Lecture Notes in Calculus

Raz Kupferman
Institute of Mathematics
The Hebrew University

July 10, 2013

Contents

Foreword

1. Texbooks: Spivak, Meizler, Hochman.

Chapter 1

Real numbers

In this course we will cover the calculus of real univariate functions, which was developed during more than two centuries. The pioneers were Isaac Newton (1642-1737) and Gottfried Wilelm Leibniz (1646-1716). Some of their followers who will be mentioned along this course are Jakob Bernoulli (1654-1705), Johann Bernoulli (1667-1748) and Leonhard Euler (1707-1783). These first two generations of mathematicians developed most of the practice of calculus as we know it; they could integrate many functions, solve many differential equations, and sum up a large number of infinite series using a wealth of sophisticated analytical techniques. Yet, they were sometimes very vague about definitions and their theory often laid on shaky grounds. The sound theory of calculus as we know it today, and as we are going to learn it in this course was mostly developed throughout the 19th century, notably by Joseph-Louis Lagrange (1736-1813), Augustin Louis Cauchy (1789-1857), Georg Friedrich Bernhard Riemann (1826-1866), Peter Gustav Lejeune-Dirichlet (1805-1859), Joseph Liouville (1804-1882), Jean-Gaston Darboux (1842-1917), and Karl Weierstrass (1815-1897). As calculus was being established on firmer grounds, the theory of functions needed a thorough revision of the concept of real numbers. In this context, we should mention Georg Cantor (1845-1918) and Richard Dedekind (1831-1916).

1.1 Axioms of field

Even though we have been using "numbers" as an elementary notion since first grade, a rigorous course of calculus should start by putting even such basic concept on axiomatic grounds.

The set of real numbers will be defined as an instance of a ***complete, ordered field*** (שדה סדור שלם).

Definition 1.1 A **field** $\mathbb{F}$ *(שָׂדֶה) is a non-empty set on which two* **binary operations** *are defined*[1]*: an operation which we call* **addition***, and denote by +, and an operation which we call* **multiplication** *and denote by · (or by nothing, as in* $a \cdot b = ab$*). The operations on elements of a field satisfy nine defining properties, which we list now.*

1. ***Addition is associative*** (קיבוצי): for all $a, b, c \in \mathbb{F}$,

$$(a + b) + c = a + (b + c).$$

[2]It should be noted that addition was defined as a *binary* operation. As such, there is no a-priori meaning to $a + b + c$. In fact, the meaning is ambiguous, as we could first add $a + b$ and then add c to the sum, or conversely, add $b + c$, and then add the sum to a. The first axiom states that in either case, the result is the same.

What about the addition of four elements, $a + b + c + d$? Does it require a separate axiom of equivalence? We would like the following additions

$$\begin{gathered}((a + b) + c) + d \\ (a + (b + c)) + d \\ a + ((b + c) + d) \\ a + (b + (c + d)) \\ (a + b) + (c + d)\end{gathered}$$

to be equivalent. It is easy to see that the equivalence follows from the axiom of associativity. Likewise (although it requires some non-trivial work), we

[1]A binary operation in $\mathbb{F}$ is a function that "takes" an ordered pair of elements in $\mathbb{F}$ and "returns" an element in $\mathbb{F}$. That is, to every $a, b \in \mathbb{F}$ corresponds one and only one $a + b \in \mathbb{F}$ and $a \cdot b \in \mathbb{F}$. In formal notation,

$$(\forall a, b \in \mathbb{F})(\exists! c \in \mathbb{F}) : (c = a + b).$$

[2]A few words about the equal sign; we will take it literally to means that "the expressions on both sides are the same". More precisely, equality is an ***equivalence relation***, a concept which will be explained further below.

can show that the n-fold addition

$$a_1 + a_2 + \cdots + a_n$$

is defined unambiguously.

2. ***Existence of an additive-neutral element*** (איבר נייטרלי לחיבור): there exists an element $0 \in \mathbb{F}$ such that for all $a \in \mathbb{F}$

$$a + 0 = 0 + a = a.$$

3. ***Existence of an additive inverse*** (איבר נגדי): for all $a \in \mathbb{F}$ there exists an element $b \in \mathbb{F}$, such that

$$a + b = b + a = 0.$$

The would like to denote the additive inverse of a, as customary, $(-a)$. There is however one problem. The axiom assumes the existence of an additive inverse, but it does not assume its uniqueness. Suppose there were two additive inverses: which one would be denote $(-a)$?

With the first three axioms there are a few things we can show. For example, that

$$a + b = a + c \qquad \text{implies} \qquad b = c.$$

The proof requires all three axioms. Denote by d an additive inverse of a. Then,

$$\begin{aligned} a + b &= a + c \\ d + (a + b) &= d + (a + c) && \text{(existence of inverse)} \\ (d + a) + b &= (d + a) + c && \text{(associativity)} \\ 0 + b &= 0 + c && \text{(property of inverse)} \\ b &= c. && \text{(0 is the neutral element)} \end{aligned}$$

It follows at once that every element $a \in \mathbb{F}$ has a unique additive inverse, for if b and c were both additive inverses of a, then $a + b = 0 = a + c$, from which follows that $b = c$. Hence we can refer to *the* additive inverse of a, which justifies the notation $(-a)$.

We can also show that there is a *unique* element in $\mathbb{F}$ satisfying the neutral property (it is not an a priori fact). Suppose there were $a, b \in \mathbb{F}$ such that

$$a + b = a.$$

Then adding $(-a)$ (on the left!) and using the law of associativity we get that $b = 0$.

Comment: We only postulate the operations of addition and multiplication. ***Subtraction*** is a short-hand notation for the addition of the additive inverse,

$$a - b \stackrel{\text{def}}{=} a + (-b).$$

Comment: A set satisfying the first three axioms is called a ***group*** (חבורה). By themselves, those three axioms have many implications, which can fill an entire course.

4. ***Addition is commutative*** (חילופי): for all $a, b \in \mathbb{F}$

$$a + b = b + a.$$

With this law we finally obtain that any (finite!) summation of elements of $\mathbb{F}$ can be re-arranged in any order.

Example: Our experience with numbers tells us that $a - b = b - a$ implies that $a = b$. There is no way we can prove it using only the first four axioms! Indeed, all we obtain from it that

$$\begin{aligned} a + (-b) &= b + (-a) && \text{(given)} \\ (a + (-b)) + (b + a) &= (b + (-a)) + (a + b) && \text{(commutativity)} \\ a + ((-b) + b) + a &= b + ((-a) + a) + b && \text{(associativity)} \\ a + 0 + a &= b + 0 + b && \text{(property of additive inverse)} \\ a + a &= b + b, && \text{(property of neutral element)} \end{aligned}$$

but how can we deduce from that that $a = b$? ▲▲▲

With very little comments, we can state now the corresponding laws for multiplication:

5. ***Multiplication is associative***: for all $a, b, c \in \mathbb{F}$,

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c.$$

6. ***Existence of a multiplicative-neutral element***: there exists an element $1 \in \mathbb{F}$ such that for all $a \in \mathbb{F}$,

$$a \cdot 1 = 1 \cdot a = a.$$

7. ***Existence of a multiplicative inverse*** (איבר הופכי): for every $a \neq 0$ there exists an element $a^{-1} \in \mathbb{F}$ such that[3]

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

The condition that $a \neq 0$ has strong implications. For example, from the fact that $a \cdot b = a \cdot c$ we can deduce that $b = c$ only if $a \neq 0$.

Comment: ***Division*** is defined as multiplication by the inverse,

$$a/b \stackrel{\text{def}}{=} a \cdot b^{-1}.$$

8. ***Multiplication is commutative***: for every $a, b \in \mathbb{F}$,

$$a \cdot b = b \cdot a.$$

The last axiom relates the addition and multiplication operations:

9. ***Distributive law*** (חוק הקיבוץ): for all $a, b, c \in \mathbb{F}$,

$$a \cdot (b + c) = a \cdot b + a \cdot c.$$

We can revisit now the $a - b = b - a$ example. We proceed,

$$\begin{aligned} a + a &= b + b \\ 1 \cdot a + 1 \cdot a &= 1 \cdot b + 1 \cdot b && \text{(property of neutral element)} \\ (1 + 1) \cdot a &= (1 + 1) \cdot b && \text{(distributive law).} \end{aligned}$$

We would be done if we knew that $1 + 1 \neq 0$, for we would multiply both sides by $(1 + 1)^{-1}$. However, this does not follow from the axioms of field!

[3] Here again, the uniqueness of the multiplicative inverse has to be proved, without which there is no justification to the notation a^{-1}.

Example: Consider the following set $\mathbb{F} = \{e, f\}$ with the following properties,

$+$	e	f
e	e	f
f	f	e

$\cdot$	e	f
e	e	e
f	e	f

It takes some explicit verification to check that this is indeed a field (in fact the smallest possible field), with e being the additive neutral and f being the multiplicative neutral (do you recognize this field?).

Note that

$$e - f = e + (-f) = e + f = f + e = f + (-e) = f - e,$$

and yet $e \neq f$, which is then no wonder that we can't prove, based only on the field axioms, that $f - e = e - f$ implies that $e = f$. ▲▲▲

With 9 axioms at hand, we can start proving theorems that are satisfied by any field.

Proposition 1.1 *For every $a \in \mathbb{F}$, $a \cdot 0 = 0$.*

Proof: Using sequentially the property of the neutral element and the distributive law:

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

Adding to both sides $-(a \cdot 0)$ we obtain

$$a \cdot 0 = 0.$$

■

Comment: Suppose it were the case that $1 = 0$, i.e., that the same element of $\mathbb{F}$ is both the additive neutral and the multiplicative neutral. It would follows that for every $a \in \mathbb{F}$,

$$a = a \cdot 1 = a \cdot 0 = 0,$$

which means that 0 is the only element of $\mathbb{F}$. This is indeed a field according to the axioms, however a very boring one. Thus, we rule it out and require the field to satisfy $0 \neq 1$.

Comment: In principle, every algebraic identity should be proved from the axioms of field. In practice, we will assume henceforth that all known algebraic manipulations are valid. For example, we will not bother to prove that

$$a - b = -(b - a)$$

(even though the proof is very easy). The only exception is (of course) if you get to prove an algebraic identity as an assignment.

✎ *Exercise 1.1* Let $\mathbb{F}$ be a field. Prove, based on the field axioms, that

① If $ab = a$ for some field element $a \neq 0$, then $b = 1$.
② $a^3 - b^3 = (a - b)(a^2 + ab + b^2)$.
③ If $c \neq 0$ then $a/b = (ac)/(bc)$.
④ If $b, d \neq 0$ then

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}.$$

Justify every step in your proof!

1.2 Axioms of order (as taught in 2009)

The real numbers contain much more structure than just being a field. They also form an ***ordered set***. The property of being ordered can be formalized by three axioms:

Definition 1.2 (שָׂדֶה סָדוּר) A field $\mathbb{F}$ is said to be ***ordered*** *if it has a distinguished subset $P \subset \mathbb{F}$, which we call the* ***positive elements****, such that the following three properties are satisfied:*

1. ***Trichotomy***: every element $a \in \mathbb{F}$ satisfies one and only one of the following properties. Either (i) $a \in P$, or (ii) $(-a) \in P$, or (iii) $a = 0$.
2. ***Closure under addition***: if $a, b \in P$ then $a + b \in P$.
3. ***Closure under multiplication***: if $a, b \in P$ then $a \cdot b \in P$.

We supplement these axioms with the following definitions:

$$\begin{array}{lll} a < b & \text{means} & b - a \in P \\ a > b & \text{means} & b < a \\ a \leq b & \text{means} & a < b \text{ or } a = b \\ a \geq b & \text{means} & a > b \text{ or } a = b. \end{array}$$

Comment: $a > 0$ means, by definition, that $a - 0 = a \in P$. Similarly, $a < 0$ means, by definition, that $0 - a = (-a) \in P$.

We can then show a number of well-known properties of real numbers:

Proposition 1.2 *For every $a, b \in \mathbb{F}$ either (i) $a < b$, or (ii) $a > b$, or (iii) $a = b$.*

Proof: This is an immediate consequence of the trichotomy law, along with our definitions. We are asked to show that either (i) $a - b \in P$, or (ii) $b - a = -(a - b) \in P$, or (iii) that $a - b = 0$. That is, we are asked to show that $a - b$ satisfies the trichotomy, which is an axiom. ■

Proposition 1.3 *If $a < b$ then $a + c < b + c$.*

Proof:

$$(b + c) - (a + c) = b - a \in P.$$

■

Proposition 1.4 (Transitivity) *If $a < b$ and $b < c$ then $a < c$.*

Proof:

$$c - a = \underbrace{\underbrace{(c - b)}_{\text{in } P} + \underbrace{(b - a)}_{\text{in } P}}_{\text{in } P \text{ by closure under addition}}.$$

■

Proposition 1.5 *If $a < 0$ and $b < 0$ then $ab > 0$.*

Proof: We have seen above that $a, b < 0$ is synonymous to $(-a), (-b) \in P$. By the closure under multiplication, it follows that $(-a)(-b) > 0$. It remains to check that the axioms of field imply that $ab = (-a)(-b)$. ■

Corollary 1.1 *If $a \neq 0$ then $a \cdot a \equiv a^2 > 0$.*

Proof: If $a \neq 0$, then by the trichotomy either $a > 0$, in which case $a^2 > 0$ follows from the closure under multiplication, or $a < 0$ in which case $a^2 > 0$ follows from the previous proposition, with $a = b$. ■

Corollary 1.2 $1 > 0$.

Proof: This follows from the fact that $1 = 1^2$. ■

Another corollary is that the field of ***complex numbers*** (yes, this is not within the scope of the present course) cannot be ordered, since $i^2 = (-1)$, which implies that (-1) has to be positive, hence 1 has to be negative, which violates the previous corollary.

Proposition 1.6 *If $a < b$ and $c > 0$ then $ac < bc$.*

Proof: It is given that $b - a \in P$ and $c \in P$. Then,

$$bc - ac = \underbrace{\underbrace{(b-a)}_{\text{in } P} \cdot \underbrace{c}_{\text{in } P}}_{\text{in } P \text{ by closure under multiplication}}.$$

■

✎ *Exercise 1.2* Let $\mathbb{F}$ be an ordered field. Prove based on the axioms that

① $a < b$ if and only if $(-b) < (-a)$.
② If $a < b$ and $c > d$, then $a - c < b - d$.
③ If $a > 1$ then $a^2 > a$.
④ If $0 < a < 1$ then $a^2 < a$.
⑤ If $0 \leq a < b$ and $0 \leq c < d$ then $ac < bd$.
⑥ If $0 \leq a < b$ then $a^2 < b^2$.

✎ *Exercise 1.3* Prove using the axioms of ordered fields and using induction on n that

① If $0 \leq x < y$ then $x^n < y^n$.
② If $x < y$ and n is odd then $x^n < y^n$.
③ If $x^n = y^n$ and n is odd then $x = y$.
④ If $x^n = y^n$ and n is even then either $x = y$ of $x = -y$.
⑤ Conclude that

$$x^n = y^n \qquad \text{if and only if} \qquad x = y \text{ or } x = -y.$$

✎ *Exercise 1.4* Let $\mathbb{F}$ be an ordered field, with multiplicative neutral term 1; we also denote $1 + 1 = 2$. Prove that $(a + b)/2$ is well defined for all $a, b \in \mathbb{F}$, and that if $a < b$, then

$$a < \frac{a+b}{2} < b.$$

✎ *Exercise 1.5* Show that there can be no ordered field that has a finite number of elements.

1.3 Axioms of order (as taught in 2010, 2011)

The real numbers contain much more structure than just being a field. They also form an ***ordered set***. The property of being ordered can be formalized by four axioms:

Definition 1.3 (שָׂדֶה סָדוּר) A field $\mathbb{F}$ is said to be **ordered** *if there exists a relation $<$ (a relation is a property that every ordered pair of elements either satisfy or not), such that the following four properties are satisfied:*

O1 ***Trichotomy***: every pair of elements $a, b \in \mathbb{F}$ satisfies one and only one of the following properties. Either (i) $a < b$, or (ii) $b < a$, or (iii) $a = b$.

O2 ***Transitivity***: if $a < b$ and $b < c$ then $a < c$.

O3 ***Invariance under addition***: if $a < b$ then $a + c < b + c$ for all $c \in \mathbb{F}$.

O4 ***Invariance under multiplication by positive numbers***: if $a < b$ then $ac < bc$ for all $0 < c$.

Comment: The set of numbers a for which $0 < a$ are called the ***positive numbers***. The set of numbers a for which $a < 0$ are called the ***negative numbers***.

We supplement these axioms with the following definitions:

$$\begin{aligned} a > b &\quad \text{means} \quad b < a \\ a \le b &\quad \text{means} \quad a < b \text{ or } a = b \\ a \ge b &\quad \text{means} \quad a > b \text{ or } a = b. \end{aligned}$$

(2 hrs, 2013)

Proposition 1.7 *The set of positive numbers is closed under addition and multiplication, namely*

$$\textit{If} \;\; 0 < a, b \quad \textit{then} \quad 0 < a + b \quad \textit{and} \quad 0 < ab.$$

Proof: Closure under addition follows from the fact that $0 < a$ implies that

$$\begin{aligned} b &= 0 + b && \text{(neutral element)} \\ &< a + b, && \text{(O3)} \end{aligned}$$

and $b < a + b$ follows from transitivity.

Closure under multiplication follows from the fact that

$$\begin{aligned} 0 &= 0 \cdot a && \text{(proved)} \\ &< b \cdot a. && \text{(O4)} \end{aligned}$$

■

(2 hrs, 2010) (2 hrs, 2011)

Proposition 1.8 $0 < a$ *if and only if* $(-a) < 0$*, i.e., a number is positive if and only if its additive inverse is negative.*

Proof: Suppose $0 < a$, then

$$\begin{aligned} (-a) &= 0 + (-a) && \text{(neutral element)} \\ &< a + (-a) && \text{(O3)} \\ &= 0 && \text{(additive inverse)} \end{aligned}$$

i.e., $(-a) < 0$. Conversely, if $(-a) < 0$ then

$$\begin{aligned} 0 &= (-a) + a && \text{(additive inverse)} \\ &< 0 + a && \text{(O3)} \\ &= a && \text{(neutral element)} \end{aligned}$$

■

Corollary 1.3 *The set of positive numbers is not empty.*

Proof: Since by assumption $1 \neq 0$, it follows by O1 that either $0 < 1$ or $1 < 0$. In the latter case $0 < (-1)$, hence either 1 or (-1) is positive. ■

Proposition 1.9 $0 < 1$.

Proof: By trichotomy, either $0 < 1$, or $1 < 0$, or $0 = 1$. The third possibility is ruled out by assumption. Suppose then, by contradiction, that $1 < 0$. Then for every $0 < a$ (and we know that at least one such exists),

$$a = a \cdot 1 < a \cdot 0 = 0,$$

i.e., $a < 0$ which violates the trichotomy axiom. Hence $0 < 1$. ■

Proposition 1.10 $0 < a$ *if and only if* $0 < a^{-1}$, *i.e., a number is positive if and only if its multiplicative inverse is also positive.*

Proof: Let $0 < a$ and suppose by contradiction that $a^{-1} < 0$ (it can't be zero because $a^{-1}a = 1$). Then,

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

which is a contradiction. ■

Proposition 1.11 *If* $a < 0$ *and* $b < 0$ *then* $ab > 0$.

Proof: We have seen above that $a, b < 0$ implies that $0 < (-a), (-b)$. By the closure under multiplication, it follows that $(-a)(-b) > 0$. It remains to check that the axioms of field imply that $ab = (-a)(-b)$. ■

Corollary 1.4 *If* $a \neq 0$ *then* $a \cdot a \equiv a^2 > 0$.

Proof: If $a \neq 0$, then by the trichotomy either $a > 0$, in which case $a^2 > 0$ follows from the closure under multiplication, or $a < 0$ in which case $a^2 > 0$ follows from the previous proposition, with $a = b$. ■

Another corollary is that the field of ***complex numbers*** (which is not within the scope of the present course) cannot be ordered, since $i^2 = (-1)$, which implies that (-1) has to be positive, hence 1 has to be negative, which is a violation.

1.4 Absolute values

Definition 1.4 *For every element a in an ordered field* $\mathbb{F}$ *we define the* ***absolute value*** *(ערך מוחלט),*

$$|a| = \begin{cases} a & a \geq 0 \\ (-a) & a < 0. \end{cases}$$

We see right away that $|a| = 0$ if and only if $a = 0$, and otherwise $|a| > 0$.

Proposition 1.12 (Triangle inequality (אי שוויון המשולש)) *For every* $a, b \in \mathbb{F}$,

$$|a + b| \leq |a| + |b|.$$

Proof: We need to examine four cases (by trichotomy):

1. $a \geq 0$ and $b \geq 0$.
2. $a \leq 0$ and $b \geq 0$.
3. $a \geq 0$ and $b \leq 0$.
4. $a \leq 0$ and $b \leq 0$.

In the first case, $|a| = a$ and $|b| = b$, hence

$$|a + b| = a + b = |a| + |b|.$$

(Hey, does this mean that $|a + b| \leq |a| + |b|$?) In the fourth case, $|a| = (-a)$, $|b| = (-b)$, and $0 < -(a + b)$, hence

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

It remains to show the second case, as the third case follows by interchanging the roles of a and b. Let $a \leq 0$ and $b \geq 0$, i.e., $|a| = (-a)$ and $|b| = b$. We need to show that

$$|a + b| \leq (-a) + b.$$

We can divide this case into two sub-cases. Either $a + b \geq 0$, in which case

$$|a + b| = a + b \leq (-a) + b = |a| + |b|.$$

If $a + b < 0$, then

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

which completes the proof[4]. ■

(3 hrs, 2013)

[4]We have used the fact that $a \geq 0$ implies $(-a) \leq a$ and $a \leq 0$ implies $a \leq (-a)$. This is very easy to show. For example,

$$a > 0 \implies a + (-a) > 0 + (-a) \implies 0 > (-a),$$

and $(-a) < a$ follows from transitivity.

Comment: By replacing b by $(-b)$ we obtain that

$$|a-b| \leq |a| + |b|.$$

The following version of the triangle inequality is also useful:

Proposition 1.13 (Reverse triangle inequality) *For every $a, b \in \mathbb{F}$,*

$$||a| - |b|| \leq |a+b|.$$

Proof: By the triangle inequality,

$$\forall a, c \in \mathbb{F}, \qquad |a-c| \leq |a| + |c|.$$

Set $c = b + a$, in which case $|b| \leq |a| + |b+a|$, or

$$|b| - |a| \leq |b+a|.$$

By interchanging the roles of a and b,

$$|a| - |b| \leq |a+b|.$$

It follows that[5]

$$||a| - |b|| \leq |a+b|,$$

■

Proposition 1.14 *For every $a, b, c \in \mathbb{F}$,*

$$|a+b+c| \leq |a| + |b| + |c|.$$

[5]Here we used the fact that any property satisfied by a and $(-a)$ is also satisfied by $|a|$.

Proof: Use the triangle inequality twice,

$$|a + b + c| = |a + (b + c)| \le |a| + |b + c| \le |a| + |b| + |c|.$$

■

(2 hrs, 2009)

✎ *Exercise 1.6* Let $\min(x, y)$ and $\max(x, y)$ denote the smallest and the largest of the two arguments, respectively. Show that

$$\max(x, y) = \frac{1}{2}(x + y + |x - y|) \qquad \text{and} \qquad \min(x, y) = \frac{1}{2}(x + y - |x - y|).$$

1.5 Special sets of numbers

What kind of sets can have the properties of an ordered field? We start by introducing the ***natural numbers*** (המספרים הטבעיים), by taking the element 1 (which exists by the axioms), and naming the numbers

$$1 + 1, \quad 1 + 1 + 1, \quad 1 + 1 + 1 + 1, \quad \text{etc.}$$

That these numbers are all distinct follows from the axioms of order (each number is greater than its predecessor because $0 < 1$). That we give them names, such as $2, 3, 4$, and may represent them using a decimal system is immaterial. We denote the set of natural numbers by $\mathbb{N}$. Clearly the natural numbers do not form a field (for example, there are no additive inverses because all the numbers are positive).

We can then augment the set of natural numbers by adding the numbers

$$0, \quad -1, \quad -(1 + 1), \quad -(1 + 1 + 1), \quad \text{etc.}$$

This forms the set of ***integers*** (השלמים), which we denote by $\mathbb{Z}$. The integers form a commutative group with respect to addition, but are not a field since multiplicative inverses do not exist[6].

[6]How do we know that $(1 + 1)^{-1}$, for example, is not an integer? We know that it is positive and non-zero. Suppose that it were true that $1 < (1 + 1)^{-1}$, then

$$1 + 1 = (1 + 1) \cdot 1 < (1 + 1) \cdot (1 + 1)^{-1} = 1,$$

which is a violation of the axioms of order.

The set of numbers can be further augmented by adding all ***integer quotients***, m/n, $n \neq 0$. This forms the set of ***rational numbers***, which we denote by $\mathbb{Q}$. It should be noted that the rational numbers are the set of integer quotients modulo an equivalence relation[7]. A rational number has infinitely many representations as the quotient of two integers.

Proposition 1.15 *Let* $b, d \neq 0$. *Then*

$$\frac{a}{b} = \frac{c}{d} \qquad \text{if and only if} \qquad ad = bc.$$

Proof: Multiply/divide both sides by bd (which cannot be zero if $b, d \neq 0$). ■

It can be checked that the rational numbers form an ordered field.

(4 hrs, 2010) (4 hrs, 2011)

In this course we will not teach (axiomatically) ***mathematical induction***, nor ***inductive definitions*** (also called ***recursive definitions***); we will assume these concepts to be understood. As an example of an inductive definition, we define for all $a \in \mathbb{Q}$ and $n \in \mathbb{N}$, the n-th power, by setting,

$$a^1 = a \qquad \text{and} \qquad a^{k+1} = a \cdot a^k.$$

As an example of an inductive proof, consider the following proposition:

Proposition 1.16 *Let* $a, b > 0$ *and* $n \in \mathbb{N}$. *Then*

$$a > b \quad \text{if and only if} \quad a^n > b^n.$$

[7] A few words about ***equivalence relations*** (יחס שקילות). An equivalence relation on a set S is a property that any two elements either have or don't. If two elements a, b have this property, we say that a is equivalent to b, and denote it by $a \sim b$. An equivalence relation has to be ***symmetric*** ($a \sim b$ implies $b \sim a$), ***reflexive*** ($a \sim a$ for all a) and ***transitive*** ($a \sim b$ and $b \sim c$ implies $a \sim c$). Thus, with every element $a \in S$ we can associate an ***equivalence class*** (מחלקת שקילות), which is $\{b : b \sim a\}$. Every element belongs to one and only one equivalence class. An equivalence relation partitions S into a collection of disjoint equivalence classes; we call this set S modulo the equivalence relation $\sim$, and denote it $S/\sim$.

Proof: We proceed by induction. This is certainly true for $n = 1$. Suppose this were true for $n = k$. Then

$$a > b \quad \text{implies} \quad a^k > b^k \quad \text{implies} \quad a^{k+1} = a \cdot a^k > a \cdot b^k > b \cdot b^k = b^{k+1}.$$

Conversely, if $a^n > b^n$ then $a \le b$ would imply that $a^n \le b^n$, hence $a > b$. ■

Another inequality that we will need occasionally is:

Proposition 1.17 (Bernoulli inequality) For all $x > (-1)$ and $n \in \mathbb{N}$,

$$(1 + x)^n \ge 1 + nx.$$

Proof: The proof is by induction. For $n = 1$ both sides are equal. Suppose this were true for $n = k$. Then,

$$(1 + x)^{k+1} = (1 + x)(1 + x)^k \ge (1 + x)(1 + kx) = 1 + (k + 1)x + kx^2 \ge 1 + (k + 1)x,$$

where in the left-most inequality we have used explicitly the fact that $1 + x > 0$. ■

✎ *Exercise 1.7* The following exercise deal with logical assertions, and is in preparation to the type of assertions we will be dealing with all along this course. For each of the following statements write its negation in hebrew, without using the word "no". Then write both the statement and its negation, using logical quantifiers like ∀ and ∃.

① For every integer n, $n^2 \ge n$ holds.

② There exists a natural number M such that $n < M$ for all natural numbers n.

③ For every integer n, m, either $n \ge m$ or $-n \ge -m$.

④ For every natural number n there exist natural numbers a, b,, such that $b < n$, $1 < a$ and $n = ab$.

(3 hrs, 2009)

1.6 The Archimedean property

We are aiming at constructing an entity that matches our notions and experiences associated with numbers. Thus far, we defined an ordered field, and showed that it must contain a set, which we called the integers, along with all numbers expressible at ratios of integers—the rational numbers $\mathbb{Q}$. Whether it contains additional elements is left for the moment unanswered. The point is that $\mathbb{Q}$ already has the property of being an ordered field.

Before proceeding, we will need the following definitions:

Definition 1.5 *A set of elements $A \subset \mathbb{F}$ is said to be* ***bounded from above*** *(חסום מלעיל) if*[8] *there exists an element $M \in \mathbb{F}$, such that $a \leq M$ for all $a \in A$, i.e.,*

$$A \text{ is bounded from above} \quad \Longleftrightarrow \quad (\exists M \in \mathbb{F}) : (\forall a \in A)(a \leq M).$$

Such an element M is called an ***upper bound*** *(חסם מלעיל) for A. A is said to be* ***bounded from below*** *(חסום מלרע) if*

$$(\exists m \in \mathbb{F}) : (\forall a \in A)(a \geq m).$$

Such an element m is called a ***lower bound*** *(חסם מלרע) for A. A set is said to be* ***bounded*** *(חסום) if it is bounded both from above and below.*

Note that if a set is upper bounded, then the upper bound is not unique, for if M is an upper bound, so are $M + 1$, $M + 2$, and so on.

Proposition 1.18 *A set A is bounded if and only if*

$$(\exists M \in \mathbb{F}) : (\forall a \in A)(|a| \leq M).$$

Proof: [9] Suppose A is bounded, then

$$(\exists M_1 \in \mathbb{F}) : (\forall a \in A)(a \leq M_1)$$
$$(\exists M_2 \in \mathbb{F}) : (\forall a \in A)(a \geq M_2).$$

[8] In definitions "if" always stands for "if and only if".

[9] Once and for all: P if Q means that "if Q holds then P holds", or Q implies P. P only if Q means that "if Q does not hold, then P does not hold either", which implies that P implies Q. Thus, "if and only if" means that each one implies the other.

That is,

$$(\exists M_1 \in \mathbb{F}) : (\forall a \in A)(a \le M_1)$$
$$(\exists M_2 \in \mathbb{F}) : (\forall a \in A)((-a) \le (-M_2)).$$

By taking $M = \max(M_1, -M_2)$ we obtain

$$(\forall a \in A)(a \le M \wedge (-a) \le M),$$

i.e.,

$$(\forall a \in A)(|a| \le M).$$

The other direction isn't much different. ■

A few notational conventions: for $a < b$ we use the following notations for an open (קטע פתוח), closed (קטע סגור), and semi-open (קטע חצי סגור) segment,

$$(a, b) = \{x \in \mathbb{F} : \; a < x < b\}$$
$$[a, b] = \{x \in \mathbb{F} : \; a \le x \le b\}$$
$$(a, b] = \{x \in \mathbb{F} : \; a < x \le b\}$$
$$[a, b) = \{x \in \mathbb{F} : \; a \le x < b\}.$$

Each of these sets is bounded (above and below).

We also use the following notations for sets that are bounded on one side,

$$(a, \infty) = \{x \in \mathbb{F} : \; x > a\}$$
$$[a, \infty) = \{x \in \mathbb{F} : \; x \ge a\}$$
$$(-\infty, a) = \{x \in \mathbb{F} : \; x < a\}$$
$$(-\infty, a] = \{x \in \mathbb{F} : \; x \le a\}.$$

The first two sets are bounded from below, whereas the last two are bounded from above. It should be emphasized that $\pm\infty$ are not members of $\mathbb{F}$! The above is purely a notation. Thus, "being less than infinity" does not mean being less than a field element called infinity.

(5 hrs, 2011)

Consider now the set of natural numbers, $\mathbb{N}$. Does it have an upper bound? Our "geometrical picture" of the number axis clearly indicates that this set is unbounded, but *can we prove it*? We can easily prove the following:

Proposition 1.19 The natural numbers are not upper bounded by a natural number.

Proof: Suppose $n \in \mathbb{N}$ were an upper bound for $\mathbb{N}$, i.e.,

$$(\forall k \in \mathbb{N}) : (k \le n).$$

However, $m = n + 1 \in \mathbb{N}$ and $m > n$, i.e.,

$$(\exists m \in \mathbb{N}) : (n < m),$$

which is a contradiction. ∎

Similarly,

Corollary 1.5 The natural numbers are not bounded from above by any rational number

Proof: Suppose that $p/q \in \mathbb{Q}$ was an upper bound for $\mathbb{N}$. Since, by the axioms of order, $p/q \le p$, it would imply the existence of a natural number p that is upper bound for $\mathbb{N}$.[10] ∎

(5 hrs, 2010)

It may however be possible that an irrational element in $\mathbb{F}$ be an upper bound for $\mathbb{N}$. Can we rule out this possibility? It turns out that this cannot be proved from the axioms of an ordered field. Since, however, the boundedness of the naturals is so basic to our intuition, we may impose it as an additional axiom, known as ***the axiom of the Archimedean field***: the set of natural numbers is unbounded from above[11].

(5 hrs, 2013)

This axiom has a number of immediate consequences:

[10] Convince yourself that $p/q \le p$ for all $p, q \in \mathbb{N}$.

[11] This additional assumption is temporary. The Archimedean property will be provable once we add our last axiom.

Proposition 1.20 *In an Archimedean ordered field $\mathbb{F}$:*

$$(\forall a \in \mathbb{F})(\exists n \in \mathbb{N}) : (a < n).$$

Proof: If this weren't the case, then $\mathbb{N}$ would be bounded. Indeed, the negation of the proposition is

$$(\exists a \in \mathbb{F}) : (\forall n \in \mathbb{N})(n \leq a),$$

i.e., a is an upper bound for $\mathbb{N}$. ■

Corollary 1.6 *In an Archimedean ordered field $\mathbb{F}$:*

$$(\forall \epsilon > 0)(\exists n \in \mathbb{N}) : (1/n < \epsilon).$$

Proof: By the previous proposition,

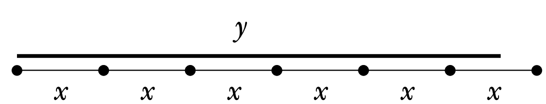
$$(\forall \epsilon > 0)(\exists n \in \mathbb{N}) : \underbrace{(0 < 1/\epsilon < n)}_{(0<1/n<\epsilon)}.$$

■

Corollary 1.7 *In an Archimedean ordered field $\mathbb{F}$:*

$$(\forall x, y > 0)(\exists n \in \mathbb{N}) : (y < nx).$$

Comment: This is really what is meant by the Archimedean property. For every $x, y > 0$, a segment of length y can be covered by a finite number of segments of length x.



Proof: By the previous corollary with $\epsilon = y/x$,

$$(\forall x, y > 0)(\exists n \in \mathbb{N}) : \underbrace{(y/x < n)}_{(y<nx)}.$$

■

1.7 Axiom of completeness

The Greeks knew already that the field of rational numbers is “incomplete”, in the sense that there is no rational number whose square equals 2 (whereas they knew by Pythagoras’ theorem that this should be the length of the diagonal of a unit square). In fact, let’s prove it:

Proposition 1.21 There is no $r \in \mathbb{Q}$ such that $r^2 = 2$.

Proof: Suppose, by contradiction, that $r = n/m$, where $n, m \in \mathbb{N}$ (we can assume that r is positive) satisfies $r^2 = 2$. Although it requires some number theoretical knowledge, we will assume that it is known that any rational number can be brought into a form where m and n have no common divisor. By assumption, $m^2/n^2 = 2$, i.e., $m^2 = 2n^2$. This means that m^2 is even, from which follows that m is even, hence $m = 2k$ for some $k \in \mathbb{N}$. Hence, $4k^2 = 2n^2$, or $n^2 = 2k^2$, from which follows that n is even, contradicting the assumption that m and n have no common divisor. ■

Proof: Another proof: suppose again that $r = n/m$ is irreducible and $r^2 = 2$. Since, $n^2 = 2m^2$, then

$$m^2 < n^2 < 4m^2,$$

it follows that

$$m < n < 2m < 2n,$$

and

$$0 < n - m < m \qquad \text{and} \qquad 0 < 2m - n < n.$$

Consider now the ratio

$$q = \frac{2m - n}{n - m},$$

whose numerator and denominator are both smaller than the respective numerator and denominator of r. By elementary arithmetic

$$q^2 = \frac{4m^2 - 4mn + n^2}{n^2 - 2nm + m^2} = \frac{4 - 4n/m + n^2/m^2}{n^2/m^2 - 2n/m + 1} = \frac{6 - 4n/m}{3 - 2n/m} = 2,$$

which is a contradiction. ■

✎ *Exercise 1.8* Prove that there is no $r \in \mathbb{Q}$ such that $r^2 = 3$.

✎ *Exercise 1.9* What fails if we try to apply the same arguments to prove that there is no $r \in \mathbb{Q}$ such that $r^2 = 4$ (an assertion that happens to be false)?

A way to cope with this "missing number" would be to add $\sqrt{2}$ "by hand" to the set of rational numbers, along with all the numbers obtained by field operations involving $\sqrt{2}$ and rational numbers (this is called in algebra a ***field extension***) (שדה הרחבה)[12]. But then, what about $\sqrt{3}$? We could add all the square roots of all positive rational numbers. And then, what about $\sqrt[3]{2}$? Shall we add all n-th roots? But then, what about a solution to the equation $x^5 + x + 1 = 0$ (it cannot be expressed in terms of roots as a result of ***Galois' theory***)?

It turns out that a single additional axiom, known as the ***axiom of completeness*** (אכסיומת השלמות), completes the set of rational numbers in one fell swoop, such to provide solutions to all those questions. Let us try to look in more detail in what sense is the field of rational numbers "incomplete". Consider the following two sets,

$$A = \{x \in \mathbb{Q} : 0 < x, x^2 \le 2\}$$
$$B = \{y \in \mathbb{Q} : 0 < y, 2 \le y^2\}.$$

Every element in B is greater or equal than every element in A (by transitivity and by the fact that $0 < x \le y$ implies $x^2 \le y^2$). Formally,

$$(\forall a \in A \wedge \forall b \in B)(a \le b).$$

Does there exist a rational number that "separates" the two sets, i.e., does there exist an element $c \in \mathbb{Q}$ such that

$$(\forall a \in A \wedge \forall b \in B)(a \le c \le b) \ \ ?$$

[12] It can be shown that this field extension of $\mathbb{Q}$ consists of all elements of the form

$$\{a + b\sqrt{2} : a, b \in \mathbb{Q}\}.$$

It can be shown that if there existed such a c it would satisfy $c^2 = 2$, hence such a c does not exist. This observation motivates the following definition:

Definition 1.6 *An ordered field $\mathbb{F}$ is said to be* ***complete*** *if for every two non-empty sets, $A, B \subseteq \mathbb{F}$ satisfying*

$$(\forall a \in A \wedge \forall b \in B)(a \le b).$$

there exists an element $c \in \mathbb{F}$ such that

$$(\forall a \in A \wedge \forall b \in B)(a \le c \le b).$$

We will soon see how this axiom "completes" the field $\mathbb{Q}$.

We next introduce more definitions. Let's start with a motivating example:

Example: For any set $[a, b) = \{x : a \le x < b\}$, b is an upper bound, but so is any larger element, e.g., $b + 1$. In fact, it is clear that b is the least upper bound. ▲▲▲

Definition 1.7 *Let $A \subseteq \mathbb{F}$ be a set. An element $M \in \mathbb{F}$ is called a* ***least upper bound*** *(חסם עליון) for A if (i) it is an upper bound for A, and (ii) if M' is also an upper bound for A then $M \le M'$. That is,*

$$M \text{ is a least upper bound for } A$$
$$\Updownarrow$$
$$(\forall a \in A)(a \le M)$$
$$(\forall M' \in \mathbb{F})(\text{ if } (\forall a \in A)(a \le M') \text{ then } (M \le M')).$$

We can see right away that a least upper bound, if it exists, is unique:

Proposition 1.22 *Let $A \subset \mathbb{F}$ be a set and let M be a least upper bound for A. If M' is also a least upper bound for A then $M = M'$.*

Proof: It follows immediately from the definition of the least upper bound. If M and M' are both least upper bounds, then both are in particular upper bounds, hence $M \le M'$ and $M' \le M$, which implies that $M = M'$. ■

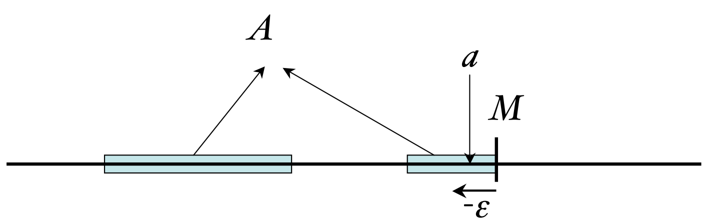
We call the least upper bound of a set (if it exists!) a ***supremum***, and denote it by $\sup A$. Similarly, the ***greatest lower bound*** (חסם תחתון) of a set is called the ***infimum***, and it is denoted by $\inf A$.[13]

(7 hrs, 2011)

We can provide an equivalent definition of the least upper bound:

Proposition 1.23 Let A be a set. A number M is a least upper bound if and only if (i) it is an upper bound, and (ii)

$$(\forall \epsilon > 0)(\exists a \in A) : (a > M - \epsilon).$$



Proof: There are two directions to be proved.

1. Suppose first that M were a least upper bound for A, i.e.,

   $$(\forall a \in A)(a \le M) \quad \text{and} \quad (\forall M' \in \mathbb{F})(\text{if } (\forall a \in A)(a \le M') \text{ then } (M \le M')).$$

   Suppose, by contradiction, that there exists an $\epsilon > 0$ such that $a \le M - \epsilon$ for all $a \in A$, i.e., that

   $$(\exists \epsilon > 0) : (\forall a \in A)(a < M - \epsilon),$$

   then $M - \epsilon$ is an upper bound for A, smaller than the least upper bound, which is a contradiction.

[13]In some books the notations ***lub*** (least upper bound) and ***glb*** (greatest lower bound) are used instead of sup and inf.

2. Conversely, suppose that M is an upper bound and

$$(\forall \epsilon > 0)(\exists a \in A) : (a > M - \epsilon).$$

Suppose that M was not a least upper bound. Then there exists a smaller upper bound $M' < M$. Take $\epsilon = M - M'$. Then, $a \leq M' = M - \epsilon$ for all $a \in A$, or in formal notation,

$$(\exists \epsilon > 0) : (\forall a \in A)(a < M - \epsilon),$$

which is a contradiction.

■

Examples: What are the least upper bounds (if they exist) in the following examples:

① $[-5, 15]$ (answer: 15).
② $[-5, 15)$ (answer: 15).
③ $[-5, 15] \cup \{20\}$ (answer: 20).
④ $[-5, 15] \cup (17, 18)$ (answer: 18).
⑤ $[-5, \infty)$ (answer: none).
⑥ $\{1 - 1/n : n \in \mathbb{N}\}$ (answer: 1).

✎ *Exercise 1.10* Let $\mathbb{F}$ be an ordered field, and $\emptyset \neq A \subseteq \mathbb{F}$. Prove that $M = \inf A$ if and only if M is a lower bound for A, and

$$(\forall \epsilon > 0)(\exists a \in A) : (a < M + \epsilon).$$

✎ *Exercise 1.11* Show that the axiom of completeness is equivalent to the following statement: for every two non-empty sets $A, B \subset \mathbb{F}$, satisfying that $a \in A$ and $b \in B$ implies $a \leq b$, there exists an element $c \in \mathbb{F}$, such that

$$a \leq c \leq b \qquad \text{for all} \qquad a \in A \quad \text{and} \quad b \in B.$$

✎ *Exercise 1.12* For each of the following sets, determine whether it is upper bounded, lower bounded, and if they are find their infimum and/or supremum:

① $\{(-2)^n : n \in \mathbb{N}\}$.

② $\{1/n : n \in \mathbb{Z}, n \neq 0\}$.
③ $\{1 + (-1)^n : n \in \mathbb{N}\}$.
④ $\{1/n + (-1)^n : n \in \mathbb{N}\}$.
⑤ $\{n/(2n+1) : n \in \mathbb{N}\}$.
⑥ $\{x : |x^2 - 1| < 3\}$.

✎ *Exercise 1.13* Let A, B be non-empty sets of real numbers. We define

$$A + B = \{a + b : a \in A, b \in B\}$$
$$A - B = \{a - b : a \in A, b \in B\}$$
$$A \cdot B = \{a \cdot b : a \in A, b \in B\}.$$

Find $A + B$, $A \cdot B$ and $A \cdot A$ in each of the following cases:

① $A = \{1, 2, 3\}$ and $B = \{-1, -2, -1/2, 1/2\}$.
② $A = B = [0, 1]$.

✎ *Exercise 1.14* Let A be a non-empty set of real numbers. Is each of the following statements necessarity true? Prove it or give a counter example:

① $A + A = \{2\} \cdot A$.
② $A - A = \{0\}$.
③ $\forall x \in A \cdot A,\ x \geq 0$.

✎ *Exercise 1.15* In each of the following items, A, B are non-empty subsets of a complete ordered field. For the moment, you can only use the axiom of completeness for the least upper bound.

① Suppose that A is lower bounded and set $B = \{x : -x \in A\}$. Prove that B is upper bounded, that A has an infimum, and $\sup B = -\inf A$.
② Show that if A is upper bounded and B is lower bounded, then $\sup(A - B) = \sup A - \inf B$.
③ Show that if A, B are lower bounded then $\inf(A + B) = \inf A + \inf B$.
④ Suppose that A, B are upper bounded. It is necessarily true that $\sup(A \cdot B) = \sup A \cdot \sup B$?
⑤ Show that if A, B only contain non-negative terms and are upper bounded, then $\sup(A \cdot B) = \sup A \cdot \sup B$.

Definition 1.8 *Let $A \subset \mathbb{F}$ be a subset of an ordered field. It is said to have a* **maximum** *if there exists an element $M \in A$ which is an upper bound for A. We denote*

$$M = \max A.$$

It is said to have a **minimum** *if there exists an element $m \in A$ which is a lower bound for A. We denote*

$$m = \min A.$$

(7 hrs, 2010) (7 hrs, 2013)

Proposition 1.24 *If a set A has a maximum, then the maximum is the least upper bound, i.e.,*

$$\sup A = \max A.$$

Similarly, if A has a minimum, then the minimum is the greatest lower bound,

$$\inf A = \min A.$$

Proof: Let $M = \max A$.[14] Then M is an upper bound for A, and for every $\epsilon > 0$,

$$A \ni M > M - \epsilon,$$

that is $(\forall \epsilon > 0)(\exists a \in A)(a > M - \epsilon)$, which proves that M is the least upper bound.
■

Examples: What are the maxima (if they exist) in the following examples:

① $[-5, 15]$ (answer: 15).
② $[-5, 15)$ (answer: none).
③ $[-5, 15] \cup \{20\}$ (answer: 20).
④ $[-5, 15] \cup (17, 18)$ (answer: none).
⑤ $[-5, \infty)$ (answer: none).
⑥ $\{1 - 1/n : n \in \mathbb{N}\}$ (answer: none).

[14] We emphasize that not every (non-empty) set has a maximum. The statement is that if a maximum exists, then it is also the supremum.

(5 hrs, 2009)

Proposition 1.25 *Every finite set in an ordered field has a minimum and a maximum.*

Proof: The proof is by induction on the size of the set. ■

Corollary 1.8 *Every finite set in an ordered field has a least upper bound and a greatest lower bound.*

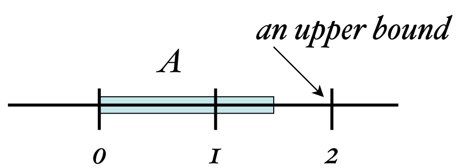
Consider now the field $\mathbb{Q}$ of rational numbers, and its subset

$$A = \{x \in \mathbb{Q} : \ 0 < x, x^2 < 2\}.$$

This set is bounded from above, as 2, for example, is an upper bound. Indeed, if $x \in A$, then

$$x^2 < 2 < 2^2,$$

which implies that $x < 2$.



But does A have a least upper bound?

Proposition 1.26 *Suppose that A has a least upper bound, then*

$$(\sup A)^2 = 2.$$

Proof: We assume that A has a least upper bound, and denote

$$\alpha = \sup A.$$

Clearly, $1 < \alpha < 2$.

Suppose, by contradiction, that $\alpha^2 < 2$. Relying on the Archimedean property, we choose n to be an integer large enough, such that

$$n > \max\left(\frac{1}{\alpha}, \frac{6}{2-\alpha^2}\right).$$

We are going to show that there exists a rational number $r > \alpha$, such that $r^2 < 2$, which means that α is not even an upper bound for A—contradiction:

$$\begin{aligned}(\alpha + 1/n)^2 &= (\alpha)^2 + \frac{2}{n}\alpha + \frac{1}{n^2} \\ &< \alpha^2 + \frac{3}{n}\alpha && \text{(use } 1/n < \alpha) \\ &< \alpha^2 + \frac{6}{n} && \text{(use } \alpha < 2) \\ &< \alpha^2 + 2 - \alpha^2 = 2, && \text{(use } 6/n < 2 - \alpha^2)\end{aligned}$$

which means that $\alpha + 1/n \in A$.

Similarly, suppose that $\alpha^2 > 2$. This time we set

$$n > \frac{2\alpha}{\alpha^2 - 2}.$$

We are going to show that there exists a rational number $r < \alpha$, such that $r^2 > 2$, which means that r is an upper bound for A smaller than α—again, a contradiction:

$$\begin{aligned}(\alpha - 1/n)^2 &= \alpha^2 - \frac{2}{n}\alpha + \frac{1}{n^2} \\ &> \alpha^2 - \frac{2}{n}\alpha && \text{(omit } 1/n^2) \\ &> \alpha^2 + 2 - \alpha^2 = 2, && \text{(use } -2\alpha/n > 2 - \alpha^2)\end{aligned}$$

which means that, indeed, $\alpha - 1/n$ is an upper bound for A. [15] ■

[15] You may ask yourself: how could I have ever guessed such a proof? Take for example the assumption that $\alpha^2 > 2$. We want to show that there exists an upper bound for A that is smaller

(8 hrs, 2013)

Since we saw that there is no $r \in \mathbb{Q}$ such that $r^2 = 2$, it follows that there is no supremum to this set in $\mathbb{Q}$. We will now see that completeness guarantees the existence of a lowest upper bound:

Proposition 1.27 In a complete ordered field every non-empty set that is upper-bounded has a least upper bound.

Proof: Let $A \subset \mathbb{F}$ be a non-empty set that is upper-bounded. Define

$$B = \{b \in \mathbb{F} : b \text{ is an upper bound for } A\},$$

which by assumption is non-empty. Clearly,

$$(\forall a \in A \wedge \forall b \in B)(a \leq b).$$

By the axiom of completeness,

$$(\exists c \in \mathbb{F}) : (\forall a \in A \wedge \forall b \in B)(a \leq c \leq b).$$

Clearly c is an upper bound for A smaller or equal than every other upper bound, hence

$$c = \sup A.$$

■

(8 hrs, 2010)

The following proposition is an immediate consequence:

---

than α, contradicting the fact that it is the least upper bound. Let's try to find an n sufficiently large such that $\alpha - 1/n$ is an upper bound for A, namely, such that $(\alpha - 1/n)^2 > 2$. We expand:

$$\left(\alpha - \frac{1}{n}\right)^2 = \alpha^2 - \frac{2\alpha}{n} + \frac{1}{n^2} > 2 + \underbrace{\left(\alpha^2 - 2\right) - \frac{2\alpha}{n}}_{\text{needs to be positive}}.$$

Similarly, when we take the assumption $\alpha^2 < 2$ we want to show that there exists an element of A that is greater than α. Let's try to find an n sufficiently large such that $\alpha + 1/n$ is in A, namely, such that $(\alpha + 1/n)^2 < 2$. We expand:

$$\left(\alpha + \frac{1}{n}\right)^2 = \alpha^2 + \frac{2\alpha}{n} + \frac{1}{n^2} < 2 + \underbrace{\left(\alpha^2 - 2\right) + \frac{2\alpha}{n} + \frac{1}{n^2}}_{\text{needs to be negative}}.$$

Proposition 1.28 *In a complete ordered field every non-empty set that is bounded from below has a greatest lower bound (an infimum).*

Proof: Let m be a lower bound for A, i.e.,

$$(\forall a \in A)(m \leq a).$$

Consider the set

$$B = \{(-a) : \ a \in A\}.$$

Since

$$(\forall a \in A)((-m) \geq (-a)),$$

it follows that $(-m)$ is a upper bound for B. By the axiom of completeness, B has a least upper bound, which we denote by M. Thus,

$$(\forall a \in A)(M \geq (-a)) \qquad \text{and} \qquad (\forall \epsilon > 0)(\exists a \in A) : ((-a) > M - \epsilon).$$

This means that

$$(\forall a \in A)((-M) \leq a) \qquad \text{and} \qquad (\forall \epsilon > 0)(\exists a \in A) : (a < (-M) + \epsilon).$$

This means that $(-M) = \inf A$. ■

With that, we define the set of real numbers as a complete ordered field, which we denote by $\mathbb{R}$. It turns out that this defines the set uniquely, up to a relabeling of its elements (i.e., up to an isomorphism). [16]

The axiom of completeness can be used to prove the Archimedean property. In other words, the axiom of completeness solves at the same time the "problems" pointed out in the previous section.

Proposition 1.29 (Archimedean property) *In a complete ordered field* $\mathbb{N}$ *is not bounded from above.*

[16] Strictly speaking, we should prove that such an animal exists. Constructing the set of real numbers from the set of rational numbers is beyond the scope of this course. You are strongly recommended, however, to read about it. See for example the construction of ***Dedekind cuts***.

Proof: Suppose $\mathbb{N}$ was bounded. Then it would have a least upper bound M, which is, in particular, and upper bound:

$$(\forall n \in \mathbb{N})\,(n \leq M).$$

Since $n \in \mathbb{N}$ implies $n + 1 \in \mathbb{N}$,

$$(\forall n \in \mathbb{N})\,((n+1) \leq M),$$

i.e.,

$$(\forall n \in \mathbb{N})\,(n \leq (M-1)).$$

This means that $M - 1$ is also an upper bound for $\mathbb{N}$, contradicting the fact that M is the least upper bound. ■

(9 hrs, 2011)

We next prove an important facts about the ***density*** (צפיפות) of rational numbers within the reals.

Proposition 1.30 (The rational numbers are dense in the reals) Let $x, y \in \mathbb{R}$, such that $x < y$. There exists a rational number $q \in \mathbb{Q}$ such that $x < q < y$. In formal notation,

$$(\forall x, y \in \mathbb{R} : \; x < y)(\exists q \in \mathbb{Q}) : (x < q < y).$$

Proof: Since the natural numbers are not bounded, there exists an $m \in \mathbb{Z} \subset \mathbb{Q}$ such that $m < x$. Also there exists a natural number $n \in \mathbb{N}$ such that $1/n < y - x$. Consider now the sequence of rational numbers,

$$r_k = m + \frac{k}{n}, \qquad k = 0, 1, 2, \dots.$$

From the Archimedean property follows that there exists a k such that

$$r_k = m + \frac{k}{n} > x.$$

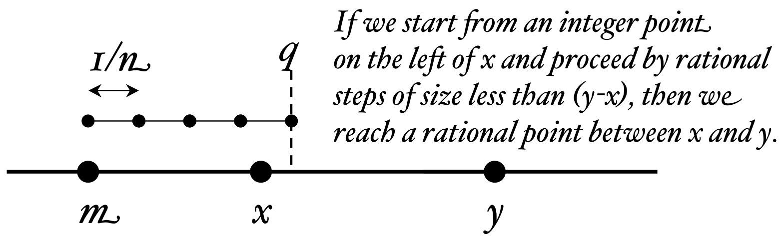
Let k^* be the smallest such number (the exists a smallest k because every finite set has a minimum), i.e.,

$$r_{k^*} > x \qquad \text{and} \qquad r_{k^*-1} \leq x.$$

Hence,

$$x < r_{k^*} = r_{k^*-1} + \frac{1}{n} \le x + \frac{1}{n} < y,$$

which completes the proof.



■

The following proposition provides a useful fact about complete ordered fields. We will use it later in the course.

Proposition 1.31 ((למת החתכים)) Let $A, B \subset \mathbb{F}$ be two non-empty sets in a complete ordered field, such that

$$(\forall a \in A \wedge \forall b \in B)(a \le b).$$

Then the following three statements are equivalent:

① $(\exists! M \in \mathbb{F}) : (\forall a \in A \wedge \forall b \in B)(a \le M \le b)$.
② $\sup A = \inf B$.
③ $(\forall \epsilon > 0)(\exists a \in A, b \in B) : (b - a < \epsilon)$.

Comment: The claim is not that the three items follow from the given data. The claim is that each statement implies the two other.

Proof: Suppose that the first statement holds. This means that M is both an upper bound for A and a lower bound for B. Thus,

$$\sup A \le M \le \inf B.$$

This implies at once that $\sup A \leq \inf B$. If $\sup A < \inf B$, then any number m in between would satisfy $a < m < b$ for all $a \in A, b \in B$, contradicting the assumption, hence ① implies ②.

By the properties of the infimum and the supremum,

$$(\forall \epsilon > 0)(\exists a \in A) : (a > \sup A - \epsilon/2)$$
$$(\forall \epsilon > 0)(\exists b \in B) : (b < \inf B + \epsilon/2).$$

That is,

$$(\forall \epsilon > 0)(\exists a \in A) : (a + \epsilon/2 > \sup A)$$
$$(\forall \epsilon > 0)(\exists b \in B) : (b - \epsilon/2 < \inf B),$$

i.e., $(\forall \epsilon > 0)(\exists a \in A, b \in B)$ such that

$$b - a < (\inf B + \epsilon/2) - (\sup A - \epsilon/2) = \epsilon.$$

Thus, ② implies ③.

It remains to show that ③ implies ①. Suppose M was not unique, i.e., there were $M_1 < M_2$ such that

$$(\forall a \in A \wedge \forall b \in B)(a \leq M_1 < M_2 \leq b).$$

Set $m = (M_1 + M_2)/2$ and $\epsilon = (M_2 - M_1)$. Then,

$$(\forall a \in A, b \in B) \underbrace{(a \leq m - \epsilon/2, b \geq m + \epsilon/2)}_{b-a\geq(m+\epsilon/2)-(m-\epsilon/2)=\epsilon},$$

i.e.,

$$(\exists \epsilon > 0) : (\forall a \in A \wedge \forall b \in B)(b - a > \epsilon),$$

i.e., if ① does not hold then ③ does not hold. This concludes the proof. ■

(7 hrs, 2009) (10 hrs, 2013)

✎ *Exercise 1.16* In this exercise you may assume that $\sqrt{2}$ is irrational.

① Show that if $a + b\sqrt{2} = 0$, where $a, b \in \mathbb{Q}$, then $a = b = 0$.

② Conclude that $a + b\sqrt{2} = c + d\sqrt{2}$, where $a, b, c, d \in \mathbb{Q}$, then $a = c$ and $b = d$.

③ Show that if $a + b\sqrt{2} \in \mathbb{Q}$, where $a, b \in \mathbb{Q}$, then $b = 0$.

④ Show that the irrational numbers are dense, namely that there exists an irrational number between every two.

⑤ Show that the set

$$\{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$$

with the standard addition and multiplication operations is a field.

✎ *Exercise 1.17* Show that the set

$$\{m/2^n : m \in \mathbb{Z}, n \in \mathbb{N}\}$$

is dense.

✎ *Exercise 1.18* Let $A, B \subseteq \mathbb{R}$ be non-empty sets. Prove or disprove each of the following statements:

① If $A \subseteq B$ and B is bounded from above then A is bounded from above.

② If $A \subseteq B$ and B is bounded from below then A is bounded from below.

③ If every $b \in B$ is an upper bound for A then every $a \in A$ is a lower bound form B.

④ A is bounded from above if and only if $A \cap \mathbb{Z}$ is bounded from above.

✎ *Exercise 1.19* Let $A \subseteq \mathbb{R}$ be a non-empty set. Determine for each of the following statements whether it is *equivalent*, *contradictory*, *implied by*, or *implying* the statement that $s = \sup A$, or *none of the above*:

① $a \leq s$ for all $a \in A$.

② For every $a \in A$ there exists a $r \in \mathbb{R}$, such that $a < t < s$.

③ For every $x \in \mathbb{R}$ such that $x > s$ there exists an $a \in A$ such that $s < a < x$.

④ For every finite subset $B \subset A$, $s \geq \max B$.

⑤ s is the supremum of $A \cup \{s\}$.

1.8 Rational powers

We defined the ***integer powers*** recursively,

$$x^1 = x \qquad x^{k+1} = x \cdot x^k.$$

We also define for $x \neq 0$, $x^0 = 1$ and $x^{-n} = 1/x^n$.

Proposition 1.32 (Properties of integer powers) Let $x, y \neq 0$ and $m, n \in \mathbb{Z}$, then

① $x^m x^n = x^{m+n}$.
② $(x^m)^n = x^{mn}$.
③ $(xy)^n = x^n y^n$.
④ $0 < \alpha < \beta$ *and* $n > 0$ *implies* $\alpha^n < \beta^n$.
⑤ $0 < \alpha < \beta$ *and* $n < 0$ *implies* $\alpha^n > \beta^n$.
⑥ $\alpha > 1$ *and* $n > m$ *implies* $\alpha^n > \alpha^m$.
⑦ $0 < \alpha < 1$ *and* $n > m$ *implies* $\alpha^n < \alpha^m$.

Proof: Do it. ■

(10 hrs, 2011)

*Definition 1.9 Let $x > 0$. An n-**th root** of x is a positive number y, such that $y^n = x$.*

It is easy to see that if x has an n-th root then it is unique (since $y^n = z^n$, would imply that $y = z$). Thus, we denote it by either $\sqrt[n]{x}$, or by $x^{1/n}$.

Theorem 1.1 (Existence and uniqueness of roots)

$$(\forall x > 0 \wedge \forall n \in \mathbb{N})(\exists! y > 0) : (y^n = x).$$

(10 hrs, 2010)

Proof: Consider the set

$$S = \{z \geq 0 : \ z^n < x\}.$$

This is a non-empty set (it contains zero) and bounded from above, since

$$\begin{aligned} &\text{if } x \leq 1 \text{ and } z \in S \text{ then } z^n < x \leq 1, \text{ which implies that } z \leq 1 \\ &\text{if } x > 1 \text{ and } z \in S \text{ then } z^n < x < x^n, \text{ which implies that } z < x, \end{aligned}$$

It follows that regardless of the value of x, $z \in S$ implies that $z \leq \max(1, x)$, and the latter is hence an upper bound for S. As a consequence of the axiom of completeness, there exists a unique $y = \sup S$, which is the "natural suspect" for being an n-th root of x. Indeed, we will show that $y^n = x$.

Claim: y is positive Indeed,

$$0 < \frac{x}{1+x} < 1,$$

hence

$$\left(\frac{x}{1+x}\right)^n < \frac{x}{1+x} < x.$$

Thus, $x/(1+x) \in S$, which implies that

$$0 < \frac{x}{1+x} \leq y.$$

Claim: $y^n \geq x$ Suppose, by contradiction that $y^n < x$. We will show by contradiction that y is not an upper bound for S by showing that there exists an element of S that is greater than y. We will do it by showing the existence of an $\epsilon > 0$ such that

$$\frac{y}{1-\epsilon} \in S \qquad \text{i.e.} \qquad \left(\frac{y}{1-\epsilon}\right)^n < x.$$

This means that we look for an $\epsilon > 0$ satisfying

$$\frac{y^n}{x} < (1-\epsilon)^n.$$

Since we can choose $\epsilon > 0$ at will, we can take it smaller than 1, in which case

$$(1-\epsilon)^n \geq (1-\epsilon n)$$

If we choose ϵ sufficiently small such that

$$1 - \epsilon n > \frac{y^n}{x},$$

then a forteriori $(1-\epsilon)^n > y^n/x$, and this will be satisfied if we choose

$$\epsilon < \frac{1 - y^n/x}{n},$$

which is possible because the right hand side is positive. Thus, we found a number greater than y, whose n-th power is less than x, i.e., in S. This contradicts the assumption that y is an upper bound for S.

Claim: $y^n \leq x$ Suppose, by contradiction that $y^n > x$. This time we will show that there exists a positive number less than y whose n-th power is greater then x, i.e., it is an upper bound for S, contradicting the assumption that y is the least upper bound for S.

To find such a number we will show that there exists an $\epsilon > 0$, such that

$$[(1-\epsilon)y]^n > x,$$

i.e.,

$$(1-\epsilon)^n > \frac{x}{y^n}.$$

Once again, we can assume that $\epsilon < 1$, in which case by the Bernoulli inequality $(1-\epsilon)^n \geq (1-\epsilon n)$, so that if

$$(1-\epsilon n) > \frac{x}{y^n},$$

then a forteriori $(1-\epsilon)^n > x/y^n$. Thus, we need $0 < \epsilon < 1$ to satisfiy

$$\epsilon < \frac{1 - x/y^n}{n},$$

which is possible because the right hand side is positive.

From the two inequalities follows (by trichotomy) that $y^n = x$. ■

Having shown that every positive number has a unique n-th root, we may proceed to define rational powers.

Definition 1.10 *Let $r = m/n \in \mathbb{Q}$, then for all $x > 0$*

$$x^r = (x^m)^{1/n}.$$

There is one little delicacy with this definition. Recall that rational numbers do not have a unique representation as the ratio of two integers. We thus need to show that the above definition is independent of the representation[17]. In other words, if $ad = bc$, with $a, b, c, d \in \mathbb{Z}$, $b, d \neq 0$, then for all $x > 0$

$$x^{a/b} = x^{c/d}.$$

[17]This is not obvious. Suppose we wanted to define the notion of even/odd for rational numbers, by saying that a rational number is even if its numerator is even. Such as definition would imply that $2/6$ is even, but $1/3$ (which is the same number!) is odd.

This is a non-trivial fact. We need to prove that $ad = bc$ implies that

$$\sup\{z > 0 : \ z^b < x^a\} = \sup\{z > 0 : \ z^d < x^c\}.$$

The arguments goes as follows:

$$(x^{a/b})^{bd} = (((x^a)^{1/b})^b)^d = (x^a)^d = x^{ad} = x^{bc} = (x^c)^b = (((x^c)^{1/d})^d)^b = (x^{c/d})^{bd},$$

hence $x^{a/b} = x^{c/d}$.

(8 hrs, 2009)

Proposition 1.33 (Properties of rational powers) *Let* $x, y > 0$ *and* $r, s \in \mathbb{Q}$, *then*

① $x^r x^s = x^{r+s}$.
② $(x^r)^s = x^{rs}$.
③ $(xy)^r = x^r y^r$.
④ $0 < \alpha < \beta$ *and* $r > 0$ *implies* $\alpha^r < \beta^r$.
⑤ $0 < \alpha < \beta$ *and* $r < 0$ *implies* $\alpha^r > \beta^r$.
⑥ $\alpha > 1$ *and* $r > s$ *implies* $\alpha^r > \alpha^s$.
⑦ $0 < \alpha < 1$ *and* $r > s$ *implies* $\alpha^r < \alpha^s$.

Proof: We are going to prove only two items. Start with the first. Let $r = a/b$ and $s = c/d$. Using the (proved!) laws for integer powers,

$$(x^{a/b}x^{c/d})^{bd} = (x^{a/b})^{bd}(x^{c/d})^{bd} = \cdots = x^{ad}x^{bc} = x^{ad+bc} = (x^{a/b+c/d})^{bd},$$

from which follows that

$$x^{a/b}x^{c/d} = x^{a/b+c/d}.$$

Take then the fourth item. Let $r = a/b$. We already know that $\alpha^a < \beta^a$. Thus it remains to show $\alpha^{1/b} < \beta^{1/b}$. This has to be because $\alpha^{1/b} \geq \beta^{1/b}$ would have implied (from the rules for integer powers!) that $\alpha \geq \beta$. ■

(12 hrs, 2013)

1.9 Real-valued powers

Delayed to much later in the course.

1.10 Addendum

Existence of non-Archimedean ordered fields Does there exist ordered fields that do not satisfy the Archimedean property? The answer is positive. We will give one example, bearing in mind that some of the arguments rely on later material. Consider the set of ***rational functions***. It is easy to see that this set forms a field with respect to function addition and multiplication, with $f(x) \equiv 0$ for additive neutral element and $f(x) \equiv 1$ for multiplicative neutral element. Ignore the fact that a rational function may be undefined at a finite collection of points. The "natural elements" are the constant functions, $f(x) \equiv 1$, $f(x) \equiv 2$, etc. We endow this set with an order relation by defining the set P of positive elements as the rational functions f that are eventually positive for x sufficiently large. We hace $f > g$ if $f(x) > g(x)$ for x sufficiently large. It is easy to see that the set of "natural elements" are bounded, say, by the element $f(x) = x$ (which obviously belongs to the set).

Connected set A set $A \subseteq \mathbb{R}$ will be said to be ***connected*** (קשיר) if $x, y \in A$ and $x < z < y$ implies $z \in A$ (that is, every point between two elements in the set is also n the set). We state that a (non-empty) connected set can only be of one of the following forms:

$$\{a\}, \quad (a,b), \quad [a,b), \quad (a,b], \quad [a,b]$$
$$(a,\infty), \quad [a,\infty), \quad (-\infty,b), \quad (-\infty,b], \quad (-\infty,\infty).$$

That is, it can only be a single point, an open, closed or semi-open segments, an open or closed ray, or the whole line.

✎ *Exercise 1.20* Prove it.

Chapter 2

Functions

2.1 Basic definitions

What is a function? There is a standard way of defining functions, but we will deliberately be a little less formal than perhaps we should, and define a function as a "machine", which when provided with a number, returns a number (only one, and always the same for the same input). In other words, *a function is a rule that assigns real numbers to real numbers*.

A function is defined by three elements:

① A ***domain*** (תחום): a subset of $\mathbb{R}$. The numbers which may be "fed into the machine".

② A ***range*** (טווח): another subset of $\mathbb{R}$. Numbers that may be "emitted by the machine". We do not exclude the possibility that some of these numbers may never be returned. We only require that every number returned by the function belongs to its range.

③ A ***transformation rule*** (העתקה). The crucial point is that to every number in its domain corresponds one and only one number in its range (חד ערכיות).

We normally denote functions by letters, like we do for real numbers (and for any other mathematical entity). To avoid ambiguities, the function has to be defined properly. For example, we may denote a function by the letter f. If $A \subseteq \mathbb{R}$ is its domain, and $B \subseteq \mathbb{R}$ is its range, we write $f : A \to B$ (f maps the set A into the set B). The transformation rule has to specify what number in B is assigned by the function to each number $x \in A$. We denote the assignment by $f(x)$ (the function f evaluated at x). That the assignment rule is "assign $f(x)$ to x" is denoted by $f : x \mapsto f(x)$ (pronounced "f maps x to $f(x)$")[1].

Definition 2.1 *Let $f : A \to B$ be a function. Its* ***image*** *(תמונה) is the subset of B of numbers that are actually assigned by the function. That is,*

$$\text{Image}(f) = \{y \in B : \ \exists x \in A : f(x) = y\}.$$

The function f is said to be ***onto*** *B (על) if B is its image. f is said to be* ***one-to-one*** *(חד חד ערכית) if to each number in its image corresponds a unique number in its domain, i.e.,*

$$(\forall y \in \text{Image}(f))(\exists! x \in A) : (f(x) = y).$$

Examples:

① A function that assigns to every real number its square. If we denote the function by f, then

$$f : \mathbb{R} \to \mathbb{R} \qquad \text{and} \qquad f : x \mapsto x^2.$$

We may also write $f(x) = x^2$.

We should not say, however, that "the function f is x^2". In particular, we may use any letter other than x as an argument for f. Thus, the functions $f : \mathbb{R} \to \mathbb{R}$, defined by the transformation rules $f(x) = x^2$, $f(t) = t^2$, $f(\alpha) = \alpha^2$ and $f(\xi) = \xi^2$ are identical[2].

Also, it turns out that the function f only returns non-negative numbers. There is however nothing wrong with the definition of the range as the whole of $\mathbb{R}$. We could limit the range to be the set $[0, \infty)$, but not to the set $[1, 2]$.

② A function that assigns to every $w \neq \pm 1$ the number $(w^3 + 3w + 5)/(w^2 - 1)$. If we denote this function by g, then

$$g : \mathbb{R} \setminus \{\pm 1\} \to \mathbb{R} \qquad \text{and} \qquad g : w \mapsto \frac{w^3 + 3w + 5}{w^2 - 1}.$$

[1] Programmers: think of $f : A \to B$ as defining the "type" or "syntax" of the function, and of $f : x \mapsto f(x)$ as defining the "action" of the function.

[2] It takes a great skill not to confuse the Greek letters ζ (zeta) and ξ (xi).

③ A function that assigns to every $-17 \le x \le \pi/3$ its square. This function differs from the function in the first example because the two functions do not have the same domain (different "syntax" but same "routine").

④ A function that assigns to every real number the value zero if it is irrational and one if it is rational[3]. This function is known as the ***Dirichlet function***. We have $f : \mathbb{R} \to \{0, 1\}$, with

$$f : x \mapsto \begin{cases} 0 & x \text{ is irrational} \\ 1 & x \text{ is rational.} \end{cases}$$

⑤ A function defined on the domain

$$A = \{2, 17, \pi^2/17, 36/\pi\} \cup \{a + b\sqrt{2} : \ a, b \in \mathbb{Q}\},$$

such that

$$x \mapsto \begin{cases} 5 & x = 2 \\ 36/\pi & x = 17 \\ 28 & x = \pi^2/17 \text{ or } 36/\pi \\ 16 & \text{otherwise.} \end{cases}$$

The range may be taken to be $\mathbb{R}$, but the image is $\{6, 16, 28, 36/\pi\}$.

⑥ A function defined on $\mathbb{R} \setminus \mathbb{Q}$ (the irrational numbers), which assigns to x the number of 7's in its decimal expansion, if this number is finite. If this number is infinite, then it returns $-\pi$. This example differs from the previous ones in that we do not have an assignment rule in closed form (how the heck do we compute $f(x)$?). Nevertheless it provides a legal assignment rule[4].

⑦ For every $n \in \mathbb{N}$ we may define the n-th power function $f_n : \mathbb{R} \to \mathbb{R}$, by $f_n : x \mapsto x^n$. Here again, we will avoid referring to "the function x^n". The function $f_1 : x \mapsto x$ is known as the ***identity function***, often denoted by Id, namely

$$\mathrm{Id} : \mathbb{R} \to \mathbb{R}, \qquad \mathrm{Id} : x \mapsto x.$$

[3]This function is going to be the course's favorite to display counter examples.

[4]We limited the range of the function to the irrationals because rational numbers may have a non-unique decimal representation, e.g.,

$$0.7 = 0.6999999\dots.$$

⑧ There are many functions that you all know since high school, such as the ***sine***, the ***cosine***, the ***exponential***, and the ***logarithm***. These function require a careful, sometimes complicated, definition. It is our choice in the present course to assume that the meaning of these functions (along with their domains and ranges) are well-understood.

Given several functions, they can be combined together to form new functions. For example, functions form a ***vector space*** over the reals. Let $f : A \to \mathbb{R}$ and $g : B \to \mathbb{R}$ be given functions, and $a, b \in \mathbb{R}$. We may define a new function[5]

$$a\,f + b\,g : A \cap B \to \mathbb{R}, \qquad a\,f + b\,g : x \mapsto a\,f(x) + b\,g(x).$$

This is a vector field whose zero element is the zero function, $x \mapsto 0$; given a function f, its inverse is $-f = (-1)f$.

Moreover, functions form an ***algebra***. We may define the product of two functions,

$$f \cdot g : A \cap B \to \mathbb{R}, \qquad f \cdot g : x \mapsto f(x)g(x),$$

as well as their quotient,

$$f/g : A \cap \{z \in B : \; g(z) \neq 0\} \to \mathbb{R}, \qquad f/g : x \mapsto f(x)/g(x),$$

(12 hrs, 2011) (13 hrs, 2013)

A third operation that combines two functions is ***composition***. Let $f : A \to B$ and $g : B \to C$. We define

$$g \circ f : A \to C, \qquad g \circ f : x \mapsto g(f(x)).$$

For example, if f is the sine function and g is the square function, then[6]

$$g \circ f : \xi \mapsto \sin^2 \xi \quad \text{and} \quad f \circ g : \zeta \mapsto \sin \zeta^2,$$

i.e., composition is non-commutative. On the other hand, composition is associative, namely,

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

[5]This is not trivial. We are adding "machines", not numbers.

[6]It is definitely a habit to use the letter x as generic argument for functions. Beware of fixations! Any letter including Ξ is as good.

Note that for every function f,

$$\text{Id} \circ f = f \circ \text{Id} = f,$$

so that the identity is the neutral element with respect to function composition. This should not be confused with the fact that $x \mapsto 1$ is the neutral element with respect to function multiplication.

Example: Consider the function f that assigns the rule

$$f : x \mapsto \frac{x + x^2 + x \sin x^2}{x \sin x + x \sin^2 x}.$$

This function can be written as

$$f = \frac{\text{Id} + \text{Id} \cdot \text{Id} + \text{Id} \cdot \sin \circ (\text{Id} \cdot \text{Id})}{\text{Id} \cdot \sin + \text{Id} \cdot \sin \cdot \sin}.$$

▲▲▲

Example: Recall the n-th power functions f_n. A function P is called a ***polynomial*** of degree n if there exist real numbers $(a_i)_{i=0}^n$, with $a_n \neq 0$, such that

$$P = \sum_{k=0}^{n} a_k f_k.$$

The union over all n's of polynomials of degree n is the set of polynomials. A function is called ***rational*** if it is the ratio of two polynomials. ▲▲▲

(10 hrs, 2009)

✎ *Exercise 2.1* The following exercise deals with the algebra of functions. Consider the three following functions,

$$s(x) = \sin x \qquad P(x) = 2^x \quad \text{and} \quad S(x) = x^2,$$

all three defined on $\mathbb{R}$.

① Express the following functions without using the symbols s, P, and S: (i) $(s \circ P)(y)$, (ii) $(S \circ s)(\xi)$, and (iii) $(S \circ P \circ s)(t) + (s \circ P)(t)$.

② Write the following functions using only the symbols s, P, and S, along with the binary operators $+$, $\cdot$, and $\circ$: (i) $f(x) = 2^{\sin x}$, (ii) $f(t) = 2^{2^x}$, (iii) $f(u) = \sin(2^u + 2^{u^2})$, (iv) $f(y) = \sin(\sin(\sin(2^{2^{\sin y}})))$, and (v) $f(a) = 2^{\sin^2 a} + \sin(a^2) + 2^{\sin(a^2+\sin a)}$.

✎ *Exercise 2.2* An open segment I is called ***symmetric*** if there exists an $a > 0$ such that $I = (-a, a)$. Let $f : I \to \mathbb{R}$ where I is a symmetric segment; the function f is called ***even*** if $f(-x) = f(x)$ for every $x \in I$, and ***odd*** if $f(-x) = -f(x)$ for every $x \in I$.

Let $f, g : I \to R$, where I is a symmetric segments, and consider the four possibilities of f, g being either even or odd. In each case, determine whether the specified function is even, odd, or not necessarily either two. Explain your statement, and where applicable give a counter example:

① $f + g$.
② $f \cdot g$.
③ $f \circ g$ (here we assume that $g : I \to J$ and $f : J \to \mathbb{R}$, where I and J are symmetric segments).

✎ *Exercise 2.3* Show that any function $f : R \to \mathbb{R}$ can be written as

$$f = E + O,$$

where E is even and O is odd. Prove then that this decomposition is unique. That is, if $f = E + O = E' + O'$, then $E = E'$ and $O = O'$.

✎ *Exercise 2.4* The following exercise deals with polynomials.

① Prove that for every polynomial of degree ≥ 1 and for every real number a there exists a polynomial g and a real number b, such that $f(x) = (x - a)g(x) + b$ (hint: use induction on the degree of the polynomial).
② Prove that if f is a polynomial satisfying $f(a) = 0$, then there exists a polynomial g such that $f(x) = (x - a)g(x)$.
③ Conclude that if f is a polynomial of degree n, then there are at most n real numbers a, such that $f(a) = 0$ (at most n roots).

✎ *Exercise 2.5* Let $a, b, c \in \mathbb{R}$ and $a \neq 0$, and let $f : A \to B$ be given by

$$f : x \mapsto ax^2 + bx + c.$$

① Suppose $B = \mathbb{R}$. Find the largest segment A for which f is one-to-one.

② Suppose $A = \mathbb{R}$. Find the largest segment B for which f is onto.

③ Find segments A and B as large as possible for which f is one-to-one and onto.

✎ *Exercise 2.6* Let f satisfy the relation $f(x + y) = f(x) + f(y)$ for all $x, y \in \mathbb{R}$.

① Prove inductively that

$$f(x_1 + x_2 + \cdots + x_n) = f(x_1) + f(x_2) + \cdots + f(x_n)$$

for all $x_1, \ldots, x_n \in \mathbb{R}$.

② Prove the existence of a real number c, such that $f(x) = cx$ for all $x \in \mathbb{Q}$. (Note that we claim nothing for irrational arguments.)

2.2 Graphs

As you know, we can associate with every function a ***graph***. What is a graph? A drawing? A graph has to be thought of as a ***subset of the plane***. For a function $f : A \to B$, we define the graph of f to be the set

$$\text{Graph}(f) = \{(x, y) : \ x \in A, y = f(x)\} \subset A \times B.$$

The defining property of the function is that it is uniquely defined, i.e.,

$$(x, y) \in \text{Graph}(f) \quad \text{and} \quad (x, z) \in \text{Graph}(f) \quad \text{implies} \quad y = z.$$

The function is one-to-one if

$$(x, y) \in \text{Graph}(f) \quad \text{and} \quad (w, y) \in \text{Graph}(f) \quad \text{implies} \quad x = w.$$

It is onto B if

$$(\forall y \in B)(\exists x \in A) : ((x, y) \in \text{Graph}(f)).$$

The definition we gave to a function as an assignment rule is, strictly speaking, not a formal one. The standard way to define a function is via its graph. A function *is* a graph—a subset of the Cartesian product space $A \times B$.

Comment: In fact, this is the general way to define functions. Let A and B be two sets (of anything!). A function $f : A \to B$ is a subset of the Cartesian product

$$\text{Graph}(f) \subset A \times B = \{(a,b) : a \in A, b \in B\},$$

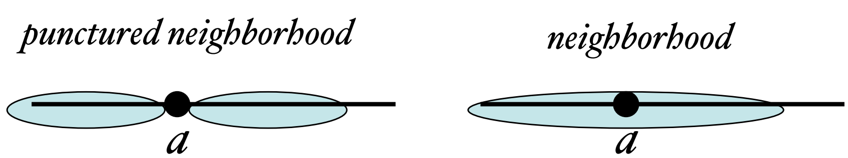
satisfying

$$(\forall a \in A)(\exists! b \in B) : ((a,b) \in \text{Graph}(f)).$$

(12 hrs, 2010)

2.3 Limits

Definition 2.2 Let $x \in \mathbb{R}$. A ***neighborhood*** *of x (סביבה) is an open segment (a,b) that contains the point x (note that since the segment is open, x cannot be a boundary point). A* ***punctured neighborhood*** *of x (סביבה מנוקֶבֶת) is a set $(a,b) \setminus \{x\}$ where $a < x < b$.*



Definition 2.3 Let $A \subset \mathbb{R}$. A point $a \in A$ is called an ***interior point*** *of A (נקודה פנימית) if it has a neighborhood contained in A.*

Notation: We will mostly deal with ***symmetric*** neighborhoods (whether punctured or not), i.e., neighborhoods of a of the form

$$\{x : |x-a| < \delta\}$$

for some $\delta > 0$. We will introduce the following notations for neighborhoods,

$$B(a,\delta) = (a-\delta, a+\delta)$$
$$B^{\circ}(a,\delta) = \{x : 0 < |x-a| < \delta\}.$$

We will also define ***one-sided neighborhoods***,

$$
\begin{aligned}
B_+(a,\delta) &= [a, a+\delta) \\
B_+^\circ(a,\delta) &= (a, a+\delta) \\
B_-(a,\delta) &= (a-\delta, a] \\
B_-^\circ(a,\delta) &= (a-\delta, a).
\end{aligned}
$$

Example: For $A = [0,1]$, the point $1/2$ is an interior point, but not the point 0. ▲▲▲

Example: The set $\mathbb{Q} \subset \mathbb{R}$ has no interior points, and neither does its complement, $\mathbb{R} \setminus \mathbb{Q}$. ▲▲▲

In this section we explain the meaning of *the limit of a function at a point*. Informally, we say that:

> The limit of a function f at a point a is ℓ, if we can make f assign a value as close to ℓ as we wish, by making its argument sufficiently close to a (excluding the value a itself).

Note that the function does not need to be equal to ℓ at the point a; in fact, it does not even need to be defined at the point a.

Example: Consider the function

$$f : \mathbb{R} \to \mathbb{R} \qquad f : x \mapsto 3x.$$

We claim that the limit of this function at the point 5 is 15. This means that we can make $f(x)$ be as close to 15 as we wish, by making x sufficiently close to 5, with 5 itself being excluded. Suppose you want $f(x)$ to differ from 15 by less than $1/100$. This means that you want

$$15 - \frac{1}{100} < f(x) = 3x < 15 + \frac{1}{100}.$$

This requirement is guaranteed if

$$5 - \frac{1}{300} < x < 5 + \frac{1}{300}.$$

Thus, if we take x to differ from 5 by less than $1/300$ (but more than zero!), then we are guaranteed to have $f(x)$ within the desired range. Since we can repeat this construction for any number other than $1/100$, then we conclude that the limit of f at 5 is 15. We write,

$$\lim_{5} f = 15.$$

although the more common notation is[7]

$$\lim_{x\to 5} f(x) = 15.$$

We can be more precise. Suppose you want $f(x)$ to differ from 15 by less than ϵ, for some $\epsilon > 0$ of your choice. In other words, you want

$$|f(x) - 15| = |3x - 15| < \epsilon.$$

This is guaranteed if $|x-5| < \epsilon/3$, thus given $\epsilon > 0$, choosing x within a symmetric punctured neighborhood of 5 of radius $\epsilon/3$ guarantees that $f(x)$ is within a distance of ϵ from 15. ▲▲▲

This example insinuates what would be a formal definition of the limit of a function at a point:

Definition 2.4 *Let $f : A \to B$ with $a \in A$ an interior point. We say that the limit of f at a is ℓ, if for every $\epsilon > 0$, there exists a $\delta > 0$, such that*[8]

$$|f(x) - \ell| < \epsilon \qquad \text{for all } x \in B^\circ(a, \delta).$$

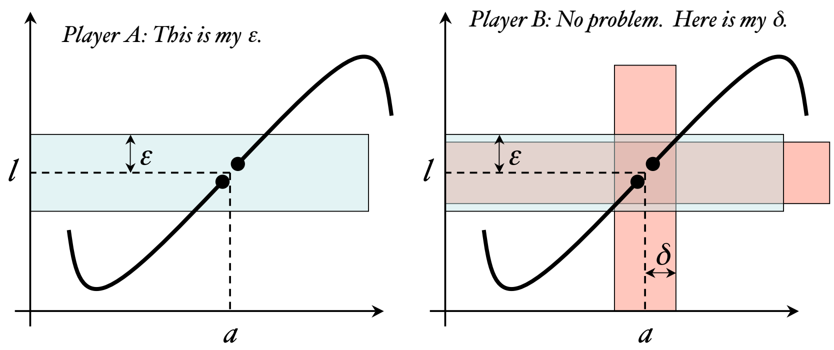
In formal notation,

$$(\forall \epsilon > 0)(\exists \delta > 0) : (\forall x \in B^\circ(a, \delta))(|f(x) - \ell| < \epsilon).$$

[7] The choice of not using the common notation becomes clear when you realize that you could as well write

$$\lim_{\aleph\to 5} f(\aleph) = 15 \qquad \text{or} \qquad \lim_{\xi\to 5} f(\xi) = 15.$$

[8] The game is "you give me ϵ and I give you δ in return".



Since Player B can find a δ for every choice of ε made by Player A, it follows that the limit of f at a is l.

Example: Consider the square function $f : \mathbb{R} \to \mathbb{R}$, $f : x \mapsto x^2$. We will show that
$$\lim_3 f = 9.$$
By definition, we need to show that for every $\epsilon > 0$ we can find a $\delta > 0$, such that
$$|f(x) - 9| < \epsilon \qquad \text{whenever} \qquad x \in B^\circ(3, \delta).$$
Thus, the game it to respond with an appropriate δ for every ϵ.

To find the appropriate δ, we examine the condition that needs to be satisfied:
$$|f(x) - 9| = |x - 3|\,|x + 3| < \epsilon.$$
If we impose that $|x - 3| < \delta$ then
$$|f(x) - 9| < |x + 3|\,\delta.$$
If the $|x+3|$ wasn't there we would have set $\delta = \epsilon$ and we would be done. Instead, we note that by the triangle inequality,
$$|x + 3| = |x - 3 + 6| \le |x - 3| + 6,$$
so that for all $|x - 3| < \delta$,
$$|f(x) - 9| \le |x - 3|\,(|x - 3| + 6) < (6 + \delta)\delta.$$

Now recall: given ϵ we have the freedom to choose $\delta > 0$ such to make the right-hand side less or equal than ϵ. We may freely require for example that $\delta \le 1$, in which case

$$|f(x) - 9| < (6 + \delta)\delta \le 7\delta.$$

If we furthermore take $\delta \le \epsilon/7$, then we reach the desired goal.

To summarize, given $\epsilon > 0$ we take $\delta = \min(1, \epsilon/7)$, in which case $(\forall x : 0 < |x - 3| < \delta)$,

$$|f(x) - 9| = |x - 3|\,|x + 3| \le |x - 3|\,(|x - 3| + 6) < (6 + \delta)\delta \le (6 + 1)\frac{\epsilon}{7} = \epsilon.$$

This proves (by definition) that $\lim_3 f = 9$.

We can show, in general, that

$$\lim_a f = a^2$$

for all $a \in \mathbb{R}$. [do it!] ▲▲▲

(13 hrs, 2010) (15 hrs, 2013)

Example: The next example is the function $f : (0, \infty)$, $f : x \mapsto 1/x$. We are going to show that for every $a > 0$,

$$\lim_a f = \frac{1}{a}.$$

First fix a; it is not a variable. Let $\epsilon > 0$ be given. We first observe that

$$\left| f(x) - \frac{1}{a} \right| = \left| \frac{1}{x} - \frac{1}{a} \right| = \frac{|x - a|}{ax}.$$

We need to be careful that the domain of x does not include zero. We start by requiring that $|x - a| < a/2$, which at once implies that $x > a/2$, hence

$$\left| f(x) - \frac{1}{a} \right| < \frac{2|x - a|}{a^2}.$$

If we further require that $|x - a| < a^2\epsilon/2$, then $|f(x) - 1/a| < \epsilon$. To conclude,

$$\left| f(x) - \frac{1}{a} \right| < \epsilon \qquad \text{whenever} \qquad x \in B^\circ(a, \delta),$$

for $\delta = \min(a/2, a^2\epsilon/2)$. ▲▲▲

(14 hrs, 2011)

Example: Consider next the Dirichlet function, $f : \mathbb{R} \to \mathbb{R}$,

$$f : x \mapsto \begin{cases} 0 & x \text{ is irrational} \\ 1 & x \text{ is rational.} \end{cases}$$

We will show that f does not have a limit at zero. To show that for all ℓ,

$$\lim_0 f \neq \ell,$$

we need to show the negation of

$$(\exists \ell \in \mathbb{R}) : (\forall \epsilon > 0)(\exists \delta > 0) : (\forall x \in B^\circ(0, \delta))(f(x) \in B(\ell, \epsilon)),$$

i.e.,

$$(\forall \ell \in \mathbb{R})(\exists \epsilon > 0) : (\forall \delta > 0)(\exists x \in B^\circ(0, \delta)) : (|f(x) - \ell| \geq \epsilon).$$

Indeed, take $\epsilon = 1/4$. Then no matter what δ is, there exist points $0 < |x| < \delta$ for which $f(x) = 0$ and points for which $f(x) = 1$, i.e.,

$$(\forall \delta > 0)(\exists x, y \in B^\circ(0, \delta)) : (f(x) = 0, f(y) = 1).$$

No ℓ can satisfy both

$$|0 - \ell| < \frac{1}{4} \qquad \text{and} \qquad |1 - \ell| < \frac{1}{4}.$$

I.e.,

$$(\forall \ell \in \mathbb{R}, \forall \delta > 0)(\exists x \in B^\circ(0, \delta)) : (f(x) \notin B(\ell, \tfrac{1}{4})).$$

▲▲▲

(12 hrs, 2009)

Comment: It is important to stress what is the negation that the limit f at a is ℓ:

> There *exists* an $\epsilon > 0$, such that *for all* $\delta > 0$, there *exists* an x, which satisfies $0 < |x - a| < \delta$, but not $|f(x) - \ell| < \epsilon$.

Example: Consider, in contrast, the function $f : \mathbb{R} \to \mathbb{R}$,

$$f : x \mapsto \begin{cases} 0 & x \text{ is irrational} \\ x & x \text{ is rational.} \end{cases}$$

Here we show that

$$\lim_{0} f = 0.$$

Indeed, let $\epsilon > 0$ be given, then by taking $\delta = \epsilon$,

$$x \in B^{\circ}(0, \delta) \qquad \text{implies} \qquad |f(x) - 0| < \epsilon.$$

▲▲▲

Having a formal definition of a limit, and having seen a number of example, we are in measure to prove general theorems about limits. The first theorem states that a limit, if it exists, is unique.

✎ *Exercise 2.7* Prove that if

$$(\forall \epsilon > 0)(\exists \delta > 0) : (\forall x : 0 < |x - 3| < \delta)(|f(x) - 14| < \epsilon),$$

then

$$(\forall \epsilon > 0)(\exists \delta > 0) : (\forall x : 0 < |x - 3| < \delta)(|f(x) - 14| < \epsilon/2),$$

and also

$$(\forall \epsilon > 0)(\exists \delta > 0) : (\forall x : 0 < |x - 3| < \delta)(|f(x) - 14| < \sqrt{\epsilon}).$$

Theorem 2.1 (Uniqueness of the limit) A function $f : A \to B$ has at most one limit at any interior point $a \in A$.

Proof: Suppose, by contradiction, that

$$\lim_{a} f = \ell \qquad \text{and} \qquad \lim_{a} f = m,$$

with $\ell \neq m$. By definition, there exists a $\delta_1 > 0$ such that

$$|f(x) - \ell| < \frac{1}{4}|\ell - m| \qquad \text{whenever } x \in B^{\circ}(a, \delta_1),$$

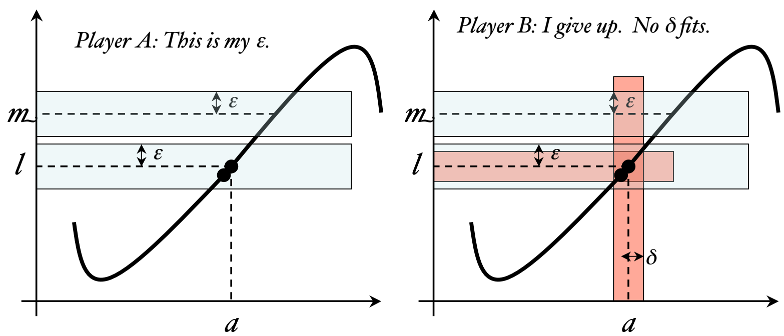
and there exists a $\delta_2 > 0$ such that

$$|f(x) - m| < \frac{1}{4}|\ell - m| \qquad \text{whenever } x \in B^{\circ}(a, \delta_2).$$

Take $\delta = \min(\delta_1, \delta_2)$. Then, whenever $x \in B^{\circ}(a, \delta)$,

$$|\ell - m| = |(f(x) - m) - (f(x) - \ell)| \leq |f(x) - \ell| + |f(x) - m| < \frac{1}{2}|\ell - m|,$$

which is a contradiction.



Since Player B cannot find a δ for a particular choice of ε made by Player A, it follows that the limit of f at a cannot be both l and m.

■

✎ *Exercise 2.8* In each of the following items, determine whether the limit exists, and if it does calculate its value. You have to calculate the limit directly from its definition (without using, for example, limits arithmetic—see below).

① $\lim_2 f$, where $f(x) = 7x$.

② $\lim_0 f$, where $f(x) = x^2 + 3$.

③ $\lim_0 f$, where $f(x) = 1/x$.

④ $\lim_a f$, where

$$f(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0, \end{cases}$$

and $a \in \mathbb{R}$.

⑤ $\lim_a f$, where $f(x) = |x|$.

⑥ $\lim_a f$, where $f(x) = \sqrt{|x|}$.

✎ *Exercise 2.9* Let I be an open interval in $\mathbb{R}$, with $a \in I$. Let $f, g : I \setminus \{a\} \to \mathbb{R}$ (i.e., defined on a punctured neighborhood of a). Let $J \subset I$ be an open interval containing the point a, such that $f(x) = g(x)$ for all $x \in J \setminus \{a\}$. Show that $\lim_a f$ exists if and only if $\lim_a g$ exists, and if they do exist they are equal. (This exercise shows that the limit of a function at a point only depends on its properties in a neighborhood of that point).

✎ *Exercise 2.10* The ***Riemann function*** $R : \mathbb{R} \to \mathbb{R}$ is defined as

$$R(x) = \begin{cases} 1/q & x = p/q \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q}, \end{cases}$$

where the representation p/q is in reduced form (for example $R(4/6) = R(2/3) = 1/3$). Note that $R(0) = R(0/1) = 1$. Show that $\lim_a R$ exists everywhere and is equal to zero.

(17 hrs, 2013)

We have seen above various examples of limits. In each case, we had to "work hard" to prove what the limit was, by showing that the definition was satisfied. This becomes impractical when the functions are more complex. Thus, we need to develop theorems that will make our task easier. But first, a lemma:

Lemma 2.1 Let $\ell, m \in \mathbb{R}$ and let $\epsilon > 0$. Then,

①

$$|x - \ell| < \frac{\epsilon}{2} \quad \textit{and} \quad |y - m| < \frac{\epsilon}{2} \quad \textit{imply} \quad |(x + y) - (\ell + m)| < \epsilon.$$

②

$$|x - \ell| < \min\left(1, \frac{\epsilon}{2(|m| + 1)}\right) \quad \textit{and} \quad |y - m| < \min\left(1, \frac{\epsilon}{2(|\ell| + 1)}\right)$$

imply

$$|xy - \ell m| < \epsilon.$$

③ *If* $m \neq 0$ *and*

$$|y - m| < \min\left(\frac{|m|}{2}, \frac{|m|^2\epsilon}{2}\right),$$

then $y \neq 0$ *and*

$$\left|\frac{1}{y} - \frac{1}{m}\right| < \epsilon.$$

(15 hrs, 2011)

Proof: The first item is obvious (the triangle inequality). Consider the second item. If the two conditions are satisfied then

$$|x| = |x - \ell + \ell| < |x - \ell| + |\ell| < |\ell| + 1,$$

hence,

$$\begin{aligned} |xy - \ell m| &= |x(y - m) + m(x - \ell)| \leq |x||y - m| + |m||x - \ell| \\ &< \epsilon\frac{|\ell| + 1}{2(|\ell| + 1)} + \epsilon\frac{|m|}{2(|m| + 1)} < \epsilon. \end{aligned}$$

For the third item it is clear that $|y - m| < |m|/2$ implies that $|y| > |m|/2$, and in particular, $y \neq 0$. Then

$$\left|\frac{1}{y} - \frac{1}{m}\right| = \frac{|y - m|}{|y||m|} < \frac{|m|^2\epsilon}{2|m|(|m|/2)} = \epsilon.$$

■

(13 hrs, 2009)

(15 hrs, 2010)

Theorem 2.2 (Limits arithmetic) Let f *and* g *be two functions defined in a neighborhood of* a*, such that*

$$\lim_a f = \ell \qquad \textit{and} \qquad \lim_a g = m.$$

Then,

$$\lim_a (f+g) = \ell + m \qquad \textit{and} \qquad \lim_a (f \cdot g) = \ell \cdot m.$$

If, moreover, $\ell \neq 0$*, then*

$$\lim_a \left(\frac{1}{f}\right) = \frac{1}{\ell}.$$

Proof: This is an immediate consequence of the above lemma. Since, given $\epsilon > 0$, we know that there exists a $\delta > 0$, such that[9]

$$|f(x) - \ell| < \frac{\epsilon}{2} \quad \text{and} \quad |g(x) - m| < \frac{\epsilon}{2} \quad \text{whenever} \quad x \in B^\circ(a, \delta).$$

It follows that

$$|(f(x) + g(x)) - (\ell + m)| < \epsilon \quad \text{whenever} \quad x \in B^\circ(a, \delta).$$

There also exists a $\delta > 0$, such that

$$|f(x) - \ell| < \min\left(1, \frac{\epsilon}{2(|m|+1)}\right) \quad \text{and} \quad |g(x) - m| < \min\left(1, \frac{\epsilon}{2(|\ell|+1)}\right)$$

whenever $x \in B^\circ(a, \delta)$. It follows that

$$|f(x)g(x) - \ell m| < \epsilon \quad \text{whenever} \quad x \in B^\circ(a, \delta).$$

Finally, there exists a $\delta > 0$, such that

$$|f(x) - \ell| < \min\left(\frac{|\ell|}{2}, \frac{|\ell|^2 \epsilon}{2}\right) \quad \text{whenever} \quad x \in B^\circ(a, \delta).$$

It follows that

$$\left|\frac{1}{f(x)} - \frac{1}{\ell}\right| < \epsilon \quad \text{whenever} \quad x \in B^\circ(a, \delta).$$

■

Comment: It is important to point out that the fact that $f + g$ has a limit at a does not imply that either f or g has a limit at that point; take for example the functions $f(x) = 5 + 1/x$ and $g(x) = 6 - 1/x$ at $x = 0$.

[9] This $\delta > 0$ is already the smallest of the two δ's corresponding to each of the functions.

Example: What does it take now to show that[10]

$$\lim_{x\to a} \frac{x^3 + 7x^5}{x^2 + 1} = \frac{a^3 + 7a^5}{a^2 + 1}.$$

We need to show that for all $c \in \mathbb{R}$

$$\lim_a c = c,$$

and that for all $k \in \mathbb{N}$,

$$\lim_{x\to a} x^k = a^k,$$

which follows by induction once we show that for $k = 1$. ▲▲▲

✎ *Exercise 2.11* Show that

$$\lim_{x\to a} x^n = a^n, \qquad a \in \mathbb{R}, n \in \mathbb{N},$$

directly from the definition of the limit, i.e., without using limits arithmetic.

✎ *Exercise 2.12* Let f, g be defined in a punctured neighborhood of a point a. Prove or disprove:

① If $\lim_a f$ and $\lim_a g$ do not exist, then $\lim_a (f + g)$ does not exist either.
② If $\lim_a f$ and $\lim_a g$ do not exist, then $\lim_a (f \cdot g)$ does not exist either
③ If both $\lim_a f$ and $\lim_a (f + g)$ exist, then $\lim_a g$ also exists.
④ If both $\lim_a f$ and $\lim_a (f \cdot g)$ exist, then $\lim_a g$ also exists.
⑤ If $\lim_a f$ exists and $\lim_a (f + g)$ does not exist, then $\lim_a (f + g)$ does no exist.
⑥ If g is bounded in a punctured neighborhood of a and $\lim_a f = 0$, then $\lim_a (f \cdot g) = 0$.

We conclude this section by extending the notion of a limit to that of a ***one-sided limit***.

[10]This is the standard notation to what we write in these notes as

$$f : \xi \mapsto \frac{\xi^3 + 7\xi^5}{\xi^2 + 1} \qquad \text{and} \qquad \lim_a f = \frac{a^3 + 7a^5}{a^2 + 1}.$$

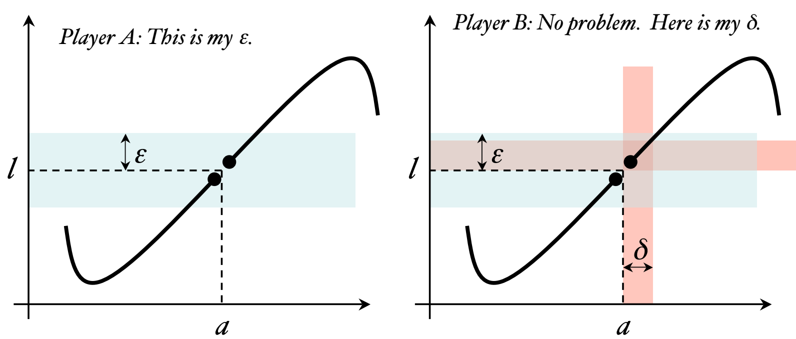
Definition 2.5 *Let $f : A \to B$ with $a \in A$ an interior point. We say that the* ***limit on the right*** *(גבול מימין) of f at a is ℓ, if for every $\epsilon > 0$, there exists a $\delta > 0$, such that*

$$|f(x) - \ell| < \epsilon \qquad \text{whenever} \qquad x \in B^{\circ}_{+}(a, \delta).$$

We write,

$$\lim_{a^+} f = \ell.$$

An analogous definition is given for the ***limit on the left*** *(גבול משמאל).*



Since Player B can find a δ for every choice of ε made by Player A, it follows that the right-limit of f at a is l.

✎ *Exercise 2.13* Prove that

$$\lim_{a} f \qquad \text{exists}$$

if and only

$$\lim_{a^-} f = \lim_{a^+} f.$$

✎ *Exercise 2.14* We denote by $f : x \mapsto \lfloor x \rfloor$ the lower integer value of x.

① Draw the graph of this function.

② Find

$$\lim_{a^-} f \qquad \text{and} \qquad \lim_{a^+} f$$

for $a \in \mathbb{R}$. Prove it based on the definition of one-sided limits.

✎ *Exercise 2.15* Show that $f : (0, \infty) \to \mathbb{R}$, $f : x \mapsto \sqrt{x}$ satisfies

$$\lim_{a} f = \sqrt{a},$$

for all $a > 0$.

2.4 Limits and order

In this section we will prove a number of properties pertinent to limits and order. First a definition:

Definition 2.6 *Let $f : A \to B$ and $a \in A$ an interior point. We say that f is* ***locally bounded*** *(חסומה מקומית) near a if there exists a punctured neighborhood U of a, such that the set*

$$\{f(x) : x \in U\}$$

is bounded. Equivalently, there exists a $\delta > 0$ such that the set

$$\{f(x) : x \in B^{\circ}(a, \delta)\}$$

is bounded.

Comment: As in the previous section, we consider punctured neighborhoods of a point, and we don't even care if the function is defined at the point.

Proposition 2.1 *Let f and g be two functions defined in a punctured neighborhood U of a point a. Suppose that*

$$f(x) < g(x) \qquad \forall x \in U,$$

and that

$$\lim_a f = \ell \qquad \text{and} \qquad \lim_a g = m.$$

Then $\ell \le m$.

Comment: The very fact that f and g have limits at a is an assumption.

Comment: Note that even though $f(x) < g(x)$ is a strict inequality, the resulting inequality of the limits is in a weak sense. To see what this must be the case, consider the example

$$f(x) = |x| \qquad \text{and} \qquad g(x) = 2|x|.$$

Even though $f(x) < g(x)$ in an open neighborhood of 0,

$$\lim_0 f = \lim_0 g = 0.$$

Proof: Suppose, by contradiction, that $\ell > g$ and set $\epsilon = \frac{1}{2}(\ell - m)$. Thus, there exists a $\delta > 0$ such that

$$(\forall x \in B^\circ(a,\delta)) : (|f(x) - \ell| < \epsilon \quad \text{and} \quad |g(x) - m| < \epsilon).$$

In particular, for every such x,

$$f(x) > \ell - \epsilon = \ell - \frac{\ell - m}{2} = m + \frac{\ell - m}{2} = m + \epsilon > g(x),$$

which is a contradiction. ■

(19 hrs, 2013)

Proposition 2.2 Let f and g be two functions defined in a punctured neighborhood U of a point a. Suppose that

$$\lim_a f = \ell \qquad \textit{and} \qquad \lim_a g = m,$$

with $\ell < m$. Then there exists a $\delta > 0$ such that

$$(\forall x \in B^\circ(a,\delta)) : (f(x) < g(x)).$$

Comment: This time both inequalities are strong.

Proof: Let $\epsilon = \frac{1}{2}(m - \ell)$. There exists a $\delta > 0$ such that

$$(\forall x \in B^\circ(a,\delta)) : (|f(x) - \ell| < \epsilon \quad \text{and} \quad |g(x) - m| < \epsilon).$$

For every such x,

$$f(x) < \ell + e = \ell + \frac{1}{2}(m - \ell) = \frac{1}{2}(\ell + m)$$
$$g(x) > m - e = m - \frac{1}{2}(m - \ell) = \frac{1}{2}(\ell + m)$$

hence,

$$f(x) < g(x).$$

■

Proposition 2.3 *Let f, g, h be defined in a punctured neighborhood U of a. Suppose that*

$$(\forall x \in U) : (f(x) \leq g(x) \leq h(x)),$$

and

$$\lim_a f = \lim_a h = \ell.$$

Then

$$\lim_a g = \ell.$$

Proof: It is given that for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$(\forall x \in B^\circ(a, \delta)) : (\ell - \epsilon < f(x) \quad \text{and} \quad h(x) < \ell + \epsilon).$$

Then for every such x,

$$\ell - \epsilon < f(x) \leq g(x) \leq h(x) < \ell + \epsilon,$$

i.e.,

$$(\forall x \in B^\circ(a, \delta)) : (|g(x) - \ell| < \epsilon).$$

■

(17 hrs, 2011)

Proposition 2.4 *Let f be defined in a punctured neighborhood U of a. If the limit*

$$\lim_a f = \ell$$

exists, then f is locally bounded near a.

Proof: Immediate from the definition of the limit. ■

Proposition 2.5 *Let f be defined in a punctured neighborhood of a point a. Then,*

$$\lim_a f = \ell \qquad \textit{if and only if} \qquad \lim_a (f - \ell) = 0.$$

Proof: Very easy. ■

Proposition 2.6 *Let f and g be defined in a punctured neighborhood of a point a. Suppose that*

$$\lim_a f = 0$$

whereas g is locally bounded near a. Then,

$$\lim_a (fg) = 0.$$

Proof: Very easy. ■

Comment: Note that we do not require g to have a limit at a.

With the above tools, here is another way of proving the product property of the arithmetic of limits (this time without ϵ and δ).

Proposition 2.7 (Arithmetic of limits, product) *Let f, g be defined in a punctured neighborhood of a point a, and*

$$\lim_a f = \ell \qquad \text{and} \qquad \lim_a g = m.$$

Then

$$\lim_a (fg) = \ell m.$$

Proof: Write

$$fg - \ell m = \underbrace{\underbrace{(f-\ell)}_{\text{limit zero}}\underbrace{g}_{\text{loc. bdd}}}_{\text{limit zero}} - \underbrace{\ell \underbrace{(g-m)}_{\text{limit zero}}}_{\text{limit zero}}.$$

■

(17 hrs, 2010)

2.5 Continuity

Definition 2.7 *A function $f : A \to \mathbb{R}$ is said to be* ***continuous*** *at an inner point $a \in A$, if it has a limit at a, and*

$$\lim_a f = f(a).$$

Comment: If f has a limit at a and the limit differs from $f(a)$ (or that $f(a)$ is undefined) , then we say that f has a ***removable discontinuity*** (אי רציפות סליקה) at a.

Examples:

① We saw that the function $f(x) = x^2$ has a limit at 3, and that this limit was equal 9. Hence, f is continuous at $x = 3$.

② We saw that the function $f(x) = 1/x$ has a limit at any $a > 0$ and that this limit equals $1/a$. By a similar calculation we could have shown that it has a limit at any $a < 0$ and that this limit also equals $1/a$. Hence, f is continuous at all $x \neq 0$. Since f is not even defined at $x = 0$, then it is not continuous at that point.

③ The function $f(x) = x \sin 1/x$ (it was seen in the tutoring session) is continuous at all $x \neq 0$. At zero it is not defined, but if we define $f(0) = 0$, then it is continuous for all $x \in \mathbb{R}$ (by the bounded times limit zero argument).

④ The function

$$f : x \mapsto \begin{cases} x & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q}, \end{cases}$$

is continuous at $x = 0$, however it is not continuous at any other point, because the limit of f at $a \neq 0$ does not exist.

⑤ The elementary functions sin and cos are continuous everywhere. To prove that sin is continuous, we need to prove that for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|\sin x - \sin a| < \epsilon \qquad \text{whenever} \qquad x \in B^\circ(a, \delta).$$

We will take it for granted at the moment.

Our theorems on limit arithmetic imply right away similar theorems for continuity:

Theorem 2.3 If $f : A \to \mathbb{R}$ and $g : A \to \mathbb{R}$ are continuous at $a \in A$ then $f + g$ and $f \cdot g$ are continuous at a. Moreover, if $f(a) \neq 0$ then $1/f$ is continuous at a.

Proof: Obvious. ■

Comment: Recall that in the definition of the limit, we said that the limit of f at a is ℓ if for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - \ell| < \epsilon \qquad \text{whenever} \qquad x \in B^\circ(a, \delta).$$

There was an emphasis on the fact that the point x itself is excluded. Continuity is defined as that for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - f(a)| < \epsilon \qquad \text{whenever} \qquad x \in B^\circ(a, \delta).$$

Note that we may require that this be true whenever $x \in B(a, \delta)$. There is no need to exclude the point $x = a$.

(19 hrs, 2010)

With that, we have all the tools to show that a function like, say,

$$f(x) = \frac{\sin^2 x + x^2 + x^4 \sin x}{1 + \sin^2 x}$$

is continuous everywhere in $\mathbb{R}$. But what about a function like $\sin x^2$. Do we have the tools to show that it is continuous. No. We don't yet have a theorem for the composition of continuous functions.

Theorem 2.4 Let $g : A \to B$ and $f : B \to C$. Suppose that g is continuous at an inner point $a \in A$, and that f is continuous at $g(a) \in B$, which is an inner point. Then, $f \circ g$ is continuous at a.

Proof: Let $\epsilon > 0$. Since f is continuous at $g(a)$, then there exists a $\delta_1 > 0$, such that

$$|f(y) - f(g(a))| < \epsilon \qquad \text{whenever} \qquad y \in B(g(a), \delta_1).$$

Since g is continuous at a, then there also exists a $\delta_2 > 0$, such that

$$|g(x) - g(a)| < \delta_1 \quad \text{or} \quad g(x) \in B(g(a), \delta_1) \qquad \text{whenever} \qquad x \in B(a, \delta_2).$$

Combining the two we get that

$$|f(g(x)) - f(g(a))| < \epsilon \qquad \text{whenever} \qquad B(a, \delta_2).$$

■

Definition 2.8 *A function f is said to be* **right-continuous** *(רציפה מימין) at a if*

$$\lim_{a^+} f = f(a).$$

A similar definition is given for **left-continuity** *(רציפה משמאל).*

So far, we have only dealt with continuity at points. Usually, we are interested in continuity on intervals.

Definition 2.9 *A function f is said to be continuous on an open interval (a, b) if it is continuous at all $x \in (a, b)$. It is said to be continuous on a closed interval $[a, b]$ if it is continuous on the open interval, and in addition it is right-continuous at a and left-continuous at b.*

Comment: We usually think of continuous function as "well-behaved". One should be careful with such interpretations; see for example

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

✎ *Exercise 2.16* Suppose $f : \mathbb{R} \to \mathbb{R}$ satisfies $f(a + b) = f(a) + f(b)$ for all $a, b \in \mathbb{R}$.)We have already seen that this implies that $f(x) = cx$, for some constant c for all $x \in \mathbb{Q}$.) Suppose that in addition f is continuous at zero. Prove that f is continuous everywhere.

✎ *Exercise 2.17* ① Prove that if $f : A \to B$ is continuous in A, then so is $|f|$ (where $|f|(x) \stackrel{\text{def}}{=} |f(x)|$).

② Prove that if $f : A \to \mathbb{R}$ and $g : A \to \mathbb{R}$ are continuous then so are $\max(f, g)$ and $\min(f, g)$.

✎ *Exercise 2.18* Suppose that $g, h : \mathbb{R} \to \mathbb{R}$ are continuous at a and satisfy $g(a) = h(a)$. We define a new function that "glues" g and h at a:

$$f(x) = \begin{cases} g(x) & x \le a \\ h(x) & x > a. \end{cases}$$

Prove that f is continuous at a.

✎ *Exercise 2.19* Here is another "gluing" problem: Suppose that g is continuous in $[a, b]$, that h is continuous is $[b, c]$, and $g(b) = h(b)$. Define

$$f(x) = \begin{cases} g(x) & x \in [a, b] \\ h(x) & x \in [b, c]. \end{cases}$$

Prove that f is continuous in $[a, c]$.

Example: Here is another "crazy function" due to Johannes Karl Thomae (1840-1921):

$$r(x) = \begin{cases} 1/q & x = p/q \\ 0 & x \notin \mathbb{Q}, \end{cases}$$

where $x = p/q$ assumes that x is rational in reduced form. This function has the wonderful property of being continuous at all $x \notin \mathbb{Q}$ and discontinuous at all $x \in \mathbb{Q}$ (this is because its limit is everywhere zero). It took some more time until Vito Volterra proved in 1881 that there can be no function that is continuous on $\mathbb{Q}$ and discontinuous on $\mathbb{R} \setminus \mathbb{Q}$. ▲▲▲

(21 hrs, 2013)

2.6 Theorems about continuous functions

Theorem 2.5 Suppose that f is continuous at a and $f(a) > 0$. Then there exists a neighborhood of a in which $f(x) > 0$. That is, there exists a $\delta > 0$ such that

$$f(x) > 0 \qquad \text{whenever} \qquad x \in B(a, \delta).$$

Comment: An analogous theorem holds if $f(a) < 0$. Also, a one-sided version can be proved, whereby if f is right-continuous at a and $f(a) > 0$, then

$$f(x) > 0 \qquad \text{whenever} \qquad x \in B_+(a,\delta).$$

Comment: Note that continuity is only required at the point a.

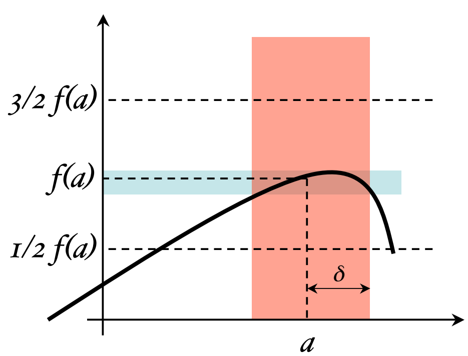
Proof: Since f is continuous at a, there exists a $\delta > 0$ such that

$$|f(x) - f(a)| < \frac{f(a)}{2} \qquad \text{whenever} \qquad x \in B(a,\delta).$$

But, $|f(x) - f(a)| < \frac{f(a)}{2}$ implies

$$f(a) - f(x) \le |f(a) - f(x)| < \frac{f(a)}{2},$$

i.e., $f(x) > f(a)/2 > 0$.



■

This theorem can be viewed as a lemma for the following important theorem:

(15 hrs, 2009)

(20 hrs, 2010)

Theorem 2.6 (Intermediate value theorem משפט ערך הביניים) Suppose f is continuous on an interval $[a,b]$, with $f(a) < 0$ and $f(b) > 0$. Then, there exists a point $c \in (a,b)$, such that $f(c) = 0$.

(19 hrs, 2011)

Comment: Continuity is required on the whole intervals, for consider $f : [0,1] \to \mathbb{R}$:

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

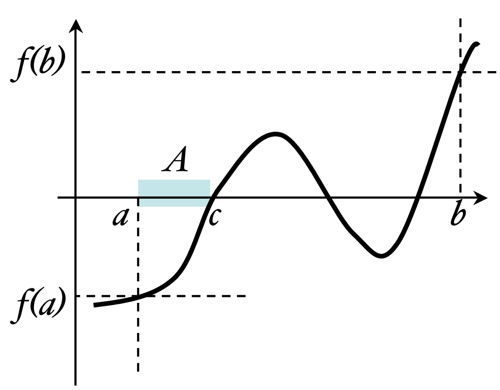
Comment: If $f(a) > 0$ and $f(b) < 0$ the same conclusion holds for just replace f by $(-f)$.

Proof: Consider the set

$$A = \{x \in [a,b] : \ f(y) < 0 \text{ for all } y \in [a,x]\}.$$

The set is non empty for it contains a. Actually, the previous theorems guarantees the existence of a $\delta > 0$ such that $[a, a+\delta) \subseteq A$. This set is also bounded by b. The previous theorem guarantees even the existence of a $\delta > 0$ such that $b - \delta$ is an upper bound for A. By the axiom of completeness, there exists a number $c \in (a,b)$ such that

$$c = \sup A.$$



Suppose it were true that $f(c) < 0$. By the previous theorem, there exists a $\delta > 0$ such that

$$f(x) < 0 \qquad \text{whenever} \qquad c \le x \le c + \delta,$$

Because c is the least upper bound of a it follows that also

$$f(x) < 0 \qquad \text{whenever} \qquad x < c,$$

i.e., $f(x) < 0$ for all $x \in [a, c + \delta]$, which means that $c + \delta \in A$, contradicting the fact that c is an upper bound for A.

Suppose it were true that $f(c) > 0$. By the previous theorem, there exists a $\delta > 0$ such that

$$f(x) > 0 \qquad \text{whenever} \qquad c - \delta \le x \le c,$$

which implies that $c - \delta$ is an upper bound for A, i.e., c cannot be the least upper bound of A. By the trichotomy property, we conclude that $f(c) = 0$. ■

Corollary 2.1 *If f is continuous on a closed interval $[a, b]$, such that*

$$f(a) < \alpha < f(b),$$

then there exists a point $c \in (a, b)$ at which $f(c) = \alpha$.

Proof: Apply the previous theorem for $g(x) = f(x) - \alpha$. ■

(23 hrs, 2013)

✎ *Exercise 2.20* Suppose that f is continuous on $[a, b]$ with $f(x) \in \mathbb{Q}$ for all $x \in [a, b]$. What can be said about f?

✎ *Exercise 2.21* (i) Suppose that $f, g : [a, b] \to \mathbb{R}$ are continuous such that $f(a) < g(a)$ and $f(b) > g(b)$. Show that there exists a point $c \in [a, b]$ such that $f(c) = g(c)$. (ii) A fly flies from Tel-Aviv to Or Akiva (along the shortest path). He starts its way at 10:00 and finishes at 16:00. The following day are makes his way back along the very same path, starting again at 10:00 and finishing at 16:00. Show that there will be a point along the way that he will cross a the exact same time on both days.

✎ *Exercise 2.22* (i) Let $f : [0,1] \to \mathbb{R}$ be continuous, $f(0) = f(1)$. Show that there exists an $x_0 \in [0, \frac{1}{2}]$, such that $f(x_0) = f(x_0 + \frac{1}{2})$. (Hint: consider $g(x) = f(x + \frac{1}{2}) - f(x)$.) (ii) Show that there exists an $x_0 \in [0, 1 - \frac{1}{n}]$, such that $f(x_0) = f(x_0 + \frac{1}{n})$.

Lemma 2.2 If f is continuous at a, then there is a $\delta > 0$ such that f has an upper bound on the interval $(a - \delta, a + \delta)$.

Proof: This is obvious, for by the definition of continuity, there exists a $\delta > 0$, such that

$$f(x) \in B(f(a), 1) \qquad \text{whenever} \qquad x \in B(a, \delta),$$

i.e., $f(x) < f(a) + 1$ on this interval. ■

Theorem 2.7 (Karl Theodor Wilhelm Weierstraß) If f is continuous on $[a, b]$ then it is bounded from above of that interval, that is, there exists a number M, such that $f(x) < M$ for all $x \in [a, b]$.

Comment: Here too, continuity is needed on the whole interval, for look at the "counter example" $f : [-1, 1] \to \mathbb{R}$,

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

Comment: It is crucial that the interval $[a, b]$ be closed. The function $f : (0, 1] \to \mathbb{R}$, $f : x \mapsto 1/x$ is continuous on the semi-open interval, but it is not bounded from above.

Proof: Let

$$A = \{x \in [a, b] : \ f \text{ is bounded from above on } [a, x]\}.$$

By the previous lemma, there exists a $\delta > 0$ such that $a + \delta \in A$. Also A is upper bounded by b, hence there exists a $c \in (a, b]$, such that

$$c = \sup A.$$

We claim that $c = b$, for if $c < b$, then by the previous lemma, there exists a neighborhood of c in which f is upper bounded, and c cannot be an upper bound for A. ■

CORRECT THE PROOF. NEED TO CONNECT TO CONTINUITY AT b.

Comment: In essence, this theorem is based on the fact that if a continuous function is bounded up to a point, then it is bounded up to a little farther. To be able to take such increments up to b we need the axiom of completeness.

(17 hrs, 2009)

(22 hrs, 2010)

(20 hrs, 2011)

Theorem 2.8 (Weierstraß, Maximum principle עקרון המקסימום) If f is continuous on $[a, b]$, then there exists a point $c \in [a, b]$ such that

$$f(x) \le f(c) \quad \textit{for all} \quad x \in [a, b].$$

Comment: Of course, there is a corresponding ***minimum principle***.

Proof: We have just proved that f is upper bounded on $[a, b]$, i.e., the set

$$A = \{f(x) : \; a \le x \le b\}$$

is upper bounded. This set is non-empty for it contains the point $f(a)$. By the axiom of completeness it has a least upper bound, which we denote by

$$\alpha = \sup A.$$

We need to show that this supremum is in fact a maximum; that there exists a point $c \in [a, b]$, for which $f(c) = \alpha$.

Suppose, by contradiction, that this were not the case, i.e., that $f(x) < \alpha$ for all $x \in [a, b]$. We define then a new function $g : [a, b] \to \mathbb{R}$,

$$g = \frac{1}{\alpha - f}.$$

This function is defined everywhere on $[a, b]$ (since we assumed that the denominator does not vanish), it is continuous and positive. We will show that g is not upper bounded on $[a, b]$, contradicting thus the previous theorem. Indeed, since $\alpha = \sup A$:

For all $M > 0$ there exists a $y \in [a, b]$ such that $f(y) > \alpha - 1/M$.

For this y,

$$g(y) > \frac{1}{\alpha - (\alpha - 1/M)} = M,$$

i.e., for every $M > 0$ there exists a point in $[a, b]$ at which f takes a value greater than M. ■

Comment: Here too we needed continuity on a closed interval, for consider the "counter examples"

$$f : [0, 1) \to \mathbb{R} \qquad f(x) = x^2,$$

and

$$f : [0, 1] \to \mathbb{R} \qquad f(x) = \begin{cases} x^2 & x < 1 \\ 0 & x = 1 \end{cases}.$$

These functions do not attain a maximum in $[0, 1]$.

✎ *Exercise 2.23* Let $f : [a, b] \to \mathbb{R}$ be continuous, with $f(a) = f(b) = 0$ and

$$M = \sup\{f(x) : a \le x \le b\} > 0.$$

Prove that for every $0 < c < M$, the set $\{x \in [a, b] : f(x) = c\}$ has at least two elements.

✎ *Exercise 2.24* Let $f : [a, b] \to \mathbb{R}$ be continuous. Prove or disprove each of the following statements:

① It is possible that $f(x) > 0$ for all $x \in [a, b]$, and that for every $\epsilon > 0$ there exists an $x \in [a, b]$ for which $f(x) < \epsilon$.

② There exists an $x_0 \in [a, b]$ such that for every $[a, b] \ni x \neq x_0$, $f(x) < f(x_0)$.

Theorem 2.9 *If $n \in \mathbb{N}$ is odd, then the equation*

$$f(x) = x^n + a_{n-1}x^{n-1} + \dots a_1 x + a_0 = 0$$

has a root (a solution) for any set of constants $a_0, \dots, a_{n-1}$.

Proof: The idea it to show that existence of points a, b for which $f(b) > 0$ and $f(a) < 0$, and apply the intermediate value theorem, based on the fact that polynomials are continuous. The only technical issue is to find such points a, b in a way that works for all choices of $a_0, \dots, a_{n-1}$.

Let

$$M = \max(1, 2n|a_0|, \dots, 2n|a_{n-1}|).$$

Then for $|x| > M$,

$$\begin{aligned}
\frac{f(x)}{x^n} &= 1 + \frac{a_{n-1}}{x} + \dots + \frac{a_0}{x^n} \\
(u + v \geq u - |v|) \qquad &\geq 1 - \frac{|a_{n-1}|}{|x|} - \dots - \frac{|a_0|}{|x^n|} \\
(|x^n| > |x|) \qquad &\geq 1 - \frac{|a_{n-1}|}{|x|} - \dots - \frac{|a_0|}{|x|} \\
(|x| > |a_i|) \qquad &\geq 1 - \frac{|a_{n-1}|}{2n|a_{n-1}|} - \dots - \frac{|a_0|}{2n|a_0|} \\
&= 1 - n \cdot \frac{1}{2n} > 0.
\end{aligned}$$

Since the sign of x^n is the same as the sign of x, it follows that $f(x)$ is positive for $x \geq M$ and negative for $x \leq -M$, hence there exists a $a < c < b$ such that $f(c) = 0$. ■

(18 hrs, 2009)

The next theorem deals with the case where n is even:

Theorem 2.10 *Let n be even. Then the function*

$$f(x) = x^n + a_{n-1}x^{n-1} + \dots a_1 x + a_0 = 0$$

has a minimum. Namely, there exists a $c \in \mathbb{R}$, such that $f(c) \leq f(x)$ for all $x \in \mathbb{R}$.

Proof: The idea is simple. We are going to show that there exists a $b > 0$ for which $f(x) > f(0)$ for all $|x| \geq b$. Then, looking at the function f in the interval $[-b, b]$, we know that it assumes a minimum in this interval, i.e., that there exists a $c \in [-b, b]$ such that $f(c) \leq f(x)$ for all $x \in [-b, b]$. In particular, $f(c) \leq f(0)$, which in turn is less than $f(x)$ for all $x \notin [-b, b]$. It follows that $f(c) \leq f(x)$ for all $x \in \mathbb{R}$.

It remains to find such a b. Let M be defined as in the previous theorem. Then, for all $|x| > M$, using the fact that n is even,

$$
\begin{aligned}
f(x) &= x^n \left(1 + \frac{a_{n-1}}{x} + \cdots + \frac{a_0}{x^n}\right) \\
&\geq x^n \left(1 - \frac{|a_{n-1}|}{|x|} - \cdots - \frac{|a_0|}{|x^n|}\right) \\
&\geq x^n \left(1 - \frac{|a_{n-1}|}{|x|} - \cdots - \frac{|a_0|}{|x|}\right) \\
&\geq x^n \left(1 - \frac{|a_{n-1}|}{2n|a_{n-1}|} - \cdots - \frac{|a_0|}{2n|a_0|}\right) \\
&= \frac{1}{2} x^n.
\end{aligned}
$$

Let then $b > \max(M, \sqrt[n]{2|f(0)|})$, from which follows that $f(x) > f(0)$ for $|x| > b$. This concludes the proof. ■

Corollary 2.2 Consider the equation

$$f(x) = x^n + a_{n-1}x^{n-1} + \ldots a_1 x + a_0 = \alpha$$

with n even. Then there exists a number m such that this equation has a solution for all $\alpha \geq m$ but has no solution for $\alpha < m$.

Proof: According to the previous theorem there exists a $c \in \mathbb{R}$ such that $f(c)$ is the minimum of f. If $\alpha < f(c)$ then for all x, $f(x) \geq f(c) > \alpha$, i.e., there is no solution. For $\alpha = f(c)$, c is a solution. For $\alpha > f(c)$ we have $f(c) < \alpha$ and for large enough x, $f(x) > \alpha$, hence, by the intermediate value theorem a root exists. ■

(25 hrs, 2013)

2.7 Infinite limits and limits at infinity

In this section we extend the notions of limits discussed in previous sections to two cases: (i) the limit of a function at a point is infinite, and (ii) the limit of a function at infinity. Before we start recall: ***infinity is not a real number!***.

Definition 2.10 *Let $f : A \to B$ be a function. We say that the limit of f at a is infinity, denoted*

$$\lim_a f = \infty,$$

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

$$f(x) > M \qquad \text{whenever} \qquad 0 < |x - a| < \delta.$$

Similarly,

$$\lim_a f = -\infty,$$

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

$$f(x) < M \qquad \text{whenever} \qquad 0 < |x - a| < \delta.$$

Example: Consider the function

$$f : x \mapsto \begin{cases} 1/|x| & x \neq 0 \\ 17 & x = 0. \end{cases}$$

The limit of f at zero is infinity. Indeed given M we take $\delta = 1/M$, then

$$0 < |x| < \delta \qquad \text{implies} \qquad f(x) > M.$$

▲▲▲

Definition 2.11 *Let $f : \mathbb{R} \to \mathbb{R}$. We say that*

$$\lim_\infty f = \ell,$$

if for every $\epsilon > 0$ there exists an M, such that

$$|f(x) - \ell| < \epsilon \qquad \text{whenever} \qquad x > M.$$

Similarly,

$$\lim_{-\infty} f = \ell,$$

if for every $\epsilon > 0$ there exists an M, such that

$$|f(x) - \ell| < \epsilon \qquad \text{whenever} \qquad x < M.$$

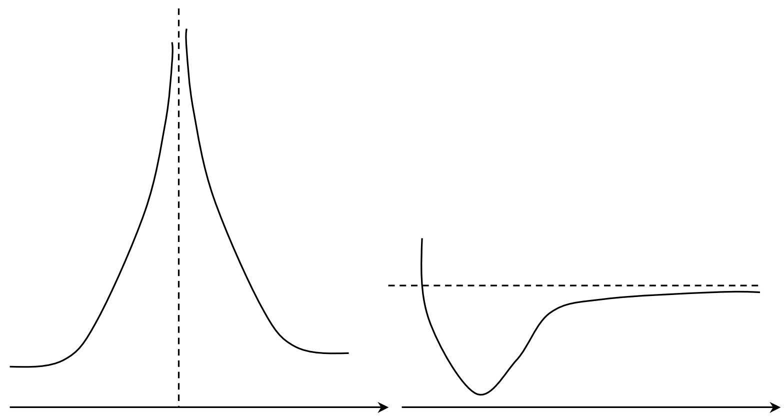


Figure 2.1: Infinite limit and limit at infinity.

Example: Consider the function $f : x \mapsto 3 + 1/x^2$. Then the limit of f at infinity is 3, since given $\epsilon > 0$ we take $M = 1/\sqrt{\epsilon}$, and

$$x > M \qquad \text{implies} \qquad |f(x) - 3| < \epsilon.$$

▲▲▲

Comment: These definitions are in full agreement with all previous definitions of limits, if we adopt the idea that "neighborhoods of infinity" are sets of the form (M, ∞).

✎ *Exercise 2.25* Prove directly from the definition that

$$\lim_{x \to 3^+} \frac{x^3 + 2x + 1}{x - 3} = \infty.$$

✎ *Exercise 2.26* Calculate the following limit using limit arithmetic,

$$\lim_{x \to \infty} \left(\sqrt{x^2 + 2x} - x \right).$$

✎ *Exercise 2.27* Prove that

$$\lim_{x \to \infty} f(x) \quad \text{exists if and only if} \quad \lim_{x \to 0^+} f(1/x) \quad \text{exist,}$$

in which case they are equal.

✎ *Exercise 2.28* Let g be defined in a punctured neighborhood of $a \in \mathbb{R}$, with $\lim_a g = \infty$. Show that $\lim_a 1/g = 0$.

✎ *Exercise 2.29*

① Let f, g be defined in a punctured neighborhood of $a \in \mathbb{R}$, with $\lim_a f = \lim_a g = -\infty$. Show that $\lim_a (f \cdot g) = \infty$.

② Let f, g be defined in a punctured neighborhood of $a \in \mathbb{R}$, with $\lim_a f = K < 0$ and $\lim_a g = -\infty$. Show that $\lim_a (f/g) = 0$.

(24 hrs, 2010)

2.8 Inverse functions

Suppose that $f : A \to B$ is one-to-one and onto. This means that for every $b \in B$ there exists a unique $a \in A$, such that $f(a) = b$. This property defines a function from B to A. In fact, this function is also one-to-one and onto. We call this function the ***function inverse to*** f (פונקציה הופכית), and denote it by f^{-1} (not to be mistaken with $1/f$). Thus, $f^{-1} : B \to A$,

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

[11] Note also that $f^{-1} \circ f$ is the identity in A whereas $f \circ f^{-1}$ is the identity in B.

Comment: Recall that a function $f : A \to B$ can be defined as a subset of $A \times B$,

$$\text{Graph}(f) = \{(a, b) : \ f(a) = b\}.$$

The inverse function can be defined as a subset of $B \times A$: it is all the pairs (b, a) for which $(a, b) \in \text{Graph}(f)$, or

$$\text{Graph}(f^{-1}) = \{(b, a) : \ f(a) = b\}.$$

[11] Strictly speaking, the right-hand side is a set of real numbers, and not a real number. However, the one-to-one and onto properties of f guarantee that it contains one and only one number, hence we identify this singleton (יחידון) with the unique element it contains.

Example: Let $f : [0, 1] \to \mathbb{R}$ be given by $f : x \mapsto x^2$. This function is one-to-one and its image is the segment $[0, 1]$. Thus we can define an inverse function $f^{-1} : [0, 1] \to [0, 1]$ by

$$f^{-1}(y) = \{x \in [0, 1] : \ x^2 = y\},$$

i.e., it is the (positive) square root. ▲▲▲

(22 hrs, 2011)

Theorem 2.11 If $f : I \to \mathbb{R}$ is continuous and one-to-one (I is an interval), then f is either monotonically increasing, or monotonically decreasing.

Comment: Continuity is crucial. The function $f : [0, 2] \to \mathbb{R}$,

$$f(x) = \begin{cases} x & 0 \le x \le 1 \\ 6 - x & 1 < x \le 2 \end{cases}$$

is one-to-one with image $[0, 1] \cup [4, 5)$, but it is not monotonic.

Proof: We first prove that for every three points $a < b < c$, either

$$f(a) < f(b) < f(c), \qquad \text{or} \qquad f(a) > f(b) > f(c).$$

Suppose that $f(a) < f(c)$ and that, by contradiction, $f(b) < f(a)$ (see Figure 2.2). By the intermediate value theorem, there exists a point $x \in (b, c)$ such that $f(x) = f(a)$, contradicting the fact that f is one-to-one. Similarly, if $f(b) > f(c)$, then there exists a point $y \in [a, b]$ such that $f(y) = f(c)$. Thus, $f(a) < f(c)$ implies that $f(a) < f(b) < f(c)$. We proceed similarly if $f(a) > f(c)$.

It follows as once that for every four points $a < b < c < d$, either

$$f(a) < f(b) < f(c) < f(d) \qquad \text{or} \qquad f(a) > f(b) > f(c) > f(d).$$

Fix now a and b, and suppose w.l.o.g that $f(a) < f(b)$ (they can't be equal). Then, for every x, y such that $x < y$, we apply the above arguments to the four points a, b, x, y (whatever their order is) and conclude that $f(x) < f(y)$.

■

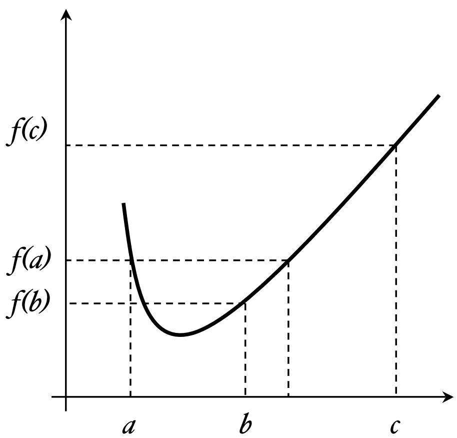


Figure 2.2: Illustration of proof

Suppose that the function $f : [a, b] \to \mathbb{R}$ is one-to-one and monotonically increasing (without loss of generality). Then, its image is the interval $[f(a), f(b)]$. Indeed, $f(a)$ and $f(b)$ are in the range, and by the intermediate value theorem, so is any point in this interval. By monotonicity, no point outside this interval can be in the range of f. It follows that the inverse function has range and image $f^{-1} : [f(a), f(b)] \to [a, b]$. We have proved, *en passant*, that continuous one-to-one functions map closed segments into closed segments[12].

(26 hrs, 2013)

Theorem 2.12 If $f : [a, b] \to \mathbb{R}$ *is continuous and one-to-one then so is* f^{-1}.

Proof: Without loss of generality, let us assume that f is monotonically increasing, i.e.,

$$x < y \qquad \text{implies} \qquad f(x) < f(y).$$

Then, so is f^{-1}. We need to show that for every $y_0 \in (f(a), f(b))$,

$$\lim_{y_0} f^{-1} = f^{-1}(y_0).$$

[12]More generally, a continuous one-to-one function maps connected sets into connected sets, and retains the open/closed properties.

Equivalently, we need to show that for every $\epsilon > 0$ there exists a $\delta > 0$, such that

$$|f^{-1}(y) - f^{-1}(y_0)| < \epsilon \qquad \text{whenever} \qquad |y - y_0| < \delta.$$

Now, every such y_0 is equal to $f(x_0)$ for a unique x_0. Since f (and f^{-1}) are both monotonically increasing, it is clear how to set δ. Let

$$y_1 = f(x_0 - \epsilon) \qquad \text{and} \qquad y_2 = f(x_0 + \epsilon).$$

By monotonicity, $y_1 < y_0 < y_2$. We let then $\delta = \min(y_0 - y_1, y_2 - y_0)$. Every $y_0 - \delta < y < y_0 + \delta$ is in the interval (y_1, y_2), and therefore,

$$f^{-1}(y_0) - \epsilon < f^{-1}(y) < f^{-1}(y_0) + \epsilon.$$

We can do it more formally. Let

$$\delta = \min(f(f^{-1}(y_0) + \epsilon) - y_0, y_0 - f(f^{-1}(y_0) - \epsilon)).$$

Then, $|y - y_0| < \delta$ implies

$$y_0 - \delta < y < y_0 + \delta,$$

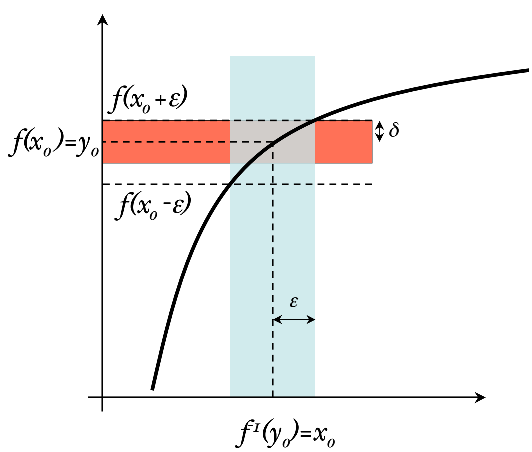
or,

$$f(f^{-1}(y_0)-\epsilon) = y_0-[y_0-f(f^{-1}(y_0)-\epsilon)] < y < y_0+[f(f^{-1}(y_0)+\epsilon)-y_0] = f(f^{-1}(y_0)+\epsilon).$$

Since f^{-1} is monotonically increasing, we may apply f^{-1} on all terms, giving

$$f^{-1}(y_0) - \epsilon < f^{-1}(y) < f^{-1}(y_0) + \epsilon,$$

or $|f^{-1}(y_0) - f^{-1}(y)| < \epsilon$.



■

(20 hrs, 2009)

(26 hrs, 2010)

2.9 Uniform continuity

Recall the definition of a continuous function: a function $f : (a, b) \to \mathbb{R}$ is continuous on the interval (a, b), if it is continuous at every point in the interval. That is, for every $x \in (a, b)$ and every $\epsilon > 0$, there exists a $\delta < 0$, such that

$$|f(y) - f(x)| < \epsilon \qquad \text{whenever} \qquad |y - x| < \delta \quad \text{and} \quad y \in (a, b).$$

In general, we expect δ to depend both on x and on ϵ. In fact, the way we set it here, this definition applies for continuity on a closed interval as well (i.e, it includes one-sided continuity as well).

Let us re-examine two examples:

Example: Consider the function $f : (0, 1) \to \mathbb{R}$, $f : x \mapsto x^2$. This function is continuous on $(0, 1)$. Why? Because for every $x \in (0, 1)$, and every $y \in (0, 1)$,

$$|f(y) - f(x)| = |y^2 - x^2| = |y - x||y + x| \le (x + 1)|y - x|.$$

Thus, given x and $\epsilon > 0$, if we choose $\delta = \delta(\epsilon, x) = \epsilon/(x + 1)$, then

$$|f(y) - f(x)| < \epsilon \qquad \text{whenever} \qquad |x - y| < \delta(\epsilon, x) \quad \text{and} \quad y \in (0, 1).$$

Thus, δ depends both on ϵ and x. However, it is always legitimate to replace δ by a smaller number. If we take $\delta = \epsilon/2$, then the same δ fits all points x. ▲▲▲

Example: Consider next the function $f : (0, 1) \to \mathbb{R}$, $f : x \mapsto 1/x$. This function is also continuous on $(0, 1)$. Why? Given $x \in (0, 1)$ and $y \in (0, 1)$,

$$|f(x) - f(y)| = \frac{|x - y|}{xy} = \frac{|x - y|}{x[x + (y - x)]}.$$

If we take $\delta(\epsilon, x) = \min(x/2, \epsilon x^2/2)$, then

$$|x - y| < \delta \qquad \text{implies} \qquad |f(x) - f(y)| < \frac{\delta}{x(x - \delta)} < \epsilon.$$

Here, given ϵ, the closer x is to zero, the smaller is $\delta(\epsilon, x)$. There is no way we can find a $\delta = \delta(\epsilon)$ that would fit all x. ▲▲▲

These two examples motivate the following definition:

Definition 2.12 *$f : A \to B$ is said to be* **uniformly continuous** *on A (*רציפה במידה שווה*) if for every $\epsilon > 0$ corresponds a $\delta = \delta(\epsilon) > 0$, such that*

$$|f(y) - f(x)| < \epsilon \qquad \textit{whenever} \qquad |y - x| < \delta \quad \textit{and} \quad x, y \in A.$$

Note that x and y play here symmetric roles[13].

What is the essential difference between the two above examples? The function $x \mapsto x^2$ could have been defined on the closed interval $[0, 1]$ and it would have been continuous there too. In contrast, there is no way we could have defined the function $x \mapsto 1/x$ as a continuous function on $[0, 1]$, even if we took care of its value at zero. We have already seen examples where continuity on a closed interval had strong implications (ensures boundedness and the existence of a maximum). This is also the case here. We will prove that *continuity on a closed interval implies uniform continuity*.

We proceed by steps:

Lemma 2.3 *Let $f : [a, c] \to \mathbb{R}$ be continuous. Let $a < b < c$, and suppose that for a given $\epsilon > 0$ there exist $\delta_1, \delta_2 > 0$, such that*

$$\begin{aligned} x, y \in [a, b] \quad &\textit{and} \quad |x - y| < \delta_1 \quad \textit{implies} \quad |f(x) - f(y)| < \epsilon \\ x, y \in [b, c] \quad &\textit{and} \quad |x - y| < \delta_2 \quad \textit{implies} \quad |f(x) - f(y)| < \epsilon. \end{aligned}$$

Then there exists a $\delta > 0$ such that

$$x, y \in [a, c] \quad \textit{and} \quad |x - y| < \delta \quad \textit{implies} \quad |f(x) - f(y)| < \epsilon.$$

Proof: We have to take care at the "gluing" at the point b. Since the function is continuous at b, then there exists a $\delta_3 > 0$ such that

$$|f(x) - f(b)| < \frac{\epsilon}{2} \qquad \text{whenever} \qquad |x - b| < \delta_3.$$

[13] The term "uniform" means that the same number can be used for all points.

Hence,

$$|x-b|<\delta_3 \quad \text{and} \quad |y-b|<\delta_3 \qquad \text{implies} \qquad |f(x)-f(y)|<\epsilon. \tag{2.1}$$

Take now $\delta = \min(\delta_1, \delta_2, \delta_3)$. We claim that

$$|f(x)-f(y)|<\epsilon \qquad \text{whenever} \qquad |x-y|<\delta \quad \text{and} \quad x,y \in [a,c].$$

Why is that true? One of three: either $x, y \in [a,b]$, in which case it follows from the fact that $\delta \le \delta_1$, or $x, y \in [b,c]$, in which case it follows from the fact that $\delta \le \delta_2$. Remains the third case, where, without loss of generality, $x \in [a,b]$ and $y \in [b,c]$. Since $|x-y|<\delta_3$, it follows that $|x-b|<\delta_3$ and $|y-b|<\delta_3$ and we use (2.1). ■

(24 hrs, 2011)

Theorem 2.13 If f is continuous on $[a,b]$ then it is uniformly continuous on that interval.

Proof: Given $\epsilon > 0$, we will say that f is ϵ-good on $[a,b]$ if there exists a $\delta > 0$, such that

$$|f(x)-f(y)|<\epsilon \qquad \text{whenever} \qquad |x-y|<\delta \quad \text{and} \quad x,y \in [a,b]$$

(i.e., if for that particular ϵ there exists a corresponding δ). We need to show that f is ϵ-good on $[a,b]$ for all $\epsilon > 0$. Note that we proved in Lemma 2.3 that if f is ϵ-good on $[a,b]$ and on $[b,c]$, then it is ϵ-good on $[a,c]$.

Given ϵ, define

$$A = \{x \in [a,b] : \ f \text{ is } \epsilon\text{-good on } [a,x]\}.$$

The set A is non-empty, as it includes the point a (trivial) and it is bounded by b. Hence it has a supremum, which we denote by α. Suppose that $\alpha < b$. Since f is continuous at α, then there exists a $\delta_0 > 0$ such that

$$|f(y)-f(\alpha)|<\frac{\epsilon}{2} \qquad \text{whenever} \qquad |y-\alpha|<\delta_0 \quad \text{and} \quad y \in [a,b].$$

Thus,

$$\forall y,z \in B(\alpha,\delta_0), \quad |f(y)-f(z)| \le |f(y)-f(\alpha)| + |f(z)-f(\alpha)| < \epsilon,$$

from which follows that f is ϵ-good in $(\alpha-\delta_0, \alpha+\delta_0)\cap[a,b]$ and hence in the closed interval $[\alpha - \delta_0/2, \alpha + \delta_0/2] \cap [a,b]$. By the definition of α, it is also ϵ-good on $[a, \alpha-\delta_0/2]\cap[a,b]$, and by the previous lemma, f is ϵ-good on $[a, \alpha+\delta_0/2]\cap[a,b]$, contradicting the fact that α is an upper bound for A. Thus f is ϵ-good on $[a,b]$. Since this argument holds independently of ϵ we conclude that f is ϵ-good in $[a,b]$ for all $\epsilon > 0$. ■

(28 hrs, 2013)

Comment: In this course we will make use of uniform continuity only once, when we study integration.

(27 hrs, 2010)

Examples:

① Consider the function $f(x) = \sin(1/x)$ defined on $(0,1)$. Even though it is continuous, it is not uniformly continuous.

② Consider the function $\mathrm{Id} : \mathbb{R} \to \mathbb{R}$. It is easy to see that it is uniformly continuous. The function $\mathrm{Id}\cdot\mathrm{Id}$, however, is continuous on $\mathbb{R}$, but not uniformly continuous. Thus, the product of uniformly continuous functions is not necessarily uniformly continuous.

③ Finally the function $f(x) = \sin x^2$ is continuous and bounded on $\mathbb{R}$, but it is not uniformly continuous.

(22 hrs, 2009)

✎ *Exercise 2.30* Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. Prove or disprove:

① If
$$\lim_{\infty} f = \lim_{-\infty} f = 0,$$
then f attains a maximum value.

② If $\lim_\infty f = \infty$ and $\lim_{-\infty} f = -\infty$, then f assumes all real values; that is f is on $\mathbb{R}$.

③ If on any interval $[n, n+1]$, $n \in \mathbb{Z}$, f is uniformly continuous, then f is uniformly continuous on $\mathbb{R}$.

✎ *Exercise 2.31* Suppose $A \subseteq \mathbb{R}$ is non-empty.

① Prove that if f, g are uniformly continuous on A then $\alpha f + \beta g$ is also uniformly continuous on A, where $\alpha, \beta \in \mathbb{R}$.

② Is in the above case $f \cdot g$ necessarily uniformly continuous?

✎ *Exercise 2.32* Let $f : (0, \infty) \to \mathbb{R}$ be given by $f : x \mapsto 1/\sqrt{x}$:

① Let $a > 0$. Prove that f is uniformly continuous on $[a, \infty)$.

② Show that f is not uniformly continuous on $(0, \infty)$.

✎ *Exercise 2.33* Let I be an open connected domain (i.e., an open segment, an open ray, or the whole line) and let $f, g : I \to \mathbb{R}$ be uniformly continuous on I. Prove that $f \cdot g$ is uniformly continuous on I. (Hint: read the proof about the limit of a product of functions being the product of their limits.)

✎ *Exercise 2.34* Let $f : A \to B$ and $g : B \to C$ be uniformly continuous. Show that $g \circ f$ is uniformly continuous.

✎ *Exercise 2.35* Let $a > 0$. Show that $f(x) = \sin(1/x)$ is not uniformly continuous on $(0, a)$.

✎ *Exercise 2.36* For which values of $\alpha > 0$ is the function $f(x) = x^\alpha$ uniformly continuous on $[0, \infty)$?

✎ *Exercise 2.37* Let $\alpha > 0$. We say that $f : A \to B$ is Hölder continuous of order α if there exists a constant $M > 0$, such that

$$|f(x) - f(y)| < M|x - y|^\alpha, \qquad \forall x, y \in A.$$

① Show that $f(x) = \sqrt{x}$ is Hölder continuous of order $1/2$ on $(0, \infty)$.

② Prove that if f is Hölder continuous of order α on A then it is uniformly continuous on A.

③ Prove that if f is Hölder continuous of order $\alpha > 1$, then it is a constant function.

Chapter 3

Derivatives

3.1 Definition

In this chapter we define and study the notion of ***differentiable functions***. Derivatives were historically introduced in order to answer the need of measuring the "rate of change" of a function. We will actually adopt this approach as a prelude to the formal definition of the derivative.

Before we start, one technical clarification. In the previous chapter, we introduced the limit of a function $f : A \to B$ at an (interior) point a,

$$\lim_a f.$$

Consider the function g defined by

$$g(x) = f(a + x),$$

on the range,

$$\{x : \; a + x \in A\}.$$

In particular, zero is an interior point of this set. We claim that

$$\lim_0 g = \lim_a f,$$

provided, of course, that the right hand side exists. It is a good exercise to prove it formally. Suppose that the right hand side equals ℓ. This means that

$$(\forall \epsilon > 0)(\exists \delta > 0) : (\forall x \in B^\circ(a, \delta))(|f(x) - \ell| < \epsilon).$$

Setting $x - a = y$,

$$(\forall \epsilon > 0)(\exists \delta > 0) : (\forall y \in B^\circ(0, \delta))(|f(a + y) - \ell| < \epsilon),$$

and by the definition of g,

$$(\forall \epsilon > 0)(\exists \delta > 0) : (\forall y \in B^\circ(0, \delta))(|g(y) - \ell| < \epsilon),$$

i.e., $\lim_0 g = \ell$. Often, one does not bother to define the function g and simply writes

$$\lim_{x \to 0} f(a + x) = \lim_a f,$$

although this in not in agreement with our notations; $f(a + x)$ stands for the composite function,

$$f \circ (x \mapsto a + x).$$

Let now f be a function defined on some interval I, and let a and x be two points inside this interval. The variation in the value of f between these two points is $f(x) - f(a)$. We define the ***mean rate of change*** (קצב שינוי ממוצע) of f between the points a and x to be the ratio

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

This quantity can be attributed with a number of interpretations. First, if we consider the graph of f, then the mean rate of change is the slope of the secant line (מיתר) that intersects the graph at the points a and x (see Figure 3.1). Second, it has a meaning in many physical situations. For example, f can be the position along a line (say, in meters relative to the origin) as function of time (say, in seconds relative to an origin of time). Thus, $f(a)$ is the distance from the origin in meters a seconds after the time origin, and $f(x)$ is the distance from the origin in meters x seconds after the time origin. Then $f(x) - f(a)$ is the ***displacement*** (העתק) between time a and time x, and $[f(x) - f(a)]/(x - a)$ is the mean displacement per unit time, or the ***mean velocity***.

This physical example is a good preliminary toward the definition of the derivative. The mean velocity, or mean rate of displacement, is an average quantity between two instants, but there is nothing to guarantee that within this time interval, the body was in a "fixed state". For example, we could intersect this time interval into two equal sub-intervals, and measure the mean velocity in each half. Nothing guarantees that these mean velocities will equal the mean velocity over

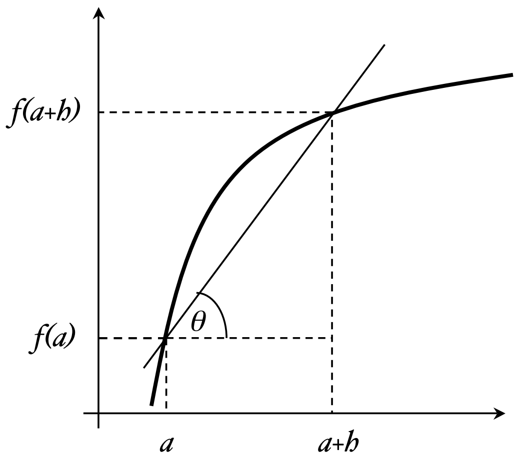


Figure 3.1: Relation between the mean rate of change of a function and the slope of the corresponding secant.

the whole interval. Physicists aimed to define an “instantaneous velocity”, and the way to do it was to make the length of the interval very small. Of course, having h small (what does small mean?) does not change the average nature of the measured velocity.

An instantaneous rate of change can be defined by using limits. Fixing the point a, we define a function

$$\Delta_{f,a} : x \mapsto \frac{f(x) - f(a)}{x - a}.$$

The function $\Delta_{f,a}$ can be defined on the same domain as f, but excluding the point a. We say that f ***is differentiable at*** a (גזירה או דיפרנציאבילית) if $\Delta_{f,a}$ has a limit at a. We write

$$\lim_{a} \Delta_{f,a} = f'(a).$$

We call the limit $f'(a)$ the ***derivative*** (נגזרת) of f at a.

Comment: In the more traditional notation,

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

At this stage, $f'(a)$ is nothing but a notation. It is not (yet!) a function evaluated at the point a. Another standard notation, due to Leibniz, is

$$\frac{df}{dx}(a).$$

Having identified the mean rate of change as the slope of the secant line, the derivative has a simple geometrical interpretation. As x tends to a, the corresponding family of secants tends to a line which is ***tangent*** to f at the point a. For us, these are only hand-waving arguments, as we have never assigned any meaning to the limit of a family of lines.

Example: Consider the constant function, $f : \mathbb{R} \to \mathbb{R}$, $f : x \mapsto c$. We are going to calculate its derivative at the point a. We construct the function

$$\Delta_{f,a}(y) = \frac{f(y) - f(a)}{y - a} = \frac{c - c}{y - a} = 0.$$

The limit of this function at a is clearly zero, hence $f'(a) = 0$. ▲▲▲

In the above example, we could have computed the derivative at *any* point. In other words, we can define a *function* that given a point x, returns the derivative of f at that point, i.e., returns $f'(x)$. A function $f : A \to B$ is called ***differentiable*** in a subset $U \subseteq A$ of its domain if it has a derivative at every point $x \in U$. We then define the derivative function, $f' : U \to \mathbb{R}$, as the function,

$$f'(x) = \lim_{x} \Delta_{f,x}.$$

Again, the more traditional notation is,

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

Example: Consider the function $f : x \mapsto x^2$. We calculate its derivative at a point x by first observing that

$$\Delta_{f,x}(y) = \frac{y^2 - x^2}{y - x} = x + y \qquad \text{or} \qquad \Delta_{f,x} = \mathrm{Id} + x,$$

so that

$$f'(x) = \lim_{x} \Delta_{f,x} = 2x.$$

▲▲▲

(23 hrs, 2009)

(25 hrs, 2011)

Example: Consider now the function $f : x \mapsto |x|$. Let's first calculate its derivative at a point $x > 0$. For $x > 0$,

$$\Delta_{f,x}(y) = \frac{|y| - x}{y - x}.$$

Since we are interested in the limit of $\Delta_{f,x}$ at $x > 0$ we may well assume that $y > 0$, in which case $\Delta_{f,x}(y) = 1$. Hence

$$f'(x) = \lim_x \Delta_{f,x} = 1.$$

For negative x, $|x| = -x$ and we may consider $y < 0$ as well, so that

$$\Delta_{f,x}(y) = \frac{(-y) - (-x)}{y - a} = -1,$$

so that

$$f'(x) = \lim_x \Delta_{f,x} = -1.$$

Remains the point 0 itself,

$$\Delta_{f,0}(y) = \frac{|y| - |0|}{y - 0} = \operatorname{sgn}(y).$$

The limit at zero does not exist since every neighborhood of zero has points where this function equals one and points where this function equals minus one. Thus f is differentiable everywhere except for the origin[1]. ▲▲▲

Comment: If f is differentiable at a, then $\Delta_{f,a}$ has a limit at a, hence it has a removable discontinuity at that point. The function

$$x \mapsto \begin{cases} \Delta_{f,a}(x) & x \neq a \\ f'(a) & x = a \end{cases}$$

is continuous.

(29 hrs, 2010)

[1] We use here a general principle, whereby $\lim_a f$ does not exist if there exist $\ell_1 \neq \ell_2$, such that every neighborhood of a has points x, y, such that $f(x) = \ell_1$ and $f(y) = \ell_2$.

One-sided differentiability This last example shows a case where a limit does not exist, but one-sided limits do exist. This motivates the following definitions:

Definition 3.1 A function f is said to be ***differentiable on the right*** *(גזירה מימין) at a if the one-sided limit*

$$\lim_{a^+} \Delta_{f,a}$$

exists. We denote the ***right-hand derivative*** *(נגזרת ימנית) by $f'(a^+)$. A similar definition holds for left-hand derivatives.*

Example: Here is one more example that practices the calculation of the derivative directly from the definition. Consider first the function

$$f : x \mapsto \begin{cases} x \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0. \end{cases}$$

We have already seen that that this function is continuous at zero, but is it differentiable at zero? We construct the function

$$\Delta_{f,0}(y) = \frac{f(y) - f(0)}{y - 0} = \sin \frac{1}{y}.$$

The function $\Delta_{f,0}$ does not have a limit at 0, because every neighborhood of zero has both points where $\Delta_{f,0} = 1$ and points where $\Delta_{f,0} = -1$. ▲▲▲

Example: In contrast, consider the function

$$f : x \mapsto \begin{cases} x^2 \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0. \end{cases}$$

We construct

$$\Delta_{f,0}(y) = \frac{f(y) - f(0)}{y - 0} = y \sin \frac{1}{y}.$$

Now $\lim_0 \Delta_{f,0} = 0$, hence f is differentiable at zero. ▲▲▲

Differentiability and continuity The following theorem shows that differentiable functions are a subclass ("better behaved") of the continuous functions.

Theorem 3.1 If f is differentiable at a then it is continuous at a.

Proof: Note that for $y \neq a$,

$$f(y) = f(a) + (y - a) \cdot \frac{f(y) - f(a)}{(y - a)}.$$

Viewing $f(a)$ as a constant, we have a functional identity,

$$f = f(a) + (\mathrm{Id} - a)\, \Delta_{f,a}$$

valid in a punctured neighborhood of a. By limit arithmetic,

$$\lim_a f = f(a) + \lim_a (\mathrm{Id} - a) \lim_a \Delta_{f,a} = f(a) + 0 \cdot f'(a) = f(a).$$

■

✎ *Exercise 3.1* Show, based on the definition of the derivative, that if f is differentiable at $a \in \mathbb{R}$ with $f(x) \neq 0$, then $|f|$ is also differentiable at that point. (Hint, use the fact that a function differentiable at a point is also continuous at that point.)

✎ *Exercise 3.2* Let I be a connected domain, and suppose that f is differentiable at every interior point of I and continuous on I. Suppose furthermore that $|f'(x)| < M$ for all $x \in I$. Show that f is uniformly continuous on I.

Higher order derivatives If a function f is differentiable on an interval A, we can construct a new function—its derivative, f'. The derivative f' can have various properties, For example, it may be continuous, or not, and in particular, it may be differentiable on A, or on a subset of A. Then, we can define a new function—the derivative of the derivative, or the ***second derivative*** (נגזרת שניה), which we denote by f''. By definition

$$f''(x) = \lim_x \Delta_{f',x}.$$

(Note that this is a limit of limits.) Likewise, the second derivative may be differentiable, in which case we may define the third derivative f''', and so on. For derivatives higher than the third, it is customary to use, for example, the notation $f^{(4)}$ rather than f''''. The k-th derivative, $f^{(k)}$, is defined recursively,

$$f^{(k)}(x) = \lim_{x} \Delta_{f^{(k-1)},x}.$$

For the recursion to hold from $k = 1$, we also set $f^{(0)} = f$.

(30 hrs, 2013)

3.2 Rules of differentiation

Recall that when we studied limits, we first calculated limits by using the definition of the limit, but very soon this became impractical, and we proved a number of theorems (arithmetic of limits), with which we were able to easily calculate a large variety of limits. The same exactly will happen with derivatives. After having calculated the derivatives of a small number of functions, we develop tools that will enable us to (easily) compute derivatives, without having to go back to the definitions.

Theorem 3.2 *If f is a constant function, then $f' = 0$ (by this we mean that $f' : x \mapsto 0$).*

Proof: We proved this is in the previous section. ■

Theorem 3.3 *If $f = \mathrm{Id}$ then $f' = 1$ (by this we mean that $f' : x \mapsto 1$).*

Proof: Immediate from the definition. ■

Theorem 3.4 *Let f, g be functions defined on the same domain (it suffices that the domains have a non-empty intersection). If both f and g are differentiable at a then $f + g$ is differentiable at a, and*

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

Proof: We consider the function

$$\begin{aligned}\Delta_{f+g,a} &= \frac{(f+g)-(f+g)(a)}{\operatorname{Id}-a} \\ &= \frac{f+g-f(a)-g(a)}{\operatorname{Id}-a} \\ &= \Delta_{f,a}+\Delta_{g,a},\end{aligned}$$

and by the arithmetic of limits,

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

■

Theorem 3.5 (Leibniz rule) Let f, g be functions defined on the same domain. If both f and g are differentiable at a then $f \cdot g$ is differentiable at a, and

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

Proof: We consider the function,

$$\Delta_{fg,a} = \frac{fg-(fg)(a)}{\operatorname{Id}-a} = \frac{fg-fg(a)+fg(a)-f(a)g(a)}{\operatorname{Id}-a} = f\Delta_{g,a} + g(a)\Delta_{f,a},$$

and it remains to apply the arithmetic rules of limits. ■

Corollary 3.1 The derivative is a linear operator,

$$(\alpha f + \beta g)' = \alpha f' + \beta g'.$$

Proof: We only need to verify that if f is differentiable at a then so is αf and $(\alpha f)' = \alpha f'$. This follows from the derivative of the product. ■

(31 hrs, 2013)

Example: Let $n \in \mathbb{N}$ and consider the functions $f_n : x \mapsto x^n$. Then,

$$f_n'(x) = n\, x^{n-1} \qquad \text{or equivalently} \qquad f_n' = n\, f_{n-1}.$$

We can show this inductively. We know already that this is true for $n = 0, 1$. Suppose this were true for $n = k$. Then, $f_{k+1} = f_k \cdot \mathrm{Id}$, and by the differentiation rule for products,

$$f_{k+1}' = f_k' \cdot \mathrm{Id} + f_k \cdot \mathrm{Id}',$$

i.e.,

$$f_{k+1}'(x) = k\, x^{k-1} \cdot x + x_k \cdot 1 = (k+1)\, x^k.$$

▲▲▲

Theorem 3.6 If g is differentiable at a and $g(a) \neq 0$, then $1/g$ is differentiable at a and

$$(1/g)'(a) = -\frac{g'(a)}{g^2(a)}.$$

Proof: Since g is continuous at a (it is continuous since it is differentiable) and it is non-zero, then there exists a neighborhood U of a in which g does not vanish (by Theorem 2.5). In $U \setminus \{a\}$ we consider the function

$$\Delta_{1/g,a} = \frac{1/g - (1/g)(a)}{\mathrm{Id} - a} = \frac{1}{\mathrm{Id} - a}\left(\frac{1}{g} - \frac{1}{g(a)}\right) = \frac{-1}{g \cdot g(a)} \cdot \Delta_{g,a}.$$

It remains to take the limit at a and apply the arithmetic laws of limits. ■

(31 hrs, 2010)

Theorem 3.7 If f and g are differentiable at a and $g(a) \neq 0$, then the function f/g is differentiable at a and

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

Proof: Apply the last two theorems. ■

(25 hrs, 2009)

With these theorems in hand, we can differentiate a large number of functions. In particular there is no difficulty in differentiating a product of more than two functions. For example,

$$(fgh)' = ((fg)h)' = (fg)'h + (fg)h' = (f'g + fg')h = (fg)h' = f'gh + fg'h + fgh'.$$

Suppose we take for granted that

$$\sin' = \cos \qquad \text{and} \qquad \cos' = -\sin.$$

Then we have no problem calculating the derivative of, say, $x \mapsto \sin^3 x$, as a product of three functions. But what about the derivative of $x \mapsto \sin x^3$? Here we need a rule for how to differentiate compositions.

Consider the composite function $g \circ f$, i.e.,

$$(g \circ f)(x) = g(f(x)).$$

Let's try to calculate its derivative at a point a, assuming for the moment that both f and g are differentiable everywhere. We then need to look at the function

$$\Delta_{g\circ f,a}(y) = \frac{(g \circ f)(y) - (g \circ f)(a)}{y - a} = \frac{g(f(y)) - g(f(a))}{y - a},$$

and calculate its limit at a. When y is close to a, we expect the arguments $f(y)$ and $f(a)$ of g to be very close. This suggest the following treatment,

$$\Delta_{g\circ f,a}(y) = \frac{g(f(y)) - g(f(a))}{f(y) - f(a)} \cdot \frac{f(y) - f(a)}{y - a} = \frac{g(f(y)) - g(f(a))}{f(y) - f(a)} \Delta_{f,a}(y).$$

It looks that as $y \to a$, since $f(y) - f(a) \to 0$, this product tends to $g'(f(a)) \cdot f'(a)$. The problem is that while the limit $y \to a$ means that the case $y = a$ is not to be considered, there is nothing to prevent the denominator $f(y) - f(a)$ from vanishing, rendering this expression meaningless. Yet, the result is correct, and it only takes a little more subtlety to prove it.

(27 hrs, 2011)

Theorem 3.8 *Let $f : A \to B$ and $g : B \to \mathbb{R}$. If f is differentiable at $a \in A$, and g is differentiable at $f(a)$, then*

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

Proof: We introduce the following function, defined in a neighborhood of $f(a)$,

$$\psi : z \mapsto \begin{cases} \Delta_{g,f(a)}(z) & z \neq f(a) \\ g'(f(a)) & z = f(a). \end{cases}$$

The fact that g is differentiable at $f(a)$ implies that ψ is continuous at $f(a)$.

We next claim that for $y \neq a$

$$\Delta_{g\circ f,a} = (\psi \circ f) \cdot \Delta_{f,a}.$$

Why that? If $f(y) \neq f(a)$, then this equation reads

$$\frac{g(f(y)) - g(f(a))}{y - a} = \frac{g(f(y)) - g(f(a))}{f(y) - f(a)} \cdot \frac{f(y) - f(a)}{y - a},$$

which holds indeed, whereas if $f(y) = f(a)$ it reads,

$$\frac{g(f(y)) - g(f(a))}{y - a} = g'(f(a)) \cdot \frac{f(y) - f(a)}{y - a},$$

and both sides are zero. Consider now the limit of the right hand side at a. By limits arithmetic,

$$\lim_a \left[(\psi \circ f) \cdot \Delta_{f,a}\right] = \psi(f(a)) \cdot \lim_a \Delta_{f,a} = g'(f(a)) \cdot f'(a).$$

■

(32 hrs, 2010)

3.3 Another look at derivatives

In this short section we provide another (equivalent) characterization of differentiability and the derivative. Its purpose is to offer a slightly different angle of view on the subject, and how clever definitions can sometimes greatly simplify proofs.

Our definition of derivatives states that a function f is differentiable at an interior point a, if the function $\Delta_{f,a}$ has a limit at a, and we denote this limit by $f'(a)$. The function $\Delta_{f,a}$ is not defined at a, hence not continuous at a, but this is a removable discontinuity. We could say that f is differentiable at a if there exists a real number, ℓ, such that the function

$$\mathcal{S}_{f,x}(y) = \begin{cases} \Delta_{f,a}(y) & y \neq a \\ \ell & y = a \end{cases}$$

is continuous at a, and $\mathcal{S}_{f,a}(a) = \ell$ is called the derivative of f at a. For $y \neq a$,

$$\mathcal{S}_{f,a}(y) = \Delta_{f,a}(y) = \frac{f(x) - f(a)}{x - a},$$

or equivalently,

$$f(y) = f(a) + \mathcal{S}_{f,a}(y)(y - a).$$

This equation holds also for $y = a$. This suggests the following alternative definition of the derivative:

Definition 3.2 *A function f defined in on open neighborhood of a point a is said to be differentiable at a if there exists a function $\mathcal{S}_{f,a}$ continuous at a, such that*

$$f(y) = f(a) + \mathcal{S}_{f,a}(y)(y - a).$$

$\mathcal{S}_{f,a}(a)$ is called the derivative of f at a and is denoted by $f'(a)$.

Of course, $\mathcal{S}_{f,a}$ coincides with $\Delta_{f,a}$ in some punctured neighborhood of a.

A first example Take the function $f : x \to x^2$. Then

$$f(y) - f(a) = (y + a)(y - a),$$

or

$$f = f(a) + (\mathrm{Id} + a)(\mathrm{Id} - a).$$

Since the function $\mathrm{Id} + a$ is continuous at a, it follows that f is differentiable at a and

$$f'(a) = \lim_a (\mathrm{Id} + a) = 2a.$$

Leibniz' rule Let's now see how this alternative definition simplifies certain proofs. Suppose for example that both f and g are differentiable at a. This implies the existence of two continuous functions, $\mathcal{S}_{f,a}$ and $\mathcal{S}_{g,a}$, such that

$$
\begin{aligned}
f(y) &= f(a) + \mathcal{S}_{f,a}(y)(y-a) \\
g(y) &= g(a) + \mathcal{S}_{g,a}(y)(y-a)
\end{aligned}
$$

in some neighborhood of a and $f'(a) = \mathcal{S}_{f,a}(a)$ and $g'(a) = \mathcal{S}_{g,a}(a)$. Then,

$$
\begin{aligned}
f(y)g(y) &= \big(f(a) + \mathcal{S}_{f,a}(y)(y-a)\big)\big(g(a) + \mathcal{S}_{g,a}(y)(y-a)\big) \\
&= f(a)g(a) + \big(f(a)\mathcal{S}_{g,a}(y) + g(a)\,\mathcal{S}_{f,a}(y) + \mathcal{S}_{f,a}(y)\mathcal{S}_{g,a}(y)(y-a)\big)\,(y-a).
\end{aligned}
$$

By limit arithmetic, the function in the brackets,

$$
F = f(a)\mathcal{S}_{g,a} + g(a)\,\mathcal{S}_{f,a} + \mathcal{S}_{f,a}\mathcal{S}_{g,a}(\mathrm{Id} - a)
$$

is continuous at a, hence fg is differentiable at a and

$$
(fg)'(a) = \lim_a F = f(a)g'(a) + g(a)f'(a).
$$

Derivative of a composition Suppose now that f is differentiable at a and g is differentiable at $f(a)$. Then, there exist continuous functions, $\mathcal{S}_{f,a}$ and $\mathcal{S}_{g,f(a)}$, such that

$$
\begin{aligned}
f(y) &= f(a) + \mathcal{S}_{f,a}(y)(y-a) \\
g(z) &= g(f(a)) + \mathcal{S}_{g,f(a)}(z)(z - f(a)).
\end{aligned}
$$

Now,

$$
\begin{aligned}
g(f(y)) &= g(f(a)) + \mathcal{S}_{g,f(a)}(f(y))(f(y) - f(a)) \\
&= g(f(a)) + \mathcal{S}_{g,f(a)}(f(y))\mathcal{S}_{f,a}(y)(y-a),
\end{aligned}
$$

or,

$$
(g \circ f)(y) = (g \circ f)(a) + \underbrace{\mathcal{S}_{g,f(a)}(f(y))\mathcal{S}_{f,a}(y)}_{[\mathcal{S}_{g,f(a)} \circ f]\cdot \mathcal{S}_{f,a}(y)}(y-a).
$$

By the properties of continuous functions, the function in the square brackets is continuous at a, and

$$
(g \circ f)'(a) = \lim_a \left[\mathcal{S}_{g,f(a)} \circ f\right] \cdot \mathcal{S}_{f,a} = g'(f(a))\,f'(a).
$$

(34 hrs, 2010)

(28 hrs, 2011)

3.4 The derivative and extrema

In high-school calculus, one of the main uses of differential calculus is to find extrema of functions. We are going to put this practice on solid grounds. First, recall the definition,

Definition 3.3 *Let $f : A \to B$. A point $a \in A$ (not necessarily an internal point) is said to be a* ***maximum point*** *of f in A, if*

$$f(x) \le f(a) \qquad \forall x \in A.$$

We define similarly a minimum point.

Comment: By no means a maximum point has to be unique, nor to exist. For example, in a constant function all points are minima and maxima. On the other hand, the function $f : (0,1) \to \mathbb{R}$, $f : x \mapsto x^2$ does not have a maximum in $(0,1)$. Also, we proved that a continuous function defined on a closed interval always has a maximum (and a minumum, both, not necessarily unique).

Here is a first connection between maximum (and minimum) points and derivatives:

Theorem 3.9 *Let $f : (a,b) \to \mathbb{R}$. If $x \in (a,b)$ is a maximum point of f and f is differentiable at x, then $f'(x) = 0$.*

Proof: Since x is, by assumption, a maximum point, then for every $y \in (a,b)$,

$$f(y) - f(x) \le 0.$$

In particular, for $y \ne x$,

$$\Delta_{f,x}(y) = \frac{f(y) - f(x)}{y - x},$$

satisfies

$$\begin{cases} \Delta_{f,x}(y) \le 0 & y > x \\ \Delta_{f,x}(y) \ge 0 & y < x. \end{cases}$$

It is given that $\Delta_{f,x}$ has a limit at x (it is $f'(x)$). We will show that this limit is necessarily zero.

Suppose the limit, ℓ, were positive. Then, taking $\epsilon = \ell/2$, there has to be a $\delta > 0$, such that

$$0 < \frac{\ell}{2} < \Delta_{f,x}(y) < \frac{3\ell}{2} \qquad \text{whenever} \qquad 0 < |y - x| < \delta,$$

but this can't hold, since every neighborhood of x has points y where $\Delta_{f,x}(y) \leq 0$. Similarly, ℓ can't be negative, hence $\ell = 0$. ■

(33 hrs, 2013)

Comment: This is a uni-directional theorem. It does not imply that if $f'(x) = 0$ then x is a maximum point (nor a minimum point).

Comment: The open interval cannot be replaced by the closed interval $[a, b]$, because at the end points we can only consider one-sided limits.

Definition 3.4 *Let $f : A \to B$. An interior point $a \in A$ is called a* ***local maximum*** *of f (מקסימום מקומי), if there exists a $\delta > 0$ such that a is a maximum of f in $(a - \delta, a + \delta)$ ("local" always means "in a sufficiently small neighborhood").*

Theorem 3.10 (Pierre de Fermat) If a is a local maximum (or minimum) of f in some open interval and f is differentiable at a then $f'(a) = 0$.

Proof: There is actually nothing to prove, as the previous theorem applies verbatim for f restricted to the interval $(a - \delta, a + \delta)$. ■

Comment: The converse is not true. Take for example the function $f : x \to x^3$. If is differentiable in $\mathbb{R}$ and $f'(0) = 0$, but zero is not a local minimum, nor a local maximum of f.

Comment: Although this theorem is due to Fermat, its proof is slightly shorter than the proof of Fermat's last theorem.

Definition 3.5 *Let $f : A \to B$ be differentiable. A point $a \in A$ is called a* ***critical point*** *of f if $f'(a) = 0$. The value $f(a)$ is called a* ***critical value****.*

Locating extremal points Let $f : [a, b] \to \mathbb{R}$ and suppose we want to find a maximum point of f. If f is continuous, then a maximum is guaranteed to exist. There are three types of candidates: (1) critical points of f in (a, b), (ii) the end points a and b, and (iii) points where f is not differentiable. If f is differentiable, then the task reduces to finding all the critical points and comparing all the critical values,

$$\{f(x) : \ f'(x) = 0\}.$$

Finally, the largest critical value has to be compared to $f(a)$ and $f(b)$.

Example: Consider the function $f : [-1, 2] \to \mathbb{R}$, $f : x \mapsto x^3 - x$. This function is differentiable everywhere, and its critical points satisfy

$$f'(x) = 3x^2 - 1 = 0,$$

i.e., $x = \pm 1/\sqrt{3}$, which are both in the domain. It is easily checked that

$$f(1/\sqrt{3}) = -\frac{2}{3\sqrt{3}} \qquad \text{and} \qquad f(-1/\sqrt{3}) = \frac{2}{3\sqrt{3}}.$$

Finally, $f(-1) = 0$ and $f(2) = 6$, hence 2 is the (unique) maximum point. ▲▲▲

So far, we have derived properties of the derivative given the function. What about the reverse direction? Take the following example: we know that the derivative of a constant function is zero. Is it also true that if the derivative is zero then the function is a constant? A priori, it is not clear how to show it. How can we go from the knowledge that

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

to showing that f is a constant function. The following two theorems will provide us with the necessary tools:

Theorem 3.11 (Michel Rolle, 1691) If f is continuous on $[a, b]$, differentiable in (a, b), and $f(a) = f(b)$, then there exists a point $c \in (a, b)$ where $f'(c) = 0$.

Proof: It follows from the continuity of f that it has a maximum and a minimum point in $[a, b]$. If the maximum occurs at some $c \in (a, b)$ then $f'(c) = 0$ and we are done. If the minimum occurs at some $c \in (a, b)$ then $f'(c) = 0$ and we are done. The only remaining alternative is that a and b are both minima and maxima, in which case f is a constant and its derivative vanishes at some interior point (well, at all of them). ■

Comment: The requirement that f be differentiable everywhere in (a, b) is imperative, for consider the function

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

Even though f is continuous in $[0, 1]$ and $f(0) = f(1) = 0$, there is no interior point where $f'(x) = 0$.

Theorem 3.12 (Mean-value theorem (משפט הערך הממוצע)) If f is continuous on $[a, b]$ and differentiable in (a, b), then there exists a point $c \in (a, b)$ where

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

that is, a point at which the derivative equals to the mean rate of change of f between a and b.

Comment: Rolle's theorem is a particular case.

Proof: This is almost a direct consequence of Rolle's theorem. Define the function $g : [a, b] \to \mathbb{R}$,

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

g is continuous on $[a, b]$ and differentiable in (a, b). Moreover, $g(a) = g(b) = f(a)$. Hence by Rolle's theorem there exists a point $c \in (a, b)$, such that

$$0 = g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}.$$

■

Corollary 3.2 Let $f : [a, b] \to \mathbb{R}$. If $f'(x) = 0$ for all $x \in (a, b)$ then f is a constant.

Proof: Let $c, d \in (a, b)$, $c < d$. By the mean-value theorem there exists some $e \in (c, d)$, such that

$$f'(e) = \frac{f(d) - f(c)}{d - c},$$

however $f'(e) = 0$, hence $f(d) = f(c)$. Since this holds for all pair of points, then f is a constant. ■

Comment: This is only true if f is defined on an interval. Take a domain of definition which is the union of two disjoint sets, and this is no longer true (f is only constant in every "connected component" (רכיב קשירות)).

(27 hrs, 2009)

Corollary 3.3 If f and g are differentiable on an interval with $f'(x) = g'(x)$, then there exists a real number c such that $f = g + c$ on that interval.

Proof: Apply the previous corollary for $f - g$. ■

Corollary 3.4 If f is continuous on a closed interval $[a, b]$ and $f'(x) > 0$ in (a, b), then f is increasing on that interval.

Proof: Let $x < y$ belong to that interval. By the mean-value theorem there exists a $c \in (x, y)$ such that

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

however $f'(c) > 0$, hence $f(y) > f(x)$. ■

Comment: The converse is not true. If f is increasing and differentiable then $f'(x) \geq 0$, but equality may hold, as in $f(x) = x^3$ at zero.

We have seen that at a local minimum (or maximum) the derivative (if it exists) vanishes, but that the opposite is not true. The following theorem gives a sufficient condition for a point to be a local minimum (with a corresponding theorem for a local maximum).

Theorem 3.13 If $f'(a) = 0$ and $f''(a) > 0$ then a is a local minimum of f.

(30 hrs, 2011)

Comment: The fact that f'' exists at a implies:

1. f' exists in some neighborhood of a.
2. f' is continuous at a.
3. f is continuous in some neighborhood of a.

Proof: By definition,

$$f''(a) = \lim_a \Delta_{f',a} > 0 \quad \text{where} \quad \Delta_{f',a}(y) = \frac{f'(y) - f'(a)}{y - a} = \frac{f'(y)}{y - a}.$$

Thus there exists a $\delta > 0$ such that

$$\frac{f'(y)}{y - a} > 0 \qquad \text{whenever} \qquad 0 < |y - a| < \delta,$$

or,

$$\begin{aligned} f'(y) > 0 &\qquad \text{whenever} \qquad a < y < a + \delta \\ f'(y) < 0 &\qquad \text{whenever} \qquad a - \delta < y < a. \end{aligned}$$

It follows that f is increasing in a right-neighborhood of a and decreasing in a left-neighborhood of a, i.e., a is a local minimum.

Another way of completing the argument is as follows: for every $y \in (a, a + \delta)$:

$$\frac{f(y) - f(a)}{y - a} = f'(c_y) > 0 \qquad \text{i.e.,} \qquad f(y) > f(a),$$

where $c_y \in (a, y)$. Similarly, for every $y \in (a - \delta, a)$:

$$\frac{f(y) - f(a)}{y - a} = f'(c_y) < 0 \qquad \text{i.e.,} \qquad f(y) > f(a),$$

where here $c_y \in (y, a)$.

■

Example: Consider the function $f(x) = x^3 - x$. There are three critical points $0, \pm 1/\sqrt{3}$: one local maximum, one local minimum, and one which is neither. ▲▲▲

Example: Consider the function $f(x) = x^4$. Zero is a minimum point, even though $f''(0) = 0$. ▲▲▲

(36 hrs, 2010)

Theorem 3.14 *If f has a local minimum at a and $f''(a)$ exists, then $f''(a) \geq 0$.*

Proof: Suppose, by contradiction, that $f''(a) < 0$. By the previous theorem this would imply that a is both a local minimum and a local maximum. This in turn implies that there exists a neighborhood of a in which f is constant, i.e., $f'(a) = f''(a) = 0$, a contradiction. ■

The following theorem states that the derivative of a continuous function cannot have a removable discontinuity.

Theorem 3.15 *Suppose that f is continuous at a and*

$$\lim_a f' \text{ exists},$$

then $f'(a)$ exists and f' is continuous at a.

Proof: The assumption that f' has a limit at a implies that f' exists in some punctured neighborhood U of a. Thus, for $y \in U \setminus \{a\}$ the function f is continuous on the closed segment that connects y and a and differentiable on the corresponding open segment. By the mean-value theorem, there exists a point c_y between y and a, such that

$$f'(c_y) = \frac{f(y) - f(a)}{y - a} = \Delta_{f,a}(y).$$

Since f' has a limit at a,

$$(\forall \epsilon > 0)(\exists \delta > 0) : (\forall y \in B^\circ(a, \delta)) \left(|f'(y) - \lim_a f'| < \epsilon \right).$$

Since $y \in B^\circ(a,\delta)$ implies that $c_y \in B^\circ(a,\delta)$,

$$(\forall \epsilon > 0)(\exists \delta > 0) : (\forall y \in B^\circ(a,\delta))(|\underbrace{f'(c_y)}_{\Delta_{f,a}(y)} - \lim_a f'| < \epsilon),$$

which precisely means that

$$f'(a) = \lim_a f'.$$

■

Example: The function

$$f(x) = \begin{cases} x^2 & x < 0 \\ x^2 + 5 & x \geq 0, \end{cases}$$

is differentiable in a neighborhood of 0 and

$$\lim_0 f' \text{ exists},$$

but it is not continuous, and therefore not differentiable at zero. ▲▲▲

Example: The function

$$f(x) = \begin{cases} x^2 \sin 1/x & x \neq 0 \\ 0 & x = 0, \end{cases}$$

is continuous and differentiable in a neighborhood of zero. However, the derivative does not have a limit at zero, so that the theorem does not apply. Note, however, that f is differentiable at zero. ▲▲▲

(29 hrs, 2009) (37 hrs, 2010)

Theorem 3.16 (Augustin Louis Cauchy; intermediate value theorem) Suppose that f, g are continuous on $[a,b]$ and differentiable in (a,b). Then, there exists a $c \in (a,b)$, such that

$$[f(b) - f(a)]g'(c) = [g(b) - g(a)]f'(c).$$

Comment: If $g(b) \neq g(a)$ then this theorem states that

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

(provided also that $g'(c) \neq 0$). This may seem like a corollary of the mean-value theorem. We know that there exist c,d, such that

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

The problem is that there is no reason for the points c and d to coincide.

Proof: Define

$$h(x) = f(x)[g(b) - g(a)] - g(x)[f(b) - f(a)],$$

and apply Rolle's theorem. Namely, we verify that h is continuous on $[a, b]$ and differentiable in (a, b), and that $h(a) = h(b)$, hence there exists a point $c \in (a, b)$, such that

$$h'(c) = f'(c)[g(b) - g(a)] - g'(c)[f(b) - f(a)] = 0.$$

■

Theorem 3.17 (L'Hôpital's rule) Suppose that

$$\lim_a f = \lim_a g = 0,$$

and that

$$\lim_a \frac{f'}{g'} \textit{ exists}.$$

Then the limit $\lim_a (f/g)$ *exists and*

$$\lim_a \frac{f}{g} = \lim_a \frac{f'}{g'}.$$

Comment: This is a very convenient tool for obtaining a limit of a fraction, when both numerator and denominator vanish in the limit[2].

[2] L'Hôpital's rule was in fact derived by Johann Bernoulli who was giving to the French nobleman, le Marquis de l'Hôpital, lessons in calculus. L'Hôpital was the one to publish this theorem, acknowledging the help of Bernoulli.

Proof: It is implicitly assumed that f and g are differentiable in a neighborhood of a (except perhaps at a) and that g' is non-zero in a neighborhood of a (except perhaps at a); otherwise, the limit $\lim_a(f'/g')$ would not have existed. Note that f and g are not necessarily defined at a; since they have a limit there, there will be no harm assuming that they are continuous at a, i.e., $f(a) = g(a) = 0$[3].

First, we claim that g does not vanish in some neighborhood U of a, for by Rolle's theorem, it would imply that g' vanishes somewhere in this neighborhood. More formally, if

$$g'(x) \neq 0 \qquad \text{whenever} \qquad 0 < |x - a| < \delta,$$

then

$$g(x) \neq 0 \qquad \text{whenever} \qquad 0 < |x - a| < \delta,$$

for if $g(x) = 0$, with, say, $0 < x < \delta$, then there exists a $0 < c_x < x < \delta$, where $g'(c_x) = 0$.

Both f and g are continuous and differentiable in some neighborhood U that contains a, hence by Cauchy's mean-value theorem, there exists for every $x \in U$, a point c_x between x and a, such that

$$\frac{f(x) - 0}{g(x) - 0} = \frac{f(x)}{g(x)} = \frac{f'(c_x)}{g'(c_x)}. \tag{3.1}$$

Since the limit of f'/g' at a exists,

$$(\forall \epsilon > 0)(\exists \delta > 0) : x \in B^\circ(a, \delta) \ \text{ implies } \ \left(\left| \frac{f'(x)}{g'(x)} - \lim_a \frac{f'}{g'} \right| < \epsilon \right),$$

and since $x \in B^\circ(a, \delta)$ implies $c_x \in B^\circ(a, \delta)$,

$$(\forall \epsilon > 0)(\exists \delta > 0) : x \in B^\circ(a, \delta) \ \text{ implies } \ \left(\Big| \underbrace{\frac{f'(c_x)}{g'(c_x)}}_{f(x)/g(x)} - \lim_a \frac{f'}{g'} \Big| < \epsilon \right),$$

---

[3] What we really do is to define continuous functions

$$\tilde{f}(x) = \begin{cases} f(x) & x \neq a \\ 0 & x = a \end{cases} \qquad \text{and} \qquad \tilde{g}(x) = \begin{cases} g(x) & x \neq a \\ 0 & x = a \end{cases},$$

and prove the theorem for the "corrected" functions $\tilde{f}$ and $\tilde{g}$. At the end, we observe that the result must apply for f and g as well.

which means that

$$\lim_{a} \frac{f}{g} = \lim_{a} \frac{f'}{g'}.$$

■

Example:

① The limit

$$\lim_{x \to 0} \frac{\sin x}{x} = 1.$$

② The limit

$$\lim_{x \to 0} \frac{\sin^2 x}{x^2} = 1,$$

with two consecutive applications of l'Hôpital's rule.

▲▲▲

(32 hrs, 2011)

Comment: Purposely, we only prove one variant among many other of l'Hôpital's rule. Another variants is: Suppose that

$$\lim_{\infty} f = \lim_{\infty} g = 0,$$

and that

$$\lim_{\infty} \frac{f'}{g'} \text{ exists.}$$

Then

$$\lim_{\infty} \frac{f}{g} = \lim_{\infty} \frac{f'}{g'}.$$

Yet another one is: Suppose that

$$\lim_{a} f = \lim_{a} g = \infty,$$

and that

$$\lim_{a} \frac{f'}{g'} \text{ exists.}$$

Then

$$\lim_{a} \frac{f}{g} = \lim_{a} \frac{f'}{g'}.$$

The following theorem is very reminiscent of l'Hôpital's rule, but note how different it is:

Theorem 3.18 *Suppose that f and g are differentiable at a, with*

$$f(a) = g(a) = 0 \qquad \text{and} \qquad g'(a) \neq 0.$$

Then,

$$\lim_a \frac{f}{g} = \frac{f'(a)}{g'(a)}.$$

Proof: Without loss of generality, assume $g'(a) > 0$. That is,

$$\lim_{y\to a} \frac{g(y) - g(a)}{y - a} = \lim_{y\to a} \frac{g(y)}{y - a} = g'(a) > 0,$$

hence there exists a punctured neighborhood of a where $g(x)/(x-a) > 0$, and in particular, in which the numerator $g(x)$ is non-zero. Thus, we can divide by $g(x)$ in a punctured neighborhood of a, and

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{\frac{f(x)-f(a)}{x-a}}{\frac{g(x)-g(a)}{x-a}} = \frac{\Delta_{f,a}(x)}{\Delta_{g,a}(x)}.$$

Using the arithmetic of limits, we obtain the desired result. ■

(33 hrs, 2011)

(30 hrs, 2009)

✎ *Exercise 3.3* Show that the function

$$f(x) = \begin{cases} \frac{1}{2}x + x^2 \sin\frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

satisfies $f'(0) > 0$, and yet, is not increasing in any neighborhood of 0 (i.e., there is no neighborhood of zero in which $x < y$ implies $f(x) < f(y)$).

✎ *Exercise 3.4* Let I be an open segment. Any suppose that $f'(x) = f(x)$ for all $x \in I$. Show that there exists a constant c, such that $f(x) = c\,e^x$. (Hint: consider the function $g(x) = f(x)/e^x$.)

✎ *Exercise 3.5*

① Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable. Show that if f vanishes at m points, then f' vanishes at at least $m-1$ points.

② Let f be n times continuously differentiable in $[a,b]$, and $n+1$ times differentiable in (a,b). Suppose furthermore that $f(b) = 0$ and $f^{(k)}(a) = 0$ for $k = 0, 1, \dots, n$. Show that there exists a $c \in (a,b)$ such that $f^{(n+1)}(c) = 0$.

✎ *Exercise 3.6* Let

$$f(x) = \begin{cases} \sin(\omega x) & x < 0 \\ \gamma & x = 0 \\ \alpha x^2 + \beta x & x > 0. \end{cases}$$

Find all the values of $\alpha, \beta, \gamma, \omega$ for which f is (i) continuous at zero, (ii) differentiable at zero, (iii) twice differentiable at zero.

✎ *Exercise 3.7* Calculate the following limits:

① $\lim_{x\to 0} \frac{\sin(ax)}{\sin(bx)\cos(cx)}$.

② $\lim_{x\to 0} \frac{\tan(x)-x}{x-\sin(x)}$.

③ $\lim_{x\to 0} \frac{x^2\sin(1/x)}{\sin(x)}$.

④ $\lim_{x\to 1}(1-x)\tan\left(\frac{\pi x}{2}\right)$.

✎ *Exercise 3.8* Suppose that $f : (0,\infty) \to \mathbb{R}$ is differentiable on $(0,\infty)$ and $\lim_\infty f' = L$. Prove, without using l'Hopital's theorem that $\lim_{x\to\infty} f(x)/x = L$. Hint: write

$$\frac{f(x)}{x} = \frac{f(x)-f(x_0)}{x-x_0} \cdot \frac{x-x_0}{x},$$

and choose x_0 appropriately in order to use Lagrange's theorem.

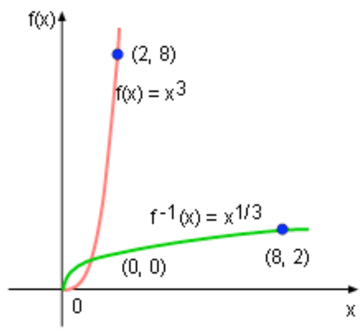


Figure 3.2: The derivative of the inverse function

3.5 Derivatives of inverse functions

We have already shown that if $f : A \to B$ is one-to-one and onto, then it has an inverse $f^{-1} : B \to A$. We then say that f is ***invertible*** (הפיכה). We have seen that an invertible function defined on an interval must be monotonic, and that the inverse of a continuous function is continuous. In this section, we study under what conditions is the inverse of a differentiable function differentiable.

Recall also that

$$f^{-1} \circ f = \mathrm{Id},$$

with $\mathrm{Id} : A \to A$. If both f and f^{-1} were differentiable at a and $f(a)$, respectively, then by the composition rule,

$$(f^{-1})'(f(a))\, f'(a) = 1.$$

Setting $f(a) = b$, or equivalently, $a = f^{-1}(b)$, we get

$$(f^{-1})'(b) = \frac{1}{f'(f^{-1}(b))}.$$

There is one little flaw with this argument. We have assumed $(f^{-1})'$ to exist. If this were indeed the case, then we have just derived an expression for the derivative of the inverse.

Corollary 3.5 If $f'(f^{-1}(b)) = 0$ then f^{-1} is not differentiable at b.

Example: The function $f(x) = x^3 + 2$ is continuous and invertible, with $f^{-1}(x) = \sqrt[3]{x-2}$. Since, $f'(f^{-1}(2)) = 0$, then f^{-1} is not differentiable at 2. ▲▲▲

(39 hrs, 2010)

***Theorem 3.19** Let $f : I \to \mathbb{R}$ be continuous and invertible. If f is differentiable at $f^{-1}(b)$ and $f'(f^{-1}(b)) \neq 0$, then f^{-1} is differentiable at b.*

Proof: Let $b = f(a)$, i.e., $a = f^{-1}(b)$. It is given

$$\lim_a \Delta_{f,a} = f'(a) \neq 0.$$

To each $x \neq a$ corresponds a $y \neq b$ such that

$$y = f(x) \qquad \text{and} \qquad x = f^{-1}(y).$$

Then,

$$\Delta_{f,a}(x) = \frac{f(x) - f(a)}{x - a} = \frac{y - b}{f^{-1}(y) - f^{-1}(b)} = \frac{1}{\Delta_{f^{-1},b}(y)},$$

which we can rewrite as

$$\Delta_{f^{-1},b}(y) = \frac{1}{\Delta_{f,a}(x)} = \frac{1}{\Delta_{f,a}(f^{-1}(y))}.$$

Since f^{-1} is continuous at b and $\Delta_{f,a}$ is continuous at $a = f^{-1}(b)$, it follows from limit artithmetic that

$$\lim_b \Delta_{f^{-1},b} = \frac{1}{f'(f^{-1}(b))}.$$

■

Example: Consider the function $f(x) = \tan x$ defined on $(-\pi/2, \pi/2)$. We denote its inverse by $\arctan x$. Then[4],

$$(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))} = \cos^2(\arctan x) = \frac{1}{1 + \tan^2(\arctan x)} = \frac{1}{1 + x^2}.$$

▲▲▲

(35 hrs, 2013)

[4]We used the identity $\cos^2 = 1/(1 + \tan^2)$.

Example: Consider the function $f(x) = x^n$, with n integer. This proves that the derivative of $f^{-1}(x) = x^{1/n}$ is

$$(f^{-1})'(x) = \frac{1}{n} x^{1/n-1},$$

and together with the composition rule gives the derivative for all rational powers. ▲▲▲

Comment: So far, we have not defined real-valued powers. We will get to it after we study integration theory.

✎ *Exercise 3.9* Let $f : \mathbb{R} \to \mathbb{R}$ be one-to-one and onto, and furthermore differentiable with $f'(x) \neq 0$ for all $x \in \mathbb{R}$. Suppose that there exists a function $F : \mathbb{R} \to \mathbb{R}$ such that $F' = f$. Define $G(x) = x f^{-1}(x) - F(f^{-1}(x))$. Prove that $G'(x) = f^{-1}(x)$. Note that this statement says that if we know a function whose derivative is one-to-one and onto, then we know the function whose derivative is its inverse.

✎ *Exercise 3.10* Let f be invertible and twice differentiable in (a, b). Furthermore, $f'(x) \neq 0$ in (a, b). Prove that $(f^{-1})''$ exists and find an expression for it in terms of f and its derivatives.

✎ *Exercise 3.11* Suppose that $f : \mathbb{R} \to \mathbb{R}$ is continuous, one-to-one and onto, and satisfies $f = f^{-1}$. (i) Show that f has a ***fixed point***, i.e., there exists an $x \in \mathbb{R}$ such that $f(x) = x$. (ii) Find such an f that has a unique fixed point. Hint: start by figuring out how do such functions look like.

3.6 Complements

Theorem 3.20 Suppose that $f : A \to B$ is monotonic (say, increasing), then it has one-sided limits at every point that has a one-sided neighborhood that belongs to A. (In particular, if A is a connected set then f has one-sided limits everywhere.)

Proof: Consider a segment $(b, a] \in A$. We need to show that f has a left-limit at a. The set

$$K = \{f(x) : b < x < a\}$$

is non-empty and upper bounded by $f(a)$ (since f is increasing). Hence we can define $\ell = \sup K$. We are going to show that

$$\ell = \lim_{a^-} f.$$

Let $\epsilon > 0$. By the definition of the supremum, there exists an $m \in K$ such that $m > \ell - \epsilon$. Let y be such $f(y) = m$. By the monotonicity of f,

$$\ell - \epsilon < f(x) \leq \ell \qquad \text{whenever} \qquad y < x < a.$$

This concludes the proof. We proceed similarly for right-limits. ■

Theorem 3.21 *Let $f : [a, b] \to \mathbb{R}$ be monotonic (say, increasing). Then the image of f is a segment if and only if f is continuous.*

Proof: (i) Suppose first that f is continuous. By monotonicity,

$$\text{Image}(f) \subseteq [f(a), f(b)].$$

That every point of $[f(a), f(b)]$ is in the image of f follows from the intermediate-value theorem.

(ii) Suppose then that $\text{Image}(f) = [f(a), f(b)]$. Suppose, by contradiction, that f is not continuous. This implies the existence of a point $c \in [a, b]$, where

$$f(c) > \lim_{c^-} f \qquad \text{and/or} \qquad f(c) < \lim_{c^+} f.$$

(We rely on the fact that one-sided derivatives exist.)

Say, for example, that $f(c) < \lim_{c^+} f$. Then,

$$f(c) < \frac{f(c) + \lim_{c^+} f}{2} < \lim_{c^+} f$$

is not in the image of f, which contradicts the fact that $\text{Image}(f) = [f(a), f(b)]$. ■

Theorem 3.22 (Jean-Gaston Darboux) *Let $f : A \to B$ be differentiable on (a, b), including a right-derivative at a and a left-derivative at b. Then, f' has the "intermediate-value property", whereby for every t between $f'(a^+)$ and $f'(b^-)$ there exists an $x \in [a, b]$, such that $f'(x) = t$.*

Comment: The statement is trivial when f is continuously differentiable (by the intermediate-value theorem). It is nevertheless correct even if f' has discontinuities. This theorem thus limits the functions that are derivatives of other functions—not every function can be a derivative.

Proof: Without loss of generality, let us assume that $f'(a^+) > f'(b^-)$. Let $f'(a^+) > t > f'(b^-)$, and consider the function

$$g(x) = f(x) - tx.$$

(Here t is a fixed parameter; for every t we can construct such a function.) We have

$$g'(a^+) = f'(a^+) - t > 0 \qquad \text{and} \qquad g'(b^-) = f'(b^-) - t < 0,$$

and we wish to show the existence of an $x \in [a, b]$ such that $g'(x) = 0$.

Since g is continuous on $[a, b]$ then it must attain a maximum. The point a cannot be a maximum point because g is increasing at a. Similarly, b cannot be a maximum point because g is decreasing at b. Thus, there exists an interior point c which is a maximum, and by Fermat's theorem, $g'(c) = f'(c) - t = 0$. ■

(32 hrs, 2009)

Example: The function

$$f(x) = \begin{cases} x^2 \sin 1/x & x \neq 0 \\ 0 & x = 0 \end{cases}$$

is differentiable everywhere, and its derivative is

$$f'(x) = \begin{cases} 2x \sin 1/x - \cos 1/x & x \neq 0 \\ 0 & x = 0 \end{cases}.$$

The derivative is not continuous at zero, but nevertheless satisfes the Darboux theorem. ▲▲▲

Example: How crazy can a continuous function be? It was Weierstrass who shocked the world by constructing a continuous function that it nowhere differentiable! As instance of his class of examples is

$$f(x) = \sum_{k=0}^{\infty} \frac{\cos(21^k \pi x)}{3^k}.$$

▲▲▲

3.7 Taylor's theorem

Recall the mean-value theorem. If $f : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and differentiable in (a,b), then for every point $x \in (a,b)$ there exists a point c_x, such that

$$\frac{f(x) - f(a)}{x - a} = f'(c_x).$$

We can write it alternatively as

$$f(x) = f(a) + f'(c_x)(x - a).$$

Suppose, for example, that we knew that $|f'(y)| < M$ for all y. Then, we would deduce that $f(x)$ cannot differ from $f(a)$ by more than M times the separation $|x - a|$. In particular,

$$\lim_a [f - f(a)] = 0.$$

It turns out that if we have more information, about higher derivatives of f, then we can refine a lot the estimates we have on the variation of f as we get away from the point a.

Theorem 3.23 *Let f be a given function, and a an interior point of its domain. Suppose that the derivatives*

$$f'(a), f''(a), \dots, f^{(n)}(a)$$

all exist. Let $a_k = f^{(k)}(a)/k!$ and set the ***Taylor polynomial*** *of degree n,*

$$P_{f,n,a}(x) = a_0 + a_1(x - a) + a_2(x - a)^2 + \cdots + a_n(x - a)^n.$$

Then,

$$\lim_a \frac{f - P_{f,n,a}}{(\mathrm{Id} - a)^n} = 0.$$

Comment: Loosely speaking, this theorem states that, as x approaches a, $P_{f,n,a}(x)$ is closer to $f(x)$ than $(x - a)^n$ (or that f and $P_{f,n,a}$ are "equal up to order n"; see below). For $n = 0$ it states that

$$\lim_a [f - f(a)] = 0,$$

whereas for $n = 1$,

$$\lim_a \left(\frac{f - f(a)}{(\mathrm{Id} - a)} - f'(a) \right) = \lim_a \Delta_{f,a} - f'(a) = 0,$$

which holds by definition.

Proof: Consider the ratio

$$\frac{f - P_{f,n,a}}{(\mathrm{Id} - a)^n},$$

which is a ratio of n-times differentiable functions at a. Both numerator and denominator vanish at a, so we will attempt to use l'Hôpital's rule. For that we note that

$$P'_{f,n,a} = P_{f',n-1,a}.$$

Thus,

$$\lim_a \frac{f - P_{f,n,a}}{(\mathrm{Id} - a)^n} = \lim_a \frac{f' - P_{f',n-1,a}}{n\,(\mathrm{Id} - a)^{n-1}},$$

provided that the right hand side exists. We then proceed to consider the right hand side, which is the limit of the ratio of two differentiable functions that vanish at a. Hence,

$$\lim_a \frac{f' - P_{f',n-1,a}}{n\,(\mathrm{Id} - a)^{n-1}} = \lim_a \frac{f'' - P_{f'',n-2,a}}{n(n-1)\,(\mathrm{Id} - a)^{n-2}},$$

provided that the right hand side exists. We proceed inductively, until we get that

$$\lim_a \frac{f - P_{f,n,a}}{(\mathrm{Id} - a)^n} = \lim_a \frac{f^{(n-1)} - P_{f^{(n-1)},1,a}}{n!\,(\mathrm{Id} - a)},$$

provided that the right hand side exists. Now write the right hand side explicitly,

$$\frac{1}{n!} \lim_a \frac{f^{(n-1)} - f^{(n-1)}(a) - f^{(n)}(a)(\mathrm{Id} - a)}{\mathrm{Id} - a} = \frac{1}{n!} \left[\lim_a \Delta_{f^{(n-1)},a} - f^{(n)}(a) \right],$$

which vanishes by the definition of $f^{(n)}(a)$. This concludes the proof. ∎

Comment: Why couldn't we use Taylor's theorem one more time at the end? Because it is not given that the n-th derivative has a limit at a. It is only given that the n-derivative exists at that point.

(35 hrs, 2011)

The polynomial $P_{f,n,a}(x)$ is known as the ***Taylor polynomial*** of degree n of f about a. Its existence relies on f being differentiable n times at a. If we denote by Π_n the set of polynomials of degree at most n, then the defining property of $P_{f,n,a}$ is:

1. $P_{f,n,a} \in \Pi_n$.
2. $P^{(k)}_{f,n,a}(a) = f^{(k)}(a)$ for $k = 0, 1, \dots, n$.

Definition 3.6 Let f, g be defined in a neighborhood of a. We say that f and g are ***equal up to order*** *n* ***at*** *a if*

$$\lim_a \frac{f - g}{(\mathrm{Id} - a)^n} = 0.$$

It is easy to see that "equal up to order n" is an ***equivalence relation***, which defines an ***equivalence class***.

Thus, the above theorem proves that $f(x)$ and $P_{f,n,a}$ (provided that the latter exists) are equal up to order n at a.

(36 hrs, 2013)

The next task is to obtain an expression for the difference between the function f and its Taylor polynomial. We define the ***remainder*** (שארית) of order n, $R_n(x)$, by

$$f(x) = P_{f,n,a}(x) + R_{f,n,a}(x).$$

Proposition 3.1 Let $P, Q \in \Pi_n$ be equal up to order n at a. Then $P = Q$.

Proof: We can always express these polynomials as

$$\begin{aligned} P(x) &= a_0 + a_1(x - a) + \cdots + a_n(x - a)^n \\ Q(x) &= b_0 + b_1(x - a) + \cdots + b_n(x - a)^n. \end{aligned}$$

We know that

$$\lim_a \frac{P - Q}{(\mathrm{Id} - a)^k} = 0 \qquad \text{for } k = 0, 1, \dots, n.$$

For $k = 0$

$$0 = \lim_a (P - Q) = a_0 - b_0,$$

i.e., $a_0 = b_0$. We proceed similarly for each k to show that $a_k = b_k$ for $k = 0, 1, \dots, n$. ■

Corollary 3.6 If f is n times differentiable at a and f is equal to $Q \in \Pi_n$ up to order n at a, then

$$P_{f,n,a} = Q.$$

Example: We know from high-school math that for $|x| < 1$

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

Thus,

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

Since the right hand side tends to zero as $x \to 0$ it follows that f and $\sum_{k=0}^{n} x^k$ are equal up to order n at 0, which means that

$$P_{f,n,0}(x) = \sum_{k=0}^{n} x^k.$$

▲▲▲

Example: Similarly for $|x| < 1$

$$f(x) = \frac{1}{1+x^2} = \sum_{k=0}^{n} (-1)^k x^{2k} + (-1)^{n+1} \frac{x^{2n+2}}{1+x^2}.$$

Thus,

$$\frac{f(x) - \sum_{k=0}^{n} (-1)^k x^{2k}}{x^{2n}} = \frac{x^2}{1+x^2}.$$

Since the right hand side tends to zero as $x \to 0$ it follows that

$$P_{f,2n,0}(x) = \sum_{k=0}^{n} (-1)^k x^{2k}.$$

▲▲▲

Example: Let $f = \arctan$. Then,

$$f(x) = \int_0^x \frac{dt}{1+t^2} = \int_0^x \left(\sum_{k=0}^n (-1)^k t^{2k} + (-1)^{n+1} \frac{t^{2n+2}}{1+t^2} \right) dt$$
$$= \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{2k+1} + \int_0^x (-1)^{n+1} \frac{t^{2n+2}}{1+t^2}\, dt.$$

Since

$$\left| f(x) - \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{2k+1} \right| \le |x|^{2n+3},$$

it follows that

$$P_{f,2n+2,0} = \sum_{k=0}^n (-1)^k \frac{x^{2k+1}}{2k+1}.$$

▲▲▲

(36 hrs, 2011)

Theorem 3.24 Let f, g be n times differentiable at a. Then,

$$P_{f+g,n,a} = P_{f,n,a} + P_{g,n,a},$$

and

$$P_{fg,n,a} = [P_{f,n,a} P_{g,n,a}]_n,$$

where $[]_n$ stands for a truncation of the polynomial in $(\mathrm{Id} - a)$at the n-th power.

Proof: We use again the fact that if a polynomial "looks" like the Taylor polynomial then it is. The sum $P_{f,n,a} + P_{g,n,a}$ belongs to Π_n and

$$\lim_a \frac{(f+g) - (P_{f,n,a} + P_{g,n,a})}{(\mathrm{Id} - a)^n} = \lim_a \frac{f - P_{f,n,a}}{(\mathrm{Id} - a)^n} + \lim_a \frac{g - P_{g,n,a}}{(\mathrm{Id} - a)^n} = 0,$$

which proves the first statement.

Second, we note that $P_{f,n,a} P_{g,n,a} \in \Pi_{2n}$, and that

$$\frac{fg - P_{f,n,a} P_{g,n,a}}{(\mathrm{Id} - a)^n} = \frac{fg - P_{f,n,a} g + P_{f,n,a} g - P_{f,n,a} P_{g,n,a}}{(\mathrm{Id} - a)^n}$$
$$= g \frac{f - P_{f,n,a}}{(\mathrm{Id} - a)^n} + P_{f,n,a} \frac{g - P_{g,n,a}}{(\mathrm{Id} - a)^n},$$

whose limit at a is zero. This proves that $P_{f,n,a}P_{g,n,a}$ and fg are equivalent up to order n at a. All that remains it to truncate the product at the n-th power of $(\mathrm{Id} - a)$. ■

(38 hrs, 2013)

Theorem 3.25 (Taylor) *Suppose that f is $(n+1)$-times differentiable on the interval $[a,x]$. Then,*

$$f(x) = P_{f,n,a}(x) + R_{n,f,a}(x),$$

where the remainder $R_{f,n,a}(x)$ can be represented in the **Lagrange form**,

$$R_{f,n,a}(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1},$$

for some $c \in (a,x)$. It can also be represented in the **Cauchy form**,

$$R_{f,n,a}(x) = \frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n(x-a),$$

for some $\xi \in (a,x)$.

Comment: For $n = 0$ this is simply the mean-value theorem.

Proof: Fix x, and consider the function $\phi : [a,x] \to \mathbb{R}$,

$$\begin{aligned}\phi(z) &= f(x) - P_{f,n,z}(x) \\ &= f(x) - f(z) - f'(z)(x-z) - \frac{f''(z)}{2!}(x-z)^2 - \cdots - \frac{f^{(n)}(z)}{n!}(x-z)^n.\end{aligned}$$

We note that

$$\phi(a) = R_{f,n,a}(x) \qquad \text{and} \qquad \phi(x) = 0.$$

Moreover, ϕ is differentiable, with

$$\begin{aligned}\phi'(z) &= -f'(z) - [f''(z)(x-z) - f'(z)] - \left(\frac{f'''(z)}{2!}(x-z)^2 - f''(z)(x-z)\right) \\ &\quad - \cdots - \left(\frac{f^{(n+1)}(z)}{n!}(x-z)^n - \frac{f^{(n)}(z)}{(n-1)!}(x-z)^{n-1}\right) \\ &= -\frac{f^{(n+1)}(z)}{n!}(x-z)^n.\end{aligned}$$

Let now ψ be any function defined on $[a, x]$, differentiable on (a, x) with non-vanishing derivative. By Cauchy's mean-value theorem, there exists a point $a < \xi < x$, such that

$$\frac{\phi(x) - \phi(a)}{\psi(x) - \psi(a)} = \frac{\phi'(\xi)}{\psi'(\xi)}.$$

That is,

$$R_{f,n,a}(x) = \frac{\psi(x) - \psi(a)}{\psi'(\xi)} \frac{f^{(n+1)}(\xi)}{n!}(x - \xi)^n$$

Set for example $\psi(z) = (x - z)^p$ for some p. Then,

$$R_{f,n,a}(x) = (x - a)^p \frac{f^{(n+1)}(\xi)}{p\, n!}(x - \xi)^{n-p+1}$$

For $p = 1$ we retrieve the Cauchy form. For $p = n + 1$ we retrieve the Lagrange form. ■

Comment: Later in this course we will study series and ask questions like "does the Taylor polynomial $P_{f,n,a}(x)$ tend to $f(x)$ as $n \to \infty$?". This is the notion of ***converging series***. Here we deal with a different beast: the degree of the polynomial n is fixed and we consider how does $P_{f,n,a}(x)$ approach $f(x)$ as $x \to a$. The fact that $P_{f,n,a}$ approaches f faster than $(x - a)^n$ as $x \to a$ means that $P_{f,n,a}$ is an ***asymptotic series*** of f.

(34 hrs, 2009)

✎ *Exercise 3.12* Write the polynomial $p(x) = x^3 - 3x^2 + 3x - 6$ as a polynomial in $x + 1$. That is, find coefficients $a_0, a_1, \dots, a_n$, such that $p(x) = \sum_{k=0}^{n} a_k (x + 1)^k$. Solve this problem using two methods: (i) Newton's binomial, and (ii) Taylor's polynomial.

✎ *Exercise 3.13* Let $f : \mathbb{R} \to \mathbb{R}$ be n times differentiable, such that $f^{(n)}(x) = 0$ for all $x \in \mathbb{R}$. Prove that f is a polynomial whose degree is at most $n - 1$.

✎ *Exercise 3.14* Write the Taylor polynomials of degree n at the point x_0 for the following functions, as well as the corresponding remainders, both in the Lagrange and Cauchy forms:

① $f(x) = 1/(1 + x)$, $x_0 = 0$, $n \in \mathbb{N}$.
② $f(x) = \cosh(x)$, $x_0 = 0$, $n \in \mathbb{N}$.

③ $f(x) = x^5 - x^4 + 7x^3 - 2x^2 + 3x + 1$, $x_0 = 1$, $n = 5$.

④ $f(x) = (x-2)^6 + 9(x+1)^4 + (x-2)^2 + 8$, $x_0 = 0$, $n = 6$.

⑤ $f(x) = (1+x)^\alpha$, $\alpha \notin \mathbb{N}$, $x_0 = 0$, $n \in \mathbb{N}$. Use the notation

$$\binom{\alpha}{n} \stackrel{\text{def}}{=} \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!},$$

which generalizes Newton's binomial.

⑥ $f(x) = \sin(x)$, $x_0 = \pi/2$, $n = 2k$, $k \in \mathbb{N}$.

✎ *Exercise 3.15* Show using different methods that of $f''(a)$ exists, then

$$f'(a) = \lim_{h\to 0} \frac{f(a+h) + f(a-h) - 2f(a)}{h}.$$

① Using l'Hopital's rule.

② Using Taylor's polynomial of order two f at a. Note that you can't use neither Lagrange or Cauchy's form of the remainder, since it is only given that f is twice differentiable at a.

③ Show that if

$$f(x) = \begin{cases} x^2 & x \geq 0 \\ -x^2 & x < 0, \end{cases}$$

then $f''(0)$ does not exist, but the limit of the above fraction does exist. Why isn't it a contradiction?

✎ *Exercise 3.16* Evaluate, using Taylor's polynomial, the following expressions, up to an accuracy of at least 0.01:

① $e^{1/2}$.

② $\cos(1/2)$.

✎ *Exercise 3.17*

① For every polynomial $P(x) = \sum_{k=0}^{n} a_k x^k$ and evert $q \in \mathbb{N}$, we denote

$$P_q(x) = \begin{cases} \sum_{k=0}^{q} a_k x^k & q < n \\ P(x) & q \geq n. \end{cases}$$

Suppose that f, g have Taylor polynomials of order n at zero, P and Q. Suppose furthermore that $g(0) = 0$, and that g does not vanish in a punctured neighborhood of zero. Show that the Taylor polynomial of order n at zero of $f \circ g$ is given by $(P \circ Q)_n$. Hint: write $f(g(x)) = f(g(x))/g(x) \cdot g(x)$.

② Given the Taylor polynomial P of order n at zero of the function f, write the Taylor polynomial of order n at zero of the function $g(x) = f(2x)$.

③ Find the Taylor polynomial of order three at zero of the function $f(x) = \exp(\sin(x))$.

✎ *Exercise 3.18* Let f be n times differentiable at x_0 and let $P(x) = \sum_{k=0}^{n} a_k x^k$ be its Taylor polynomial at zero. What is the Taylor polynomial of f' of order $n-1$ at zero. Justify your answer.

✎ *Exercise 3.19* Show that if f and g are n times differentiable at x_0, then

$$(f \cdot g)^{(n)}(x_0) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x_0)\, g^{(n-k)}(x_0).$$

Hint: think of the Taylor polynomial of a product of functions.

✎ *Exercise 3.20* Suppose that f and g have derivatives of all orders in $\mathbb{R}$ and that $P_n(x) = Q_n(x)$ for all $n \in \mathbb{N}$ and every $x \in \mathbb{R}$, where P_n and Q_n are the corresponding Taylor polynomials at zero. Does it imply that $f(x) = g(x)$?

Chapter 4

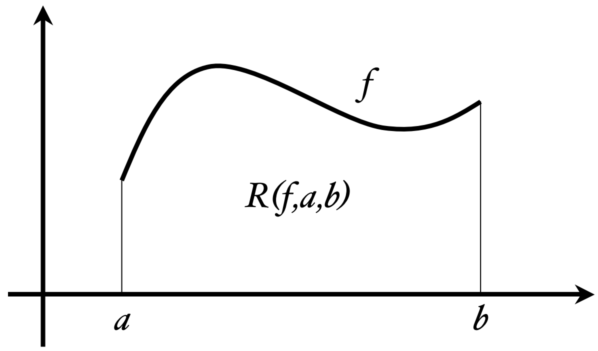
Integration theory

4.1 Definition of the integral

Derivatives were about calculating the rate of change of a function. Integrals were invented in order to calculate areas. The fact that the two are intimately related—that derivatives are in a sense the inverse of integrals—is a wonder.

When we learn in school about the area of a region, it is defined as "the number of squares with sides of length one that fit into the region". It is of course hard to make sense with such a definition. We will learn how to calculate the area of a region bounded by the x-axis, the graph of a function f, and two vertical lines passing though points $(a, 0)$ and $(b, 0)$. For the moment let us assume that $f \geq 0$. This region will be denoted by $R(f, a, b)$, and its area (once defined) will be called the ***integral*** of f on $[a, b]$[1].

[1]There are multiple non-equivalent ways to define integrals. We study integration theory in the sense of Riemann; another classical integration theory is that of Lebesgue (studied in the context of measure theory). Within the Riemann integration theory, we adopt a particular derivation due to Darboux.



Definition 4.1 *A* **partition** *(חלוקה) of the interval* $[a, b]$ *is a partition (yes, I know, you can't use the word to be defined in a definition) of the interval into a finite number of segments,*

$$[a, b] = [x_0, x_1] \cup [x_1, x_2] \cup \cdots \cup [x_{n-1}, x_n].$$

We will identify a partition with a finite number of points,

$$a = x_0 < x_1 < x_2 < \cdots < x_n = b.$$

Definition 4.2 *Let* $f : [a, b] \to \mathbb{R}$ *be a bounded function, and let* $P = \{x_0, x_1, \ldots, x_n\}$ *be a partition of* $[a, b]$*. Let*

$$m_i = \inf\{f(x) : x_{i-1} \le x \le x_i\}$$
$$M_i = \sup\{f(x) : x_{i-1} \le x \le x_i\}.$$

We define the **lower sum** *(סכום תחתון) of* f *for* P *to be*

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

We define the **upper sum** *(סכום עליון) of* f *for* P *to be*

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

Clearly, any sensible definition of the area A of $R(f, a, b)$ must satisfy

$$L(f, P) \le A \le U(f, P)$$

for all partitions P.

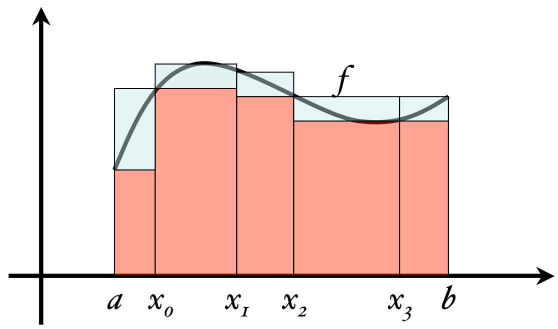


Figure 4.1: The lower and upper sums of f for P.

Comments:

① The boundedness of f is essential in these definitions.

② It is always true that $L(f, P) \le U(f, P)$ because $m_i \le M_i$ for all i.

③ Is it however true that $L(f, P_1) \le U(f, P_2)$ where P_1 and P_2 are arbitrary partitions? It ought to be true.

(35 hrs, 2009)

(38 hrs, 2011)

✎ *Exercise 4.1* Let $f : [-1, 1] \to \mathbb{R}$ be given by $f : x \mapsto x^2$ and let $P = \{-1, -1/2, -1/3, 2/3, 8/9, 1\}$ be a partition of $[-1, 1]$. Calculate $L(f, P)$ and $U(f, P)$.

Definition 4.3 A partition Q is called a **refinement** *(עידון) of a partition P if $P \subset Q$.*

Lemma 4.1 Let Q be a refinement of P. Then

$$L(f, Q) \ge L(f, P) \qquad \text{and} \qquad U(f, Q) \le U(f, P).$$

Proof: Suppose first that Q differs from P by exactly one point. That is,

$$P = \{x_0, x_1, \dots, x_n\}$$
$$Q = \{x_0, x_1, \dots, x_{j-1}, y, x_j, \dots, x_n\}.$$

Then,

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

and

$$L(f, Q) = \sum_{i=1}^{j-1} m_i(x_i - x_{i-1}) + \tilde{m}_j(y - x_{j-1}) + \hat{m}_j(x_j - y) + \sum_{i=j+1}^{n} m_i(x_i - x_{i-1}),$$

where

$$\tilde{m}_j = \inf\{f(x) : x_{j-1} \le x \le y\}$$
$$\hat{m}_j = \inf\{f(x) : y \le x \le x_j\}.$$

Thus,

$$L(f, Q) - L(f, P) = (\tilde{m}_j - m_j)(y - x_{j-1}) + (\hat{m}_j - m_j)(x_j - y).$$

It remains to note that $\tilde{m}_j \ge m_j$ and $\hat{m}_j \ge m_j$, since both are infimums over a smaller set[2]. Thus, $L(f, Q) - L(f, P) \ge 0$.

Consider now two partitions, $P \subset Q$. There exists a sequence of partitions,

$$P = P_0 \subset P_1 \subset \cdots \subset P_k = Q,$$

such that each P_j differs from P_{j-1} by one point. Hence,

$$L(f, P) \le L(f, P_1) \le \cdots \le L(f, Q).$$

A similar argument works for the upper sums. ∎

Theorem 4.1 For every two partitions P_1,P_2 of $[a, b]$,

$$L(f, P_1) \le U(f, P_2).$$

[2]It is always true that if $A \subset B$ then

$$\inf A \ge \inf B \qquad \text{and} \qquad \sup A \le \sup B.$$

Proof: Let $Q = P_1 \cup P_2$. Since Q is finer than both P_1 and P_2 it follows that

$$L(f, P_1) \le L(f, Q) \le U(f, Q) \le U(f, P_2).$$

■

(40 hrs, 2013)

✎ *Exercise 4.2* Prove or disprove:

① A partition Q is a refinement of a partition P if and only if every segment in Q is contained in a segment of P.
② If a partition P is a refinement of a partition Q then every segment in Q is a segment in P.
③ If R is a partition of both P and Q then it is also a refinement of $P \cup Q$.

✎ *Exercise 4.3* Characterize the functions f for which there exists a partition P such that $L(f, P) = U(f, P)$.

We may now consider the collection of lower and upper sums. Denote by

$$\text{Part}(a, b) = \{\text{All partitions of } [a, b]\}.$$

$$\mathscr{L}(f) = \{L(f, P) : P \in \text{Part}(a, b)\}$$
$$\mathscr{U}(f) = \{U(f, P) : P \in \text{Part}(a, b)\}.$$

Note that both $\mathscr{L}(f)$ and $\mathscr{U}(f)$ are sets of real numbers, and not sets of partitions. The set $\mathscr{L}(f)$ is non-empty and bounded from above by any element of $\mathscr{U}(f)$. Similarly, the set $\mathscr{U}(f)$ is non-empty and bounded from below by any element of $\mathscr{L}(f)$. Thus, we may define the least upper bound and greatest lower bound,

$$L(f) = \sup \mathscr{L}(f) \qquad \text{and} \qquad U(f) = \inf \mathscr{U}(f).$$

The real number $U(f)$ is called the ***upper integral*** (אינטגרל עליון) of f on $[a, b]$ and $L(f)$ is called the ***lower integral*** (אינטגרל תחתון) of f on $[a, b]$

We once proved a theorem stating that the following assertions are equivalent:

① $\sup \mathscr{L}(f) = \inf \mathscr{U}(f)$.
② There is a unique number c, such that $\ell \le c \le u$ for all $\ell \in \mathscr{L}(f)$ and $u \in \mathscr{U}(f)$.

③ For every $\epsilon > 0$ there exist an $\ell \in \mathscr{L}(f)$ and a $u \in \mathscr{U}(f)$, such that $u - \ell < \epsilon$.

Definition 4.4 *If $L(f) = U(f)$ we say that f is* ***integrable*** *on $[a,b]$. We call this number the* ***integral*** *of f on $[a,b]$, and denote it by*

$$\int_a^b f.$$

That is, f is integrable on $[a,b]$ if the upper and lower integrals coincide.

Comment: The integral of f is the unique number such that

$$L(f,P) \le \int_a^b f \le U(f,Q),$$

for all partitions P and Q.

Comment: You are all accustomed to the notation

$$\int_a^b f(x)\,dx.$$

This notation is indeed very suggestive about the meaning of the integral, however, it is very important to realize that the integral only involves f, a and b, and it is not a function of any x.

Comment: At this stage we only consider integrals of bounded functions over bounded segments—so-called ***proper integrals***. Integrals like

$$\int_a^\infty f \qquad \text{and} \qquad \int_0^1 \log x\,dx,$$

are instances of ***improper integrals***. We will only briefly talk about them (only due to a lack of time).

Comment: What about the case where the sign of f is not everywhere positive? The definition remains the same. The geometric interpretation is that areas under the x-axis count as "negative areas".

Two questions arise:

① How do we know whether a function is integrable?

② How to compute the integral of an integrable function?

(40 hrs, 2011)

The following lemma will prove useful in checking whether a function is integrable:

Theorem 4.2 A bounded function f is integrable if and only if there exists for every $\epsilon > 0$ a partition P, such that

$$U(f, P) < L(f, P) + \epsilon.$$

Proof: If for every $\epsilon > 0$ there exists a partition P, such that

$$U(f, P) < L(f, P) + \epsilon,$$

then f is integrable, by the equivalent criteria described above.

Conversely, suppose f is integrable. Then, given $\epsilon > 0$, there exist partitions P, P', such that

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

Let $P'' = P \cup P'$. Since P'' is a refinement of both, then

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

■

Example: Let $f : [a, b] \to \mathbb{R}$, $f : x \mapsto c$. Then for every partition P,

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

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

It follows that $L(f) = U(f) = c(b - a)$, hence f is integrable and

$$\int_a^b f = c(b - a).$$

▲▲▲

Example: Let $f : [0,1] \to \mathbb{R}$,

$$f : x \mapsto \begin{cases} 0 & x \notin \mathbb{Q} \\ 1 & x \in \mathbb{Q}. \end{cases}$$

For every partition and every j,

$$m_j = 0 \qquad \text{and} \qquad M_j = 1,$$

hence

$$L(f, P) = 0 \quad \text{and} \quad U(f, P) = 1,$$

and $L(f) = 0$ and $U(f) = 1$. Since $L(f) \neq U(f)$, f is not integrable. ▲▲▲

(41 hrs, 2010)

Example: Consider next the function $f : [0,2] \to \mathbb{R}$,

$$f : x \mapsto \begin{cases} 0 & x \neq 1 \\ 1 & x = 1. \end{cases}$$

Let P_n, $n \in \mathbb{N}$ denote the family of ***uniform partitions***,

$$x_i = \frac{2i}{n}.$$

For every such partition,

$$L(f, P) = 0 \qquad \text{and} \qquad U(f, P) = 2/n.$$

Set now,

$$\mathscr{L}_{\text{unif}}(f) = \{L(f, P_n) : n \in \mathbb{N}\}$$
$$\mathscr{U}_{\text{unif}}(f) = \{U(f, P_n) : n \in \mathbb{N}\}$$

Clearly, $\mathscr{L}_{\text{unif}}(f) \subseteq \mathscr{L}(f)$ and $\mathscr{U}_{\text{unif}}(f) \subseteq \mathscr{U}(f)$, hence

$$L(f) = \sup \mathscr{L}(f) \geq \sup \mathscr{L}_{\text{unif}}(f) = 0$$
$$U(f) = \inf \mathscr{U}(f) \leq \inf \mathscr{U}_{\text{unif}}(f) = 0,$$

thus

$$0 \leq L(f) \leq U(f) \leq 0,$$

which implies that f is integrable, and

$$\int_0^2 f = 0.$$

This example shows that integrals are not affected by the value of a function at an isolated point. ▲▲▲

The following example motivate the following useful lemma:

Lemma 4.2 *Let* $\mathscr{A} \subset \text{Part}(a, b)$ *be a subset of the set of partitions. Suppose that*

$$\sup\{L(f, P) : \; P \in \mathscr{A}\} = \inf\{U(f, P) : \; P \in \mathscr{A}\} = \ell.$$

Then f *is integrable on* $[a, b]$ *and*

$$\int_a^b f = \ell.$$

Proof: Since

$$\{L(f, P) : \; P \in \mathscr{A}\} \subseteq \mathscr{L}(f) \quad \text{and} \quad \{U(f, P) : \; P \in \mathscr{A}\} \subseteq \mathscr{U}(f),$$

if follows that

$$\ell = \sup\{L(f, P) : P \in \mathscr{A}\} \leq \sup \mathscr{L}(f) \leq \inf \mathscr{U}(f) \leq \inf\{U(f, P) : P \in \mathscr{A}\} = \ell,$$

which completes the proof. ■

We are next going to examine a number of examples, which on the one hand, will practice our understanding of integrability, and on the other hand, will convince us that we must develop tools to deal with more complicated functions.

Example: Let $f : [0, b] \to \mathbb{R}$, $f = \text{Id}$. Let P be any partition, then $m_i = x_{i-1}$ and $M_i = x_i$, hence

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

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

Using again uniform partition,

$$P_n = \{0, b/n, 2b/n, \ldots, b\},$$

$$L(f, P_n) = \sum_{i=1}^{n} \frac{(i-1)b}{n}\frac{b}{n} = \frac{b^2}{n^2}\frac{n(n-1)}{2} = \frac{b^2}{2}\left(1 - \frac{1}{n}\right)$$

$$U(f, P_n) = \sum_{i=1}^{n} \frac{ib}{n}\frac{b}{n} = \frac{b^2}{n^2}\frac{n(n+1)}{2} = \frac{b^2}{2}\left(1 + \frac{1}{n}\right).$$

Since

$$\sup\{L(f, P_n) : \; n \in \mathbb{N}\} = \frac{b^2}{2} \quad \text{and} \quad \inf\{U(f, P_n) : \; n \in \mathbb{N}\} = \frac{b^2}{2},$$

if follows from Lemma 4.2 that

$$\int_0^b f = \frac{b^2}{2}.$$

▲▲▲

(37 hrs, 2009)

Example: Let $f : [0, b] \to \mathbb{R}$, $f(x) = x^2$. Let P be any partition, then $m_i = x_{i-1}^2$ and $M_i = x_i^2$, hence

$$L(f, P) = \sum_{i=1}^{n} x_{i-1}^2 (x_i - x_{i-1})$$

$$U(f, P) = \sum_{i=1}^{n} x_i^2 (x_i - x_{i-1}).$$

Take the same uniform partition, then[3]

$$L(f, P_n) = \sum_{i=1}^{n} \frac{(i-1)^2 b^2}{n^2}\frac{b}{n} = \frac{b^3}{n^3}\frac{(n-1)n(2n-1)}{6} = \frac{b^3}{3}\left(1 - \frac{1}{n}\right)\left(1 - \frac{1}{2n}\right)$$

$$U(f, P_n) = \sum_{i=1}^{n} \frac{i^2 b^2}{n^2}\frac{b}{n} = \frac{b^3}{n^3}\frac{n(n+1)(2n+1)}{6} = \frac{b^3}{3}\left(1 + \frac{1}{n}\right)\left(1 + \frac{2}{n}\right).$$

[3]

$$\sum_{k=1}^{n} k^2 = \frac{1}{6} n(n+1)(2n+1).$$

Since

$$\sup\{L(f,P_n) : n \in \mathbb{N}\} = \frac{b^3}{3} \quad \text{and} \quad \inf\{U(f,P_n) : n \in \mathbb{N}\} = \frac{b^3}{3},$$

if follows from Lemma 4.2 that

$$\int_0^b f = \frac{b^3}{3}.$$

▲▲▲

This game can't go much further, i.e., we need stronger tools. We start by verifying which classes of functions are integrable.

(42 hrs, 2010)

4.2 Integration theorems

Theorem 4.3 If $f[a,b] \to \mathbb{R}$ is monotone, then it is integrable.

Proof: Assume, without loss of generality, that f is monotonically increasing. Then, for every partition P,

$$L(f,P) = \sum_{i=1}^n f(x_{i-1})(x_i - x_{i-1}) \quad \text{and} \quad U(f,P) = \sum_{i=1}^n f(x_i)(x_i - x_{i-1}),$$

hence,

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

Take now a family of equally-spaced partitions, $P_n = \{a + j(b-a)/n\}_{j=0}^n$. For such partitions,

$$U(f,P_n) - L(f,P_n) = \frac{b-a}{n} \sum_{i=1}^n [f(x_i) - f(x_{i-1})] = \frac{(b-a)[f(b) - f(a)]}{n}.$$

Thus, for every $\epsilon > 0$ we can choose $n > (b-a)[f(b) - f(a)]/\epsilon$, which guarantees that $U(f,P_n) - L(f,P_n) < \epsilon$ (we used the fact that the sum was "telescopic"). ■

Theorem 4.4 *If $f[a,b] \to \mathbb{R}$ is continuous, then it is integrable.*

Proof: First, a continuous function on a segment is bounded. We want to show that for every $\epsilon > 0$ there exists a partition P, such that

$$U(f,P) < L(f,P) + \epsilon.$$

We need to somehow control the difference between m_i and M_i. Recall that a function that is continuous on a closed interval is uniformly continuous. That is, there exists a $\delta > 0$, such that

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

Take then a partition P such that

$$x_i - x_{i-1} < \delta \qquad \text{for all } i = 1, \dots, n.$$

Since continuous functions on closed intervals assume minima and maxima in the interval, it follows that $M_i - m_i < \epsilon/(b-a)$, and

$$U(f,P) - L(f,P) = \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1}) < \frac{\epsilon}{b-a} \sum_{i=1}^{n} (x_i - x_{i-1}) = \epsilon.$$

■

(42 hrs, 2011)

Theorem 4.5 *Let f be a bounded function on $[a,b]$, and let $a < c < b$. Then, f is integrable on $[a,b]$ if and only if it is integrable on both $[a,c]$ and $[c,b]$, in which case*

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

Proof: (1) Suppose first that f is integrable on $[a,b]$. It is sufficient to show that it is integrable on $[a,c]$. For every $\epsilon > 0$ there exists a partition P of $[a,b]$, such that

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

If the partition P does not include the point c then let $Q = P \cup \{c\}$, and

$$U(f, Q) - L(f, Q) \le U(f, P) - L(f, P) < \epsilon.$$

Suppose that $x_j = c$, then $P' = \{x_0, \dots, x_j\}$ is a partition of $[a, c]$, and $P'' = \{x_j, \dots, x_n\}$ is a partition of $[c, b]$. We have

$$L(f, Q) = L(f, P') + L(f, P'') \qquad \text{and} \qquad U(f, Q) = U(f, P') + U(f, P''),$$

from which we deduce that

$$[U(f, P') - L(f, P')] = \underbrace{[U(f, Q) - L(f, Q)]}_{<\epsilon} - \underbrace{[U(f, P'') - L(f, P'')]}_{\ge 0} < \epsilon,$$

hence f is integrable on $[a, c]$.

(2) Conversely, suppose that f is integrable on both $[a, c]$ and $[c, b]$. Then there a exists a partition P' of $[a, c]$ and a partition P'' of $[c, b]$, such that

$$U(f, P') - L(f, P') < \frac{\epsilon}{2} \qquad \text{and} \qquad U(f, P'') - L(f, P'') < \frac{\epsilon}{2}.$$

Since $P = P' \cup P''$ is a partition of $[a, b]$, summing up, we obtain the desired result.

Thus, we have shown that f is integrable on $[a, b]$ if and only if it is integrable on both $[a, c]$ and $[c, b]$. Next, since

$$L(f, P') \le \int_a^c f \le U(f, P') \qquad \text{for every partition } P' \text{ of } [a, c]$$
$$L(f, P'') \le \int_c^b f \le U(f, P'') \qquad \text{for every partition } P'' \text{ of } [c, b],$$

it follows that

$$L(f, P) \le \int_a^c f + \int_c^b f \le U(f, P) \qquad \text{for every partition } P \text{ of } [a, b].$$

(More precisely, for every partition P that is a union of partitions of P' and P'', but it is easy to see that this implies for every partition.) Since there is only one such number, it is equal to $\int_a^b f$. ■

(39 hrs, 2009)

Comment: So far, we have only defined the integral on $[a, b]$ with $b > a$. We now add the following conventions,

$$\int_b^a f = -\int_a^b f \qquad \text{and} \qquad \int_a^a f = 0,$$

so that (as can easily be checked) the addition rule always holds.

Theorem 4.6 *If f and g are integrable on $[a, b]$, then so is $f + g$, and*

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

Proof: First of all, we observe that if f and g are bounded functions on a set A, then[4]

$$\inf\{f(x) + g(x) : x \in A\} \geq \inf\{f(x) : x \in A\} + \inf\{g(x) : x \in A\}$$
$$\sup\{f(x) + g(x) : x \in A\} \leq \sup\{f(x) : x \in A\} + \sup\{g(x) : x \in A\}.$$

This implies that for every partition P,

$$L(f+g, P) \geq L(f, P) + L(g, P)$$
$$U(f+g, P) \leq U(f, P) + U(g, P),$$

hence,

$$U(f+g, P) - L(f+g, P) \leq [U(f, P) - L(f, P)] + [U(g, P) - L(g, P)]. \tag{4.1}$$

We are practically done. Given $\epsilon > 0$ there exist partitions P', P'', such that

$$U(f, P') - L(f, P') < \frac{\epsilon}{2} \qquad \text{and} \qquad U(g, P'') - L(g, P'') < \frac{\epsilon}{2}.$$

[4]This is because for every x,

$$f(x) + g(x) \geq \inf\{f(x) : x \in A\} + \inf\{g(x) : x \in A\},$$

so that the right-hand side is a lower bound of $f + g$.

Take $P = P' \cup P''$, then

$$U(f,P) - L(f,P) \le U(f,P') - L(f,P') < \frac{\epsilon}{2}$$
$$U(g,P) - L(g,P) \le U(g,P'') - L(g,P'') < \frac{\epsilon}{2},$$

and it remains to substitute into (4.1), to prove that $f+g$ is integrable.
Now, for every partition P,

$$L(f,P) + L(g,P) \le L(f+g,P) \le \int_a^b (f+g) \le U(f+g,P) \le U(f,P) + U(g,P),$$

and also

$$L(f,P) + L(g,P) \le \int_a^b f + \int_a^b g \le U(f,P) + U(g,P),$$

It follows that

$$\left| \int_a^b (f+g) - \int_a^b f - \int_a^b g \right| \le [U(f,P) + U(g,P)] - [L(f,P) + L(g,P)].$$

Since the right-hand side can be made less than ϵ for every ϵ, it follows that it is zero[5]. ■

Corollary 4.1 *If f is integrable on $[a,b]$ and g differs from f only at a finite number of points then*

$$\int_a^b f = \int_a^b g.$$

Proof: We saw that $(g-f)$, which is non-zero only at a finite number of points, is integrable and its integral is zero. Then we only need to use the additivity of the integral for $g = f + (g-f)$. ■

✎ *Exercise 4.4* Denote by $\overline{\int}$ the upper integral. Let f,g be bounded on $[a,b]$. Show that

$$\overline{\int_a^b}(f+g) \le \overline{\int_a^b} f + \overline{\int_a^b} g.$$

Also, find two functions f,g for which the inequality is strong.

[5]If $|x| < \epsilon$ for every $\epsilon > 0$ then $x = 0$.

✎ *Exercise 4.5* Recall the Riemann function,

$$R(x) = \begin{cases} 1/q & x = p/q \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q}. \end{cases}$$

Let $[a, b]$ be a gievn interval.

① Show that $\underline{\int_a^b} R = 0$ (the lower integral).

② Let $\epsilon > 0$ and define

$$Z_\epsilon = \{x \in [a, b] : R(x) > \epsilon\}.$$

Show that Z_ϵ is a finite set.

③ Show that there exists a partition P such that $U(R, P) < \epsilon$.

④ Conclude that R is integrable on $[a, b]$ and that $\int_a^b R = 0$.

(44 hrs, 2010)

Theorem 4.7 If f is integrable on $[a, b]$ and $c \in \mathbb{R}$, then cf is integrable on $[a, b]$ and

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

Proof: Start with the case where $c > 0$. For every partition P,

$$L(cf, P) = c\, L(f, P) \qquad \text{and} \qquad U(cf, P) = c\, U(f, P).$$

Since f is integrable on $[a, b]$, then there exists for every $\epsilon > 0$ a partition P, such that

$$U(f, P) - L(f, P) < \frac{\epsilon}{c},$$

hence

$$U(cf, P) - L(cf, P) < \epsilon,$$

which proves that cf is integrable on $[a, b]$. Then, we note that for every partition P,

$$L(cf, P) = c\, L(f, P) \le c \int_a^b f \le c\, U(f, P) = U(cf, P).$$

Since there is only one number ℓ such that $L(cf, P) \le \ell \le U(cf, P)$ for all partitions P, and this number is by definition $\int_a^b (cf)$, then

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

Consider then the case where $c < 0$; in fact, it is sufficient to consider the case of $c = -1$, since for arbitrary $c < 0$ we write $cf = (-1)|c|f$. For every partition P,

$$L(-f, P) = -U(f, P) \qquad \text{and} \qquad U(-f, P) = -L(f, P).$$

Since f is integrable on $[a, b]$, then there exists for every $\epsilon > 0$ a partition P, such that

$$U(f, P) - L(f, P) < \epsilon,$$

hence

$$U(-f, P) - L(-f, P) = -(L(f, P) - U(f, P)) < \epsilon,$$

which proves that $(-f)$ is integrable on $[a, b]$. Then, we note that for every partition P,

$$L(-f, P) = -U(f, P) \le -\int_a^b f \le -L(f, P) = U(-f, P),$$

and the rest is as above. ■

(45 hrs, 2010)

This last theorem is in fact a special case of a more general theorem:

Theorem 4.8 *Suppose that both f and g are integrable on $[a, b]$, then so is their product $f \cdot g$.*

Proof: We will prove this theorem for the case where $f, g \ge 0$. The generalization to arbitrary functions is not hard (for example, since f and g are bounded, we can add to each a constant to make it positive and then use the "arithmetic of integrability").

Let P be a partition. It is useful to introduce the following notations:

$$\begin{aligned} m_j^f &= \inf\{f(x) : x_{j-1} \le x \le x_j\} \\ m_j^g &= \inf\{g(x) : x_{j-1} \le x \le x_j\} \\ m_j^{fg} &= \inf\{f(x)g(x) : x_{j-1} \le x \le x_j\} \\ M_j^f &= \sup\{f(x) : x_{j-1} \le x \le x_j\} \\ M_j^g &= \sup\{g(x) : x_{j-1} \le x \le x_j\} \\ M_j^{fg} &= \sup\{f(x)g(x) : x_{j-1} \le x \le x_j\}. \end{aligned}$$

We also define

$$\begin{aligned} M^f &= \sup\{|f(x)| : a \le x \le b\} \\ M^g &= \sup\{|g(x)| : a \le x \le b\}. \end{aligned}$$

For every $x_{j-1} \le x \le x_j$,

$$m_j^f m_j^g \le f(x)g(x) \le M_j^f M_j^g,$$

hence

$$m_j^f m_j^g \le m_j^{fg} \qquad \text{and} \qquad M_j^{fg} \le M_j^f M_j^g.$$

It further follows that

$$M_j^{fg} - m_j^{fg} \le M_j^f M_j^g - m_j^f m_j^g \le (M_j^f - m_j^f) M_j^g + m_j^f (M_j^g - m_j^g) \le M^g (M_j^f - m_j^f) + M^f (M_j^g - m_j^g).$$

This inequality implies that

$$U(fg, P) - L(fg, P) \le M^g [U(f, P) - L(f, P)] + M^f [U(g, P) - L(g, P)].$$

It is then a simple task to show that the integrability of f and g implies the integrability of fg. ■

Theorem 4.9 *If f is integrable on $[a, b]$ and $m \le f(x) \le M$ for all x, then*

$$m(b - a) \le \int_a^b f \le M(b - a).$$

Proof: Let $P_0 = \{x_0, x_n\}$, then

$$L(f, P_0) \geq m(b-a) \qquad \text{and} \qquad U(f, P_0) \leq M(b-a),$$

and the proof follows from the fact that

$$L(f, P_0) \leq \int_a^b f \leq U(f, P_0).$$

■

Suppose now that f is integrable on $[a, b]$. We proved that it is integrable on every sub-segment. Define for every x,

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

Theorem 4.10 *The function F defined above is continuous on* $[a, b]$.

Proof: Since f is integrable it is bounded. Define

$$M = \sup\{|f(x)| : a \leq x \leq b\}.$$

Given $\epsilon > 0$ take $\delta = \epsilon/M$. Then for $|x - y| < \delta$ (suppose wlog that $y > x$),

$$F(y) - F(x) = \int_a^y f - \int_a^x f = \int_x^y f.$$

By the previous theorem, since $-M \leq f(x) \leq M$ for all x,

$$-M(y-x) \leq F(y) - F(x) \leq M(y-x),$$

i.e., $|F(y) - F(x)| \leq M(y-x) < \epsilon$. That is, given $\epsilon > 0$, we take $\delta = \epsilon/M$, and then

$$|F(y) - F(x)| < \epsilon \qquad \text{whenever} \qquad |x - y| < \delta,$$

which proves that F is (uniformly) continuous on $[a, b]$. ■

(41 hrs, 2009)

✎ *Exercise 4.6* Let f be bounded and integrable in $[a, b]$ and denote $M = \sup\{|f(x)| : a \le x \le b\}$. Show that for every $p > 0$,

$$\left(\int_a^b |f|^p\right)^{1/p} \le M|b-a|^{1/p},$$

and in particular, that the integral on the left hand side exists.

✎ *Exercise 4.7* Show that if f is bounded and integrable on $[a, b]$, then $|f|$ is integrable on $[a, b]$ and

$$\left|\int_a^b f\right| \le \int_a^b |f|.$$

✎ *Exercise 4.8* Let $a < b$ and $c > 0$ and let f be integrable on $[ca, cb]$. Show that

$$\int_{ca}^{cb} f = c\int_a^b g,$$

where $g(x) = f(cx)$.

✎ *Exercise 4.9* Let f be integrable on $[a, b]$ and $g(x) = f(x - c)$. Show that g is integrable on $[a + c, b + c]$ and

$$\int_a^b f = \int_{a+c}^{b+c} g.$$

✎ *Exercise 4.10* Let $f : [a, b] \to \mathbb{R}$ be bounded, and define the variation of f in $[a, b]$ as

$$\omega_f = \sup\{f(x) : a \le x \le b\} - \inf\{f(x) : a \le x \le b\}.$$

① Prove that

$$\omega_f = \sup\{f(x_1)-f(x_2) : a \le x_1, x_2 \le b\} = \sup\{|f(x_1)-f(x_2)| : a \le x_1, x_2 \le b\}.$$

② Show that f is integrable on $[a, b]$ if and only if there exists for every $\epsilon > 0$ a partition $P = \{x_0, \dots, x_n\}$, such that $\sum_{i=1}^n \omega_i < \epsilon$, where ω_i is the variation of f in the segment $[x_{i-1}, x_i]$.

③ Prove that if f is integrable on $[a, b]$, and there exists an $m > 0$ such that $|f(x)| \ge m$ for all $x \in [a, b]$, then $1/f$ is integrable on $[a, b]$.

✎ *Exercise 4.11* A function $s : [a, b] \to \mathbb{R}$ is called a ***step function*** (פונקצית מדרגות) if there exists a partition p such that s is constant on every open interval (we do not care for the values at the nodes),

① Show that if f is integrable on $[a, b]$ then there exist for every $\epsilon > 0$ two step functions, $s_1 \le f$ and $s_2 \ge f$, such that

$$\int_a^b f - \int_a^b s_1 < \epsilon \quad \text{and} \quad \int_a^b s_2 - \int_a^b f < \epsilon.$$

② Show that if f is integrable on $[a, b]$ then there exists for every $\epsilon > 0$ a continuous function g, $g \le f$, such that

$$\int_a^b f - \int_a^b g < \epsilon.$$

Discussion: A question that occupied many mathematicians in the second half of the 19th century is what characterizes integrable functions. Hermann Hankel (1839-1873) proved that if a function is integrable, then it is necessarily continuous on a dense set of points (he also proved the opposite, but had an error in his proof). The ultimate answer is due to Henri Lebesgue (1875-1941) who proved that a bounded function is integrable if and only if its points of discontinuity have measure zero (a notion not defined in this course).

(44 hrs, 2011)

4.3 The fundamental theorem of calculus

Theorem 4.11 (המשפט היסודי של החשבון הדיפרנציאלי) Let f be integrable on $[a, b]$ and define for every $x \in [a, b]$,

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

If f is continuous at $c \in (a, b)$, then F is differentiable at c and

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

Comment: The theorem states that in a certain way the integral is an operation inverse to the derivative. Note the condition on f being continuous. For example, if $f(x) = 0$ everywhere except for the point $x = 1$ where it is equal to one, then $F(x) = 0$, and it is not true that $F'(1) = f(1)$.

Proof: We need to show that

$$\lim_{c} \Delta_{F,c} = f(c).$$

We are going to show that this holds for the right derivative; the same arguments can then be applied for the left derivative.

For $x > c$,

$$F(x) - F(c) = \int_x^c f.$$

If we define

$$m(c,x) = \inf\{f(y) : x < y < c\}$$
$$M(c,x) = \sup\{f(y) : x < y < c\},$$

then,

$$m(c,x)(x-c) \le F(x) - F(c) \le M(c,x)(x-c),$$

or

$$m(c,x) \le \Delta_{F,c}(x) \le M(c,x).$$

Then,

$$\left|\Delta_{F,c}(x) - f(c)\right| \le \max\{|m(c,x) - f(y)|, |M(c,x) - f(y)|\}.$$

It remains that the right-hand side as a limit zero at c. Since f is continuous at c,

$$(\forall \epsilon > 0)(\exists \delta > 0) : (\forall c < y < c+\delta)(|f(y) - f(c)| < \epsilon/2).$$

For $c < x < c + \delta$,

$$(\forall \epsilon > 0)(\exists \delta > 0) : (\forall c < y < x)(|f(y) - f(c)| < \epsilon/2),$$

hence

$$|m(c,x) - f(c)| \le \epsilon/2 < \epsilon \qquad \text{and} \qquad |M(c,x) - f(c)| \le \epsilon/2 < \epsilon.$$

To conclude, for every $\epsilon > 0$ there exists a $\delta > 0$ such that for all $c < x < c + \delta$,

$$\left|\Delta_{F,c}(x) - f(c)\right| < \epsilon,$$

which concludes the proof. ∎

Corollary 4.2 (Newton-Leibniz formula) *If f is continuous on $[a,b]$ and $f = g'$ for some function g, then*

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

Proof: We have seen that

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

satisfies the identity $F'(x) = f(x)$ for all x (since f is continuous on $[a,b]$). We have also seen that it implies that $F = g + c$, for some constant c, hence

$$\int_a^x f = F(x) = g(x) + c,$$

but since $F(a) = 0$ it follows that $c = -g(a)$. Setting $x = b$ we get the desired result. ■

This is a very useful result. It means that whenever we know a pair of functions f, g, such that $g' = f$, we have an easy way to calculate the integral of f. On the other hand, it is totally useless if we want to calculate the integral of f, but no ***primitive function*** (פונקציה קדומה) g is known.

(43 hrs, 2009) (47 hrs, 2010)

✎ *Exercise 4.12* Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and $g, h : (a,b) \to \mathbb{R}$ differentiable. Define $F : (a,b) \to \mathbb{R}$ by

$$F(x) = \int_{g(x)}^{h(x)} f.$$

① Show that F is well-defined for all $x \in (a,b)$.

② Show that F is differentiable, and calculate its derivative.

✎ *Exercise 4.13* Differentiate the following functions:

① $\int_0^1 \cos(t^2)\,dt$.

② $\int_0^{e^x} \cos(t)\,dt$.

③ $\int_x^{x^2} e^{\sin(t)}\, dt$.

✎ *Exercise 4.14* Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. For every $\delta > 0$ we define

$$F_\delta(x) = \frac{1}{2\delta} \int_{x-\delta}^{x+\delta} f.$$

Show that for every $x \in \mathbb{R}$,

$$f(x) = \lim_{\delta \to 0^+} F_\delta(x).$$

✎ *Exercise 4.15* Let $f, g : [a, b] \to \mathbb{R}$ be integrable, such that $g(x) \geq 0$ for all $x \in [a, b]$.

① Show that there exists a $\lambda \in \mathbb{R}$, such that

$$\inf\{f(x) : a \leq x \leq b\} \leq \lambda \leq \sup\{f(x) : a \leq x \leq b\},$$

and

$$\int_a^b (fg) = \lambda \int_a^b g.$$

Hint: distinguish between the cases where $\int_a^b g > 0$ and $\int_a^b g = 0$.

② Show that if f is continuous, then there exists a $c \in [a, b]$, such that

$$\int_a^b (fg) = f(c) \int_a^b g.$$

(This is known as the ***integral mean value theorem***.)

✎ *Exercise 4.16* Let $f : [a, b] \to \mathbb{R}$ be monotonic and let $g : [a, b] \to \mathbb{R}$ be non-negative and integrable. Show that there exists a $\xi \in [a, b]$, such that

$$\int_a^b (fg) = f(a) \int_a^\xi g + f(b) \int_\xi^b g.$$

✎ *Exercise 4.17* Let $f : [a, b] \to \mathbb{R}$ be integrable, $c \in (a, b)$, and $F(x) = \int_a^x f$. Prove or disprove:

① If f is differentiable at c then F is differentiable at c.

② If f is differentiable at c then F' is continuous at c.

③ If f is differentiable in a neighborhood of c and f' is continuous at c then F is twice differentiable at c.

4.4 Riemann sums

Let's ask a practical question. Let f be a given integrable function on $[a, b]$, and suppose we want to calculate the integral of f on $[a, b]$, however we do not know any primitive function of f. Suppose that we are content with an ***approximation*** of this integral within a maximum error of ϵ. Take a partition P: since

$$L(f, P) \le \int_a^b f \le U(f, P),$$

it follows that if the difference between the upper and lower sum is less than ϵ, then their average differs from the integral by less that $\epsilon/2$. If the difference is greater than ϵ, then it can be reduced by refining the partition. Thus, up to the fact that we have to compute lower and upper sums, we can keep refining until they differ by less than ϵ, and so we obtain an approximation to the integral within the desired accuracy.

There is one drawback in this plan. Infimums and supremums may be hard to calculate. On the other hand, since $m_i \le f(x) \le M_i$ for all $x_{i-1} \le x \le x_i$, it follows that if in every interval in the partition we take some point t_i (a ***sample point***), then

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

The sum in the middle is called a ***Riemann sum***. Its evaluation only requires function evaluations, without any inf or sup. If the partition is sufficiently fine, such that $U(f, P) - L(f, P) < \epsilon$, then

$$\left| \sum_{i=1}^n f(t_i)(x_{i-1} - x_i) - \int_a^b f \right| < \epsilon.$$

Note that we don't know a priori whether a Riemann is greater or smaller than the integral. What we are going to show is that Riemann sums converge to the integral as we refine the partition, irrespectively on how we sample the points t_i.

*Definition 4.5 The **diameter** of a partition P is*

$$\text{diam}\, P = \max\{x_i - x_{i-1} : 1 \le i \le n\}.$$

(45 hrs, 2011)

Theorem 4.12 Let f be an integrable function on $[a,b]$. For every $\epsilon > 0$ there exists a $\delta > 0$, such that

$$\left| \sum_{i=1}^{n} f(t_i)(x_{i-1} - x_i) - \int_a^b f \right| < \epsilon$$

for every partition P such that $\operatorname{diam} P < \delta$ and any choice of points t_i.

Proof: It is enough to show that given $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|U(f,P) - L(f,P)| < \epsilon \qquad \text{whenever} \qquad \operatorname{diam} P < \delta,$$

for any Riemann sum with the same partition will differ from the integral by less than ϵ.

Let

$$M = \sup\{|f(x)| : a \le x \le b\},$$

and take some partition $P' = \{y_0, \dots, y_k\}$ for which

$$|U(f,P') - L(f,P')| < \frac{\epsilon}{2}.$$

We set

$$\delta < \frac{\epsilon}{4Mk}.$$

For every partition P of diameter less than δ,

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

can be divided into two sums. One for which each of the intervals $[x_{i-1}, x_i]$ is contained in an interval $[y_{j-1}, y_j]$. The contribution of these terms is at most $\epsilon/2$. The second sum includes those intervals for which $x_{i-1} < y_j < x_i$. There are at most k such terms and each contributes at most $M\delta$. This concludes the proof. ■

4.5 The trigonometric functions

Discussion: Describe the way the trigonometric functions are defined in geometry, and explain that we want definitions that only rely on the 13 axioms of real numbers.

Definition 4.6 *The number π is defined as*

$$\pi = 2 \int_{-1}^{1} \sqrt{1 - x^2}\, dx.$$

Comment: We have used here the standard notation of integrals. In our notation, we have a function $\varphi : [-1, 1] \to \mathbb{R}$ defined as $\varphi : x \mapsto \sqrt{1 - x^2}$, and we define

$$\pi = 2 \int_{-1}^{1} \varphi.$$

This is well defined because φ is continuous, hence integrable.

Consider now a unit circle. For $-1 \le x \le 1$ we look at the point $(x, \sqrt{1 - x^2})$, connect it to the origin, and express the area of the sector bounded by this segment and the positive x-axis. This geometric construction is just the motivation; formally, we define a function $A : [-1, 1] \to \mathbb{R}$,

$$A(x) = \frac{x\sqrt{1 - x^2}}{2} + \int_{x}^{1} \varphi.$$

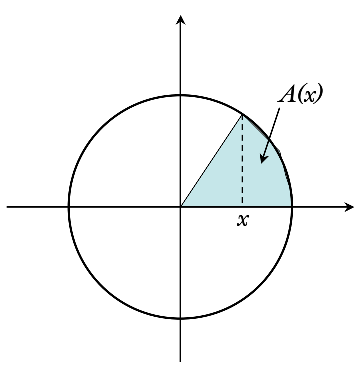
If θ is the angle of this sector, then we know that $x = \cos\theta$, and that $A(x) = \theta/2$. Thus,

$$A(\cos\theta) = \frac{\theta}{2},$$

or,

$$\cos\theta = A^{-1}\left(\frac{\theta}{2}\right).$$

This observation motivates the way we are going to define the cosine function.



We now examine some properties of the function A. First, $A(-1) = \pi/2$ and $A(1) = 0$. Second,

$$A'(x) = -\frac{1}{2\sqrt{1-x^2}},$$

i.e., the function A is monotonically decreasing. It follows that it is invertible, with its inverse defined on $[0, \pi/2]$.

Definition 4.7 *We define the functions* $\cos : [0, \pi] \to \mathbb{R}$ *and* $\sin : [0, \pi] \to \mathbb{R}$ *by*

$$\cos x = A^{-1}(x/2) \qquad \text{and} \qquad \sin x = \sqrt{1-\cos^2 x}.$$

With these definitions in hand we define the other trigonometric functions, tan*,* cot*, etc.*

We now need to show that the sine and cosine functions satisfy all known differential and algebraic properties.

Theorem 4.13 *For* $0 < x < \pi$*,*

$$\sin' x = \cos x \qquad \text{and} \qquad \cos' x = -\sin x.$$

Proof: Since A is differentiable and $A'(x) \neq 0$ for $0 < x < \pi$, we have by the chain rule and the derivative of the inverse function formula,

$$\cos' x = \frac{1}{2}(A^{-1})'(x/2) = \frac{1}{2}\frac{1}{A'(A^{-1}(x/2))} = -\frac{1}{2}\frac{1}{\frac{1}{2\sqrt{1-(A^{-1}(x/2))^2}}} = \cdots = -\sin x.$$

The derivative of the sine function is then obtained by the composition rule. ■

(45 hrs, 2009)

We are now in measure to start picturing the graphs of these two functions. First, by definition,

$$\sin x > 0 \qquad 0 < x < \pi.$$

It thus follows that cos is decreasing in this domain. We also have

$$\cos \pi = A^{-1}(\pi/2) = -1 \qquad \text{and} \qquad \cos 0 = A^{-1}(0) = 1,$$

from which follows that

$$\sin \pi = 0 \qquad \text{and} \qquad \sin 0 = 0.$$

By the intermediate value theorem, cos must vanish for some $0 < y < \pi$. Since

$$A(\cos y) = \frac{y}{2},$$

it follows that

$$y = 2A(0) = 2\int_0^1 \varphi = \frac{\pi}{2},$$

where we have used the fact that by symmetry,

$$\int_{-1}^0 \varphi = \int_0^1 \varphi.$$

Thus,

$$\cos\frac{\pi}{2} = 0 \qquad \text{and} \qquad \sin\frac{\pi}{2} = 1.$$

We can then extend the definition of the sine and cosine to all of $\mathbb{R}$ by setting

$$\sin x = -\sin(2\pi - x) \quad \text{and} \quad \cos x = \cos(2\pi - x) \qquad \text{for } \pi \le x \le 2\pi,$$

and

$$\sin(x + 2\pi k) = \sin x \quad \text{and} \quad \cos(x + 2\pi k) = \cos x \qquad \text{for all } k \in \mathbb{Z},$$

It is easy to verify that the expressions for the derivatives of sine and cosine remain the same for all $x \in \mathbb{R}$.

Since we defined the trigonometric functions (more precisely their inverses) via integrals, it is not surprising that it was straightforward to differentiate them. A little more subtle is to derive their algebraic properties.

Lemma 4.3 Let $f : \mathbb{R} \to \mathbb{R}$ be twice differentiable, and satisfy the ***differential equation***

$$f'' + f = 0,$$

with $f(0) = f'(0) = 0$. Then, f is identically zero.

Proof: Multiply the equation by $2f'$, then

$$2f''f' + 2ff' = [(f')^2 + f^2]' = 0,$$

from which we deduce that

$$(f')^2 + f^2 = c,$$

for some $c \in \mathbb{R}$. The conditions at zero imply that $c = 0$, which implies that f is identically zero. ■

Theorem 4.14 *Let $f : \mathbb{R} \to \mathbb{R}$ be twice differentiable, and satisfy the differential equation*

$$f'' + f = 0,$$

with $f(0) = a$ and $f'(0) = b$. Then,

$$f = a\cos + b\sin.$$

Proof: Define

$$g = f - a\cos - b\sin,$$

then,

$$g'' + g = 0,$$

and $g(0) = g'(0) = 0$. It follows from the lemma that g is identically zero, hence $f = a\cos + b\sin$. ■

Theorem 4.15 *For every $x, y \in \mathbb{R}$,*

$$\sin(x + y) = \sin x \cos y + \cos x \sin y$$
$$\cos(x + y) = \cos x \cos y - \sin x \sin y.$$

Proof: Fix y and define

$$f : x \mapsto \sin(x + y),$$

from which follows that

$$f' : x \mapsto \cos(x+y)$$
$$f'' : x \mapsto -\sin(x+y).$$

Thus, $f'' + f = 0$ and

$$f(0) = \sin y \qquad \text{and} \qquad f'(0) = \cos y.$$

It follows from the previous theorem that

$$f(x) = \sin y \, \cos x + \cos y \, \sin x.$$

We proceed similarly for the cosine function. ■

4.6 The logarithm and the exponential

Our (starting) goal is to define the function

$$f : x \mapsto 10^x$$

for $x \in \mathbb{R}$, along with its inverse, which we denote by $\log_{10}$.
For $n \in \mathbb{N}$ we have an inductive definition, which in particular, satisfies the identity

$$10^n 10^m = 10^{n+m}.$$

We then extend the definition of the exponential (with base 10) to integers, by requiring the above formula to remain correct. Thus, we must set

$$10^0 = 1 \qquad \text{and} \qquad 10^{-n} = \frac{1}{10^n}.$$

The next extension is to fractional powers of the form $10^{1/n}$. Since we want

$$\underbrace{10^{1/n} \times 10^{1/n} \times \cdots \times 10^{1/n}}_{n \text{ times}} = 10^1 = 10,$$

this forces the definition $10^{1/n} = \sqrt[n]{10}$. Then, we extend the definition to rational powers, by

$$\underbrace{10^{1/n} \times 10^{1/n} \times \cdots \times 10^{1/n}}_{m \text{ times}} = 10^{m/n}.$$

We have proved in the beginning of this course that the definition of a rational power is independent of the representation of that fraction.

The question is how to extend the definition of 10^x when x is a real number? This can be done with lots of "sups" and "infs", but here we shall adopt another approach. We are looking for a function f with the property that

$$f(x+y) = f(x)\, f(y),$$

and $f(1) = 10$. Can we find a *differentiable* function that satisfies these properties? If there was such a function, then

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h\to 0} \frac{f(x)f(h) - f(x)}{h} = \left(\lim_{h\to 0} \frac{f(h) - 1}{h}\right) f(x).$$

Suppose that the limit on the right hand side exists, and denote it by α. Then,

$$f'(x) = \alpha\, f(x) \qquad \text{and} \qquad f(1) = 10.$$

This is not very helpful, as the derivative of f is expressed in terms of the (unknown) f.

It is clear that f should be monotonically increasing (it is so for rational powers), hence invertible. Then,

$$(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))} = \frac{1}{\alpha f(f^{-1}(x))} = \frac{1}{\alpha x}.$$

From this we deduce that

$$f^{-1}(x) = \int_1^x \frac{dt}{\alpha t} = \log_{10}(x).$$

We seem to have succeeded in defining (using integration) the logarithm-base-10 function, but there is a caveat. We do not know the value of α. Since, in any case, the so far distinguished role of the number 10 is unnatural, we define an alternative logarithm function:

Definition 4.8 *For all $x > 0$,*

$$\log(x) = \int_1^x \frac{dt}{t}.$$

Comment: Since $1/x$ is continuous on $(0, \infty)$, this integral is well-defined for all $0 < x < \infty$.

Theorem 4.16 *For all $x, y > 0$,*

$$\log(xy) = \log(x) + \log(y).$$

Proof: Fix y and consider the function

$$f : x \mapsto \log(xy).$$

Then,

$$f'(x) = \frac{1}{xy} \cdot y = \frac{1}{x}.$$

Since $\log'(x) = 1/x$, it follows that

$$\log(xy) = \log(x) + c.$$

The value of the constant is found by setting $x = 1$. ■

Corollary 4.3 *For integer n,*

$$\log(x^n) = n \log(x).$$

Proof: By induction on n. ■

Corollary 4.4 *For all $x, y > 0$,*

$$\log\left(\frac{x}{y}\right) = \log(x) - \log(y).$$

Proof:

$$\log(x) = \log\left(\frac{x}{y}y\right) = \log\left(\frac{x}{y}\right) + \log(y).$$

■

Definition 4.9 *The exponential function is defined as*

$$\exp(x) = \log^{-1}(x).$$

Comment: The logarithm is invertible because it is monotonically increasing, as its derivative is strictly positive. What is less easy to see is that the domain of the exponential is the whole of $\mathbb{R}$. Indeed, since

$$\log(2^n) = n\,\log(2),$$

it follows that the logarithm is not bounded above. Similarly, since

$$\log(1/2^n) = -n\,\log(2),$$

it is neither bounded below. Thus, its range is $\mathbb{R}$.

Theorem 4.17

$$\exp' = \exp.$$

Proof: We proceed as usual,

$$\exp'(x) = (\log^{-1})'(x) = \frac{1}{\log'(\log^{-1}(x))} = \log^{-1}(x) = \exp(x).$$

■

Theorem 4.18 *For every x, y,*

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

Proof: Set $t = \exp(x)$ and $s = \exp(y)$. Thus, $x = \log(t)$ and $y = \log(s)$, and

$$x + y = \log(t) + \log(s) = \log(ts).$$

Exponentiating both sides, we get the desired result. ■

Definition 4.10

$$e = \exp(1).$$

There is plenty more to be done with exponentials and logarithms, like for example, show that for rational numbers r, $e^r = \exp(r)$, and the passage between bases. Due to the lack of time, we will stop at this point.

(47 hrs, 2011)

4.7 Integration methods

Given a function f, a question of practical interest is how to calculate integrals,

$$\int_a^b f.$$

If we know a function F, which is a primitive of f, in the sense that $F' = f$, then the answer is known,

$$\int_a^b f = F(b) - F(a) \stackrel{\text{def}}{=} F|_a^b.$$

To effectively use this formula, we must know many pairs of functions f, F. Here is a table of pairs f, F, which we already know:

$f(x)$	$F(x)$
1	x
x^n	$x^{n+1}/(n+1)$
$\cos x$	$\sin x$
$\sin x$	$-\cos x$
$1/x$	$\log(x)$
$\exp(x)$	$\exp(x)$
$\sec^2(x)$	$\tan(x)$
$1/(1+x^2)$	$\arctan(x)$
$1/\sqrt{1-x^2}$	$\arcsin(x)$

Sums of such functions and multiplication by constants are easily treated with our integration theorems, for

$$\int_a^b (f+g) = \int_a^b f + \int_a^b g \quad \text{and} \quad \int_a^b (cf) = c\int_a^b f.$$

Functions that can be expressed in terms of powers, trigonometric, exponential and logarithmic functions and called ***elementary***. In this section we study techniques that in some cases allow to express a primitive function in terms of elementary functions. There are basically two techniques—integration by parts, and substitution—both we will study.

Comment: Every continuous function f has a primitive,

$$F(x) = \int_a^x f,$$

where a is arbitrary. It is however not always the case that the primitive of an elementary function is itself and elementary function.

Example: Consider the function

$$F(x) = x\arctan(x) - \frac{1}{2}\log(1+x^2).$$

It is easily checked that

$$f(x) = F'(x) = \arctan(x).$$

As a result of this observation,

$$\int_a^b \arctan(x)\,dx = F|_a^b.$$

The question is how could have we guessed this answer if the function F had not been given to us. ▲▲▲

Theorem 4.19 (Integration by parts (אינטגרציה בחלקים)) Let f, g have continuous first derivatives, then

$$\int_a^b fg' = (fg)|_a^b - \int_a^b f'g.$$

Comment: When we use this formula, f and g' are given. Then, g on the right-hand side may be any primitive of g'. Also, setting $b = x$ and noting that primitives are defined up to constants, we have that

$$\text{Primitive}(fg') = f\,\text{Primitive}(g') - \text{Primitive}(f'\text{Primitive}(g)) + \text{const}.$$

Example:

$$\int_a^b x\,e^x\,dx.$$

▲▲▲

Example: The 1-trick:

$$\int_a^b \log(x)\,dx.$$

▲▲▲

Example: The back-to-original-problem-trick

$$\int_a^b (1/x)\log(x)\,dx.$$

▲▲▲

Once we have used this method to get new f, F pairs, we use this knowledge to obtain new such pairs:

Example: Using the fact that a primitive of $\log(x)$ is $x\,\log(x) - x$, we solve

$$\int_a^b \log^2(x)\,dx.$$

▲▲▲

Theorem 4.20 (Substitution formula (שיטת ההצבה)) For continuously differentiable f and g,

$$\int_a^b (g\circ f)\,f' = \int_{f(a)}^{f(b)} g.$$

Proof: Let G be a primitive of g, then

$$(G \circ f)' = (g \circ f) f',$$

hence

$$\int_a^b (g \circ f) f' = \int_a^b (G \circ f)' = (G \circ f)\big|_a^b = G\big|_{f(a)}^{f(b)} = \int_{f(a)}^{f(b)} g.$$

■

The simplest use of this formula is by recognizing that a function is of the form $(g \circ f) f'$.

Example: Take the integral

$$\int_a^b \sin^5(x)\,\cos(x)\,dx.$$

If $f(x) = \sin(x)$ and $g(x) = x^5$, then

$$\sin^5(x)\,\cos(x) = g(f(x))\,f'(x),$$

hence

$$\int_a^b \sin^5(x)\,\cos(x)\,dx = \int_{\sin(a)}^{\sin(b)} t^5\,dt = \frac{t^6}{6}\bigg|_{\sin(a)}^{\sin(b)}.$$

▲▲▲

(46 hrs, 2009)

(48 hrs, 2011)

Chapter 5

Sequences

5.1 Basic definitions

An (infinite) sequence (סדרה) is an infinite ordered set of real numbers. We denote sequences by a letter which labels the sequence, and a subscript which labels the position in the sequence, for example,

$$a_1, a_2, a_3, \dots .$$

The essential fact is that to each natural number corresponds a real number. Having said that, we can define:

Definition 5.1 *An infinite sequence is a function* $\mathbb{N} \to \mathbb{R}$*.*

Comment: Let $a : \mathbb{N} \to \mathbb{R}$. Rather than writing $a(3), a(17)$, we write a_3, a_{17}.

Comment: As for functions, where we often refer to "the function x^2", we often refer to "the sequence $1/n$".

Notation: We usually denote a sequence $a : \mathbb{N} \to \mathbb{R}$ by (a_n). Another ubiquitous notation is $(a_n)_{n=1}^{\infty}$.

Examples:

① $a : n \mapsto n$ or $a_n = n$.

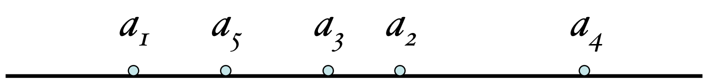
② $b : n \mapsto (-1)^n$ or $b_n = (-1)^n$.

③ $c : n \mapsto 1/n$ or $c_n = 1/n$.

④ $(d_n) = \{2, 3, 5, 7, 11, \dots\}$, the sequence of primes.

In the first and fourth cases we would say that "the sequence goes (שואפת) to infinity". In the second case we would say that "it jumps between -1 and 1". In the third case we would say that "it goes to zero". We will make such statements precise momentarily.

Like functions, sequences can be graphed. More useful often is to mark the elements of the sequence on the number axis.



5.2 Limits of sequences

Definition 5.2 *A sequence (a_n)* ***converges*** *(מתכנסת) to ℓ, denoted*

$$\lim_{n\to\infty} a_n = \ell,$$

if

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\forall n > N)(|a_n - \ell| < \epsilon).$$

Definition 5.3 *A sequence is called* ***convergent*** *(מתכנסת) if it converges to some (real) number; otherwise it is called* ***divergent*** *(מתבדרת) (there is no such thing as convergence to infinity).*

Comment: There is a close connection between the limit of a sequence and the limit of a function at infinity. We will elaborate on that later on.

Notation: Like for functions, a "cleaner" notation for the limit of a sequence (a_n) would be

$$\lim_{\infty} a.$$

Example: Let $a : n \mapsto 1/n$, or

$$(a_n) = 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$$

It is easy to show that

$$\lim_{\infty} a = 0,$$

since,

$$(\forall \epsilon > 0)(\text{set } N = \lceil 1/\epsilon \rceil) \quad \text{then} \quad (\forall n > N)\left(|a_n| = \frac{1}{n} < \frac{1}{N} < \epsilon \right).$$

▲▲▲

Example: Let $a : n \mapsto \sqrt{n+1} - \sqrt{n}$, or

$$(a_n) = \sqrt{2} - \sqrt{1}, \sqrt{3} - \sqrt{2}, \sqrt{4} - \sqrt{3}, \dots .$$

It seems reasonable that

$$\lim_{\infty} a = 0,$$

however showing it requires some manipulations. One way is by an algebraic trick,

$$a_n = \frac{(\sqrt{n+1} - \sqrt{n})(\sqrt{n+1} + \sqrt{n})}{\sqrt{n+1} + \sqrt{n}} = \frac{1}{\sqrt{n+1} + \sqrt{n}} < \frac{1}{2\sqrt{n}},$$

from which it is easy to proceed. The same inequality can be obtained by defining a function $f : x \mapsto \sqrt{x}$, and using the mean-value theorem,

$$\sqrt{y} - \sqrt{x} = \frac{y - x}{2\sqrt{\xi}},$$

for some $x < \xi < y$. Substituting $x = n$ and $y = n + 1$, we obtain the desired inequality. ▲▲▲

Example: Let now

$$a : n \mapsto \frac{3n^3 + 7n^2 + 1}{4n^3 - 8n + 63}.$$

It seems reasonable that

$$\lim_{\infty} a = \frac{3}{4},$$

but once again, proving it is not totally straightforward. One way to proceed is to note that

$$a_n = \frac{3 + 7/n + 1/n^3}{4 - 8/n^2 + 63/n^3},$$

and use limit arithmetic as stated below. ▲▲▲

✎ *Exercise 5.1* For each of the following sequence determine whether it is converging or diverging. Prove your claim using the definition of the limit:

① $a_n = \frac{7n^2+\sqrt{n}}{n^2+3n}$.
② $a_n = \frac{n^4+1}{n^3-3}$.
③ $a_n = \sqrt{n^2+1} - n$.

Theorem 5.1 Let a and b be convergent sequences. Then the sequences $(a+b)_n = a_n + b_n$ *and* $(a \cdot b)_n = a_n b_n$ *are also convergent, and*

$$\lim_{\infty}(a+b) = \lim_{\infty} a + \lim_{\infty} b$$
$$\lim_{\infty}(a \cdot b) = \lim_{\infty} a \cdot \lim_{\infty} b.$$

If furthermore $\lim_{\infty} b \neq 0$, *then there exists an* $N \in \mathbb{N}$ *such that* $b_n \neq 0$ *for all* $n > N$, *the sequence* (a/b) *converges and*

$$\lim_{\infty} \frac{a}{b} = \frac{\lim_{\infty} a}{\lim_{\infty} b}.$$

Comment: Since the limit of a sequence is only determined by the "tail" of the sequence, we do not care if a *finite* number of elements is not defined.

Proof: Easy once you've done it for functions. ■

✎ *Exercise 5.2* Show that if a and b are converging sequences, such that

$$\lim_{\infty} a = L \qquad \text{and} \qquad \lim_{\infty} b = M,$$

then

$$\lim_{\infty}(a + b) = L + M.$$

We now establish the analogy between limits of sequences and limits of functions. Let a be a sequence, and define a piecewise-constant function $f : [1, \infty) \to \mathbb{R}$ by

$$f(x) = a_n \qquad n \le x \le n + 1.$$

Proposition 5.1

$$\lim_{\infty} a = \ell \quad \textit{if and only if} \quad \lim_{\infty} f = \ell.$$

Proof: Suppose first that the sequence converges. Thus,

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\forall n > N)(|a_n - \ell| < \epsilon).$$

Since for every x, $f(x) = f(\lfloor x \rfloor)$, it follows that for all $x > N + 1$,

$$|f(x) - \ell| \le |a_{\lfloor x \rfloor} - \ell| < \epsilon,$$

hence

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\forall x > N)(|f(x) - \ell| < \epsilon).$$

Conversely, if f converges to ℓ as $x \to \infty$, then,

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\forall x > N)(|f(x) - \ell| < \epsilon),$$

and this is in particular true for $x \in \mathbb{N}$. ■

Conversely, given a function $f : \mathbb{R} \to \mathbb{R}$, it is often useful to define a sequence $a_n = f(n)$. If $f(x)$ converges as $x \to \infty$, then so does the sequence (a_n).

Example: Let $0 < a < 1$, then

$$\lim_{n\to\infty} a^n = ?$$

We define $f(x) = a^x$. Then

$$f(x) = e^{x \log a},$$

and $a_n = f(n)$. Since $\log a < 0$ it follows that $\lim_{x\to\infty} f(x) = 0$, hence the sequence converges to zero. ▲▲▲

Comment: Like for functions, we may define convergence of sequences to $\pm\infty$, only that we don't use the term convergence, but rather say that a sequence ***approaches*** (שואפת) infinity.

Theorem 5.2 Let f be defined on a (punctured) neighborhood of a point c. Suppose that

$$\lim_c f = \ell.$$

If a is a sequence whose entries belong to the domain of f, such that $a_n \neq c$, and

$$\lim_\infty a = c,$$

then

$$\lim_\infty f(a) = \ell,$$

where the composition of a function and a sequence is defined by

$$[f(a)]_n = f(a_n).$$

Conversely, if $\lim_\infty f(a) = \ell$ *for* every *sequence a that converges to c, then* $\lim_c f = \ell$.

Comment: This theorem provides a characterization of the limit of a function at a point in terms of sequences. This is known as ***Heine's characterization*** of the limit.

Proof: Suppose first that

$$\lim_c f = \ell.$$

Thus,

$$(\forall \epsilon > 0)(\exists \delta > 0) : (\forall x : 0 < |x - c| < \delta)(|f(x) - \ell| < \epsilon).$$

Let a be a sequence as specified in the theorem, namely, $a_n \neq c$ and $\lim_\infty a = c$. Then

$$(\forall \delta > 0)(\exists N \in \mathbb{N}) : (\forall n > N)(0 < |a_n - c| < \delta).$$

Combining the two,

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\forall n > N)(|f(a_n) - \ell| < \epsilon),$$

i.e., $\lim_\infty f(a) = \ell$.

Conversely, suppose that for any sequence a as above $\lim_\infty f(a) = \ell$. Suppose, by contradiction that

$$\lim_c f \neq \ell$$

(i.e., either the limit does not exist or it is not equal to ℓ). This means that

$$(\exists \epsilon > 0) : (\forall \delta > 0)(\exists x : 0 < |x - c| < \delta) : (|f(x) - \ell| \geq \epsilon).$$

In particular, setting $\delta_n = 1/n$,

$$(\exists \epsilon > 0) : (\forall n \in \mathbb{N})(\exists a_n : 0 < |a_n - c| < \delta_n) : (|f(a_n) - \ell| \geq \epsilon).$$

The sequence a converges to c but $f(a)$ does not converge to ℓ, which contradicts our assumptions. ■

Example: Consider the sequence

$$a : n \mapsto \sin\left(13 + \frac{1}{n^2}\right).$$

This sequence can be written as $a = \sin(b)$, where $b_n = 13 + 1/n^2$. Since

$$\lim_{13} \sin = \sin(13) \qquad \text{and} \qquad \lim_\infty b = 13,$$

it follows that

$$\lim_\infty a = \sin(13).$$

▲▲▲

(48 hrs, 2009)

Theorem 5.3 A convergent sequence is bounded.

Proof: Let a be a sequence that converges to a limit α. We need to show that there exists L_1, L_2 such that

$$L_1 \le a_n \le L_2 \qquad \forall n \in \mathbb{N}.$$

By definition, setting $\epsilon = 1$,

$$(\exists N \in \mathbb{N}) : (\forall n > N)(|a_n - \alpha| < 1),$$

or,

$$(\exists N \in \mathbb{N}) : (\forall n > N)(\alpha - 1 < a_n < \alpha + 1).$$

Set

$$M = \max\{a_n : 1 \le n \le N\}$$
$$m = \min\{a_n : 1 \le n \le N\}.$$

then for all n,

$$\min(m, \alpha - 1) \le a_n \le \max(M, \alpha + 1).$$

■

Theorem 5.4 Let a be a bounded sequence and let b be a sequence that converges to zero. Then

$$\lim_{n\to\infty}(ab) = 0.$$

Proof: Let M be a bound on a, namely,

$$|a_n| \le M \qquad \forall n \in \mathbb{N}.$$

Since b converges to zero,

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\forall n > N)\left(|b_n| < \frac{\epsilon}{M}\right).$$

Thus,

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\forall n > N)\left(|a_n b_n| \le M|b_n| < \epsilon\right),$$

which implies that the sequence ab converges to zero. ■

Example: The sequence

$$a : n \mapsto \frac{\sin n}{n}$$

converges to zero. ▲▲▲

Theorem 5.5 *Let a be a sequence with non-zero elements, such that*

$$\lim_{\infty} a = 0.$$

Then

$$\lim_{\infty} \frac{1}{|a|} = \infty.$$

Proof: By definition,

$$(\forall M > 0)(\exists N \in \mathbb{N}) : (\forall n > N)(0 < |a_n| < 1/M),$$

hence

$$(\forall M > 0)(\exists N \in \mathbb{N}) : (\forall n > N)(1/|a_n| > M),$$

which concludes the proof. ∎

Theorem 5.6 *Let a be a sequence such that*

$$\lim_{\infty} |a| = \infty.$$

Then

$$\lim_{\infty} \frac{1}{a} = 0.$$

Proof: By definition,

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\forall n > N)(|a_n| > 1/\epsilon),$$

hence

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\forall n > N)(0 < 1/|a_n| < \epsilon),$$

which concludes the proof. ∎

Theorem 5.7 *Suppose that a and b are convergent sequences,*

$$\lim_{\infty} a = \alpha \qquad \text{and} \qquad \lim_{\infty} b = \beta,$$

and $\beta > \alpha$. Then there exists an $N \in \mathbb{N}$, such that

$$b_n > a_n \qquad \text{for all} \qquad n > N,$$

i.e., the sequence b is eventually greater (term-by-term) than the sequence a.

Proof: Since $(\beta - \alpha)/2 > 0$,

$$(\exists N_1 \in \mathbb{N}) : (\forall n > N_1) \left(a_n < \alpha + \tfrac{1}{2}(\beta - \alpha) \right),$$

and

$$(\exists N_2 \in \mathbb{N}) : (\forall n > N_2) \left(b_n > \beta - \tfrac{1}{2}(\beta - \alpha) \right).$$

Thus,

$$(\exists N_1, N_2 \in \mathbb{N}) : (\forall n > \max(N_1, N_2)) \, (b_n - a_n > 0).$$

■

(50 hrs, 2011)

Corollary 5.1 *Let a be a sequence and $\alpha, \beta \in \mathbb{R}$. If the sequence a converges to α and $\alpha > \beta$ then eventually $a_n > \beta$.*

Theorem 5.8 *Suppose that a and b are convergent sequences,*

$$\lim_{\infty} a = \alpha \qquad \text{and} \qquad \lim_{\infty} b = \beta,$$

and there exists an $N \in \mathbb{N}$, such that $a_n \leq b_n$ for all $n > N$. Then $\alpha \leq \beta$.

Proof: By contradiction, using the previous theorem. ■

Comment: If instead $a_n < b_n$ for all $n > N$, then we still only have $\alpha \leq \beta$. Take for example the sequences $a_n = 1/n$ and $b_n = 2/n$. Even though $a_n < b_n$ for all n, both converge to the same limit.

Theorem 5.9 (Sandwich) Suppose that a and b are sequences that converge to the same limit ℓ. Let c be a sequence for which there exists an $N \in \mathbb{N}$ such that

$$a_n \leq c_n \leq b_n \qquad \textit{for all} \qquad n > N.$$

Then

$$\lim_{\infty} c = \ell.$$

Proof: It is given that

$$(\forall \epsilon > 0)(\exists N_1 \in \mathbb{N}) : (\forall n > N_1)(|a_n - \ell| < \epsilon),$$

and

$$(\forall \epsilon > 0)(\exists N_2 \in \mathbb{N}) : (\forall n > N_2)(|b_n - \ell| < \epsilon).$$

Thus,

$$(\forall \epsilon > 0)(\exists N_1, N_2 \in \mathbb{N}) : (\forall n > \max(N_1, N_2))(-\epsilon < a_n - \ell \leq c_n - \ell \leq b_n - \ell < \epsilon),$$

or

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\forall n > N)(|c_n - \ell| < \epsilon).$$

■

Example: Since for all n,

$$1 < \sqrt{1 + 1/n} < \sqrt{1 + 2/n + 1/n^2} = 1 + 1/n,$$

it follows that

$$\lim_{n\to\infty} \sqrt{1 + 1/n} = 1.$$

We can also deduce this from the fact that

$$\lim_{x\to 1} \sqrt{x} = 1 \qquad \text{and} \qquad \lim_{n\to\infty} (1 + 1/n) = 1.$$

▲▲▲

✎ *Exercise 5.3*

① Find sequences a and b, such that a converges to zero, b is unbounded, and ab converges to zero.

② Find sequences a and b, such that a converges to zero, b is unbounded, and ab converges to a limit other then zero.

③ Find sequences a and b, such that a converges to zero, b is unbounded, and ab is bounded but does not converge.

Definition 5.4 *A sequence a is called* ***increasing*** *(עולה) if $a_{n+1} > a_n$ for all n. It is called* ***non-decreasing*** *(לא יורדת) if $a_{n+1} \geq a_n$ for all n. We define similarly a decreasing and a non-increasing sequence.*

Theorem 5.10 (Bounded + monotonic = convergent) Let a be a non-decreasing sequence bounded from above. Then it is convergent.

Proof: Set

$$\alpha = \sup\{a_n : n \in \mathbb{N}\}.$$

(Note that a supremum is a property of a set, i.e., the order in the set does not matter.) By the definition of the supremum,

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\alpha - \epsilon < a_N \leq \alpha).$$

Since the sequence is non-decreasing,

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\forall n > N)(\alpha - \epsilon < a_N \leq a_n \leq \alpha),$$

and in particular,

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\forall n > N)(|a_n - \alpha| < \epsilon).$$

■

Example: Consider the sequence

$$a : n \mapsto \left(1 + \frac{1}{n}\right)^n.$$

We will first show that $a_n < 3$ for all n. Indeed[1],

$$\begin{aligned}\left(1+\frac{1}{n}\right)^n &= 1 + \binom{n}{1}\frac{1}{n} + \binom{n}{2}\frac{1}{n^2} + \cdots + \binom{n}{n}\frac{1}{n^n} \\ &= 1 + 1 + \frac{n(n-1)}{2!\,n^2} + \cdots + \frac{n!}{n!\,n^n} \\ &\le 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{n!} \\ &\le 1 + 1 + \frac{1}{2} + \cdots + \frac{1}{2^{n-1}} < 3.\end{aligned}$$

Second, we claim that this sequence is increasing, as

$$a_n = 1 + 1 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right) + \cdots .$$

When n grows, both the number of elements grows as well as their values. It follows that

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \quad \text{exists}$$

(and equals to $2.718\cdots \equiv e$). ▲▲▲

Definition 5.5 *Let a be a sequence. A* ***subsequence*** *(תת סדרה) of a is any sequence*

$$a_{n_1}, a_{n_2}, \dots,$$

such that

$$n_1 < n_2 < \cdots .$$

More formally, b is a subsequence of a if there exists a monotonically increasing sequence of natural numbers (n_k), such that $b_k = a_{n_k}$ for all $k \in \mathbb{N}$.

Example: Let a by the sequence of natural numbers, namely $a_n = n$. The subsequence b of all even numbers is

$$b_k = a_{2k},$$

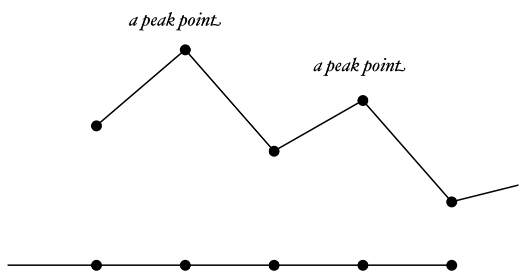
i.e., $n_k = 2k$. ▲▲▲

[1] The binomial formula is

$$(a+b)^n = \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}.$$

Lemma 5.1 *Any sequence contains a subsequence which is either non-decreasing or non-increasing.*

Proof: Let a be a sequence. Let's call a number n a peak point (נקודת שיא) of the sequence a if $a_m < a_n$ for all $m > n$.



There are now two possibilities.

There are infinitely many peak points: If $n_1 < n_2 < \cdots$ are a sequence of peak points, then the subsequence a_{n_k} is decreasing.

There are finitely many peak points: Then let n_1 be greater than all the peak points. Since it is not a peak point, there exists an n_2, such that $a_{n_2} \geq a_{n_1}$. Continuing this way, we obtain a non-decreasing subsequence. ■

Corollary 5.2 (Bolzano-Weierstraß) *Every bounded sequence has a convergent subsequence.*

(50 hrs, 2009)

In many cases, we would like to know whether a sequence is convergent even if we do not know what the limit is. We will now provide such a convergence criterion.

Definition 5.6 *A sequence a is called a* **Cauchy sequence** *if*

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\forall m, n > N)(|a_n - a_m| < \epsilon).$$

Comment: A common notation for the condition satisfied by a Cauchy sequence is

$$\lim_{n,m\to\infty} |a_n - a_m| = 0.$$

(52 hrs, 2011)

Theorem 5.11 A sequence converges if and only if it is a Cauchy sequence.

Proof: One direction is easy[2]. If a sequence a converges to a limit α, then

$$(\forall\epsilon > 0)(\exists N \in \mathbb{N}) : (\forall n > N)(|a_n - \alpha| < \epsilon/2),$$

hence

$$(\forall\epsilon > 0)(\exists N \in \mathbb{N}) : (\forall m, n > N)(|a_n - a_m| \le |a_n - \alpha| + |a_m - \alpha| < \epsilon),$$

i.e., the sequence is a Cauchy sequence.

Suppose now that a is a Cauchy sequence. We first show that the sequence is bounded. Taking $\epsilon = 1$,

$$(\exists N \in \mathbb{N}) : (\forall n > N)(|a_n - a_{N+1}| < 1),$$

which proves that the sequence is bounded.

By the Bolzano-Weierstraß theorem, it follows that a has a converging subsequence. Denote this subsequence by b with $b_k = a_{n_k}$ and its limit by β. We will show that the whole sequence converges to β.

Indeed, by the Cauchy property

$$(\forall\epsilon > 0)(\exists N \in \mathbb{N}) : (\forall m, n > N)(|a_n - a_m| < \epsilon/2),$$

and whereas by the convergence of the sequence b,

$$(\forall N \in \mathbb{N})(\exists K \in \mathbb{N} : n_K > N) : (\forall k > K)(|b_k - \beta| < \epsilon/2).$$

[2]There is something amusing about calling sequences satisfying this property a Cauchy sequence. Cauchy assumed that sequences that get eventually arbitrarily close converge, without being aware that this is something that ought to be proved.

Combining the two,

$$(\forall \epsilon > 0)(\exists N, K \in \mathbb{N} : n_K > N) : (\forall n > N)(|a_n - \beta| \le |a_n - b_k| + |b_k - \beta| < \epsilon).$$

This concludes the proof. ■

(51 hrs, 2009)

(53 hrs, 2011)

5.3 Infinite series

Let a be a sequence. We define the sequence of ***partial sums*** (סכומים חלקיים) of a by

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

Note that S^a does not depend on the order of summation (a finite summation). If the sequence S^a converges, we will interpret its limit as the infinite sum of the sequence.

Definition 5.7 A sequence a is called ***summable*** *(סכימה) if the sequence of partial sums S^a converges. In this case we write*

$$\sum_{k=1}^{\infty} a_k \stackrel{def}{=} \lim_{\infty} S^a.$$

The sequence of partial sum is called a ***series*** *(טור). Often, the term series refers also to the limit of the sequence of partial sums.*

Example: The canonical example of a series is the ***geometric series***. Let a be a geometric sequence

$$a_n = q^n \qquad |q| < 1.$$

The geometric series is the sequence of partial sums, for which we have an explicit expresion:

$$S^a_n = q\frac{1 - q^n}{1 - q}$$

This series converges, hence this geometric sequence is summable and

$$\sum_{n=1}^{\infty} q^n = \frac{q}{1-q}.$$

▲▲▲

Comment: Note the terminology:

sequence is summable = series is convergent.

Theorem 5.12 Let a and b be summable sequences and $\gamma \in \mathbb{R}$. Then the sequences $a+b$ and γa are summable and

$$\sum_{n=1}^{\infty}(a_n + b_n) = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n,$$

and

$$\sum_{n=1}^{\infty}(\gamma a_n) = \gamma \sum_{n=1}^{\infty} a_n.$$

Proof: Very easy. For example,

$$S_n^{a+b} = \sum_{k=1}^{n}(a_k + b_k) = S_n^a + S_n^b,$$

and the rest follows from limits arithmetic. ■

An important question is how to identify summable sequences, even in cases where the sum is not known. This brings us to discuss ***convergence criteria***.

Theorem 5.13 (Cauchy's criterion) The sequence a is summable if and only if for every $\epsilon > 0$ there exists an $N \in \mathbb{N}$, such that

$$|a_{n+1} + a_{n+2} + \cdots + a_m| < \epsilon \qquad \text{for all} \qquad m > n > N.$$

Proof: This is just a restatement of Cauchy's convergence criterion for sequences. We know that the series converges if and only if

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\forall m > n > N)(|S^a_m - S^a_n| < \epsilon),$$

but

$$S^a_m - S^a_n = a_{n+1} + a_{n+2} + \cdots + a_m.$$

■

Corollary 5.3 If a is summable then

$$\lim_{\infty} a = 0.$$

Proof: We apply Cauchy's criterion with $m = n + 1$. That is,

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\forall n > N)(|a_{n+1}| < \epsilon).$$

■

The vanishing of the sequence is only a *necessary* condition for its summability. It is not sufficient as proves the ***harmonic series***, $a_n = 1/n$,

$$S^a_n = 1 + \frac{1}{2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{\geq 1/2} + \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}}_{\geq 1/2} + \cdots .$$

This grouping of terms shows that the sequence of partial sums is unbounded.

For a while we will deal with non-negative sequences. The reason will be made clear later on.

Theorem 5.14 A non-negative sequence is summable if and only if the series (i.e., the sequence of partial sums) is bounded.

Proof: Consider the sequence of partial sums, which is monotonically increasing. Convergence implies boundedness. Monotonicity and boundedness imply convergence. ■

Theorem 5.15 (Comparison test) If $0 \le a_n \le b_n$ for every n and the sequence b is summable, then the sequence a is summable.

Proof: Since b is summable then the corresponding series S^b is bounded, namely.

$$(\exists M > 0) : (\forall n \in \mathbb{N}) \left(0 \le S^b_n \le M\right).$$

The bound M bounds the series S^a, which therefore converges. ■

Example: Consider the following sequence:

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

Since

$$0 < a_n \le \frac{3}{2^n} \stackrel{\text{def}}{=} b_n,$$

and the series associated with b converges, it follows that the series of a converges. ▲▲▲

Example: Consider now the sequence

$$a_n = \frac{1}{2^n - 1 + \sin^3(n^2)}.$$

Here we have

$$0 < \frac{1}{2^n} \le a_n \le \frac{1}{2^n - 2} \le \frac{2}{2^n}, \qquad n \ge 2,$$

so again, based on the comparison test and the convergence of the geometric series we conclude that the series of a converges. ▲▲▲

Example: The next example is

$$a_n = \frac{n+1}{n^2+1}.$$

It is clear that the corresponding series diverges, since a_n "roughly behaves" like $1/n$. This can be shown by the following inequality

$$\frac{1}{n} = \frac{n+1}{n^2+n} \le a_n,$$

which proves that the series S^a is unbounded. The following theorem provides an even easier way to show that. ▲▲▲

(53 hrs, 2009)

Theorem 5.16 (Limit comparison) *Let a and b be positive sequences, such that*

$$\lim_{\infty} \frac{a}{b} = \gamma \neq 0.$$

Then a is summable if and only if b is summable.

Example: Back to the previous example, setting $b_n = 1/n$ we find

$$\lim_{\infty} \frac{a}{b} = \lim_{n\to\infty} \frac{n^2+n}{n^2+1} = 1,$$

hence the series of a diverges. ▲▲▲

Proof: Suppose that b is summable. Since

$$\lim_{\infty} \frac{a}{\gamma b} = 1,$$

it follows that

$$(\exists N \in \mathbb{N}) : (\forall n > N)\left(\frac{a_n}{\gamma b_n} < 2\right),$$

or

$$(\exists N \in \mathbb{N}) : (\forall n > N)\,(a_n < 2\gamma b_n)\,.$$

Now,

$$S^a_n = S^a_N + \sum_{k=N+1}^{n} a_k < S^a_N + 2\gamma \sum_{k=N+1}^{n} b_k \leq S^a_N + 2\gamma S^b_n.$$

Since the right hand sides is uniformly bounded, then a is summable. The roles of a and b can then be interchanged. ■

Theorem 5.17 (Ratio test (מבחן המנה)) *Let $a_n > 0$ and suppose that*

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = r.$$

(i) If $r < 1$ then (a_n) is summable. (ii) If $r > 1$ then (a_n) is not summable.

Proof: Suppose first that $r < 1$. Then,

$$(\exists s : r < s < 1)(\exists N \in \mathbb{N}) : (\forall n > N)\left(\frac{a_{n+1}}{a_n} \le s\right).$$

In particular, $a_{N+1} \le s\, a_N$, $a_{N+2} \le s^2\, a_N$, and so on. Thus,

$$\sum_{k=1}^{n} a_k = \sum_{k=1}^{N} a_k + a_N \sum_{k=N+1}^{n} s^{k-N}.$$

The second term on the right hand side is a partial sum of a convergent geometric series, hence the partial sums of (a_n) are bounded, which proves their convergence.

Conversely, if $r > 1$, then

$$(\exists s : 1 < s < r)(\exists N \in \mathbb{N}) : (\forall n > N)\left(\frac{a_{n+1}}{a_n} \ge s\right),$$

from which follows that $a_n \ge s^{n-N} a_N$ for all $n > N$, i.e., the sequence (a_n) is not even bounded (and certainly not summable). ■

Examples: Apply the ratio test for

$$a_n = \frac{1}{n!}$$
$$b_n = \frac{r^n}{n!}$$
$$c_n = n\, r^n.$$

Comment: What about the case $r = 1$ in the above theorem? This does not provide any information about convergence, as show the two examples[3],

$$\sum_{n=1}^{\infty} \frac{1}{n^2} < \infty \qquad \text{and} \qquad \sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

These examples motivate, however, another comparison test. Note that

$$\lim_{x\to\infty} \int_1^x \frac{dt}{t} = \infty \quad \text{whereas} \quad \lim_{x\to\infty} \int_1^x \frac{dt}{t^2} < \infty.$$

Indeed, there is a close relation between the convergence of the series and the convergence of the corresponding integrals.

[3]This notation means that the sequence $a_n = 1/n^2$ is summable, whereas the sequence $b_n = 1/n$ is not.

Definition 5.8 *Let f be integrable on any segment $[a, b]$ for some a and $b > a$. Then*

$$\int_a^\infty f \stackrel{def}{=} \lim_{x\to\infty} \int_a^x f.$$

Theorem 5.18 (Integral test) Let f be a positive decreasing function on $[1, \infty)$ and set $a_n = f(n)$. Then $\sum_{n=1}^\infty a_n$ converges if and only if $\int_1^\infty f$ exists.

Proof: The proof hinges on showing that

$$\sum_{k=2}^{\lfloor x\rfloor} a_k \le \int_1^x f \le \sum_{k=1}^{\lceil x\rceil} a_k.$$

TO COMPLETE. ■

Example: Consider the series

$$\sum_{n=1}^\infty \frac{1}{n^p}$$

for $p > 0$. We apply the integral text, and consider the integral

$$\int_1^x \frac{dt}{t^p} = \begin{cases} \log x & p = 1 \\ (1-p)^{-1}(x^{-p+1} - 1) & p \ne 1. \end{cases}$$

This integral converges as $x \to \infty$ if and only if $p > 1$, hence so does the series.
▲▲▲

(55 hrs, 2009)

Example: One can take this further and prove that[4]

$$\sum_{n=1}^\infty \frac{1}{n(\log n)^p}$$

[4]Use the fact that

$$F(x) = \frac{(\log x)^{-p+1}}{-p+1} \qquad \text{implies} \qquad F'(x) = \frac{1}{x\,(\log x)^p},$$

and

$$F(x) = \frac{(\log\log x)^{-p+1}}{-p+1} \qquad \text{implies} \qquad F'(x) = \frac{1}{x\,\log x(\log\log x)^p}.$$

converges if and only if $p > 1$, and so does

$$\sum_{n=1}^{\infty} \frac{1}{n \log n (\log \log n)^p},$$

and so on. This is known as ***Bertrand's ladder***. ▲▲▲

We have dealt with series of non-negative sequences. The case of non-positive sequences is the same as we can define $b_n = (-a_n)$. Sequences of non-fixed sign are a different story.

Definition 5.9 *A series $\sum_{n=1}^{\infty} a_n$ is said to be* ***absolutely convergent*** *(מתכנסת בהחלט) if $\sum_{n=1}^{\infty} |a_n|$ converges. A series that converges but does not converge absolutely is called* ***conditionally convergent*** *(מתכנסת בתנאי).*

Example: The series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots,$$

converges conditionally. ▲▲▲

Example: The series

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots,$$

converges conditionally as well. It is a well-known alternating series, which Leibniz showed (without any rigor) to be equal to $\pi/4$. It is known as the ***Leibniz series***. ▲▲▲

Theorem 5.19 *Every absolutely convergent series is convergent. Also, a series is absolutely convergent, if and only if the two series formed by its positive elements and its negative elements both converge.*

Proof: The fact that an absolutely convergent series converges follows from Cauchy's criterion. Indeed, for every $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that

$$|a_{n+1} + \cdots + a_m| \leq |a_{n+1}| + \cdots + |a_m| < \epsilon$$

whenever $m, n > N$.

Define now[5]

$$a_n^+ = \begin{cases} a_n & a_n > 0 \\ 0 & a_n \le 0 \end{cases} \qquad \text{and} \qquad a_n^- = \begin{cases} a_n & a_n < 0 \\ 0 & a_n \ge 0 \end{cases}.$$

Then, $a_n = a_n^+ + a_n^-$, $|a_n| = a_n^+ - a_n^-$, and conversely,

$$a_n^+ = \frac{1}{2}(a_n + |a_n|) \qquad \text{and} \qquad a_n^- = \frac{1}{2}(a_n - |a_n|).$$

Thus,

$$\sum_{k=1}^n |a_k| = \sum_{k=1}^n a_n^+ - \sum_{k=1}^n a_n^-,$$

hence if the two series of positive and negative terms converge so does the series of absolute values. On the other hand,

$$\sum_{k=1}^n a_n^+ = \frac{1}{2}\sum_{k=1}^n a_n + \frac{1}{2}\sum_{k=1}^n |a_n|$$
$$\sum_{k=1}^n a_n^- = \frac{1}{2}\sum_{k=1}^n a_n - \frac{1}{2}\sum_{k=1}^n |a_n|$$

so that the other direction holds as well. ■

(56 hrs, 2009)

Theorem 5.20 (Leibniz) Suppose that a_n is a non-increasing sequence of non-negative numbers, i.e.,

$$a_1 \ge a_2 \ge \cdots \ge 0,$$

and

$$\lim_{n\to\infty} a_n = 0.$$

[5]For example, if $a_n = (-1)^{n+1}/n$, then

$$a_n^+ = 1, 0, \frac{1}{3}, 0, \frac{1}{5}, 0, \cdots$$
$$a_n^- = 0, -\frac{1}{2}, 0, -\frac{1}{4}, 0, -\frac{1}{6}, \cdots.$$

Then the series

$$\sum_{n=1}^{\infty}(-1)^{n+1}a_n = a_1 - a_2 + a_3 - a_4 + \cdots$$

converges.

Proof: Simple considerations show that

$$s_2 \le s_4 \le s_6 \le \cdots \le s_5 \le s_3 \le s_1.$$

Indeed,

$$s_{2n} \le s_{2n} + (a_{2n+1} - a_{2n+2}) = s_{2n+2} = s_{2n+1} - a_{2n+2} \le s_{2n+1}.$$

Thus, the sequence $b_n = s_{2n}$ is non-decreasing and bounded from above, whereas the sequence $c_n = s_{2n+1}$ is non-increasing and bounded from below. It follows that both subsequences converge,

$$\lim_{n\to\infty} b_n = b \qquad \text{and} \qquad \lim_{n\to\infty} c_n = c.$$

By limit arithmetic

$$\lim_{n\to\infty}(c_n - b_n) = \lim_{n\to\infty}(s_{2n+1} - s_{2n}) = \lim_{n\to\infty} a_{2n+1} = c - b,$$

however a_{2n+1} converges to zero, hence $b = c$. It remains to show that if both the odd and even elements of a sequence converge to the same limit then the whole sequence converges to this limit.

Indeed,

$$(\forall \epsilon > 0)(\exists N_1 \in \mathbb{N}) : (\forall n > N_1)(|s_{2n+1} - b| < \epsilon)$$

and

$$(\forall \epsilon > 0)(\exists N_2 \in \mathbb{N}) : (\forall n > N_2)(|s_{2n} - b| < \epsilon),$$

i.e.,

$$(\forall \epsilon > 0)(\exists N_1, N_2 \in \mathbb{N}) : (\forall n > \max(2N_1 + 1, 2N_2))(|s_n - b| < \epsilon).$$

■

Definition 5.10 *A* **rearrangement** *(סידור מחדש) of a sequence a_n is a sequence*

$$b_n = a_{f(n)},$$

where $f : \mathbb{N} \to \mathbb{N}$ is one-to-one and onto.

Theorem 5.21 (Riemann) If $\sum_{n=1}^{\infty} a_n$ is conditionally convergent then for every $\alpha \in \mathbb{R}$ there exists a rearrangement of the series that converges to α.

Proof: I will not give a formal proof, but only a sketch. ■

Example: Consider the series

$$S = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} + \cdots .$$

We proved that this series does indeed converge to some $\frac{1}{2} < S < 1$. By limit arithmetic for sequences (or series),

$$\frac{1}{2}S = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{2n} = \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} + \cdots ,$$

which we can also pad with zeros,

$$\frac{1}{2}S = 0 + \frac{1}{2} + 0 - \frac{1}{4} + 0 + \frac{1}{6} + 0 - \frac{1}{8} + 0 + \frac{1}{10} + \cdots .$$

Using once more sequence arithmetic,

$$\frac{3}{2}S = 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \ldots ,$$

which directly proves that a rearrangement of the alternating harmonic series converges to three halves of its value. In this example the rearrangement is

$$b_{3n+1} = a_{4n+1} \qquad b_{3n+2} = a_{4n+3} \qquad b_{3n} = a_{2n}.$$

▲▲▲

Comment: Here is an important fact about Cauchy sequences: let (a_n) be a Cauchy sequence with limit a. Then,

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N}) : (\forall m, n > N)(|a_n - a_m| < \epsilon).$$

Think now of the left hand side as a sequence with index m with n fixed. As $m \to \infty$ this sequence converges to $|a_n - a|$. This implies that

$$|a_n - a| \leq \epsilon \qquad \text{for all} \qquad n > N.$$

Theorem 5.22 *If $\sum_{n=1}^{\infty} a_n$ converges absolutely then any rearrangement $\sum_{n=1}^{\infty} b_n$ converges absolutely, and*

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} b_n.$$

In other word, the order of summation does not matter in an absolutely convergent series.

Proof: Let $b_n = a_{f(n)}$ and set

$$s_n = \sum_{k=1}^{n} a_k \qquad \text{and} \qquad t_n = \sum_{k=1}^{n} b_k.$$

Let $\epsilon > 0$ be given. Since $\sum_{n=1}^{\infty} a_n$ converges, then there exists an N such that

$$\left| s_N - \sum_{k=1}^{\infty} a_k \right| < \frac{\epsilon}{2}.$$

Moreover, since the sequence is absolutely convergent, we can choose N large enough such that

$$\sum_{k=1}^{\infty} |a_k| - \sum_{k=1}^{N} |a_k| = \sum_{k=N+1}^{\infty} |a_k| < \frac{\epsilon}{2}.$$

Take now M be sufficiently large, so that each of the $a_1, \dots, a_N$ appears among the $b_1, \dots, b_M$. For all $m > M$,

$$|t_m - s_N| = \left| \sum_{k>N, f(k) \leq m} a_k \right| \leq \sum_{k=N+1}^{\infty} |a_k| < \frac{\epsilon}{2},$$

i.e.,

$$\left|\sum_{k=1}^{\infty} a_k - t_m\right| \leq \left|\sum_{k=1}^{\infty} a_k - s_N\right| + |s_N - t_m| \leq \epsilon.$$

This proves that

$$\lim_{m\to\infty} t_m = \sum_{k=1}^{\infty} a_k,$$

i.e.,

$$\sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} a_n.$$

That $\sum_{n=1}^{\infty} b_n$ converges absolutely follows form the fact that $|b_n|$ is a rearrangement of $|a_n|$ and the first part of this proof. ■

Absolute convergence is also necessary in order to multiply two series. Consider the product

$$\left(\sum_{n=1}^{\infty} a_n\right)\left(\sum_{k=1}^{\infty} b_k\right) = (a_1 + a_2 + \cdots)(b_1 + b_2 + \cdots).$$

This product seems to be a sum of all products $a_n b_k$, i.e.,

$$\sum_{k,n=1}^{\infty} (a_n b_k),$$

except that these numbers to be added form a "matrix", i.e., a set that can be ordered in different ways.

Theorem 5.23 Suppose that $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ converge absolutely, and that c_n is any sequence that contains all terms $a_n b_k$. Then

$$\sum_{n=1}^{\infty} c_n = \left(\sum_{n=1}^{\infty} a_n\right)\left(\sum_{k=1}^{\infty} b_k\right).$$

Proof: Consider the sequence

$$s_n = \left(\sum_{k=1}^{\infty} |a_k|\right)\left(\sum_{\ell=1}^{\infty} |b_\ell|\right).$$

This sequence converges as $n \to \infty$ (by limit arithmetic), which implies that it is a Cauchy sequence. For every $\epsilon > 0$ there exists an $N \in \mathbb{N}$, such that

$$\left|\left(\sum_{k=1}^{n} |a_k|\right)\left(\sum_{\ell=1}^{n} |b_\ell|\right) - \left(\sum_{k=1}^{m} |a_k|\right)\left(\sum_{\ell=1}^{m} |b_\ell|\right)\right| < \frac{\epsilon}{2}$$

for all $m, n > N$. By the comment above,

$$\left|\left(\sum_{k=1}^{m} |a_k|\right)\left(\sum_{\ell=1}^{m} |b_\ell|\right) - \left(\sum_{k=1}^{\infty} |a_k|\right)\left(\sum_{\ell=1}^{\infty} |b_\ell|\right)\right| \le \frac{\epsilon}{2}$$

for all $m > N$. Thus,

$$\sum_{k \text{ or } \ell > N} |a_k||b_\ell| \le \frac{\epsilon}{2} < \epsilon.$$

Let now (c_n) be a sequence comprising of all pairs $a_n b_\ell$. There exists an M sufficiently large, such that $m > M$ implies that either k or ℓ is greater than N, where $c_m = a_k b_\ell$. I.e.,

$$\sum_{n=1}^{M} c_n - \left(\sum_{k=1}^{N} a_k\right)\left(\sum_{\ell=1}^{N} b_\ell\right)$$

only consists of terms $a_k b_\ell$ with either k or ℓ greater than N, hence for all $m > M$,

$$\left|\sum_{n=1}^{m} c_n - \left(\sum_{k=1}^{N} a_k\right)\left(\sum_{\ell=1}^{N} b_\ell\right)\right| \le \sum_{k \text{ or } \ell > N} |a_k||b_\ell| < \epsilon.$$

Letting $N \to \infty$,

$$\left|\sum_{n=1}^{m} c_n - \left(\sum_{k=1}^{\infty} a_k\right)\left(\sum_{\ell=1}^{\infty} b_\ell\right)\right| \le \epsilon,$$

for all $m > M$, which proves the desires result. ∎