limit of functions

The essential notion required to study the continuity of a function is the **limit of a function** at some point on $[a, b]$.

Definition 7.1. *Let $f : [a, b] \to \mathbb{R}$ be a given function. We say that $L \in \mathbb{R}$ is the limit of f at $\bar{x} \in [a, b]$ if* **for any** *sequence $\{x_n\}_{n=1}^{\infty}$ contained in $[a, b]$ that converges to $\bar{x}$, we have that $f(x_n) \to L$. We denote the limit of f at $\bar{x}$ by*

$$\lim_{x \to \bar{x}} f(x).$$

Let us consider the function $f(x) = x^m$ defined on any interval of the form $[a, b]$, where $m \in \mathbb{N}$. We claim that $\lim_{x \to \bar{x}} f(x)$ exists and equals $f(\bar{x})$, regardless the values of a, b or m. To see this, we use an inductive argument:

- **case** $m = 0$**:** Under these circumstances, $f(x) = 1$ for any $x \in [a, b]$ and so clearly, $\lim_{x \to \bar{x}} f(x) = 1$ for each $\bar{x} \in [a, b]$. Indeed, for any sequence that converges to $\bar{x} \in [a, b]$ we must have $f(x_n) = 1$, which means that $f(x_n) \to 1$.

- **case** $m \implies m + 1$**:** We assume that the limit of $x \mapsto x^m$ at $\bar{x} \in [a, b]$ is $\bar{x}^m$. By the algebraic properties of the limit we know that if $x_n \to \bar{x}$ and $y_n \to y$, then $x_n \cdot y_n \to \bar{x} \cdot y$. Let us take $y_n = x_n^m$ and $y = \bar{x}^m$, which is a suitable choice by the induction hypothesis. Hence, $x_n^{m+1} = x_n \cdot x_n^m \to \bar{x} \cdot \bar{x}^m = x^{m+1}$. Therefore, the limit of the mapping $x \mapsto x^{m+1}$ at $\bar{x}$ is $\bar{x}^m$, and the conclusion follows by the induction principle.

7.2 Continuous function

Suppose now that we want to compute the value of $(\sqrt{2})^m$ for some $m \in \mathbb{N}$. We may agree that, in practical terms, this value cannot be computed in an exact way although it can be approximated. One way to do it, is to use the arguments exhibited above. We know that for any sequence $\{x_n\}_{n=1}^{\infty}$ we may take that converges to $\sqrt{2}$, we will get that x_n^m can be as close as we desire from $(\sqrt{2})^m$. If we now want to compute $f(\sqrt{2})$ for another function, it would be very useful to have a similar property as above, that is, that the value $f(\sqrt{2})$ can be approximated by a sequence of terms of the form $f(x_n)$, where $x_n \to \sqrt{2}$. When this happens we say that f is continuous at $\bar{x} = \sqrt{2}$. In general terms we have the following definition.

Definition 7.2. *Let $f : [a, b] \to \mathbb{R}$ be a given function. We say that f is continuous at $\bar{x} \in [a, b]$ if the limit of f at $\bar{x}$ exists and*

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

We say that f is continuous, if it is continuous on each point on its domain $[a, b]$.

In the light of the discussion above, we have that the map $x \mapsto x^m$ is continuous on any interval $[a, b]$ and for any $m \in \mathbb{N}$. It is also, easy to see that $x \mapsto |x|$ is continuous.

Continuity is a property that can be conserved under several operations. The following result is a direct consequence of Theorem 4.4 about algebraic combinations of sequences. We leave the details of the proof as exercise for the reader.

Theorem 7.1. *Let $f : [a, b] \to \mathbb{R}$ and $g : [a, b] \to \mathbb{R}$ be two given functions. Suppose that f and g are continuous at $\bar{x} \in [a, b]$. Then $f + g$ and $f \cdot g$ are continuous at $\bar{x}$. Furthermore, if $g(\bar{x}) \neq 0$, then f/g is also continuous at $\bar{x}$.*

Likewise, continuity is preserved under the composition of functions.

Theorem 7.2. *Let $f : [a, b] \to \mathbb{R}$ and $g : [c, d] \to \mathbb{R}$ be two given functions. Suppose that $f(x) \in [c, d]$ for any $x \in [a, b]$. If f is continuous at $\bar{x} \in [a, b]$ and g is continuous at $f(\bar{x})$, then $g \circ f : [a, b] \to \mathbb{R}$ is continuous at $\bar{x}$.*

Proof. Let $\{x_n\}_{n=1}^{\infty}$ be a sequence contained in $[a, b]$ that converges to $\bar{x}$. Since f is continuous at $\bar{x}$, we get that $y_n := f(x_n) \to f(\bar{x})$. Now, since $f(x) \in [c, d]$ for any $x \in [a, b]$ we have that $\{y_n\}_{n=1}^{\infty}$ is a sequence contained in $[c, d]$ that converges to $f(\bar{x}) \in [c, d]$. Therefore, since g is continuous at $f(\bar{x})$, we get that

$$g \circ f(x_n) = g(f(x_n)) = g(y_n) \to g(f(\bar{x})) = g \circ f(\bar{x}).$$

□

Remark 7.1. *Given two continuous functions $f : [a, b] \to \mathbb{R}$ and $g : [a, b] \to \mathbb{R}$, the maps $x \mapsto \max\{f(x), g(x)\}$ and $x \mapsto \min\{f(x), g(x)\}$ are also continuous on $[a, b]$. Indeed, this is a consequence of Theorem 7.1 and Theorem 7.2. To see this, it is enough to check that both maps can be written as the sum and composition of continuous functions in the following way*

$$\max\{f(x), g(x)\} = \frac{1}{2}\left(f(x) + g(x) + |f(x) - g(x)|\right),$$
$$\min\{f(x), g(x)\} = \frac{1}{2}\left(f(x) + g(x) - |f(x) - g(x)|\right).$$

7.3 The range of a continuous function

We discuss now about the shape of the range of a continuous function. We recall that the range of a function $f : [a, b] \to \mathbb{R}$ is the set defined by

$$f([a, b]) := \{y \in \mathbb{R} \mid \ \exists x \in [a, b], \ y = f(x)\}.$$

Our goal now is to show that is a bounded closed interval. This result will be a consequence of two important theorems for continuous functions, namely, the Intermediate Value Theorem and the Weierstrass Theorem for Extremal Points.

7.3.1 The Intermediate Value Theorem

This result says that $f([a, b])$ is actually an interval and reads as follows.

Theorem 7.3. *Consider a continuous function $f : [a, b] \to \mathbb{R}$. If $c, d \in f([a, b])$ with $c < d$, then $[c, d] \subseteq f([a, b])$.*

Proof. Let $y \in (c, d)$, we need to prove that there is $x \in [a, b]$ such that $y = f(x)$. To do this, we use the same argument we have used to prove the Bolzano-Weierstrass Theorem.

Let $a_0, b_0 \in [a, b]$ such that $c = f(a_0)$ and $d = f(b_0)$, and define $e_0 = \frac{1}{2}(a_0 + b_0)$. Without loss of generality we assume that $a_0 < b_0$. We know then that either $y \le f(e_0)$ or $f(e_0) < y$. So we set

$$a_1 = \begin{cases} a_0 & \text{if } y \le f(e_0) \\ e_0 & \text{otherwise} \end{cases} \quad \text{and} \quad b_1 = \begin{cases} e_0 & \text{if } y \le f(e_0) \\ b_0 & \text{otherwise} \end{cases}$$

Note that in any case we have $f(a_1) \le y$ and $y \le f(b_1)$.

We define then inductively $e_n = \frac{1}{2}(a_n + b_n)$ and set

$$a_{n+1} = \begin{cases} a_n & \text{if } y \le f(e_n) \\ e_n & \text{otherwise} \end{cases} \quad \text{and} \quad b_{n+1} = \begin{cases} e_n & \text{if } y \le f(e_n) \\ b & \text{otherwise} \end{cases}$$

Thus, in this way we have constructed two sequences $\{a_n\}_{n=1}^\infty$ and $\{b_n\}_{n=1}^\infty$, having the following features:

- $\{a_n\}_{n=1}^\infty$ is increasing and bounded above by b.
- $\{b_n\}_{n=1}^\infty$ is decreasing and bounded below by a.
- $b_n - a_n = \frac{1}{2^n}(b - a)$ for any $n \in \mathbb{N} \setminus \{0\}$.
- $f(a_n) \le y$ and $y \le f(b_n)$ for any $n \in \mathbb{N} \setminus \{0\}$.

The first three points implies that $\{a_n\}_{n=1}^\infty$ and $\{b_n\}_{n=1}^\infty$ converges to the same limit, some $x \in [a, b]$. The last point yields, thanks to the continuity of f, to $f(x) \le y$ and $y \le f(x)$, from where we deduce that $y = f(x)$, and thus $y \in f([a, b])$. $\square$

7.3.2 The Weierstrass Theorem for Extremal Points

We have seen that $f([a, b])$ must be an interval (by the Intermediate Value Theorem), we now prove that it must be bounded. Actually, we prove something stronger, that the $\max(f([a, b]))$ and $\min(f([a, b]))$ are well defined, for any continuous function defined on a closed bounded interval. The points where $\max(f([a, b]))$ or $\min(f([a, b]))$ are attained are called the **extremal points** of f.

Definition 7.3. *Let $f : [a, b] \to \mathbb{R}$ be a given function. We say that $x_* \in [a, b]$ is a* **minimum point** *of f on $[a, b]$ if*

$$\forall x \in [a, b], \quad f(x_*) \leq f(x).$$

Similarly, we say that $x^* \in [a, b]$ is a **maximum point** *of f on $[a, b]$ if*

$$\forall x \in [a, b], \quad f(x) \leq f(x^*).$$

Example 7.1. *Consider the function $x \mapsto |x|$ defined on $[-1, 1]$. Since $0 \leq |x|$ for any $x \in [-1, 1]$ and $|x| = 0$ if and only if $x = 0$, we get that $x_* = 0$ is the unique minimum point of the function on $[-1, 1]$. On the other hand, since $|x| \leq 1$ for any $x \in [-1, 1]$ and $|x| = 1$ if and only if $x = 1$ or $x = -1$, we get that $x_1^* = 1$ and $x_2^* = -1$ are both maximum points of the function on $[-1, 1]$.*

Theorem 7.4. *Any continuous function $f : [a, b] \to \mathbb{R}$ has a minimum and a maximum point on $[a, b]$.*

Proof. $m = \inf\{f(x) \mid x \in [a, b]\}$, then there is a sequence $\{x_n\}_{n=1}^{\infty}$ contained in $[a, b]$ such that $f(x_n) \to m$. By the Bolzano-Weierstrass Theorem, that sequence has a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ that converges to some $x_* \in [a, b]$. But, by the continuity of f, the sequence whose terms are $f(x_{n_k})$ converges to $f(x_*)$. However, $\{f(x_{n_k})\}_{k=1}^{\infty}$ is a subsequence of $\{f(x_n)\}_{n=1}^{\infty}$, so it must converge to the same limit, which means that $f(x_*) = m$, and so the infimum is attained at x_* and so

$$\forall x \in [a, b], \quad f(x_*) \leq f(x).$$

Using similar arguments, we can provide the existence of a maximum point $x^* \in [a, b]$. We leave the details to the reader, and so the proof is complete. □

We are now in position to prove that the range of a continuous function defined on a bounded closed interval, is a bounded closed interval as well.

Theorem 7.5. *For any continuous function $f : [a, b] \to \mathbb{R}$, there are $c, d \in \mathbb{R}$ with $c < d$ such that $f([a, b]) = [c, d]$.*

Proof. Let $c = \min(f([a, b]))$ and $d = \max(f([a, b]))$. Thanks to Theorem (7.4) these are well-defined Real numbers.

On the one hand, by definition we get that for any $x \in [a, b]$, $c \leq f(x)$ and $f(x) \leq d$, which leads to $f([a, b]) \subseteq [c, d]$. On the other hand, since $c, d \in f([a, b])$ and $c < d$, by Theorem 7.3 we get that $[c, d] \subseteq f([a, b])$. Therefore, $f([a, b]) = [c, d]$ and the proof is complete. □

7.4 Continuity of the inverse function

One of the consequences of Theorem 7.6 is that, given a continuous function, $f : [a, b] \to \mathbb{R}$ the map $g : [a, b] \to [c, d]$ defined via

$$g(x) := f(x), \quad \text{for any } x \in [a, b]$$

is surjective, where $c, d \in \mathbb{R}$ are such that $[c, d] = f([a, b])$. Hence, if $f : [a, b] \to \mathbb{R}$ is also injective we get that $g : [a, b] \to [c, d]$ is a bijection and its inverse function $g^{-1} : [c, d] \to [a, b]$ is well-defined. Gathering all these facts, we can define a unique function $f^{-1} : [c, d] \to \mathbb{R}$ so that

$$\forall x \in [a, b],\ f^{-1} \circ f(x) = x \quad \text{and} \quad \forall y \in [c, d],\ f \circ f^{-1}(y) = y.$$

For sake of definition, we call the function $f^{-1} : [c, d] \to \mathbb{R}$ the **inverse function of** f.

We now prove that the inverse function of a continuous and injective map is also continuous.

Theorem 7.6. *Let $f : [a, b] \to \mathbb{R}$ be a continuous and injective function, and let $c, d \in \mathbb{R}$ such that $[c, d] = f([a, b])$. Then, $f^{-1} : [c, d] \to \mathbb{R}$, the inverse function f, is continuous on $[c, d]$.*

Proof. Let $y \in [c, d]$ and take a sequence $\{y_n\}_{n=1}^{\infty}$ contained in $[c, d]$ that converges to y. By defintion, there are $x, x_1, x_2, \ldots \in [a, b]$ for which $y = f(x)$ and $y_n = f(x_n)$ for any $n \in \mathbb{N} \setminus \{0\}$. Note as well that $f^{-1}(y) = x$ and $f^{-1}(y_n) = x_n$ for any $n \in \mathbb{N} \setminus \{0\}$. So, to prove the continuity of f^{-1} at y we need to prove that $x_n \to x$. To do so, we show that $\{x_n\}_{n=1}^{\infty}$ has a unique accumulation point, namely x, and since it's bounded, in the light of Theorem 5.3 we obtain that $x_n \to x$.

Let $\{x_{n_k}\}_{k=1}^{\infty}$ a subsequence of $\{x_n\}_{n=1}^{\infty}$ that converges to some $\bar{x} \in [a, b]$. By the continuity of f on $[a, b]$ we get that $f(x_{n_k}) \to f(\bar{x})$ as $k \to +\infty$. However, since $\{f(x_{n_k})\}_{k=1}^{\infty}$ a subsequence of $\{f(x_n)\}_{n=1}^{\infty}$ and $f(x_n) = y_n \to y = f(x)$, by the uniqueness of the limit $f(x) = f(\bar{x})$. Finally, since f is injective, we must have $x = \bar{x}$, and so the conclusion follows. $\square$

7.5 Exercises

1. Let $f : [-1, 1] \to \mathbb{R}$ be a function that satisfies

$$f(x) \leq 0 \text{ if } x \in [-1, 0] \quad \text{and} \quad 1 \leq f(x) \text{ if } x \in (0, 1].$$

Determine whether f is continuous at $x = 0$.

2. Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Show that for any $e \in \mathbb{R}$, the set

$$f^{-1}(e) := \{x \in [a, b] \mid\ f(x) = e\}$$

is closed.

3. Let $a \in (0, \infty)$ and consider a function $f : [-a, a] \to \mathbb{R}$ that satisfies

$$\forall x, y \in [-a, a],\ \forall \lambda \in \mathbb{R},\ \lambda x + y \in [-a, a] \implies f(\lambda x + y) = \lambda f(x) + f(y).$$

Prove that f is continuous at $x = 0$, and then show that it is also continuous on $[-a, a]$.

4. We say that a function $f : [a,b] \to \mathbb{R}$ has a **fixed point** if there is $x \in [a,b]$ such that $f(x) = x$. Prove that if f is continuous on $[a,b]$ such that $a = \min(f([a,b]))$ and $b = \max(f([a,b]))$, then f has a fixed point.

   Hint: Study the sign of the function $x \mapsto f(x) - x$.

5. Let $f : [a,b] \to \mathbb{R}$ be a function that satisfies:

   (1) For any $[a_0, b_0] \subseteq [a,b]$, if $c, d \in f([a_0, b_0])$ with $c < d$, then $[c,d] \subseteq f([a_0, b_0])$.

   (2) For any $e \in \mathbb{R}$, the set $f^{-1}(e) := \{x \in [a,b] \mid f(x) = e\}$ is closed.

   The aim of this problem is to demonstrate that f is continuous on $[a,b]$. To do so, assume by contradiction that f is not continuous at some $x \in [a,b]$ and follow the next steps:

   (a) Show that there are $\varepsilon \in (0, +\infty)$ and a sequence $\{x_n\}_{n=1}^{\infty}$ contained in $[a,b]$ that converges to x so that
   $$\forall n \in \mathbb{N} \setminus \{0\}, \quad \varepsilon < |f(x_n) - f(x)|.$$

   (b) Prove using condition (1) that for any $n \in \mathbb{N} \setminus \{0\}$ there is $y_n \in [a,b]$ such that
   $$|y_n - x| \leq |x_n - x| \quad \text{and} \quad |f(y_n) - f(x)| = \varepsilon.$$

   (c) Prove for any $n \in \mathbb{N} \setminus \{0\}$ either $y_n \in f(f(x) + \varepsilon)$ or $y_n \in f(f(x) - \varepsilon)$. Finally, get a contradiction using condition (2) and conclude.

6. Let $f : [a,b] \to \mathbb{R}$ be a continuous function. Show that there are $x_*, x^* \in [a,b]$ such that
   $$f(x_*) \leq \frac{f(x_1) + f(x_2)}{2} \leq f(x^*), \quad \text{for any } x_1, x_2 \in [a,b].$$
   Using this, prove that for any $x_1, x_2 \in [a,b]$ there is $x \in [a,b]$ such that
   $$f(x) = \frac{f(x_1) + f(x_2)}{2}.$$

7. Let $m \in \mathbb{N} \setminus \{0\}$ be a fixed Natural number and consider the function $f : [0,1] \to \mathbb{R}$ defined via
   $$f(x) = x^m, \quad \text{for any } x \in [0,1].$$
   Prove that its inverse, denoted $\sqrt[m]{\ } : [0,1] \to \mathbb{R}$, is well-defined and continuous on $[0,1]$.

INSTRUCTOR: DR. CRISTOPHER HERMOSILLA
LOUISIANA STATE UNIVERSITY - SPRING 2016

Week 8: Continuous Real-valued functions: topological definition

We continue the study of continuous Real-valued function using another point of view. So far, we have used sequences to define the continuity of a function. We now present a definition based on topological notions, which in this setting means based on open intervals around a point $\bar{x} \in \mathbb{R}$ that have the following form

$$(\bar{x} - r, \bar{x} + r) := \{x \in \mathbb{R} \mid \ \bar{x} - r < x \ \wedge \ x < \bar{x} + r\} = \{x \in \mathbb{R} \mid \ |x - \bar{x}| < r\} \tag{8.1}$$

where $r \in (0, +\infty)$. The set (8.1) is called the **open interval centered at $\bar{x}$ of radius** r.

8.1 Topological definition of continuity

The characterization of the continuity we introduce now is written in terms of $\varepsilon - \delta$.

Theorem 8.1. *Let $f : [a, b] \to \mathbb{R}$ be a given function. Then, f is continuous at $\bar{x} \in [a, b]$ if and only if*

$$\forall \varepsilon \in (0, +\infty), \ \exists \delta \in (0, \infty), \ \forall x \in [a, b], \ |x - \bar{x}| < \delta \implies |f(x) - f(\bar{x})| < \varepsilon. \tag{8.2}$$

Proof. We first see the implication ($\Longleftarrow$), that is, suppose that (8.2) holds. Let $\{x_n\}_{n=1}^{\infty}$ be an arbitrary sequence contained in $[a, b]$ that converges to $\bar{x}$, we need to prove that $f(x_n) \to f(\bar{x})$.

Let $\varepsilon \in (0, +\infty)$ be fixed but arbitrary. Let us consider $\delta \in (0, +\infty)$ given by (8.2) and associated with ε. Since $x_n \to \bar{x}$, there is $N \in \mathbb{N}$ such that $|x_n - \bar{x}| < \delta$ for any $n \in \mathbb{N}$ with $N \leq n$. Hence, by (8.2) we have that $|f(x_n) - f(\bar{x})| < \varepsilon$, or in other words, $f(x_n) \to f(\bar{x})$.

For the other implication we argue by contradiction. Assume that f is continuous at $\bar{x}$ and that (8.2) is false. Then, there is $\varepsilon \in (0, +\infty)$ such that for any $\delta \in (0, +\infty)$ there is $x \in [a, b]$ such that

$$|x - \bar{x}| < \delta \ \wedge \ \varepsilon \leq |f(x) - f(\bar{x})|.$$

By taking, $\delta = \frac{1}{n}$ we can produce a sequence that converges to $\bar{x}$ (by the condition $|x_n - \bar{x}| < \frac{1}{n}$) and such that $\varepsilon \leq |f(x_n) - f(\bar{x})|$. Since $\varepsilon \in (0, +\infty)$ doesn't depend on the sequence nor on δ, the latter implies that $\{f(x_n)\}_{n=1}^{\infty}$ cannot converge to $f(\bar{x})$, which yields to a contradiction. Therefore, (8.2) holds true and the theorem has been proved. $\square$

Remark 8.1. *Let $\bar{x} \in (a, b)$ and $\varepsilon \in (0, +\infty)$ be given. Set $B = (f(\bar{x}) - \varepsilon, f(\bar{x}) + \varepsilon)$, that is, the open interval centered at $f(\bar{x})$ of radius ε. Then, condition (8.2) says that there is another open interval $A = (\bar{x} - \delta, \bar{x} + \delta)$, centered at $\bar{x}$ of radius δ, such that*

$$A \subseteq f^{-1}(B) := \{x \in [a, b] \mid f(x) \in B\}.$$

Hence, from a topological point of view, a continuous function is a function, for which the pre-image of an open interval centered at $f(\bar{x})$ contains an open interval centered at $\bar{x}$.

In the cases $\bar{x} = a$ or $\bar{x} = b$, the same is true but with the slight modification that, instead of the open interval $A = (\bar{x} - \delta, \bar{x} + \delta)$, we need to consider the open intervals relative to $[a, b]$

$$[a, \bar{a} + \delta) \quad \wedge \quad (\bar{x} - \delta, b].$$

It's important to remark that in (8.2), δ depends in general on ε and the interval $[a, b]$, we will see later that δ can always be taken independently of $\bar{x}$. Let's see some example:

- Take $f(x) = x^m$ for some $m \in \mathbb{N}$ defined on any interval $[a, b]$. Take $\bar{x} \in [a, b]$ fixed.
  - If $m = 0$, we get $|f(x) - f(\bar{x})| = 0$, so (8.2) holds immediately for any $\delta \in (0, +\infty)$. In this case, δ doesn't depend on ε, $\bar{x}$ nor the interval $[a, b]$.
  - If $m = 1$, we have $|f(x) - f(\bar{x})| = |x - \bar{x}|$, so taking $\varepsilon = \delta$ in (8.2) the condition holds. In this case, δ depends only on ε.
  - For the case $m \in \mathbb{N} \setminus \{0, 1\}$ we use the fact that

    $$x^m - \bar{x}^m = (x - \bar{x})(x^{m-1} + x^{m-2}\bar{x} + x^{m-3}\bar{x}^2 + \ldots + x^2\bar{x}^{m-3} + x\bar{x}^{m-2} + \bar{x}^{m-1}).$$

    Note that for any $x \in [a, b]$, including $\bar{x}$, we have $|x| \leq \max\{|a|, |b|\}$. Hence,

    $$\begin{aligned} |x^{m-1} + x^{m-2}\bar{x} + \ldots + x\bar{x}^{m-2} + \bar{x}^{m-1}| &\leq |x|^{m-1} + |x|^{m-2}|\bar{x}| + \ldots + |x||\bar{x}|^{m-2} + |\bar{x}|^{m-1}| \\ &\leq m \cdot \max\{|a|, |b|\}^{m-1}. \end{aligned}$$

    Therefore, combining these inequalities we

    $$|x^m - \bar{x}^m| \leq |x - \bar{x})| \cdot m \cdot \max\{|a|, |b|\}^{m-1}.$$

    So,

    $$\delta = \frac{\varepsilon}{m \cdot \max\{|a|, |b|\}^{m-1}}$$

    works for (8.2). Note that in this case δ depends on ε and the interval $[a, b]$. Note that δ decreases its values as long as $|a|$ or $|b|$ increase their values.

- Consider a function $f : [a, b] \to \mathbb{R}$ that satisfies

  $$\forall x, y \in [a, b],\ \forall \lambda \in \mathbb{R},\ \lambda x + y \in [a, b] \implies f(\lambda x + y) = \lambda f(x) + f(y). \tag{8.3}$$

  Let see that (8.2) holds at $\bar{x} \in [a, b]$. Assume that $a \neq 0$, otherwise use b instead of a. Note that (8.3) yields to

  $$f(x) = f\left(\frac{x - \bar{x}}{a} \cdot a + \bar{x}\right) = \frac{x - \bar{x}}{a} \cdot f(a) + f(\bar{x}). \tag{8.4}$$

Hence,

$$|f(x) - f(\bar{x})| = |x - \bar{x}| \cdot \left|\frac{f(a)}{a}\right|.$$

If on the one hand, $f(a) = 0$ we get that the right-hand side equals zero, in which case, any $\delta \in (0, +\infty)$ makes (8.2) to hold. On the other hand, if $f(a) \neq 0$, then it's enough to take $\delta = \frac{\varepsilon \cdot |a|}{|f(a)|}$ and (8.2) will hold.

In the latter case, we can prove that condition (8.3) implies that $\frac{f(x)}{x}$ is constant for any $x \in [a, b] \setminus \{0\}$. Indeed, this comes from evaluating (8.4) at $\bar{x} = a$; if $a = 0$ the same arguments work with b instead of a. In particular, we get that there is $c \in [0 + \infty)$, which doesn't depend on $\bar{x}$ such that $f(a) = c \cdot a$. This means that in the last example we have

$$\forall x, y \in [a, b], \quad |f(x) - f(y)| \leq c \cdot |x - y|. \tag{8.5}$$

Functions that satisfies (8.5) receive a special name, they are called **Lipchitz continuous**. Moreover, the non-negative Real number c in (8.5) is called a Lipschitz constant of f on $[a, b]$.

It is worthy to note that Lipschitz continuous functions, are actually continuous maps, and have the remarkable property that δ in (8.2) can be always taken independently of $\bar{x}$. In fact, $\delta = \varepsilon \cdot \frac{1}{c}$ always works, where $c \in \mathbb{R}$ is a positive Lipschitz constant of f on $[a, b]$.

All the examples we have seen so far are actually Lipschitz continuous function, however, there are functions that don't have this property. For instances, the function $x \mapsto \sqrt{x}$ defined on $[0, 1]$ is not Lipschitz continuous. Indeed, if that were the case, there would be $c \in (0, +\infty)$ so that

$$\forall x, \bar{x} \in [0, 1], \quad |\sqrt{x} - \sqrt{\bar{x}}| \leq c \cdot |x - \bar{x}|.$$

Evaluating this at $\bar{x} = 0$ we get $\sqrt{x} \leq c \cdot x$ for any $x \in (0, 1]$. Since, by definition $x = \sqrt{x} \cdot \sqrt{x}$, we get then that $x \leq c^2 \cdot x^2$, so dividing by x we finally get $1 \leq c^2 \cdot x$. We know that the function $x \mapsto c^2 \cdot x$ is continuous at $x = 0$, so letting x goes to 0 we finally get $1 \leq 0$, which cannot be.

Although the $x \mapsto \sqrt{x}$ on $[0, 1]$ is not Lipschitz continuous, it does satisfies an interesting property (we leave the details of the inequality as exercise for the reader):

$$\forall x, y \in [0, 1], \quad |\sqrt{x} - \sqrt{y}| \leq \sqrt{|x - y|}.$$

Let us point out that $x \mapsto \sqrt{x}$ is then continuous on $[0, 1]$. Indeed, for any $\varepsilon \in (0, +\infty)$, we just need to take $\delta = \varepsilon^2$ and (8.2) will be true for any $\bar{x} \in [0, 1]$. Moreover, any function that satisfies a similar inequality is called **Hölder continuous**. More precisely, given $\alpha \in \mathbb{Q} \cap (0, 1)$ we say that a function $f : [a, b] \to \mathbb{R}$ is α-Hölder continuous if

$$\forall x, y \in [a, b], \quad |f(x) - f(y)| \leq c \cdot |x - y|^{\alpha}. \tag{8.6}$$

The notation x^α for $x \in [0, +\infty)$ stands for the Real number $\sqrt[q]{x^p}$, where $p \in \mathbb{Z}$ and $q \in \mathbb{N} \setminus \{0\}$ are so that $q \cdot \alpha = p$. In particular, we say that $x \mapsto \sqrt{x}$ is $\frac{1}{2}$-Hólder continuous.

The functions we have reviewed have the particularity that, to prove their continuity properties using (8.2), we have provided a $\delta \in (0, +\infty)$ that doesn't depend on the point $\bar{x}$ we are studying the continuity. So to speak, we have taken the parameter δ uniformly on the interval

$[a,b]$. This fact, is actually general for function defined on bounded closed intervals; it is not necessarily true if the domain is not bounded. Therefore, we say that a function $f:[a,b]\to\mathbb{R}$ is **uniformly continuous** on $[a,b]$ provided that

$$\forall \varepsilon\in(0,+\infty),\ \exists\delta\in(0,+\infty),\ \forall x,y\in[a,b],\ |x-y|<\delta \implies |f(x)-f(y)|<\varepsilon. \tag{8.7}$$

It is clear that a uniformly continuous function is continuous. We now will prove that the converse is also true for function defined on closed bounded intervals.

Theorem 8.2. *Let $f:[a,b]\to\mathbb{R}$ be a given continuous function. Then, f is uniformly continuous on $[a,b]$.*

Proof. Assume by contradiction that (8.7) doesn't hold. Then, there is $\varepsilon\in(0,+\infty)$ and two sequence $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ contained in $[a,b]$ such that

$$\forall n\in\mathbb{N}\setminus\{0\},\quad |x_n-y_n|<\frac{1}{n}\quad\text{and}\quad \varepsilon\le|f(x_n)-f(y_n)|.$$

By the Bolzano-Weierstrass Theorem, $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ have converging subsequences, say to $\bar{x}$ and $\bar{y}$, respectively. We can assume that both subsequences have the same numeration, that is, $x_{n_k}\to\bar{x}$ and $y_{n_k}\to\bar{y}$ as $k\to+\infty$. Since $x_n-y_n\to 0$ as $n\to+\infty$, then $x_{n_k}-y_{n_k}\to 0$ as $k\to+\infty$, which means that $\bar{x}=\bar{y}$. Moreover, by continuity of f, we get that $f(x_{n_k})-f(y_{n_k})\to f(\bar{x})-f(\bar{y})=0$, because $\bar{x}=\bar{y}$, but this implies that $\varepsilon\le 0$, which cannot be. We conclude then that f is uniformly continuous on $[a,b]$. □

8.2 Exponential functions

We now present a function that plays a fundamental role in Calculus, the so-called exponential function $\exp:\mathbb{R}\to\mathbb{R}$ define via

$$\exp(x)=\lim_{n\to+\infty}\left(1+\frac{x}{n}\right)^n,\quad \text{for any } x\in\mathbb{R}.$$

Sometimes, it is also denoted by $x\mapsto e^x$. Let us begin the study by proving that the function is well-defined, which in this case means that, given $x\in\mathbb{R}$ **fixed** the limit of $\{f_n(x)\}_{n=1}^{\infty}$ exists and it is a Real number, where

$$f_n(x):=\left(1+\frac{x}{n}\right)^n.$$

- For any $n\in\mathbb{N}$ and $h\in(-1,+\infty)$, we have

  $$1+n\cdot h\le(1+h)^n. \tag{8.8}$$

  To see we use induction. The case $n=0$ is trivial, let us just focus on the inductive step. Suppose that the result is true for $n\in\mathbb{N}$, then

  $$1+(n+1)\cdot h\le 1+(n+1)\cdot h+n\cdot h^2=(1+n\cdot h)\cdot(1+h)\le(1+h)^n\cdot(1+h)=(1+h)^{n+1}.$$

- Let $N := \max\{n \in \mathbb{N} \mid x > -N\}$, then $\{f_{N+n}(x)\}_{n=1}^{\infty}$ is increasing. Indeed, note first of all that for any $n \in \mathbb{N} \setminus \{0\}$, we have $f_{N+n}(x) > 0$. Hence, for $n \in \mathbb{N} \setminus \{0\}$ fixed we obtain
$$\frac{f_{N+n+1}(x)}{f_{N+n}(x)} = \left(1 - \frac{x}{(N+n+1)(N+n+x)}\right)^{N+n+1} \cdot \left(1 + \frac{x}{N+n}\right).$$
On the other hand, $\frac{x}{(N+n+x)} < 1$, so by (8.8) we get
$$1 = \left(\frac{N+n}{N+n+x}\right) \cdot \left(\frac{N+n+x}{N+n}\right) = \left(1 - \frac{x}{N+n+x}\right) \cdot \left(1 + \frac{x}{N+n}\right) \leq \frac{f_{N+n+1}(x)}{f_{N+n}(x)}$$

- For any $n \in \mathbb{N} \setminus \{0\}$, if $x < 1$ we have
$$f_{N+n}(x) = \left(\frac{N+n+x}{N+n}\right)^{N+n} = \left(\frac{1}{1 - \frac{x}{N+n+x}}\right)^{N+n}.$$
Since $\frac{x}{N+n+x} < 1$ we get by (8.8) that
$$f_{N+n}(x) \leq \frac{1}{1-x}.$$
Therefore, $\{f_n(x)\}_{n=1}^{\infty}$ is bounded provided $x < 1$.

- For any $n \in \mathbb{N} \setminus \{0\}$, if $1 \leq x$, by the Archimedean property, there is $m \in \mathbb{N}$ so that $\frac{x}{m} < 1$, and since $\{f_n(x)\}_{n=1}^{\infty}$ is increasing we have
$$f_{N+n}(x) \leq f_{m\cdot(N+n)}(x) = \left(1 + \frac{x}{m\cdot(N+n)}\right)^{m\cdot(N+n)} \leq \left(\frac{1}{1 - \frac{x}{m}}\right)^{m}.$$
Thus, $\{f_n(x)\}_{n=1}^{\infty}$ is also bounded above if $1 \leq x$.

Gathering all there affirmations we get that, for any $x \in \mathbb{R}$, the sequence $\{f_{N+n}(x)\}_{n=1}^{\infty}$ is increasing and bounded below, so it must converge to some positive Real number. It is also clear that the limit of $\{f_{N+n}(x)\}_{n=1}^{\infty}$ converges, then $\{f_n(x)\}_{n=1}^{\infty}$ also converges, which means that $x\exp(x)$ is well defined.

8.3 Exercises

1. Let $f : [a,b] \to \mathbb{R}$ be a continuous function. Suppose that $f(\bar{x}) \in (0,+\infty)$. Show that there is $\delta \in (0,+\infty)$ such that
$$\forall x \in (\bar{x} - \delta, \bar{x} + \delta) \cap [a,b], \quad f(x) \in (0,+\infty).$$

2. Let $f : [a,b] \to \mathbb{R}$ be a Lipschitz continuous function with Lipschitz constant $c \in (0,1)$. Prove that f has a unique fixed point, that is, there is a unique $x \in [a,b]$ such that $f(x) = x$. **Hint:** Consider a sequence defined inductively as follows:
$$\forall n \in \mathbb{N} \setminus \{0\}, \quad x_{n+1} = f(x_n),$$
where $x_1 \in [a,b]$ is any point. Show that this sequence is a Cauchy sequence.

3. $f : \mathbb{R} \to \mathbb{R}$ be a continuous function such that
$$\forall x, y \in \mathbb{R}, \quad |x - y| \leq |f(x) - f(y)|.$$
The aim of this problem is to show that $f(\mathbb{R})$, the range of f, equals $\mathbb{R}$.
   (a) Prove that $f(\mathbb{R})$ is a closed interval, that is, either $[c, d]$, $[c, +\infty)$, $(-\infty, d]$ or $\mathbb{R}$.
   (b) Prove that for any $x \in \mathbb{R}$ and $y \in (0, +\infty)$ with $x \neq y$ we have
$$f(x + y) < f(x) \quad \vee \quad f(x - y) < f(x).$$
   Hint: Assume the statement is false and find a contradiction using the Intermediate Value Theorem.
   (c) Prove that the only possible option is that $f(\mathbb{R})$ equals $\mathbb{R}$.

4. Let $f : (a, b) \to \mathbb{R}$ be a function that is uniformly continuous on (a, b), that is, satisfies (8.7) when replacing $[a, b]$ with (a, b). Prove that there is a unique continuous function $g : [a, b] \to \mathbb{R}$ such that
$$\forall x \in (a, b), \quad f(x) = g(x).$$
Hint: Study the existence of the limit of $\{f(x_n)\}_{n=1}^{\infty}$ for any sequence contained in (a, b) such that $x_n \to a$ or $x_n \to b$.

5. The goal of this problem is to prove that $x \mapsto \exp(x)$ doesn't satisfy (8.7) when replacing $[a, b]$ with $\mathbb{R}$, that is, it is not uniformly continuous on $\mathbb{R}$.
   (a) Prove using the definition of $x \mapsto \exp(x)$ that it has the **semi-group property**, that is,
$$\forall x, y \in \mathbb{R}, \quad \exp(x + y) = \exp(x) \cdot \exp(y).$$
   Hint: Prove that if $f_n(x) := \left(1 + \frac{x}{n}\right)^n$, then for some sequence $\{z_n\}_{n=1}^{\infty}$ that converges to zero, we have
$$1 - z_n \leq \frac{f_n(x + y)}{f_n(x) \cdot f_n(y)} \leq \frac{1}{1 + z_n}$$
   (b) Show that $\exp(1) \in [2, +\infty)$ and that $\exp(n) = \exp(1)^n$.
   (c) Conclude using a contradiction argument and the previous parts.

Instructor: Dr. Cristopher HERMOSILLA
Louisiana State University - Spring 2016

Week 9: Uniform convergence of Continuous Real-valued functions

9.1 Pointwise convergence

We have defined the exponential function $\exp : \mathbb{R} \to \mathbb{R}$ via

$$\exp(x) = \lim_{n\to+\infty} \left(1 + \frac{x}{n}\right)^n, \quad \text{for any } x \in \mathbb{R}.$$

We have proved that the limit exists and it is a positive Real number. In other words, we have shown that the sequence of function $f_n : \mathbb{R} \to \mathbb{R}$ given by

$$f_n(x) := \left(1 + \frac{x}{n}\right)^n$$

satisfies the following property:

$$\forall x \in \mathbb{R}, \quad f_n(x) \to \exp(x).$$

In this case we say that f_n **converges pointwise to** $\exp(\cdot)$ **on any interval** $[a, b]$**.** In general terms we have the following definition.

Definition 9.1. *Let $f : [a, b] \to \mathbb{R}$ be a given function and $\{f_n\}_{n=1}^{\infty}$ be a sequence functions on $[a, b]$, that is, $f_n : [a, b] \to \mathbb{R}$ for each $n \in \mathbb{N} \setminus \{0\}$. We say that f_n converges pointwise to f on $[a, b]$ if*

$$\forall x \in [a, b],\ \forall \varepsilon \in (0, +\infty),\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ N \leq n \implies |f_n(x) - f(x)| \leq \varepsilon, \tag{9.1}$$

or in other words, for each $x \in [a, b]$ we have $f_n(x) \to f(x)$ as a sequence of Real numbers. Under these circumstances, we write $f_n \to f$ pointwise on $[a, b]$.

Let us point out a few things about pointwise convergence.

- Consider $\{f_n\}_{n=1}^{\infty}$, a sequence of continuous functions on $[a, b]$, that is, each f_n is continuous on $[a, b]$. If $f_n \to f$ on $[a, b]$, to some function $f : [a, b] \to \mathbb{R}$, then f can be continuous but also discontinuous:

1. Consider $x \mapsto f_n(x) := \left(1 + \frac{x}{n}\right)^n$, we know that each f_n is continuous on any interval $[a, b]$. The exponential function is continuous at any $x \in \mathbb{R}$. Indeed, since for any $x \in (-1, 1)$ we have
$$1 + x \leq \exp(x) \leq \frac{1}{1-x}$$
Then, by the Squeeze theorem $x \mapsto \exp(x)$ is continuous at $\bar{x} = 0$. Moreover, since
$$\exp(\bar{x}) - \exp(x) = \exp(\bar{x}) \cdot (1 - \exp(x - \bar{x})),$$
we get that $x \mapsto \exp(x)$ is continuous at any $\bar{x} \in \mathbb{R}$. Hence, $\{f_n\}_{n=1}^{\infty}$ converges pointwise on any interval $[a, b]$ to a continuous function, namely, to $x \mapsto \exp(x)$.

2. Consider the sequence of functions defined via $x \mapsto x^n$ defined on $[0, 1]$. We have that setting $f_n(x) = x^n$
$$f_n(0) = 0 \quad \text{and} \quad f_n(1) = 1.$$
Furthermore, since for any $x \in (0, 1)$ we have that $x^n \to 0$ as $n \to +\infty$. Indeed, for any $x \in (0, 1)$ fixed, there is $k \in \mathbb{N}$ such that $x \leq \frac{k}{k+1}$, and so,
$$x^n \leq \left(\frac{k}{k+1}\right)^n = \frac{1}{(1 + \frac{1}{k})^n} \leq \frac{1}{1 + \frac{n}{k}} = \frac{k}{k+n}.$$
Since the right handside converges to zero as $n \to +\infty$, by the Squeeze Theorem we get our initial claim. Hence, the sequence of function converges to the following discontinuous function.
$$f(x) = \begin{cases} 0 & \text{if } t \in [0, 1), \\ 1 & \text{if } t = 1. \end{cases}$$

- $N \in \mathbb{N}$ given by (9.1) depends in general on $x \in [a, b]$ as well as $\varepsilon \in (0, +\infty)$. This means that we cannot expect to have in every situation the same N for all $x \in [a, b]$. To see this, consider the following sequence of functions on $[0, 1]$:
$$f_n(x) = \begin{cases} n \cdot x & \text{if } t \in [0, 1\frac{1}{n}], \\ 2 - n \cdot x & \text{if } t \in (\frac{1}{n}, \frac{2}{n}] \\ 0 & \text{if } t \in (\frac{2}{n}, 1]. \end{cases}$$
Clearly $f_n(0) \to 0$, and furthermore, since for any $x \in (0, 1]$ there is $n \in \mathbb{N}$ such that $x < \frac{2}{n}$, we get that $f_n \to 0$ pointwise on $[0, 1]$. Let $k \in \mathbb{N} \setminus \{0\}$ and take $\varepsilon \in (0, 1)$, by (9.1) there is $N \in \mathbb{N}$ so that
$$N \leq n \quad \Rightarrow \quad |f_n\left(\frac{1}{k}\right)| < \varepsilon < 1.$$
But, $f_k(\frac{1}{k}) = 1$, which means that $N \leq k$ is not possible, and so we must have $k < N$.

These remarks show that the notion of pointwise converge is not strong enough to preserve continuity of function when passing into the limit. For these reasons we introduce a new notion of convergence for functions.

Definition 9.2. *Let $f : [a,b] \to \mathbb{R}$ be a given function and $\{f_n\}_{n=1}^{\infty}$ be a sequence functions on $[a,b]$, that is, $f_n : [a,b] \to \mathbb{R}$ for each $n \in \mathbb{N} \setminus \{0\}$. We say that f_n* **converges uniformly** *to f on $[a,b]$ if*

$$\forall \varepsilon \in (0,+\infty),\ \exists N \in \mathbb{N},\ \forall n \in \mathbb{N},\ \forall x \in [a,b],\ N \leq n \implies |f_n(x) - f(x)| \leq \varepsilon, \tag{9.2}$$

Under these circumstances, we write $f_n \to f$ uniformly on $[a,b]$.

Note that on (9.2) the condition

$$\forall x \in [a,b],\ N \leq n \implies |f_n(x) - f(x)| \leq \varepsilon$$

is equivalent to

$$N \leq n \implies \|f_n - f\|_\infty \leq \varepsilon$$

where $\|g\|_\infty = \sup\{|g(x)| \mid x \in [a,b]\}$ for any function $g : [a,b] \to \mathbb{R}$, is called the **sup-norm** of g on $[a,b]$. Consequently, (9.2) simply means that each $\|f_n - f\|_\infty$ is a Real number and $\|f_n - f\|_\infty \to 0$ as $n \to +\infty$; this must be understood as convergence as Real number.

We now show that the notion of uniform convergence is more appropriate to handle continuous functions rather than the pointwise convergence.

Theorem 9.1. *Let $f : [a,b] \to \mathbb{R}$ be a given function and $\{f_n\}_{n=1}^{\infty}$ be a sequence of continuous functions on $[a,b]$, that is, each $f_n : [a,b] \to \mathbb{R}$ is continuous. Suppose that $f_n \to f$ uniformly on $[a,b]$, then f is continuous on $[a,b]$.*

Proof. Let $\bar{x} \in [a,b]$ and $\varepsilon \in (0,+\infty)$. Since $f_n \to f$ uniformly on $[a,b]$, there is $N \in \mathbb{N}$ such that

$$\|f_N - f\|_\infty \leq \frac{\varepsilon}{3}.$$

Furthermore, since $x \mapsto f_N(x)$ is continuous, there is $\delta \in (0,+\infty)$ such that

$$\forall x \in [a,b],\ |x - \bar{x}| < \delta \quad \implies \quad |f_N(x) - f_N(\bar{x})| < \frac{\varepsilon}{3}.$$

Also, note that for any $x \in [a,b]$, including $\bar{x}$ we have

$$|f(x) - f_N(x)| \leq \|f - f_N\|_\infty.$$

Hence, combining all these inequalities, for any $x \in [a,b]$ with $|x - \bar{x}| < \delta$ we get

$$\begin{aligned} |f(x) - f(\bar{x})| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(\bar{x})| + |f_N(\bar{x}) - f(\bar{x})| \\ &\leq \|f - f_N\|_\infty + \frac{\varepsilon}{3} + \|f_n - f\|_\infty \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

From where we conclude that f is continuous at $\bar{x}$, and since this is a generic point, we have proved that f is continuous on $[a,b]$. $\square$

9.2 Cauchy criterion

From this onward we start noticing that the set of Real numbers and the set of all continuous functions share several properties. For example, both are closed under:

- algebraic combinations, that is algebraic combinations of (Real numbers)/(continuous functions) are a (Real numbers)/(continuous functions).
- limiting process, that is, the limit of a convergent Sequence of (Real numbers)/(continuous functions) are (Real numbers)/(continuous functions).

In the last point we are of course talking about uniform convergence. Many others properties are common to Real numbers and continuous functions, and the key point is to let $\|f\|_\infty$ play the role absolute value.

Recall that a sequence of Real numbers $\{x_n\}_{n=1}^\infty$ converges if and only if it is a Cauchy sequence, that is,

$$\forall \varepsilon \in (0,+\infty),\ \exists N \in \mathbb{N}\setminus\{0\},\ \forall n,p \in \mathbb{N}, \quad N \le n \implies |x_{n+p} - x_n| \le \varepsilon.$$

In a similar way we can define a **Cauchy sequence of functions**.

Definition 9.3. *Let $\{f_n\}_{n=1}^\infty$ be a sequence of functions on $[a,b]$. We say that $\{f_n\}_{n=1}^\infty$ is a Cauchy sequence of functions provided that*

$$\forall \varepsilon \in (0,+\infty),\ \exists N \in \mathbb{N}\setminus\{0\},\ \forall n,p \in \mathbb{N}, \quad N \le n \implies \|f_{n+p} - f_n\|_\infty \le \varepsilon. \tag{9.3}$$

Similarly as for Real numbers, if $\{f_n\}_{n=1}^\infty$ converges uniformly to some function, then it is a Cauchy sequence of functions; we leave the details as exercise for the reader. The converse, as well as for Real numbers turns out to be true.

Theorem 9.2. *Let $\{f_n\}_{n=1}^\infty$ be a Cauchy sequence of continuous functions on $[a,b]$, then there is a function $f:[a,b]\to\mathbb{R}$ continuous such that $f_n \to f$ uniformly on $[a,b]$.*

Proof. The proof is divide into two parts, first we construct a candidate to limit, and then we prove that the candidate is actually a uniform limit of $\{f_n\}_{n=1}^\infty$.

- Note that for any $x \in [a,b]$ we have

  $$|f_{n+p}(x) - f_n(x)| \le \|f_{n+p} - f_n\|_\infty.$$

  Hence, for any $x \in [a,b]$, the sequence of Real numbers $\{f_n(x)\}_{n=1}^\infty$ is a Cauchy sequence (in $\mathbb{R}$). Hence, by the completeness of $\mathbb{R}$ the sequence $\{f_n(x)\}_{n=1}^\infty$ converges. Let us call this limit L_x (the subindex is because the limit depends on x). Note that this can be done for any $x \in [a,b]$, and so, we can define a function $f:[a,b]\to\mathbb{R}$ via

  $$\forall x \in [a,b], \quad f(x) := L_x.$$

  The function f is our candidate to limit, because if $f_n \to f$ uniformly on $[a,b]$, we must also have that $f_n \to f$ pointwise on $[a,b]$. The latter is because

  $$\forall x \in [a,b], \quad |f(x) - f_n(x)| \le \|f - f_n\|_\infty.$$

- Let $n \in \mathbb{N} \setminus \{0\}$ such that for any $n \in \mathbb{N}$ with $N \leq n$ we have

  $$\|f_{n+p} - f_n\|_\infty \leq \varepsilon$$

  In particular, we have for any $x \in [a, b]$ that

  $$|f_{n+p}(x) - f_n(x)| \leq \varepsilon$$

  Then, letting $p \to +\infty$ we get

  $$|f(x) - f_n(x)| \leq \varepsilon.$$

  But this is true for any $x \in [a, b]$ and any $n \in \mathbb{N}$ with $N \leq n$, which means that $f_n \to f$ uniformly on $[a, b]$. The fact that f is continuous on $[a, b]$ is a direct consequence of Theorem 9.1, and so the proof is complete.

□

9.3 Dini's theorem

We have seen that the notion of pointwise convergence is weaker than the notion of uniform converge, and that the first one doesn't imply the second. There are few instances in which both notions of convergence coincide, one of these is the so-called **Dini's Theorem**, which we read as follows.

Theorem 9.3. *Let $\{f_n\}_{n=1}^\infty$ be a sequence of continuous functions on $[a, b]$ and let $f : [a, b] \to \mathbb{R}$ be a continuous function on $[a, b]$. Suppose that converges $f_n \to f$ pointwise on $[a, b]$ and that one of the following holds:*

- *for any $x \in [a, b]$, the sequence $\{f_n(x)\}_{n=1}^\infty$ is decreasing.*

- *for any $x \in [a, b]$, the sequence $\{f_n(x)\}_{n=1}^\infty$ is increasing.*

Then $f_n \to f$ uniformly on $[a, b]$.

Before exhibiting the proof, let us make few comments about this theorem:

- This result is a somehow generalization of the monotonic sequences Theorem to continuous functions. In this case, the function f is playing the role of upper bound.

- The monotonic assumptions are fundamental to obtain the result, which can fail if this is not consider, even if f is continuous; see Exercise 6.

The proof of Dini's Theorem requires some tools we haven't introduced so far. We make now a stop in the presentation of continuous function to study the so-called **Heine-Borel Theorem**.

9.3.1 Heine-Borel Theorem

We explained that the underlying idea behind the topological definition of continuity $f : [a, b] \to \mathbb{R}$ at $\bar{x}$ is that any set of the form

$$\{x \in (a, b) \mid \ f(x) \in (f(\bar{x}) - \varepsilon, f(\bar{x}) + \varepsilon)$$

must contain an open interval that contains $\bar{x}$. So to speak, any set having this property is called open around $\bar{x}$. Formally speaking we say that $A \subseteq \mathbb{R}$ is **open** if for any for any $x \in A$, there is an open interval centered at x of radius $r_x \in (0, +\infty)$ contained in A, that is

$$\forall x \in A,\ \exists r_x \in (0, +\infty), \quad (x - r_x, x + r_x) \subseteq A \tag{9.4}$$

It is not difficult to see that any open interval (a, b) is actually an open set in terms of the preceding definition; we leave the details as exercise for the reader.

Let be a closed bounded interval $[a, b]$ and $r \in (0, +\infty)$, we know that we can cover the interval $[a, b]$ using all the possible open interval of the forms $(x - r, x + r)$, that is,

$$[a, b] \subseteq \bigcup_{x \in [a,b]} (x - r, x + r).$$

We readily see that, since $r \in (0, +\infty)$ is fixed, we don't need to use all the $x \in [a, b]$ to cover $[a, b]$ but a finite number of them, that is, we can select $x_1,\ x_2, \ldots, x_p \in [a, b]$ such that

$$[a, b] \subseteq \bigcup_{k=1}^{p} (x_k - r, x_k + r),$$

that is, we can pass from an **arbitrary open covering** to a **finite open covering** of $[a, b]$. What we have just described is the basic idea behind the Heine-Borel Theorem, although instead of having open intervals of fixed radius we allow them to vary.

Theorem 9.4. *Let $\{O_i\}_{i\in I}$ be a collection of open sets of $\mathbb{R}$ that covers the interval $[a, b]$, that is*

$$[a, b] \subseteq \bigcup_{i \in I} O_i.$$

Then, there are $i_1, \ldots, i_p \in I$ such that

$$[a, b] \subseteq \bigcup_{k=1}^{p} O_{i_k}.$$

Proof. Let us argue by contradiction. Suppose that there is a covering $\{O_i\}_{i\in I}$ of $[a, b]$ that doesn't have a finite family of elements that covers the interval $[a, b]$. If that is true, then either the subinterval $[a, \frac{a+b}{2}]$ or $[\frac{a+b}{2}, b]$ cannot be covered by a finite family of elements of $\{O_i\}_{i\in I}$. Hence we set

$$a_1 = \begin{cases} a & \text{if } [a, \frac{a+b}{2}] \text{ cannot be covered by a finite of } \{O_i\}_{i\in I} \\ \frac{a+b}{2} & \text{otherwise} \end{cases}$$

$$b_1 = \begin{cases} \frac{a+b}{2} & \text{if } [a, \frac{a+b}{2}] \text{ cannot be covered by a finite of } \{O_i\}_{i\in I} \\ b & \text{otherwise} \end{cases}$$

Note that in any case $b_1 - a_1 = \frac{1}{2}(b-a)$. Since the interval $[a_1, b_2]$ cannot be covered by a finite family of elements of $\{O_i\}_{i\in I}$, the same is true for one of the subintervals $[a_1, \frac{a_1+b_1}{2}]$ or $[\frac{a_1+b_1}{2}, b_1]$. Hence, repeating the preceding process we can create a a sequence of intervals

$$\ldots \subseteq [a_{n+1}, b_{n+1}] \subseteq [a_n, b_n] \ldots \subseteq [a_1, b_1] \subseteq [a, b].$$

with $[a_n, b_n] = \frac{1}{2^n}(b-a)$. We get then that there is $x \in [a,b]$ such that $\{x\} = \bigcap_{n=1}^{\infty}[a_n, b_n]$.

Now, since $\{O_i\}_{i\in I}$ cover $[a,b]$ there is $i \in I$ such that $x \in O_i$, and since each O_i is open there is $r_i \in (0, +\infty)$ such that

$$(x - r_i, x + r_i) \subseteq O_i.$$

Note that $x - a_n \leq \frac{1}{2^n}(b-a)$ and $b_n - x \leq \frac{1}{2^n}(b-a)$, Hence for taking $n \in \mathbb{N}$ such that

$$\frac{1}{2^n}(b-a) < 2 \cdot r_i$$

we get that $[a_n, b_n] \subseteq (x - r_i, x + r_i) \subseteq O_i$, which contradicts the fact that $[a_n, b_n]$ cannot be covered by a finite family of elements of $\{O_i\}_{i\in I}$. So, the conclusion follows. □

Remark 9.1. *Note that the result is still true if the collection $\{O_i\}_{i\in I}$ is open but relatively to $[a,b]$, that is, if the following holds*

$$\forall i \in I, \forall x \in O_i,\ \exists r_x \in (0, +\infty), \quad (x - r_x, x + r_x) \cap [a,b] \subseteq O_i.$$

The exact same proof works; we leave the details for the reader.

We are now in position to prove the Dini's Theorem

Proof of Theorem 9.3. Suppose that the sequences $\{f_n(x)\}_{n=1}^{\infty}$ are all increasing. Hence, in particular, $f_n(x) \leq f(x)$ for any $x \in [a,b]$. This is because $f(x)$ is the pointwise limit (and so, under these circumstances the supremum) of $\{f_n(x)\}_{n=1}^{\infty}$.

Take $\varepsilon \in (0, +\infty)$. For every $n \in \mathbb{N} \setminus \{0\}$ we set

$$O_n = \{x \in [a,b] \mid f(x) - \varepsilon < f_n(x)\}.$$

Clearly, since the sequence $\{f_n(x)\}_{n=1}^{\infty}$ are all increasing, $O_n \subseteq O_{n+1}$ for any $n \in \mathbb{N} \setminus \{0\}$.

Also $\{O_n\}_{n=1}^{n}$ is a collection of open subsets relative to $[a,b]$ whose union covers $[a,b]$:

- Let $x \in [a,b]$, since $f_n \to f$ pointwise on $[a,b]$ there is $N \in \mathbb{N}$ such that $f(x) - f_N(x) = |f(x) - f_N(x)| < \varepsilon$, and so $x \in O_N$. This leads then to say that $\{O_n\}_{n=1}^{n}$ covers $[a,b]$.
- Let $n \in \mathbb{N} \setminus \{0\}$ to be fixed and let $\bar{x} \in [a,b]$, we know that $f - f_n$ is continuous at $\bar{x}$, and so, for $\tilde{\varepsilon} = \varepsilon - f(\bar{x}) + f_n(\bar{x})$ there is $\delta_n \in (0, +\infty)$ so that

$$\forall x \in [a,b],\ |x - \bar{x}| < \delta_n \implies |(f(x) - f_n(x)) - (f(\bar{x}) - f_n(\bar{x}))| < \tilde{\varepsilon}.$$

  In other words, we have that each O_n is relatively open to $[a,b]$ because:

$$(\bar{x} - \delta_n, \bar{x} + \delta_n) \cap [a,b] \subseteq O_n.$$

By the Heine-Borel Theorem and Remark (9.1), there is an integer $N \in \mathbb{N} \setminus \{0\}$ such that $[a,b] \subseteq O_N$. This means that $f(x) - \varepsilon < f_N(x)$ for any $x \in [a,b]$. Thus, for every $n \in \mathbb{N}$ with $N \leq n$, we have $f(x) - \varepsilon < f_n(x) \leq f(x)$ for all $x \in [a,b]$. This proves $\|f - f_n\|_\infty \leq \varepsilon$ whenever $N \leq n$, and so the conclusion follows. □

9.4 Exercises

1. Let $f : \mathbb{R} \to \mathbb{R}$ be a given function. For any $a \in \mathbb{R}$, let $f_a : \mathbb{R} \to \mathbb{R}$ be the shifted function

$$\forall x \in \mathbb{R},\ f_a(x) := f(x-a).$$

   (a) Show that f is continuous if and only if, whenever $\{a_n\}_{n=1}^{\infty}$ is a sequence in $\mathbb{R}$ which converges to zero, the shifted functions f_{a_n} converge pointwise to f on $\mathbb{R}$.

   (b) Show that f is uniformly continuous if and only if, whenever $\{a_n\}_{n=1}^{\infty}$ is a sequence in $\mathbb{R}$ with $a_n \to 0$, the shifted functions f_{a_n} converge uniformly to f on $\mathbb{R}$.

2. Let $\{f_n\}_{n=1}^{\infty}$ and $\{g_n\}_{n=1}^{\infty}$ be two Cauchy sequences of continuous functions on $[a,b]$. Show, using the Cauchy criterion, that there is a continuous function $h : [a,b] \to \mathbb{R}$ such that $f_n \cdot g_n \to h$ uniformly on $[a,b]$.

   Hint: Prove that both sequences are uniformly bounded, that is,

$$\exists M \in (0,+\infty),\ \forall n \in \mathbb{N} \quad \|f_n\|_\infty \le M \ \wedge\ \|g_n\|_\infty \le M.$$

3. Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of continuous functions on $[a,b]$ and a sequence $\{c_n\}_{n=1}^{\infty}$ contained in $[0,+\infty)$ such that

$$\forall x \in [a,b], \quad |f(x)| \le c_n \quad \text{and} \quad \sum_{k=1}^{n} c_k \to S \in \mathbb{R}.$$

   Let $\{s_n\}_{n=1}^{\infty}$ be the sequence of continuous functions on $[a,b]$ defined via:

$$\forall x \in [a,b],\ \forall n \in \mathbb{N}\setminus\{0\}, \quad s_n(x) := \sum_{k=1}^{n} f_n(x).$$

   Prove that $\{s_n\}_{n=1}^{\infty}$ converges uniformly on $[a,b]$ to some function $s : [a,b] \to \mathbb{R}$. Is s uniformly continuous on $[a,b]$?

   Hint: Prove that $\{s_n\}_{n=1}^{\infty}$ is a Cauchy sequence of functions.

4. Prove that a set A is closed if and only if $\mathbb{R}\setminus A$ is open.

5. Prove that the sequence of function $\{f_n\}_{n=1}^{\infty}$ given by

$$f_n(x) := \left(1 + \frac{x}{n}\right)^n$$

   converges uniformly to $x \mapsto \exp(x)$ on any closed bounded interval $[a,b]$.

6. Show that $\{f_n\}_{n=1}^{\infty}$, the sequence of continuous functions on $[0,1]$ given by

$$\forall x \in [0,1],\ \forall n \in \mathbb{N}\setminus\{0\}, \quad f_n(x) = (n+1)\cdot x^n \cdot (1-x),$$

   converges pointwise to zero, but it fail to converge uniformly to zero. Does this contradict the Dini's Theorem?

   Hint: Evaluate f_n at $x = \frac{n}{n+1}$.

Math 4031 - Advanced Calculus I

Instructor: Dr. Cristopher HERMOSILLA
Louisiana State University - Spring 2016

Week 11: Introduction to integration theory

We begin the study of one of the pillar of calculus, namely *integration*. Along this part of the course we mainly focus on the **Riemann integral**, however at the end of the course we will present a short overview on **Lebesgue integral**.

Unless otherwise stated, we focus on bounded functions defined on a bounded closed interval and whose values belong to $\mathbb{R}$. Recall that a function is said to be bounded if

$$m_f := \inf\{f(x) \mid \ x \in [a,b]\} \in \mathbb{R} \qquad M_f := \sup\{f(x) \mid \ x \in [a,b]\} \in \mathbb{R}$$

11.1 Introductory example

From a practical point of view, the integral of a non negative function can be interpreted as the *area of the region* on the xy-plane limited by the curves

$$y = f(x), \quad y = 0, \quad x = a, \quad x = b.$$

For example, let us consider the function $f(x) = x^2$ defined on $[0,1]$. There are several way to estimate the area under the curve, but we basically use two, one that approximates it from below and another that does it from above.

We begin by taking $n \in \mathbb{N} \setminus \{0\}$ and then dividing the interval $[0,1]$ into n subintervals whose lengths are $\frac{1}{n}$, that is, setting $x_k := \frac{k}{n}$ for any $k \in \{0,\dots,n\}$, we consider the intervals

$$[x_{k-1}, x_k] \subseteq [0,1], \quad \forall k \in \{1,\dots,n\}.$$

On each of these intervals, let us denote by A_k the area of the region on the xy-plane limited by the curves

$$y = x^2, \quad y = 0, \quad x = x_{k-1}, \quad x = x_k$$

Note that $x_{k-1}^2 \leq x^2 \leq x_{k-1}^2$ for any $x \in [x_{k-1}, x_k]$, from where we may assume that

$$x_{k-1}^2(x_k - x_{k-1}) \leq A_k \leq x_k^2(x_k - x_{k-1}).$$

The latter yields to

$$\frac{(k-1)^2}{n^3} \leq A_k \leq \frac{k^2}{n^3}.$$

If we call A the area of the region limited by the curves

$$y = x^2, \quad y = 0, \quad x = 0, \quad x = 1,$$

we immediately see that $A = \sum_{k=1}^{n} A_k$, from where get the estimates

$$\frac{1}{n^3}\sum_{k=1}^{n}(k-1)^2 \leq A \leq \frac{1}{n^3}\sum_{k=1}^{n}k^2.$$

At this point we might state an intermediate result regarding sums:

Lemma 11.1. *For any $n \in \mathbb{N} \setminus \{0\}$ we have*

$$\sum_{k=1}^{n} k = \frac{n^2+n}{2} \quad \text{and} \quad \sum_{k=1}^{n} k^2 = \frac{2\cdot n^3 + 3\cdot n^2 + n}{6}$$

Hence, by Lemma 11.1 we have that

$$\frac{2\cdot n^3 + n^2 - n}{6\cdot n^3} \leq A \leq \frac{2\cdot n^3 + 3\cdot n^2 + n}{6\cdot n^3}.$$

Finally, since $n \in \mathbb{N} \setminus \{0\}$ is arbitrary, we can let $n \to +\infty$ and so, by the Squeeze Theorem we get that $A = \frac{1}{3}$.

This example shows the underlying ideas behind the concept of Riemann integral of a function. However, we need to be careful and not take this as the definition of the Riemann integral, but only as an application; the idea only works if the function has non-negative values. Note that we have talked about *areas* but we have never defined them in mathematical terms.

11.2 Riemann integral for piecewise constant functions

We turn to a formal definition of the Riemann integral of a function. We first introduce the simplest class of functions for which this integral can be defined, and later on, we give a definition that works for a wider class of mappings, which in particular includes the continuous functions.

Let $I \subseteq \mathbb{R}$ be a bounded interval, that is, for some $a, b \in \mathbb{R}$ with $a \leq b$, I agrees with one of the following sets

$$(a,b), \quad [a,b), \quad (a,b], \quad [a,b].$$

Recall that if $a = b$ then in the first three situations we get $I = \emptyset$ and in the last one $I = \{a\} = \{b\}$. In any case, we define the **length of the interval** I, denoted by $\ell(I)$, via

$$\ell(I) := b - a.$$

Note that this means that if I and J are two bounded interval, then

$$I \cap J = \emptyset \quad \Longrightarrow \quad \ell(I \cap J) = 0.$$

11.2.1 Partitions of intervals

One of the key tools required to introduce the concept of Riemann integral of a function is the notion of **partition** of an interval.

Definition 11.1. *Let $a, b \in \mathbb{R}$ with $a < b$. A partition P of $[a, b]$ is a finite collection of nonempty pairwise disjoint intervals that covers $[a, b]$. In other words, $P = \{I_1,\ I_2, \ldots, I_n\}$ where each I_k is a nonempty interval contained in $[a, b]$ such that*

$$\forall x \in [a, b],\ \exists! k \in \{1, \ldots, n\},\ x \in I_k.$$

Under these circumstances, we define the **mesh** *of P via*

$$\|P\| := \max\{\ell(I_1), \ldots, \ell(I_k), \ldots, \ell(I_n)\}.$$

Remark 11.1. *Alternatively, we can define a partition of $[a, b]$ as a finite collection of points $P = \{x_0,\ x_1, \ldots, x_n\}$ with $x_0 = a$ and $x_n = b$, so that*

$$x_k \in [a, b] \quad \text{and} \quad x_{k-1} \leq x_k, \quad \forall k = 1, \ldots, n.$$

Under these circumstances, each x_k is called a **node** *of the partition. Furthermore, the* **mesh** *of P is then given by*

$$\|P\| := \max\{x_1 - x_0, \ldots, x_k - x_{k-1}, \ldots, x_n - x_{n-1}\}.$$

We leave the details as exercise for the reader.

In the light of the preceding remark, it is not difficult to see that, given a partition $P = \{I_1,\ I_2, \ldots, I_n\}$ of $[a, b]$, the length of the interval $[a, b]$ is the sum of the length of the subintervals of the partition, that is,

$$\ell([a, b]) = \sum_{k=1}^{n} \ell(I_k).$$

Moreover, if $[c, d] \subset [a, b]$, then

$$\ell([c, b]) = \sum_{k=1}^{n} \ell([c, d] \cap I_k).$$

To see this is enough to notice that $\tilde{P} = \{[c, d] \cap I_1, \ldots, [c, d] \cap I_n\} \setminus \{\emptyset\}$ is a partition of $[c, d]$.

Given two partitions $P_1 = \{I_1,\ I_2, \ldots, I_n\}$ and $P_2 = \{J_1,\ J_2, \ldots, J_m\}$ of an interval $[a, b]$, we defined the **common refinement** of P_1 and P_2, denoted $P_1 \# P_2$, as the partition

$$P_1 \# P_2 = \bigcup_{i=1}^{n} \bigcup_{j=1}^{m} \{I_i \cap J_j\} \setminus \{\emptyset\}.$$

Example 11.1. *Let $P_1 = \{[0, 1), \{1\}, (1, 2), [2, 3]\}$ and $P_2 = \{\{0\}, (0, 1], (1, 3]\}$ be two partitions of the interval $[0, 3]$, then*

$$P_1 \# P_2 = \{\{0\}, (0, 1), \{1\}, (1, 2), [2, 3]\}.$$

11.2.2 Piecewise constant functions

Given a partition $P = \{I_1,\ I_2, \ldots, I_n\}$ of $[a, b]$, we say that a function $f : [a, b] \to \mathbb{R}$ is **piecewise constant relative to** P if there are $c_1, \ldots, c_n \in \mathbb{R}$ such that

$$f(x) = c_k \quad \text{whenever } x \in I_k.$$

Example 11.2. *Let $P = \{[0, 1), \{1\}, (1, 2]\}$ be a partition of the interval $[0, 2]$, then*

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

is a piecewise constant function relative to P defined on $[0, 2]$.

Note that if $f : [a, b] \to \mathbb{R}$ is piecewise constant relative to a partition P_1 and to another partition P_2, then it is also piecewise constant relative to their common refinement $P_1\#P_2$. This means in particular that a function can be piecewise constant relative to several partitions at the same time. For this reason, it is convenient to introduce a new definition that avoids fixing beforehand a partition.

Definition 11.2. *We say that a function $f : [a, b] \to \mathbb{R}$ is* **piecewise constant** *on $[a, b]$ if there is a partition $P = \{I_1,\ I_2, \ldots, I_n\}$ of $[a, b]$ so that f is piecewise constant relative to P.*

It turns out that for the class of functions that are piecewise constant on an interval $[a, b]$, the Riemann integral can be defined in simple terms.

Definition 11.3. *Let $f : [a, b] \to \mathbb{R}$ be a piecewise constant function, then its* **Riemann integral** *is defined via the formula:*

$$\int_a^b f := \sum_{k=1}^{n} c_k \ell(I_k),$$

where $P = \{I_1,\ I_2, \ldots, I_n\}$ is any partition of $[a, b]$ for which f is piecewise constant relative to P and each c_k is the value of f on the interval I_k.

Remark 11.2. *Note that in Definition 11.3 the value of $\int_a^b f$ is independent of the partition taken. Indeed, if $P_1 = \{I_1,\ I_2, \ldots, I_n\}$ and $P_2 = \{J_1,\ J_2, \ldots, J_m\}$ are two different partition associated with f, that is, there are $c_1, \ldots, c_n \in \mathbb{R}$ and $d_1, \ldots, d_m \in \mathbb{R}$ such that*

$$f(x) = c_i \quad \text{whenever } x \in I_i \quad \text{and} \quad f(x) = d_j \quad \text{whenever } x \in J_j.$$

Since f is also piecewise constant relative to $P_1\#P_2$ we have that

$$f(x) = e_{i,j} \quad \text{whenever } x \in I_i \cap J_j.$$

Note that if $I_i \cap J_j = \emptyset$ then $e_{i,j}$ can be any Real number, but if $I_i \cap J_j \neq \emptyset$, then $e_{i,j} = c_i = d_j$. Hence,

$$\sum_{i=1}^{n} c_i \cdot \ell(I_i) = \sum_{i=1}^{n} c_i \cdot \sum_{j=1}^{m} \ell(I_i \cap J_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} e_{i,j} \cdot \ell(I_i \cap J_j)$$

Similarly,

$$\sum_{j=1}^{m} d_j \cdot \ell(J_j) = \sum_{j=1}^{m} d_j \cdot \sum_{i=1}^{n} \ell(I_i \cap J_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} e_{i,j} \cdot \ell(I_i \cap J_j).$$

From where we get that

$$\sum_{i=1}^{n} c_i \cdot \ell(I_i) = \sum_{j=1}^{m} d_j \cdot \ell(J_j).$$

Example 11.3. *Let us pick up the data on Example 11.2, then*

$$\int_0^2 f = 2 \cdot \ell([0,1)) + 1 \cdot \ell(\{1\}) - 1 \cdot \ell((1,2]) = 2 + 0 - 1 = 1.$$

Let us point out some properties of the Riemann integral of a piecewise constant function $f : [a,b] \to \mathbb{R}$; for more general properties we refer to Exercise 4.

- If f is constant all along $[a,b]$, that is, if $f(x) = c$ for any $x \in [a,b]$, then
$$\int_a^b f = c \cdot (b - a).$$
- If f is bounded, that is, $m_f, M_f \in \mathbb{R}$, then
$$m_f \cdot (b-a) \leq \int_a^b f \leq M_f \cdot (b-a).$$
- If f is non-negative all along $[a,b]$, that is, if $f(x) \in [0,+\infty)$ for any $x \in [a,b]$, then
$$0 \leq \int_a^b f.$$
- For any $c \in [a,b]$, if there are other piecewise constant functions $f_1 : [a,c] \to \mathbb{R}$ and $f_2 : [c,b] \to \mathbb{R}$ such that
$$f(x) = f_1(x), \quad \forall x \in [a,c), \quad \text{and} \quad f(x) = f_2(x), \quad \forall x \in (c,b],$$
then
$$\int_a^b f = \int_a^c f_1 + \int_c^b f_2.$$

11.3 Riemann integrable functions

We now focus on the definition of the Riemann integral for an arbitrary bounded function $f : [a,b] \to \mathbb{R}$, not necessarily piecewise constant.

We say that a function $g : [a,b] \to \mathbb{R}$ **majorizes** f on $[a,b]$ if $f(x) \leq g(x)$ for any $x \in [a,b]$. Similarly, we say that $h : [a,b] \to \mathbb{R}$ **minorizes** f on $[a,b]$ if $h(x) \leq f(x)$ for any $x \in [a,b]$. As we have shown in the introductory example, the idea of the Riemann integral is to try to integrate a function by first majorizing or minorizing that function by a piecewise constant function, for which we already know how to define the Riemann integral.

It is worthy to notice that since $f : [a, b] \to \mathbb{R}$ is supposed to be a bounded function, then the constant functions $x \mapsto m_f$ and $x \mapsto M_f$ minorizes and majorizes f on $[a.b]$. Furthermore, these functions are in particular piecewise constant, so the sets of Real numbers

$$A_f := \left\{ \int_a^b g \;\middle|\; g : [a, b] \to \mathbb{R} \text{ is piecewise constant and majorizes } f \text{ on } [a, b] \right\}$$
$$B_f := \left\{ \int_a^b h \;\middle|\; h : [a, b] \to \mathbb{R} \text{ is piecewise constant and minorizes } f \text{ on } [a, b] \right\}$$

are non-empty. Moreover, if $g : [a, b] \to \mathbb{R}$ and $h : [a, b] \to \mathbb{R}$ are function that majorizes and minorizes f on $[a, b]$, respectively, then

$$m_f \leq g(x), \quad \text{and} \quad h(x) \leq M_f, \qquad \forall x \in [a.b].$$

In particular, A_f is bounded below and B_f is bounded above. Therefore, by the Supremum axiom, we have that their infimum and supremum are well-defined Real numbers.

Definition 11.4. *Let $f : [a, b] \to \mathbb{R}$ be a bounded function. We define the* **upper Riemann integral** *of f on $[a, b]$ by*

$$\overline{\int_a^b} f := \inf \left\{ \int_a^b g \;\middle|\; g : [a, b] \to \mathbb{R} \textit{ is piecewise constant and majorizes } f \textit{ on } [a, b] \right\}$$

and the **lower Riemann integral** *of f on $[a, b]$ by*

$$\underline{\int_a^b} f := \sup \left\{ \int_a^b h \;\middle|\; h : [a, b] \to \mathbb{R} \textit{ is piecewise constant and majorizes } f \textit{ on } [a, b] \right\}$$

Note that we alway have

$$m_f \cdot (b - a) \leq \underline{\int_a^b} f \leq \overline{\int_a^b} f \leq M_f \cdot (b - a).$$

The inequality on the middle can be strict, for example consider the function

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

If $g : [0, 1] \to \mathbb{R}$ is a piecewise function that majorizes f, then g must be bounded below by 1 except at a finite number of points, and if $h : [0, 1] \to \mathbb{R}$ is a piecewise function that minorizes f, then h is bounded above by 0 except at a finite number of points. Thus,

$$\underline{\int_a^b} f \leq 0 < 1 \leq \overline{\int_a^b} f.$$

In the special case that the upper and lower Riemann integrals agree and are Real numbers, we say that the function is **Riemann integrable** and we then define its **Riemann integral** on $[a, b]$ via

$$\int_a^b f := \overline{\int_a^b} f = \underline{\int_a^b} f.$$

Remark 11.3. *If $f : [a, b] \to \mathbb{R}$ is piecewise constant and bounded, then f majorizes and minorizes f the same time on $[a, b]$. Hence it is clear that f is Riemann integrable.*

Remark 11.4. *The definition of Riemann integrable functions we have adopted is specially suited for bounded functions, as a matter of fact only bounded functions can be Riemann integrable. Indeed, if $f : [a, b] \to \mathbb{R}$ is Riemann integrable, by the definition of the upper and lower Riemann integrals we get that there are piecewise constant functions $h : [a, b] \to \mathbb{R}$ and $h : [a, b] \to \mathbb{R}$ such that $h \le f \le g$, and since piecewise constant functions are bounded, f must be bounded too.*

11.4 Exercises

1. Prove Lemma 11.1.

   Hint: Show that

   $$\sum_{k=1}^{n+1} k^3 = \sum_{k=1}^{n} k^3 + 3 \cdot \sum_{k=1}^{n} k^2 + 3 \cdot \sum_{k=1}^{n} k + (n+1)$$

2. Let $P_1 = \{I_1, \ldots, I_n\}$ and $P_2 = \{J_1, \ldots, J_m\}$ be two given partitions of $[a, b]$. Show that

   $$\|P_1 \# P_2\| \le \min\{\|P_1\|, \|P_2\|\}$$

3. Let $f : [a, b] \to \mathbb{R}$ be a piecewise constant function on $[a, b]$. Show that $x \mapsto |f(x)|$ is piecewise constant on $[a, b]$.

4. Let $f_1 : [a, b] \to \mathbb{R}$ and $f_2 : [a, b] \to \mathbb{R}$ be two piecewise constant functions on $[a, b]$.

   (a) Use the notion of common refinement to prove that for any $\lambda \in \mathbb{R}$, the function $f_1 + \lambda \cdot f_2$ is piecewise constant on $[a, b]$ and then show that

   $$\int_a^b (f_1 + \lambda \cdot f_2) = \int_a^b f_1 + \lambda \cdot \int_a^b f_2, \quad \forall \lambda \in \mathbb{R}.$$

   (b) Use the notion of common refinement to prove that the functions $\min\{f_1, f_2\}$ and $\max\{f_1, f_2\}$ are also piecewise constant $[a, b]$.

   (c) Use the notion of common refinement to prove that the function $f_1 \cdot f_2$ is piecewise constant on $[a, b]$. Prove or give a counterexample for the following formula

   $$\int_a^b (f_1 \cdot f_2) = \int_a^b f_1 \cdot \int_a^b f_2$$

5. **(Riemann sums)** Let $f : [a,b] \to \mathbb{R}$ be a bounded function and $P = \{I_1,\ I_2, \dots, I_n\}$ be a partition of $[a,b]$. We define the **upper Riemann sum** of f relative to P by

$$U(f,P) := \sum_{k=1}^{n} \sup\{f(x) \mid\ x \in I_k\} \cdot \ell(I_k),$$

and the **lower Riemann sum** of f relative to P by

$$L(f,P) := \sum_{k=1}^{n} \inf\{f(x) \mid\ x \in I_k\} \cdot \ell(I_k).$$

   (a) Prove that $U(f,P)$ and $L(f,P)$ are well-defined Real numbers.

   (b) Let $g : [a,b] \to \mathbb{R}$ and $h : [a,b] \to \mathbb{R}$ be two function that majorizes and minorizes f on $[a,b]$, respectively. Suppose that g and h are piecewise constant relative to P. Prove that

$$U(f,P) \leq \int_a^b g \quad \text{and} \quad \int_a^b h \leq L(f,P).$$

   (c) Prove that

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

   and

$$\underline{\int_a^b} f = \sup\{L(f,P) \mid\ P \text{ is a partition of } [a,b]\}.$$

6. Suppose that $f : [a,b] \to \mathbb{R}$ is Riemann integrable.

   (a) Suppose that f is non-negative all along $[a,b]$. Show that

$$0 \leq \int_a^b f. \tag{11.1}$$

   If (11.1) holds, can we conclude that f is non-negative? Prove this or give a counterexample.

   (b) For any $c \in [a,b]$, if there are other Riemann integrable functions $f_1 : [a,c] \to \mathbb{R}$ and $f_2 : [c,b] \to \mathbb{R}$ such that

$$f(x) = f_1(x), \quad \forall x \in [a,c], \quad \text{and} \quad f(x) = f_2(x), \quad \forall x \in [c,b],$$

   Show that

$$\int_a^b f = \int_a^c f_1 + \int_c^b f_2.$$

Math 4031 - Advanced Calculus I

Instructor: Dr. Cristopher HERMOSILLA
Louisiana State University - Spring 2016

Week 12: Riemann integrable function

The class of functions that are Riemann integrable is wide and difficult to characterize; it is much larger than the set of continuous functions. For this reason, we now turn our attention into criteria that ensure that a function is Riemann integrable.

12.1 Basic properties of Riemann integrable functions

It can be proved that algebraic combinations and min/max functions of piecewise constant functions are also piecewise constant functions; see Exercise 4, Week 11's notes. Hence, it is natural to imagine that algebraic combinations and min/max functions of Riemann integrable functions are also Riemann integrable. This is true, but their proofs are not simple extensions. Furthermore, in most cases an exact formula for the value of the integral cannot be given.

12.1.1 Linear combinations of Riemann integrable functions

We begin with the simplest cases which corresponds to linear combinations of Riemann integrable functions. In this case, it is possible to provide an explicit formula for the Riemann integral's value.

Theorem 12.1. *Let $f_1 : [a, b] \to \mathbb{R}$ and $f_2 : [a, b] \to \mathbb{R}$ be two Riemann integrable functions on $[a, b]$. Then, for any $\lambda \in \mathbb{R}$ we have that $f_1 + \lambda \cdot f_2$ is Riemann integrable with*

$$\int_a^b (f_1 + \lambda \cdot f_2) = \int_a^b f_1 + \lambda \cdot \int_a^b f_2$$

Proof. The case $\lambda = 0$ is trivial, so we might either assume $\lambda > 0$ or $\lambda < 0$. We only do the case $\lambda > 0$, the other is similar and is left as exercise for the reader.

Let $\varepsilon > 0$ be given but arbitrary. Since f_1 is Riemann integrable, there are two piecewise constant function $g_1 : [a, b] \to \mathbb{R}$ and $h_1 : [a, b] \to \mathbb{R}$ such that $h_1 \leq f_1 \leq g_1$ and

$$\int_a^b g_1 - \frac{\varepsilon}{2} \leq \int_a^b f_1 \leq \int_a^b h_1 + \frac{\varepsilon}{2}.$$

In a similar way, there are $g_2 : [a, b] \to \mathbb{R}$ and $h_2 : [a, b] \to \mathbb{R}$, piecewise constant on $[a, b]$, such that $h_2 \leq f_2 \leq g_2$ and

$$\int_a^b g_2 - \frac{\varepsilon}{2\lambda} \leq \int_a^b f_2 \leq \int_a^b h_2 + \frac{\varepsilon}{2\lambda}.$$

Note that $h_1+\lambda h_2$ and $g_1+\lambda g_2$ are both piecewise constant, and the first one minorizes $f_1+\lambda\cdot f_2$ and the second one majorizes it. Hence, by definition of the lower and upper Riemann integrals and the properties of the Riemann integral over piecewise constant function we have

$$\int_a^b h_1+\lambda\cdot\int_a^b h_2 = \int_a^b (h_1+\lambda\cdot h_2) \leq \underline{\int_a^b}(f_1+\lambda\cdot f_2) \leq \overline{\int_a^b}(f_1+\lambda\cdot f_2) \leq \int_a^b (g_1+\lambda\cdot g_2) = \int_a^b g_1+\lambda\cdot\int_a^b g_2.$$

Note that the left and right hand-sides also satisfy

$$\int_a^b f_1 + \lambda\cdot\int_a^b f_2 - \varepsilon \leq \int_a^b h_1 + \lambda\cdot\int_a^b h_2 \quad \text{and} \quad \int_a^b g_1 + \lambda\cdot\int_a^b g_2 \leq \int_a^b f_1 + \lambda\cdot\int_a^b f_2 + \varepsilon.$$

From where we get that

$$\int_a^b f_1 + \lambda\cdot\int_a^b f_2 - \varepsilon \leq \underline{\int_a^b}(f_1 + \lambda\cdot f_2) \leq \overline{\int_a^b}(f_1 + \lambda\cdot f_2) \leq \int_a^b f_1 + \lambda\cdot\int_a^b f_2 + \varepsilon.$$

Finally, the conclusion follows because the latter is true for any $\varepsilon \in (0, +\infty)$, so we get that

$$\int_a^b f_1 + \lambda\cdot\int_a^b f_2 = \underline{\int_a^b}(f_1 + \lambda\cdot f_2) = \overline{\int_a^b}(f_1 + \lambda\cdot f_2).$$

□

12.1.2 Max and Min functions of Riemann integrable functions

Let us now pass to the case of max and min function.

Theorem 12.2. *Let $f_1 : [a, b] \to \mathbb{R}$ and $f_2 : [a, b] \to \mathbb{R}$ be two Riemann integrable functions on $[a, b]$. Then, $\max\{f_1, f_2\}$ and $\min\{f_1, f_2\}$ are Riemann integrable.*

Proof. We focus on the case $\max\{f_1, f_2\}$, the other one is similar and left as exercise.

The proof starts in a similar way as the one given for Theorem 12.1, that is, let $\varepsilon > 0$ be given but arbitrary. Since, for each $i = 1, 2$, the function f_i is Riemann integrable, there are two piecewise constant function $g_i : [a, b] \to \mathbb{R}$ and $h_i : [a, b] \to \mathbb{R}$ such that $h_i \leq f_i \leq g_i$ and

$$\int_a^b g_i - \frac{\varepsilon}{4} \leq \int_a^b f_1 \leq \int_a^b h_1 + \frac{\varepsilon}{4},$$

Note that $\max\{h_1, h_2\}$ and $\max\{g_1, g_2\}$ are both piecewise constant on $[a, b]$, and furthermore,

$$\max\{h_1, h_2\} \leq \max\{f_1, f_2\} \leq \max\{g_1, g_2\}.$$

Hence, by definition of the lower and upper Riemann integrals we have

$$\int_a^b \max\{h_1, h_2\} \leq \underline{\int_a^b} \max\{f_1, f_2\} \leq \overline{\int_a^b} \max\{f_1, f_2\} \leq \int_a^b \max\{g_1, g_2\}.$$

From where

$$0 \leq \overline{\int_a^b} \max\{f_1, f_2\} - \underline{\int_a^b} \max\{f_1, f_2\} \leq \int_a^b \max\{g_1, g_2\} - \int_a^b \max\{h_1, h_2\}.$$

By Theorem 12.1, we have that

$$\int_a^b \max\{g_1, g_2\} - \int_a^b \max\{h_1, h_2\} = \int_a^b (\max\{g_1, g_2\} - \max\{h_1, h_2\})$$

On the other hand, for each $i = 1, 2$,

$$g_i = h_i + (g_i - h_i) \leq \max\{h_1, h_2\} + (g_1 - h_1) + (g_2 - h_2),$$

which means that $\max\{g_1, g_2\}$ is bounded above by the right hand-side of the latter inequality. Thus, by the properties of the Riemann integral for piecewise constant functions we get

$$\int_a^b (\max\{g_1, g_2\} - \max\{h_1, h_2\}) \leq \int_a^b g_1 - \int_a^b h_1 + \int_a^b g_2 - \int_a^b h_2$$

Note that by the initial assumption, the right hand-side is less than or equal to ε, hence

$$0 \leq \overline{\int_a^b} \max\{f_1, f_2\} - \underline{\int_a^b} \max\{f_1, f_2\} \leq \varepsilon.$$

Since, $\varepsilon \in (0, +\infty)$ is arbitrary, we get that the upper and lower Riemann integrals agree, and the proof is then complete. □

Given a function $f : [a, b] \to \mathbb{R}$, we define its **positive part** by $f_+ := \max\{f, 0\}$ and its **negative part** by $f_- := \min\{f, 0\}$. Note that $|f| = f_+ - f_-$, and so, if f is Riemann integrable, then so are f_+, f_- and $|f|$.

12.1.3 Multiplication of Riemann integrable functions

We would like now to prove that if $f_1 : [a, b] \to \mathbb{R}$ and $f_2 : [a, b] \to \mathbb{R}$ are Riemann integrable on $[a, b]$, then do it is $f_1 \cdot f_2$. Note that

$$f_1 \cdot f_2 = \frac{1}{2}(f_1 + f_2)^2 - f_1^2 - f_2^2.$$

Thus, to prove that $f_1 \cdot f_2$ is Riemann integrable we only need to show that for any Riemann integrable function $f : [a, b] \to \mathbb{R}$, its square is also a Riemann integrable function. Then, the result will follows as consequence of Theorem 12.1 and the following one.

Theorem 12.3. *Let $f : [a, b] \to \mathbb{R}$ be a Riemann integrable function, then f^2 is also a Riemann integrable function.*

Proof. Since $f = f_+ + f_-$ we have that $f^2 = f_+^2 + 2 \cdot f_+ \cdot f_- + f_-^2$. But, $f_+ \cdot f_- = 0$, so to prove that f^2 is Riemann integrable we only need to prove that f_+^2 and f_-^2 are both Riemann integrable. We only exhibit the proof for f_+^2, the one for f_-^2 it is enough to replace f_+ with $-f_-$ when appropriate; recall that f_- is non positive.

Since f is Riemann integrable it is also bounded, so there is $M \in (0, +\infty)$ so that

$$0 \leq f_+ \leq M.$$

Let $\varepsilon \in (0, +\infty)$. Since f_+ is also Riemann integrable, there are two piecewise constant functions $g : [a, b] \to \mathbb{R}$ and $h : [a, b] \to \mathbb{R}$ such that $h \leq f_+ \leq g$ and

$$\int_a^b g - \frac{\varepsilon}{4 \cdot M} \leq \int_a^b f_+ \leq \int_a^b h + \frac{\varepsilon}{4 \cdot M}.$$

Note that $x \mapsto 0$ and $x \mapsto M$ are both piecewise constant function on $[a, b]$, which minorizes and majorizes, respectively, the function f_+. Hence, without loss of generality, we can assume that $0 \leq h$ and $g \leq M$. Therefore, f_+^2 is minorized and majorized by h^2 and g^2, respectively. Since g and h are both piecewise constant, h^2 and g^2 are piecewise constant too. Consequently

$$0 \leq \overline{\int_a^b} f_+^2 - \underline{\int_a^b} f_+^2 \leq \int_a^b (g^2 - h^2) = \int_a^b [(g - h) \cdot (g + h)].$$

But, $g + h \leq 2 \cdot M$ and $0 \leq g - h$, so

$$\int_a^b [(g - h) \cdot (g + h)] \leq 2 \cdot M \int_a^b (g - h) = 2 \cdot M \left(\int_a^b g - \int_a^b h \right) \leq \varepsilon.$$

Since, $\varepsilon \in (0, +\infty)$ is arbitrary, the upper and lower Riemann integrals of f_+^2 coincide, this ends the proof. □

12.2 Riemann integral and monotonic functions

Recall that a function is called monotonic if it is either increasing or decreasing all along the interval where it is defined, that is, $f : [a, b] \to \mathbb{R}$ is monotonic if one of the following holds:

- For any $x, y \in [a, b]$ with $x < y$ we have $f(x) \leq f(y)$.
- For any $x, y \in [a, b]$ with $x < y$ we have $f(y) \leq f(x)$.

Theorem 12.4. *Let $f : [a, b] \to \mathbb{R}$ be a monotonic function, then f is Riemann integrable.*

Proof. Note that since f is monotonic it is also bounded; either $f(a) \leq f(x) \leq f(b)$ or $f(b) \leq f(x) \leq f(a)$ for any $x \in [a, b]$. Let $n \in \mathbb{N} \setminus \{0\}$ be fixed but arbitrary. We consider the case that f is increasing, the other is left as exercise for the reader. Let $x_k := a + \frac{k}{n}(b - a)$ for any $k \in \{0, \dots, n\}$ and the partition $P = \{I_1, \dots, I_n\}$, where

$$\mathcal{I}_1 = [0, x_1] \quad \text{and} \quad I_k = (x_{k-1}, x_k], \quad \forall k \in \{2, \dots, n\}.$$

For each $k \in \{1, \dots, n\}$ we have $\ell(I_k) = \frac{b-a}{n}$, and since f is increasing we also have that for any $x \in \mathcal{I}_k$

$$f(x_{k-1}) \leq f(x) \leq f(x_k).$$

Therefore, the piecewise constant function $g : [a, b] \to \mathbb{R}$ given by

$$g(x) = f(x_k), \quad \text{whenever } x \in I_k$$

majorizes f and the piecewise constant function $h : [a, b] \to \mathbb{R}$ given by

$$h(x) = f(x_{k-1}), \quad \text{whenever } x \in I_k$$

minorizes it. Therefore,

$$\sum_{k=1}^{n} f(x_{k-1})\ell(I_k) = \int_a^b h \leq \underline{\int_a^b} f \leq \overline{\int_a^b} f \leq \int_a^b g = \sum_{k=1}^{n} f(x_k)\ell(I_k)$$

From this we get that

$$\overline{\int_a^b} f - \underline{\int_a^b} f \leq \sum_{k=1}^{n} (f(x_k - f(x_{k-1}))\ell(I_k) = \frac{b-a}{n} \sum_{k=1}^{n} (f(x_k - f(x_{k-1})) = \frac{b-a}{n}(f(x_n) - f(x_0)).$$

But $x_n = b$ and $x_0 = a$, from where we get that the right hand side is less than or equal to $\frac{b-a}{n}(f(b) - f(a))$. Therefore, letting $n \to +\infty$ we get that the upper and lower Riemann integrals of f agree and the function is then Riemann integrable. □

12.3 Riemann integral and continuous functions

As we have claimed at the beginning, continuous functions are also Riemann integrable. This is essentially due to the fact that continuous functions on bounded closed interval are bounded and also uniformly continuous, that is,

$$\forall \varepsilon \in (0, +\infty),\ \exists \delta \in (0, +\infty),\ \forall x, y \in [a, b],\ |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

Theorem 12.5. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function, then f is Riemann integrable.*

Proof. Let $\varepsilon \in (0, +\infty)$ and $\delta \in (0, +\infty)$ given by the uniform continuity of f on $[a, b]$ associated with $\tilde{\varepsilon} = \frac{\varepsilon}{b-a}$. By the Archimedean property, there is $n \in \mathbb{N}$ such that $\frac{1}{n}(b - a) < \delta$. Let $x_k := a + \frac{k}{n}(b - a)$ for any $k \in \{0, \dots, n\}$ and the partition $P = \{I_1, \dots, I_n\}$, where

$$\mathcal{I}_1 = [0, x_1] \quad \text{and} \quad I_k = (x_{k-1}, x_k], \quad \forall k \in \{2, \dots, n\}.$$

Clearly, for each $k \in \{1, \dots, n\}$ we have $\ell(I_k) < \delta$, and so, for any $x, y \in \mathcal{I}_k$ we must have

$$|f(x) - f(y)| < \frac{\varepsilon}{b - a},$$

but since the maximum and minimum of the function are attained, because it is continuous defined on a closed bounded interval, we also have:

$$\max\{f(x) \mid \ x \in I_k\} - \min\{f(y) \mid \ y \in I_k\} < \frac{\varepsilon}{b-a}.$$

Note that the piecewise constant function $g : [a, b] \to \mathbb{R}$ given by

$$g(x) = g_k := \max\{f(x) \mid \ x \in I_k\}, \quad \text{whenever } x \in I_k$$

majorizes f and the piecewise constant function $h : [a, b] \to \mathbb{R}$ given by

$$h(x) = h_k := \min\{f(x) \mid \ x \in I_k\}, \quad \text{whenever } x \in I_k$$

minorizes it. Therefore,

$$\sum_{k=1}^{n} h_k \ell(I_k) = \int_a^b h \le \underline{\int_a^b} f \le \overline{\int_a^b} f \le \int_a^b g = \sum_{k=1}^{n} g_k \ell(I_k)$$

From this we get that

$$\overline{\int_a^b} f - \underline{\int_a^b} f \le \sum_{k=1}^{n} (g_k - h_k)\ell(I_k).$$

But $g_k - h_k < \frac{\varepsilon}{b-a}$ and $\sum_{k=1}^{n} \ell(I_k) = b - a$, so the right hand side is less than ε, which is any positive Real number. Consequently, the upper and lower Riemann integrals of f agree and the function is then Riemann integrable. □

The fact that continuous functions are Riemann integrable provides an interesting property known as the **Mean Value Theorem for integrals**.

Theorem 12.6. *Let $f : [a, b] \to \mathbb{R}$ be a continuous function, then there is $x \in [a, b]$ such that*

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

Proof. By Theorem 12.5 we know that f is Riemann integrable, and since it is bounded below and above by m_f and M_f we have

$$m_f \cdot (b - a) \le \int_a^b f \le M_f \cdot (b - a).$$

Furthermore, since $m_f, M_f \in f([a, b])$ we have that $y = \frac{1}{b-a} \int_a^b f \in f([a, b])$. Hence, by the Intermediate Value Theorem, there is $x \in [a, b]$ such that $f(x) = y$, and so the conclusion follows. □

12.4 Exercises

1. Suppose that $f : [a, b] \to \mathbb{R}$ is Riemann integrable.

   (a) Suppose that f is non-negative all along $[a, b]$. Show that

   $$0 \le \int_a^b f. \tag{12.1}$$

   If (12.1) holds, can we conclude that f is non-negative? Prove this or give a counterexample.

   (b) For any $c \in [a, b]$, if there are other Riemann integrable functions $f_1 : [a, c] \to \mathbb{R}$ and $f_2 : [c, b] \to \mathbb{R}$ such that

   $$f(x) = f_1(x), \quad \forall x \in [a, c], \quad \text{and} \quad f(x) = f_2(x), \quad \forall x \in [c, b],$$

   Show that
   $$\int_a^b f = \int_a^c f_1 + \int_c^b f_2.$$

2. Let $f : [a, b] \to \mathbb{R}$ and $g : [a, b] \to \mathbb{R}$ be two Riemann integrable functions on $[a, b]$.

   (a) Prove the following formulas

   $$\left| \int_a^b f \right| \le \int_a^b |f| \tag{12.2}$$

   $$\max\left\{ \int_a^b f, \int_a^b g \right\} \le \int_a^b \max\{f, g\} \tag{12.3}$$

   $$\int_a^b \min\{f, g\} \le \min\left\{ \int_a^b f, \int_a^b g \right\} \tag{12.4}$$

   Hint: Use Exercise 1a.

   (b) Give an example for each inequality (12.2) - (12.4) where the equality doesn't hold.

3. Prove Theorem 12.1 for the case $\lambda < 0$.

4. Prove Theorem 12.2 for the case $\min\{f_1, f_2\}$.

5. Prove Theorem 12.4 for the case f is decreasing.

6. Let $f : [a, b] \to \mathbb{R}$ be a continuous function, Prove that for any $\bar{x} \in [a, b)$ we have

   $$\lim_{h\to 0} \frac{1}{h} \int_{\bar{x}}^{\bar{x}+h} f = f(\bar{x}).$$

Math 4031 - Advanced Calculus I

Instructor: Dr. Cristopher HERMOSILLA
Louisiana State University - Spring 2016

Week 13: Sequences of Riemann integrable functions

We now turn into the study of sequence of Riemann integrable functions and we study some criteria that allow us to interchange order in which integrals and limits are considered, that is, we are concerned with the question of when, for an appropriate notion of limit of functions, the following holds

$$(*) \qquad \lim_{n\to+\infty} \int_a^b f_n = \int_a^b \lim_{n\to+\infty} f_n.$$

13.1 Notion of limit of functions

So far, we have studied two notions of convergence for functions, namely, poitwise and uniform convergence. Recall that a sequence of functions $\{f_n\}_{n=1}^{\infty}$ defined on $[a,b]$ is said to converge pointwise to another function $f:[a,b]\to\mathbb{R}$ if

$$\forall x\in[a,b],\quad f_n(x)\to f(x) \text{ as } n\to+\infty,$$

and the sequence is said to converge uniformly to f if

$$\|f_n - f\|_\infty \to 0 \text{ as } n\to+\infty.$$

We know that uniform convergence is a stronger notion than pointwise converge; the first one implies the second one, but no vice versa. It turns out that pointwise converge by itself is too weak to allow $(*)$ to hold (without further assumptions). For example, let $\{x_n\}_{n=1}^{\infty}$ be an enumeration of the set of Rational numbers on $[0,1]$; recall that this set is infinite countable. Consider then the sequence of functions given by

$$f_n(x) = \begin{cases} 1 & \text{if } x\in\{x_1,\ldots,x_n\} \\ 0 & \text{otherwise} \end{cases} \qquad \forall x\in[0,1].$$

Clearly, the each function on the sequence is Riemann integrable. Furthermore, this sequence of functions converges pointwise to

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

But have seen that the latter is not even Riemann integral, let alone ($*$) makes any sense. We will see later that if in addition of pointwise converge, the sequence satisfies further assumptions and the limit function is also Riemann integrable, then ($*$) does hold.

The situation for uniform convergence is different, and as a matter of fact it suffices by itself to as the following theorem shows. Note that in the theorem we are not assuming that the functions are continuous; in that case the limiting function would be Riemann integrable because the uniform limit of a sequence of continuous functions is continuous too.

Theorem 13.1. *Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of Riemann integrable functions that converges uniformly to a function $f : [a, b] \to \mathbb{R}$. Then f is Riemann integrable and*

$$\lim_{n\to+\infty} \int_a^b f_n = \int_a^b f.$$

Proof. First of all, let us show that the sequence of Real numbers given by

$$y_n := \int_a^b f_n$$

converges to a Real number. To see this we use the Cauchy criterion. Note that for any $n, p \in \mathbb{N} \setminus \{0\}$ we have

$$f_{n+p}(x) - f_n(x) \leq \|f_{n+p} - f_n\|_\infty \quad \text{and} \quad f_n(x) - f_{n+p}(x) \leq \|f_{n+p} - f_n\|_\infty, \quad \forall x \in [a, b].$$

Hence

$$|y_{n+p} - y_n| \leq \|f_{n+p} - f_n\|_\infty \cdot (b - a).$$

Therefore, since $f_n \to f$ uniformly on $[a, b]$, we have that $\{f_n\}_{n=1}^{\infty}$ is a Cauchy sequence of function and so, for any $\varepsilon \in (0, +\infty)$ there is $N \in \mathbb{N}$ such that if $n, p \in \mathbb{N}$ with $N \leq n$ we have

$$\|f_{n+p} - f_n\|_\infty \leq \frac{\varepsilon}{b - a}.$$

This implies that $\{y_n\}_{n=1}^{n}$ is a Cauchy sequence of Real numbers, and thus, by the completeness of $\mathbb{R}$, it converges to some $L \in \mathbb{R}$. On the other hand, note that

$$f_n(x) - \|f_n - f\|_\infty \leq f(x) \leq f_n(x) + \|f_n - f\|_\infty, \quad \forall x \in [a, b].$$

Let $\varepsilon \in (0, +\infty)$ be fixed from now on and take $n \in \mathbb{N}$ such that

$$\|f_n - f\|_\infty \leq \frac{\varepsilon}{3 \cdot (b - a)} \quad \text{and} \quad \left| \int_a^b f_n - L \right| \leq \frac{\varepsilon}{3}.$$

Let $g_n : [a, b] \to \mathbb{R}$ and $h_n : [a, b] \to \mathbb{R}$ be two piecewise constant functions that majorizes and minorizes f_n on $[a, b]$, respectively, and such that

$$\int_a^b g_n - \frac{\varepsilon}{3} \leq \int_a^b f_n \leq \int_a^b h_n + \frac{\varepsilon}{3}$$

Note that $x \mapsto g_n(x) - \|f_n - f\|_\infty$ and $x \mapsto h_n(x) - \|f_n - f\|_\infty$ are also piecewise constant, and the first one majorizes f and the second one minorizes it. Consequently we get

$$\int_a^b h_n - \|f_n - f\|_\infty \cdot (b-a) \leq \underline{\int_a^b} f \leq \overline{\int_a^b} f \leq \int_a^b g_n(x) + \|f_n - f\|_\infty \cdot (b-a).$$

Gathering together all the preceding estimates, we get

$$L - \varepsilon \leq \int_a^b f_n - \frac{2\cdot\varepsilon}{3} \leq \underline{\int_a^b} f \leq \overline{\int_a^b} f \leq \int_a^b f_n(x) + \frac{2\cdot\varepsilon}{3} \leq L + \varepsilon.$$

But $\varepsilon \in (0, +\infty)$ is arbitrary, so we finally get that the upper and lower Riemann integrals of f agree with L, which is the limit of the Riemann integrals of f_n, and so the function is then Riemann integrable and its integral coincides with the limit of the integrals of f_n. □

13.2 Monotone convergence Theorem

In the introductory example we have exhibited shows that the mere pointwise convergence of a sequence of Riemann integrable functions is not enough to interchange limit with integral, that is, to get ($*$). The following result allows to do this change. It is worthy notice that the theorem requires the limiting function to be Riemann integrable beforehand, and so, **it is not a criterion to determine whether the limiting function is Riemann integrable.**

The theorem we review in this section is called the **Monotone Convergence Theorem** for Riemann Integrals.The proof of the theorem is based on a characterization of Riemann integrability and the so-called **Cousin's Lemma**. The latter reads as follows.

Lemma 13.1. *Let $[a, b]$ be a given bounded and closed interval and $r : [a, b] \to (0, +\infty)$ be a given positive function. Then, there exist a partition $P = \{I_1, \ldots, I_m\}$ and $y_1 < \ldots < y_m$ Real numbers such that $y_k \in I_k$ and $\ell(I_k) < 2r(y_k)$ for each $k \in \{1, \ldots, m\}$.*

Proof. Let us consider the family of open intervals $\{O_y\}_{y\in[a,b]}$ given by

$$O_y = (y - r(y), y + r(y)), \quad \forall y \in [a, b].$$

Clearly, this is an open covering of the interval $[a, b]$, and so, by the Heine-Borel Theorem, there are $y_1, \ldots, y_m \in [a, b]$ such that $\{O_{y_k}\}_{k=1}^m$ covers $[a, b]$ too. Without loss of generality we assume that $y_1 < \ldots < y_{m-1} < y_m$ and furthermore, we can also assume that no O_{y_k} contains other O_{y_l}, provided that $k \neq l$. The latter means that

$$y_{k+1} - r(y_{k+1}) < y_k + r(y_k), \quad \forall k \in \{1, \ldots, m\}.$$

Note that $y_1 - r(y_1) < a$ and $b < y_m + r(y_m)$. Hence, by defining $x_0 = a$ and $x_m = b$, we see that taking any $x_k \in (y_{k+1} - r(y_{k+1}), y_k + r(y_k)) \cap (y_k, y_{k+1})$ for any $k \in \{1, \ldots, m-1\}$ we get the desired properties. □

Before going further, we introduce an useful criterion for a function to be Riemann integrable, which is a version of the so-called **Darboux criterion** for Riemann integrability.

Definition 13.1. *We say that a function $f : [a,b] \to \mathbb{R}$ satisfies the Darboux property if for any $\varepsilon \in (0,+\infty)$ there is $\delta \in (0,+\infty)$ such that for any partition $P = \{J_1,\dots,J_m\}$ of $[a,b]$ with mesh $\|P\| < \delta$ and any collection of points $y_1,\dots,y_m \in [a,b]$ such that $y_j \in J_j$ for any $j \in \{1,\dots,m\}$ we have*

$$\sum_{j=1}^{m} \left| \int_{J_j} f - f(y_j) \cdot \ell(J_j) \right| \leq \varepsilon.$$

It's worthy to notice that the Darboux property is sometimes taken as definition for a function to be Riemann integrable. This is because a function $f : [a,b] \to \mathbb{R}$ is Riemann integrable if and only if it satisfies the Darboux criterion. We now prove the one of the implication, the other is left as exercise for the reader.

Lemma 13.2. *Any Riemann integrable function $f : [a,b] \to \mathbb{R}$ satisfies the Darboux property.*

Proof. Let $\varepsilon \in (0,+\infty)$, since f is Riemann integrable, it is bounded ($\|f\|_\infty \in \mathbb{R}$) and there are two piecewise constant functions $h : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ with $h \leq f \leq g$ such that

$$\int_a^b g - \frac{\varepsilon}{4} \leq \int_a^b f \leq \int_a^b h + \frac{\varepsilon}{4}.$$

Without loss of generality, we assume that h and g are piecewise constant relative to the same partition $P_0 = \{I_1,\dots,I_n\}$ and that for each $i \in \{1,\dots,n\}$ we have $0 < \ell(I_i)$. Let

$$\delta = \min\left\{ \frac{\varepsilon}{4 \cdot n \cdot (\|f\|_\infty + 1)}, \ell(I_1), \dots, \ell(I_1) \right\} \in (0,+\infty)$$

and take any partition $P = \{J_1,\dots,J_m\}$ of $[a,b]$ such that $\|P\| < \delta$. Let $y_1,\dots,y_m \in [a,b]$ such that $y_j \in J_j$ for any $j \in \{1,\dots,m\}$ and consider the set of indexes

$$\Lambda = \{j \in \{1,\dots,m\} \mid \ \exists i \in \{1,\dots,n\}\ J_j \subseteq I_i\}.$$

Note that for any $j \in \Lambda$ we have

$$h(x) \leq f(y_j) \leq g(x), \quad \forall x \in J_j.$$

Therefore, we obtain that

$$\int_{J_j} h \leq f(y_j) \cdot \ell(J_j) \leq \int_{J_j} g.$$

Hence, using that $f \leq g$ and that $h \leq f$, we get, respectively,

$$\int_{J_j} f - f(y_j) \cdot \ell(J_j) \leq \int_{J_j} g - \int_{J_j} h \quad \text{and} \quad f(y_j) \cdot \ell(J_j) - \int_{J_j} f \leq \int_{J_j} g - \int_{J_j} h.$$

Moreover, since $0 \leq g - h$ we obtain

$$\sum_{j \in \Lambda} \left(\int_{J_j} g - \int_{J_j} h \right) \leq \int_a^b g - \int_a^b h \leq \frac{\varepsilon}{2}.$$

Therefore

$$\sum_{j\in\Lambda}\left|\int_{J_j} f - f(y_j)\cdot \ell(J_j)\right| \leq \frac{\varepsilon}{2}$$

On the other hand, for each $j \notin \Lambda$, we have that

$$\left|\int_{J_j} f - f(y_j)\cdot \ell(J_j)\right| \leq 2\cdot \|f\|_\infty \cdot \|P\| \leq \frac{\varepsilon}{2\cdot n}$$

Let us point out that $\ell(J_j) \leq \|P\| < \ell(I_i)$ for any $i \in \{1,\dots,n\}$. Consequently, J_j is covered by exactly two subintervals of the partition P_0. This leads to state that the number of indexes that don't belong to Λ is at most n, and so the proof is complete because we have

$$\sum_{j\notin\Lambda}\left|\int_{J_j} f - f(y_j)\cdot \ell(J_j)\right| \leq \frac{\varepsilon}{2}$$

□

We are now in position to prove the Monotone Convergence Theorem.

Theorem 13.2. *Let $\{f_n\}_{n=1}^\infty$ be a sequence of Riemann integrable functions that converges pointwise to a function $f : [a,b] \to \mathbb{R}$. Suppose that f is also Riemann integrable and that $\{f_n\}_{n=1}^\infty$ is monotonic, that is, one of the following holds*

$$\forall x \in [a,b],\ \forall n \in \mathbb{N}\setminus\{0\}, \quad f_n(x) \leq f_{n+1}(x). \tag{13.1}$$

$$\forall x \in [a,b],\ \forall n \in \mathbb{N}\setminus\{0\}, \quad f_{n+1}(x) \leq f_n(x). \tag{13.2}$$

Then one has

$$\lim_{n\to+\infty}\int_a^b f_n = \int_a^b f.$$

Proof. Without loss of generality we can assume that $f = 0$ and that each f_n is nonnegative and the sequence is decreasing; it is enough to change f_n with $f - f_n$ in the first case and $f_n - f$ in the second one. Here is important the fact that f_n and f are both Riemann integrable. Also, using the change of variables $x \mapsto \frac{x-a}{b-a}$ we might also assume that $a = 0$ and $b = 1$.

First of all we note that since $0 \leq f_{n+1} \leq f_n$ we have that

$$0 \leq y_{n+1} \leq y_n, \quad \text{where } y_n := \int_0^1 f_n.$$

Therefore, the sequence of Real numbers $\{y_n\}_{n=1}^\infty$ converges to some $L \in [0,+\infty)$. Since we are in the case $f = 0$, we must show that $L = 0$ to conclude.

Let $\varepsilon \in (0,+\infty)$, by the Darboux property, for any $n \in \mathbb{N}\setminus\{0\}$ there is $\delta_n \in (0,+\infty)$ such that for any partition $P = \{I_1,\dots,I_m\}$ of $[0,1]$ with mesh $\|P\| < \delta_n$ and any collection $y_1,\dots,y_m \in [0,1]$ such that $y_k \in I_k$ we have that

$$\sum_{k=1}^m \left|\int_{I_k} f_n - f_n(y_k)\cdot \ell(I_k)\right| \leq \frac{\varepsilon}{2^{n+1}}. \tag{13.3}$$

Furthermore, we define

$$N(x) := \inf\left\{n \in \mathbb{N}\setminus\{0\} \mid \ f_n(x) \le \frac{\varepsilon}{2}\right\}, \quad \forall x \in [0,1].$$

By the Cousin's Lemma applied with $r(x) = \frac{1}{2}\cdot\delta_{N(x)}$ there exist a partition $P = \{I_1, \ldots, I_m\}$ and $y_1 < \ldots < y_m$ such that $y_k \in I_k$ and $\ell(I_k) < \delta_{N(y_k)}$ for any $k \in \{1, \ldots, m\}$. Take $n \in \mathbb{N}$ such that $N_0 := \max\{N(y_1), \ldots, N(y_m)\} \le n$. Hence

$$0 \le \int_0^1 f_n = \sum_{k=1}^m \int_{I_k} f_n \le \sum_{k=1}^m \int_{I_k} f_{N(y_k)} = \sum_{j=1}^{N_0} \sum_{k\in J(j)} \int_{I_k} f_{N(y_k)}.$$

where $J(j) = \{k \in \{1, \ldots, m\} \mid \ j = N(y_k)\}$ for any $j \in \{1, \ldots, N_0\}$. Note that some $J(j)$ can possibly be the empty set. Also, remark that $\ell(I_k) < \delta_j$ for any $k \in J(j)$. Therefore, by (13.3) and the definition of $N(y_k)$ we get

$$\sum_{k\in J(j)} \int_{I_k} f_{N(y_k)} \le \sum_{k\in J(j)} f_{N(y_k)}(y_k)\cdot\ell(I_k) + \frac{\varepsilon}{2^{j+1}} \le \frac{\varepsilon}{2}\cdot\left(\sum_{k\in J(j)} \ell(I_k) + \frac{1}{2^j}\right).$$

This leads then to

$$0 \le \int_0^1 f_n = \frac{\varepsilon}{2}\cdot\sum_{j=1}^{N_0}\left(\sum_{k\in J(j)} \ell(I_k) + \frac{1}{2^j}\right) \le \frac{\varepsilon}{2}\cdot(1+1) = \varepsilon.$$

Given that $\varepsilon \in (0, +\infty)$ is arbitrary, the conclusion follows. □

13.3 Functionals and Integral equations

We have considered so far Real-valued functions defined on subsets of $\mathbb{R}$, that is, functions of the type $f : I \subseteq \mathbb{R} \to \mathbb{R}$. In this section we introduce the idea of function of function, which we might call **functional** to make a distinction with respect to Real-valued functions.

We begin by introducing some notation, let us denote by $C([a,b])$ the set of all Real-valued continuous functions defined on the closed bounded interval $[a,b]$. We call a functional, denoted generically by $T : C([a,b]) \to C([a,b])$, to any mapping defined on $C([a,b])$ and whose value is an element in $C([a,b])$, that is,

$$\forall f \in C([a,b]), \quad \exists \varphi \in C([a,b]) \text{ such that } T(f) = \varphi.$$

The Riemann integral allows us to define in several way a functional on $C([a,b])$.

Example 13.1. *Let $f \in C([a,b])$, let us consider the functional $T : C([a,b]) \to C([a,b])$ defined via*

$$T(f)(x) = \int_a^x f, \quad \forall x \in [a,b].$$

The function $T(f)$ defines a continuous function; as a matter of fact, Lipschitz continuous function. Indeed, let us denote $\varphi = T(f)$, then

$$\varphi(x) - \varphi(y) = \int_a^x f - \int_a^y f = \begin{cases} \int_y^x f & \text{if } y < x \\ -\int_x^y f & \text{if } x < y \end{cases} \quad \forall x, y \in [a, b].$$

Since $f \in C([a, b])$ then, $|f| \leq \|f\|_\infty$ and so

$$|\varphi(x) - \varphi(y)| \leq \|f\|_\infty |x - y| \quad \forall x, y \in [a, b].$$

Example 13.2. *Let $f \in C([a, b])$ and $K : [a, b] \times [a, b] \to \mathbb{R}$ be a function such that*

1. *$\exists L \in [0, +\infty)$ such that $x \mapsto K(x, u)$ is Lipschitz continuous with constant L for any $u \in [a, b]$ (the same for any u).*

2. *$u \mapsto K(x, u)$ is Riemann integrable for any $x \in [a, b]$.*

We consider the functional

$$T(f)(x) = \int_a^b (K(x, \cdot) \cdot f).$$

Note that each $T(f)(x) \in \mathbb{R}$, this is due to the fact that for any $x \in [a, b]$ the map $u \mapsto K(x, u)$ is Riemann integrable, and then $u \mapsto K(x, u) \cdot f(u)$ is also Riemann integrable, and so integral if well defined. For sake of simplicity, let $\varphi = T(f)$ and so

$$\varphi(x) - \varphi(z) = \int_a^b (K(x, \cdot) \cdot f) - \int_a^b (K(z, \cdot) \cdot f) = \int_a^b [(K(x, \cdot) - K(z, \cdot)) \cdot f]$$

Remark that

$$(K(x, u) - K(y, u)) \cdot f(y) \leq |K(x, u) - K(y, u)| \cdot \|f\|_\infty \leq L \cdot |x - y| \cdot \|f\|_\infty$$

Therefore,

$$|\varphi(x) - \varphi(y)| \leq L|x - y| \cdot \|f\|_\infty \cdot (b - a).$$

From where we actually get that $x \mapsto \varphi(x)$ is Lipschitz continuous on $[a, b]$.

Since a functional is essentially a function between sets endowed with a norm, we can also define notions of continuity. For the scope of the exposition, the most important one is Lipschitz continuity.

Definition 13.2. *We say that a functional $T : C([a, b]) \to C([a, b])$ is Lipschitz continuous if there is $L \in [0, +\infty)$ such that*

$$\|T(f) - T(g)\|_\infty \leq L \cdot \|f - g\|_\infty$$

Example 13.3. *Let us pick up the data from Example 13.1. Let $f, g \in C([a,b])$ be given, then for any $x \in [a,b]$ fixed but arbitrary we have*

$$|T(f)(x) - T(g)(x)| = \left| \int_a^x f - \int_a^x g \right| = \left| \int_a^x (f-g) \right| \leq \int_a^x |f-g|$$

Given that $|f(y) - g(y)| \leq \|f-g\|_\infty$ for any $y \in [a,x]$ we get

$$|T(f)(x) - T(g)(x)| \leq \|f-g\|_\infty \cdot (x-a) \leq \|f-g\|_\infty \cdot (b-a)$$

Since the right hand side doesn't depend on x, we can take supremum and get

$$\|T(f) - T(g)\|_\infty \leq (b-a) \cdot \|f-g\|_\infty.$$

Hence, $f \mapsto T(f)$ is a Lipschitz continuous functional with Lipschitz constant $L = (b-a)$.

Example 13.4. *Let $F : \mathbb{R} \to \mathbb{R}$ be a Lipschitz continuous function with Lipschitz constant $L_F \in (0, +\infty)$ and let $y_0 \in \mathbb{R}$. We consider the functional defined on $C([a,b])$ given by*

$$T(f)(x) = y_0 + \int_a^x F \circ f, \qquad \forall f \in C([a,b]).$$

Note that the Riemann integral of $F \circ f$ is well-defined because that function is continuous thanks to the composition rule for continuous function. Therefore, we see that

$$|T(f)(x) - T(g)(x)| = \left| \int_a^x F \circ f - \int_a^x F \circ g \right| = \left| \int_a^x (F \circ f - F \circ g) \right| \leq \int_a^x |F \circ f - F \circ g|$$

The Lipschitz continuity of $y \mapsto F(y)$, implies that

$$|F \circ f(y) - F \circ g(y)| = |F(f(y)) - F(g(y))| \leq L_F \cdot |f(y) - g(y)| \leq L_F \cdot \|f-g\|_\infty.$$

Consequently, we get that

$$|T(f)(x) - T(g)(x)| \leq (b-a) \cdot L \cdot \|f-g\|_\infty$$

So, the functional $f \mapsto T(f)$ is Lipschitz continuous with Lipschitz constant $L_F \cdot (b-a)$.

13.3.1 Banach Fixed Point Theorem

Since we have introduced the notion of function among a set of functions, we can also introduce the notion of functional equation, that is, an equation where the unknown is a function. For example, we would like to find at least a $f \in C([a,b])$ such that

$$T(f) = f \tag{13.4}$$

for some functional $T : C([a,b]) \to C([a,b])$. Any function $f \in C([a,b])$ that satisfies (13.4) is called a **fixed point** of the functional T.

Remark 13.1. *The functional exhibited in Example 13.4 plays a key role in the theory of ordinary differential equations. Indeed, any solution to an ordinary differential equation*

$$f' = F(f), \quad f(a) = y_0$$

is by definition a function that satisfies

$$f(x) = T(f)(x) = y_0 + \int_a^x F \circ f, \quad \forall x \in [a, b].$$

Hence, the existence of solutions to an ordinary differential equation can be study by analyzing the fixed points of the functional T.

Fixed point theorems are more difficult for function defined on $C([a, b])$ than in $\mathbb{R}$, because for instance in the first set there is no result playing the role of the Intermediate Value Theorem in $\mathbb{R}$; essentially because not every bounded sequence of continuous functions has a subsequence that converges uniformly.

To find fixed point of functional we need to use the completeness of the space of continuous function. The following is one of the most classical result regarding fixed points of functional. It is worth noting that the following theorem provides more information about the fixed point, it says that it is unique.

Theorem 13.3. *Let $T : C([a, b]) \to C([a, b])$ be a Lipschitz continuous functional and suppose that its Lipschitz constant L belongs to $(0, 1)$. Then, there is a unique fixed point of T, that is*

$$\exists! f \in C([a, b]), \quad T(f) = f.$$

Proof. Let $f_1 \in C([a, b])$ be any function, and define inductively the sequence of functions

$$f_{n+1} = T(f_n), \quad \forall n \in \mathbb{N} \setminus \{0\}.$$

Then, for any $n \in \mathbb{N} \setminus \{0, 1\}$ we have

$$\|f_{n+1} - f_n\|_\infty = \|T(f_n) - T(f_{n-1})\|_\infty \leq L \cdot \|f_n - f_{n-1}\|_\infty \leq \ldots \leq L^{n-1}\|f_2 - f_1\|_\infty$$

Let any $n, p \in \mathbb{N} \setminus \{0\}$, then

$$\|f_{n+p} - f_n\|_\infty \leq \sum_{k=n}^{n+p-1} \|f_{k+1} - f_k\|_\infty \leq \sum_{k=n}^{n+p-1} L^{k-1}\|f_2 - f_1\|_\infty = L^{n-1} \cdot \frac{1 - L^{p+1}}{1 - L} \cdot \|f_2 - f_1\|_\infty$$

Thanks to the fact that $L \in (0, 1)$ we have that

$$L^{n-1} \cdot \frac{1 - L^{p+1}}{1 - L} \to 0 \quad \text{as } n \to +\infty$$

and so, it is easy to see that the sequence $\{f_n\}_{n=1}^\infty$ is a Cauchy sequence, and it converges uniformly to some $f \in C([a, b])$.

Let us see that f is a fixed point of T. Note that for any $n \in \mathbb{N} \setminus \{0\}$ we have

$$\|T(f) - f\|_\infty \leq \|T(f) - T(f_n)\|_\infty + \|T(f_n) - f_n\|_\infty + \|f_n - f\|_\infty$$

Thus, using the Lipschitz continuity of T, the definition of f_{n+1} and the estimate founded above, we get
$$\|T(f) - f\|_\infty \leq (L+1) \cdot \|f_n - f\|_\infty + L^{n-1}\|f_2 - f_1\|_\infty$$
Letting $n \to +\infty$ we finally obtain that $\|T(f) - f\|_\infty = 0$ and so, since $T(f)$ and f are both continuous, we conclude that $T(f) = f$.

The only issue remaining is the uniqueness, which comes from the following observation: If $f \in C([a,b])$ and $g \in C([a,b])$ are both fixed point of T, then

$$\|f - g\|_\infty = \|T(f) - T(g)\|_\infty \leq L\|f - g\|_\infty$$

But, since $L \in (0,1)$ the only option is that $\|f-g\|_\infty = 0$ and so $f = g$, because both functions are continuous on $[a,b]$. $\square$

13.4 Exercises

1. Consider the sequence of functions $\{f_n\}_{n=1}^\infty$ defined on $\left[0, \frac{\pi}{4}\right]$ via

$$f_n(x) = \sin^n(x), \quad x \in \left[0, \frac{\pi}{4}\right].$$

Show that
$$\lim_{n\to+\infty} \int_0^{\frac{\pi}{4}} f_n = 0$$

2. Let $\{f_n\}_{n=1}^\infty$ be a sequence of nonnegative Riemann integrable functions on $[a,b]$. Suppose that the function $f : [a,b] \to \mathbb{R}$ given below is well-defined and Riemann integrable

$$f(x) = \lim_{n\to+\infty} \sum_{k=1}^n f_k(x), \quad \forall x \in [a,b].$$

Prove that
$$\int_a^b f = \sum_{k=1}^\infty \int_a^b f_k := \lim_{n\to+\infty} \sum_{k=1}^n \int_a^b f_k$$

3. Show that there is a unique continuous function $f : [0,1] \longrightarrow \mathbb{R}$ such that

$$f(x) = \sqrt{x} + \int_0^1 (K(x,\cdot) \cdot f),$$

where $K(x,u) = \exp(-(x+u+1))$ for any $x, u \in [0,1]$.

4. Let $K : [0,1] \times [0,1] \longrightarrow \mathbb{R}$ be a continuous function so that $|K(x,u)| < 1$ $\forall x, u \in [0,1]$. Prove that there is a unique continuous function $f : [0,1] \longrightarrow \mathbb{R}$ such that that satisfies

$$f(x) + \int_0^1 (K(x,\cdot) \cdot f) = \exp(x^2).$$

Math 4031 - Advanced Calculus I

Instructor: Dr. Cristopher HERMOSILLA
Louisiana State University - Spring 2016

Week 14: Special topics on Riemann integrability

We end the exposition about the Riemann integral studying some classes of functions for which it is possible to compute their Riemann integral (in some sense) regardless the fact that they may not be Riemann integrable in the sense we have adopted in this course.

14.1 p-Riemann integrable functions

Recall that if a $f : [a, b] \to \mathbb{R}$ is Riemann integrable, then $|f|$ is also Riemann integrable. Furthermore, since the product of Riemann integrable function is also Riemann integrable, we can infer that for any $p \in \mathbb{N}$, the $|f|^p$ is also Riemann integrable.

Definition 14.1. *Let $p \in \mathbb{Q} \cap (0, +\infty)$. We say that a function $f : [a, b] \to \mathbb{R}$ is p-Riemann integrable if $|f|^p$ is Riemann integrable and*

$$\|f\|_p := \left(\int_a^b |f|^p \right)^{\frac{1}{p}} \in \mathbb{R}.$$

The value $\|f\|_p$ is called the p-norm of f.

As we have pointed out, any Riemann integrable function is also p-Riemann integrable if $p \in \mathbb{N}$; this fact is also true if $p \in \mathbb{Q} \cap (0, +\infty)$, but their proof is beyond the scope of these notes. The converse is not true, there are p-Riemann integrable functions that are not Riemann integrable, for example

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \cap [0, 1] \\ -1 & \text{if } x \in [0, 1] \setminus \mathbb{Q} \end{cases}$$

14.1.1 Properties of the p-norms

The role of the p-norms over set of Riemann integrable function can be compared with the role of the absolute value and the sup-norm over $\mathbb{R}$ and $C([a, b])$, respectively. There are some properties in common for all the three, but others that are only characteristic of absolute value and the sup-norm. For example, the p-norm of a Riemann integrable function can equal zero but, the function not be the constant function zero. However, we know that

$$\forall x \in \mathbb{R}, \quad |x| = 0 \quad \Leftrightarrow \quad x = 0 \quad \wedge \quad \forall f \in C([a, b]), \quad \|f\|_\infty = 0 \quad \Leftrightarrow \quad f(x) = 0, \ \forall x \in [a, b].$$

Recall that for any $\lambda \in \mathbb{R}$ we have that

$$|x+\lambda\cdot y| \le |x|+|\lambda|\cdot|y| \quad \forall x,y\in\mathbb{R} \quad \text{and} \quad \|f+\lambda\cdot g\|_\infty \le \|f\|_\infty+|\lambda|\cdot\|g\|_\infty \quad \forall f,g\in C([a,b]).$$

This property means that absolute value and the sup-norm are sublinear over $\mathbb{R}$ and $C([a,b])$, respectively. A similar inequality holds for the p-norm over the set of Riemann integrable functions.

Theorem 14.1. *Let $p \in \mathbb{N}\backslash\{0\}$ and let $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ be Riemann integrable, then for any $\lambda \in \mathbb{R}$, $f + \lambda \cdot g$ is p-Riemann integrable and*

$$\|f + \lambda \cdot g\|_p \le \|f\|_p + |\lambda| \cdot \|g\|_p.$$

The proof of the theorem is based on two preliminary inequalities.

Lemma 14.1 (Young's Inequality)**.** *Let $p \in \mathbb{N} \setminus \{0,1\}$ and set $q = \frac{p}{p-1}$. Then, for any $a,b \in [0,+\infty)$ we have that*

$$a\cdot b \le \frac{1}{p}\cdot a^p + \frac{1}{q}\cdot b^q.$$

Proof. First of all, if $a = 0$ or $b = 0$ then the inequality is trivial, so, let us assume that $a, b \in (0,+\infty)$.

Recall that for any $p \in \mathbb{N}$ and $h \in (-1,+\infty)$ we have the Bernouilli's inequality (see equation (8.8)), that is,

$$1 + p\cdot h \le (1+h)^p.$$

Let $h = \frac{1}{p}\cdot\left(\frac{a^p}{b^q}-1\right)$. Since $\frac{a^p}{b^q} \in (0,+\infty)$ we have that $-\frac{1}{p} < h$ and so by the Bernouilli's inequality we get that

$$\frac{a}{b^{\frac{q}{p}}} = (1+p\cdot h)^{\frac{1}{p}} \le 1+h = \frac{1}{p}\cdot\frac{a^p}{b^q}+\frac{1}{q}.$$

The multiplying by b^q and using the fact that $q - \frac{q}{p} = 1$ we get the desired result. □

The other inequality we need is known as the Hölder's inequality.

Lemma 14.2 (Hölder's Inequality)**.** *Let $p \in \mathbb{N} \setminus \{0,1\}$ and set $q = \frac{p}{p-1}$. Let $f : [a,b] \to \mathbb{R}$ be p-Riemann integrable and and $g : [a,b] \to \mathbb{R}$ be q-Riemann integrable. If $f \cdot g$ is 1-Riemann integrable then*

$$\|f\cdot g\|_1 \le \|f\|_p \cdot \|g\|_q.$$

Proof. Note that if $\|f\|_p = 0$, then $|f| = 0$ except at a finite number of points on $[a,b]$, and so $f \cdot g = 0$ except at a finite number of points on $[a,b]$. Hence $\|f \cdot g\|_1 = 0$ and so the conclusion would follow. A similar remark holds for the case $\|g\|_p = 0$. Therefore, without loss of generality, we assume that $\|f\|_p, \|g\|_q \in (0,+\infty)$.

Let $x \in [a,b]$ be fixed but arbitrary and set

$$a = \frac{|f(x)|}{\|f\|_p} \quad \text{and} \quad b = \frac{|g(x)|}{\|g\|_q}.$$

By the Young's inequality we obtain

$$\frac{|f(x)| \cdot |g(x)|}{\|f\|_p \cdot \|g\|_q} \leq \frac{1}{p} \cdot \frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{q} \cdot \frac{|g(x)|^q}{\|g\|_q^q}.$$

Since $x \in [a,b]$ is arbitrary, by integrating the last inequality we that

$$\frac{\|f \cdot g\|_1}{\|f\|_p \cdot \|g\|_q} \leq \frac{1}{p \cdot \|f\|_p^p} \cdot \int_a^b |f|^p + \frac{1}{q \cdot \|g\|_q^q} \cdot \int_a^b g|^q = \frac{1}{p} + \frac{1}{q} = 1.$$

From where the conclusion follows. □

Remark 14.1. *The Hölder's inequality is also true for the case $p = 1$, but in this case $\|g\|_q$ is replace with , $\|g\|_\infty$. This is a direct consequence of the fact that*

$$|f(x) \cdot g(x)| \leq |f(x)| \cdot \|g\|_\infty.$$

We are now in position to prove Theorem 14.1.

Proof of Theorem 14.1. Since f and g are Riemann integrable, so they are $f+\lambda \cdot g$ and $|f+\lambda \cdot g|$. In particular, for any $k \in \{2, \ldots, p\}$, the function $|f + \lambda \cdot g|^k$ is Riemann integrable.

Note that $(p-1) \cdot q = p$, thus $|f+\lambda \cdot g|^{p-1}$ is Riemann integrable, and so it is also q-Riemann integrable. By the Hölder's inequality, since $|f| \cdot |f+\lambda \cdot g|^{p-1}$ and $|g| \cdot |f+\lambda \cdot g|^{p-1}$ are Riemann integrable and nonnegative, we get that

$$\||f| \cdot |f + \lambda \cdot g|^{p-1}\|_1 \leq \|f\|_p \cdot \|f + \lambda \cdot g\|^{\frac{p}{q}} \quad \text{and} \quad \||g| \cdot |f + \lambda \cdot g|^{p-1}\|_1 \leq \|g\|_p \cdot \|f + \lambda \cdot g\|^{\frac{p}{q}}.$$

Now, since

$$|f(x) + \lambda \cdot g(x)|^p \leq (|f(x) + |\lambda| \cdot |g(x)|) \cdot |f(x) + \lambda \cdot g(x)|^{p-1}, \qquad \forall x \in [a,b],$$

integrating this inequality and using the first inequality we get that

$$\|f + \lambda \cdot g\|_p^p \leq (\|f\|_p + |\lambda| \cdot \|g\|_p) \cdot (\|f + \lambda \cdot g\|^{\frac{p}{q}}).$$

Using finally the fact that $p - \frac{p}{q} = 1$ the proof is complete. □

For a given Riemann integrable function, the value of p-norm times a given factor (depending only on p and $[a,b]$) increases with p, eventually reaching the sup-norm of f.

Theorem 14.2. *Let $f : [a,b] \to \mathbb{R}$ be Riemann integrable, then*

$$\|f\|_1 \leq (b-a)^{\frac{p-1}{p}} \cdot \|f\|_p \leq (b-a)^{\frac{p+h-1}{p+h}} \cdot \|f\|_{p+h} \leq (b-a) \cdot \|f\|_\infty, \quad \forall p, h \in \mathbb{N} \setminus \{0\}.$$

Furthermore, if f is continuous, then

$$\lim_{p \to +\infty} \|f\|_p = \|f\|_\infty.$$

Proof. For the first inequality, we use the Hölder's inequality with $g(x) = 1$ for any $x \in [a, b]$. Furthermore, note that $\|f\|_p^p = \|f^p\|_1$, hence using the Hölder's inequality f^p and g as before the second inequality follows. The last inequality is simply a consequence of the definition of the sup-norm.

On the other hand, for any $p \in \mathbb{N} \setminus \{0\}$ let us set $x_p = (b-a)^{\frac{p-1}{p}} \cdot \|f\|_p$. By the preceding part, the sequence generated by this numbers is increasing and bounded above by $(b-a)\cdot\|f\|_\infty$. Therefore, $\{x_p\}_{p=1}^\infty$ converges to some $L \in \mathbb{R}$, and furthermore, $L \leq (b-a)\cdot\|f\|_\infty$. Note that $\{(b-a)^{\frac{p-1}{p}}\}_{p=1}^\infty$ converges to $(b-a)$ as $p \to +\infty$. So, the sequence generated by

$$\|f\|_p = \frac{1}{(b-a)^{\frac{p-1}{p}}} \cdot x_p$$

also converges, and its limit is less than or equal to $\|f\|_\infty$. Note that up to this point we haven't used the fact that f is continuous.

Now, since f is assume to be continuous on $[a, b]$ there is $\bar{x} \in [a, b]$ such that $\|f\|_\infty = |f(\bar{x})|$. Furthermore, by continuity of $|f|$, for any $\varepsilon \in (0, \|f\|_\infty)$ and there is $\delta \in (0, +\infty)$ such that

$$0 \leq \|f\|_\infty - \varepsilon < |f(x)|, \quad \forall x \in (\bar{x}-\delta, \bar{x}+\delta) \cap [a, b].$$

Let $I = (\bar{x}-\delta, \bar{x}+\delta) \cap [a, b]$ and note that $\ell(I) \in (0, +\infty)$. Then,

$$(\|f\|_\infty - \varepsilon)^p \cdot \ell(I) \leq \int_I |f|^p \leq \|f\|_p^p.$$

This yields then to $(\|f\|_\infty - \varepsilon)\cdot \ell(I)^{\frac{1}{p}} \leq \|f\|_p$. So, letting $p \to +\infty$ we get that $\ell(I)^{\frac{1}{p}} \to 1$, from where we obtain

$$\|f\|_\infty - \varepsilon \leq \lim_{p\to+\infty} \|f\|_p$$

and the conclusion follows because $\varepsilon \in (0, \|f\|_\infty)$ is arbitrary and positive. ☐

Remark 14.2. *The assumption that f is continuous on Theorem 14.2 is important, otherwise the limit, which always exists, can be strictly less than $\|f\|_\infty$. For instance, consider*

$$f(x) = \begin{cases} 0 & \text{if } x \in [0,1) \cup (1,2] \\ 1 & \text{if } x = 1 \end{cases}$$

Clearly, $\|f\|_p = 0$ for any $p \in \mathbb{N} \setminus \{0\}$ but $\|f\|_\infty = 1$.

14.2 Improper integrals

Recall that to define the Riemann integral we have restrict ourselves to **bounded** functions defined on **closed and bounded** intervals.

We now show that is it possible to extend the notion of Riemann integral to functions that don't satisfy some of the boundedness assumptions described above.

Before going further, let us define the lateral limit of a function. Let $r \in (0, +\infty)$ and consider a function $F : (0, r) \to \mathbb{R}$. We say that $L \in \mathbb{R}$ is the limit from above of F at 0 if

$$\forall \varepsilon \in (0, +\infty),\ \exists \delta \in (0, r),\ \forall h \in \mathbb{R},\ (0 < h < \delta \ \Rightarrow\ |F(h) - L| < \varepsilon.$$

In this case we denote the limit from above by

$$\lim_{h \to 0^+} F(h)$$

Definition 14.2. *Let $a, b \in \mathbb{R}$ with $a < b$, and $f : (a, b) \to \mathbb{R}$ be a given function. We say that f is integrable if:*

- *for any $h \in (0, \frac{b-a}{2})$, f is Riemann integrable on $[a + h, b - h]$.*
- *the following limit exists*

$$\lim_{h \to 0^+} \int_{a+h}^{b-h} f$$

Under these circumstances we say that the improper integral converges and we denote its value in the same way as the Riemann integral, that is,

$$\int_a^b f = \lim_{h \to 0^+} \int_{a+h}^{b-h} f$$

We now evoke some results concerning previous calculus courses:

Let $\alpha \in \mathbb{Q} \cap (0, +\infty)$ and consider $f_\alpha(x) = \frac{1}{(1-x)^\alpha}$ defined on $(0, 1)$. We know that for any $h \in (0, \frac{1}{2})$ we have

$$\int_h^{1-h} f_\alpha = \begin{cases} \ln(1-h) - \ln(h) & \text{if } \alpha = 1 \\ \dfrac{1}{1-\alpha}\left(\dfrac{1}{(1-h)^{\alpha-1}} - \dfrac{1}{h^{\alpha-1}}\right) & \text{otherwise} \end{cases}$$

Since $h^{\alpha-1} \to 0$ if and only if $\alpha \in (1, +\infty)$ and $\frac{1}{h^{\alpha-1}} \to 0$ if and only if $\alpha \in (0, 1)$. We get that the improper integral of f_α converges if and only if $\alpha \in (0, 1)$. In any other case, the limit doesn't exist.

Let us now present a criterion for the convergence of improper integrals.

Theorem 14.3. *Let $f : (a, b) \to \mathbb{R}$ and $g : (a, b) \to \mathbb{R}$ be two given functions, such that*

- *for any $h \in (0, \frac{b-a}{2})$, f and g are Riemann integrable on $[a + h, b - h]$.*
- *there is $r \in (0, b - a)$ for which $0 \leq f(x) \leq g(x)$ for any $x \in (a, a + r) \cup (b - r, b)$.*

If the improper integral of g converges, then so does the improper integral of f.

Proof. Note that, thanks to the first assumption, for any $h \in (0, r)$ we have that

$$\int_{a+h}^{b-h} f = \int_{a+h}^{a+r} f + \int_{a+r}^{b-r} f + \int_{b-r}^{b-h} f$$

Let $F_1 : (0, r) \to \mathbb{R}$ and $F_2 : (0, r) \to \mathbb{R}$ be given by

$$F_1(h) = \int_{a+h}^{a+r} f \quad \text{and} \quad F_2(h) = \int_{b-r}^{b-h} f, \quad \forall h \in (0, r).$$

The fact that f is non negative on $(a, a+r) \cup (b-r, b)$ implies that both functions are decreasing.

Similarly, we define $G_1 : (0, r) \to \mathbb{R}$ and $G_2 : (0, r) \to \mathbb{R}$ via

$$G_1(h) = \int_{a+h}^{a+r} g \quad \text{and} \quad G_2(h) = \int_{b-r}^{b-h} g, \quad \forall h \in (0, r).$$

This function are as well decreasing and by the second assumption we have that $F_1 \leq G_1$ and $F_2 \leq G_2$ on $(a, a+r) \cup (b-r, b)$. Moreover, since the improper integral of g converges we have that the limit from above of G_1 and G_2 at 0 exist, this leads then to

$$F_1(h) \leq G_1(h) \leq \lim_{h \to 0^+} G_1(h) \quad \text{and} \quad F_2(h) \leq G_2(h) \leq \lim_{h \to 0^+} G_2(h), \qquad \forall h \in (0, r).$$

Hence, $\sup\{F_1(h) \mid h \in (0, r)\}$ and $\sup\{F_2(h) \mid h \in (0, r)\}$ are well defined Real numbers. Combining this with the fact that F_1 and F_2 are decreasing functions, we get that

$$\lim_{h \to 0^+} F_1(h) = \sup\{F_1(h) \mid h \in (0, r)\} \quad \text{and} \quad \lim_{h \to 0^+} F_2(h) = \sup\{F_2(h) \mid h \in (0, r)\}, \qquad \forall h \in (0, r).$$

□

14.3 Exercises

1. Consider the sequence of continuous functions $\{f_n\}_{n=1}^{\infty}$ given by

$$f_n(x) := \begin{cases} x^n & \text{if } x \in [0, 1] \\ 1 & \text{if } x \in (1, 2] \end{cases}$$

Prove that $\{f_n\}_{n=1}^{\infty}$ satisfies the Cauchy criterion for the 1-norm, that is,

$$\forall \varepsilon \in (0, +\infty),\ \exists N \in \mathbb{N} \setminus \{0\},\ \forall n, p \in \mathbb{N}, \quad N \leq n \implies \|f_{n+p} - f_n\|_1 \leq \varepsilon.$$

Does $\{f_n\}_{n=1}^{\infty}$ converges uniformly to some continuous function? Determine whether or not the set of continuous function is complete if we replace $\|\cdot\|_\infty$ with $\|\cdot\|_1$.

Hint: Show that

$$\int_0^2 |f_{n+p} - f_n| \leq \int_0^1 f_n = \frac{1}{n+1} \quad \forall n, p \in \mathbb{N}$$

2. Consider the sequence of continuous functions $\{f_n\}_{n=1}^{\infty}$ defined via

$$f_n(x) = \begin{cases} 1 - nx & \text{if } 0 \leq x \leq \frac{1}{n} \\ 0 & \text{if } \frac{1}{n} < x \leq 1 \end{cases}$$

Show that the sequence of Real numbers $\{\|f_n\|_2\}_{n=1}^{\infty}$ converge to 0. What about the sequences $\{\|f_n\|_1\}_{n=1}^{\infty}$ and $\{\|f_n\|_\infty\}_{n=1}^{\infty}$? Do they converges to to 0?

3. Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of Riemann integrable functions on $[a,b]$ and let $f:[a,b]\to\mathbb{R}$ be another Riemann integrable function. Suppose that for some $p\in\mathbb{N}\setminus\{0\}$ we have that
$$\|f_n - f\|_p \to 0 \quad \text{as } n\to+\infty.$$
Prove that for any $k\in\{1,\dots,p\}$ we have that $\|f_n-f\|_k\to 0$ as $n\to+\infty$.

4. Suppose that $f:(a,b)\to\mathbb{R}$ is uniformly continuous on (a,b). Prove that its improper integral converges.

5. Let $f:(a,b)\to\mathbb{R}$ and $g:(a,b)\to\mathbb{R}$ be two given non negative functions, such that
   - for any $h\in(0,\frac{b-a}{2})$, f and g are Riemann integrable on $[a+h,b-h]$.
   - the following limits from above exists and are positive Real number
$$\lim_{h\to 0^+}\frac{f(a+h)}{g(a+h)} \quad \text{and} \quad \lim_{h\to 0^+}\frac{f(b-h)}{g(b-h)}$$

   Prove that the improper integral of f converges if and only if the improper integral of g does it too.

   Hint: Use Theorem 14.3.